BIRKHÄUSER

Modern Birkhäuser Classics

Set-Valued Analysis

Jean-Pierre Aubin
Hélène Frankowska

Modern Birkhäuser Classics

Many of the original research and survey monographs in pure and applied mathematics published by Birkhäuser in recent decades have been groundbreaking and have come to be regarded as foundational to the subject. Through the MBC Series, a select number of these modern classics, entirely uncorrected, are being re-released in paperback (and as eBooks) to ensure that these treasures remain accessible to new generations of students, scholars, and researchers.

Set-Valued Analysis

Jean-Pierre Aubin
Hélène Frankowska

Reprint of the 1990 Edition

Birkhäuser
Boston • Basel • Berlin

Jean-Pierre Aubin and Hélène Frankowska
CEREMADE
(Centre de Recherche de Mathématiques de la Décision)
Université de Paris-Dauphine and CNRS
F-75775 Paris Cedex 16
France

Originally published in the series *Systems & Control: Foundations & Applications*

ISBN: 978-0-8176-4847-3 e-ISBN: 978-0-8176-4848-0
DOI: 10.1007/978-0-8176-4848-0

Library of Congress Control Number: 2008939076

Mathematics Subject Classification (1985): 26A27, 26A51, 26E25, 28B20, 28D05, 34A60, 49A52, 54C60, 54C65, 58C06, 58C07, 58C30, 93C15, 93C30

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

www.birkhauser.com

Jean-Pierre Aubin
Hélène Frankowska

Set-Valued Analysis

With 11 Illustrations

1990

Birkhäuser
Boston · Basel · Berlin

Jean-Pierre Aubin and Hélène Frankowska
CEREMADE
(Centre de Recherche de Mathématiques de la Décision)
Université de Paris-Dauphine and CNRS
F-75775 Paris Cedex 16
France

AMS Subject Classification (1985): 26A27, 26A51, 26E25, 28B20, 28D05, 34A60, 49A52, 54C60, 54C65, 58C06, 58C07, 58C30, 93C15, 93C30.

Library of Congress Cataloging-in-Publication Data
Aubin, Jean-Pierre.
Set-valued analysis / Jean-Pierre Aubin, Hélène Frankowska.
p. cm.—(Systems & control ; v. 2)
Includes bibliographical references.
ISBN 0-8176-3478-9 (alk. paper)
1. Set-valued maps. I. Frankowska, Hélène. II. Title.
III. Series.
QA611.3.A93 1990
514′.32—dc20 89-48854

Printed on acid-free paper.

ISBN 0-8176-3478-9
ISBN 3-7643-3478-9

Photocomposed text prepared using the authors' TeX files.
Printed and bound by R.R. Donnelley & Sons, Harrisonburg, Virginia.
Printed in the U.S.A.

9 8 7 6 5 4 3 2 1

THIS BOOK IS DEDICATED TO C. OLECH
AND POLISH MATHEMATICIANS[1]
who contributed so much to set-valued analysis.

[1] "It was a common belief that cultivation of science and the growth of its potential would somehow guarantee the maintenance of the nation" wrote Kuratowski about the situation of Poland before 1918.

I will share all [my results] with you whenever you wish and do so without any ambition, from which I am more exempt and more distant than any man in the world.

Pierre de Fermat[a]

Fermat was one of the most important innovators in the history of mathematics. Newton himself recognized explicitly that he got the hint of the differential calculus from Fermat's method of building tangents devised half a century earlier.

This same method is used in our book to build a differential calculus for set-valued maps.

Fermat was also the one who discovered that the derivative of a (polynomial) function vanishes when it reaches an extremum. (This is Fermat's Rule, which remains the main strategy for obtaining necessary conditions of optimality, from mathematical programming to calculus of variations to optimal control). Fermat also was the first to discover the "principle of least time" in optics, the prototype of the variational principles governing so many physical and mechanical laws. He shared independently with Descartes the invention of analytic geometry and with Pascal the creation of the mathematical theory of probability. His achievements in number theory overshadowed his other contributions, and the Last Fermat Theorem remains a challenge. He was on top of that a poet, a linguist, ... and made his living as a lawyer!

---

[a] in his answer to a first letter from Mersenne inviting him to share his findings with the Parisian mathematicians, which put an end to Fermat's isolation in Toulouse, in 1636.

P. DE FERMAT

Epigraph

Who needs set-valued analysis?

Everyone, we are tempted to say, and we shall state our case.

This strong conviction — born out of accumulated experience in using it in control theory and differential games, mathematical economics and game theory, biomathematics, qualitative physics and viability theory — led us to devote time and effort to share some basic material which is used over and over.

One can no longer afford the luxury of studying only *well posed problems* in Hadamard's sense: *Ill posed problems, inverse problems* and many other unorthodox problems under other names are popping up in every domain of activity, whenever the existence of a solution may fail for some data, whenever uniqueness of the solution is at stake. Requiring that maps should be always single-valued, and even bijective, is too costly an attitude, above all in many applied fields, where we are not free to make such assumptions. This was indeed recognized during the three first decades of this century by the founders of "Functional Calculus": Painlevé, Hausdorff, Bouligand, Kuratowski to quote only a few. In his important book TOPOLOGIE, Kuratowski gave set-valued maps their proper status.

Set-valued maps were abandoned by the authors of Bourbaki's volume TOPOLOGIE GÉNÉRALE, who chose to restrict their study to single-valued maps, regarding set-valued maps as single-valued maps from a set to the power set of another set, or factorizing single-valued maps to make them bijective.

This is not always the solution, for, by so doing, many important structural properties may be unfortunately lost; others are useless artifacts, making life more difficult rather than more simple. These points of view, which were widely disseminated all over the world after World War II, misled many of us into unnecessary detours (often towards culs de sac), encouraging the perception that direct routes were too arduous, or worse, that they did not exist.

Hence, set-valued analysis inherited the undeserved image of being something difficult and mysterious and, consequently, was regarded as a mathematical curiosity, to be left in the hands of mathematicians who like to generalize for the sake of generalizing, without proper motivations.

In contrast, as it turned out, the need for set-valued analysis in solving problems arising in other fields of knowledge — control theory, economics and management, biology and systems sciences, artificial intelligence, etc. — was pressing enough to help mathematicians overcome the kind of recalcitrance felt towards set-valued analysis.

In view of such a wide variety of *motivating applications*, it is fortunate that most of the basic results of the chapters of "single-valued" analysis can be adapted to the set-valued realm. These include:

- Limits and Continuity
- Linear Functional Analysis
- Nonlinear Functional Analysis (existence and approximation of solutions to equations and inclusions)
- Tangents and Normals
- Differentiation of Maps
- Gradients of Functions and the Fermat Rule
- Convergence of Maps
- Measures and Integration
- Differential Equations

The set-valued version of this list is nothing other than the outline of this book.

Our account of set-valued analysis is by no means exhaustive: it is just an introduction. The choice of the material has been dictated by our experience in applying these results in control and viability theory. However, we tried very hard to make this presentation as clear as possible, to help the reader become familiar with the main tools.

We did not restrict our exposition to the framework of finite dimensional vector-spaces, since set-valued analysis is also useful for solving problems involving partial differential equations or inclusions. But whenever the proofs of the finite-dimensional and infinite-dimensional statements are quite different, two proofs are provided.

This book presents only the tools, without mentioning applications. Some applications (to control theory and viability theory in particular) are presented in companion texts[2].

Jean-Pierre Aubin

Hélène Frankowska

Paris
September 21, 1989

[2] VIABILITY THEORY [50] by Aubin, and the forthcoming books SET-VALUED ANALYSIS AND CONTROL THEORY [195] by Frankowska and SET-VALUED ANALYSIS AND SUBDIFFERENTIAL CALCULUS [355] by Rockafellar and Wets.

Acknowledgments

We are indebted to **CEREMADE** (Centre de Recherches de Mathématiques de la Décision), UNIVERSITÉ DE PARIS-DAUPHINE, the SCUOLA NORMALE DI PISA[3] and the **Systems and Decision Sciences**[4] of IIASA (International Institute for Applied Systems Analysis) and their members[5], who provided us with the intellectual environment which made possible the writing of this book[6].

We are also grateful to C. Hess, M. Valadier, O. Dordan, R. Duluc, M. Quincampoix and C. Vigneron for their numerous suggestions for improving the manuscript[7].

We are happy to publish this monograph in the new series *Systems and Control: Foundations and Applications* of Birkhäuser.

[3]and in particular, G. da Prato.

[4]especially, its head, A. Kurzhanski, as well as the dedicated and competent librarians of IIASA.

[5]C. Byrnes, R. Rockafellar and R. Wets among many other colleagues

[6]"...[a mathematician] does need a proper atmosphere; this proper atmosphere can be created only by the cultivation of common topics." Janiszewski wrote in 1918 in *On the Needs of Mathematics in Poland*, a program which in our mind can still be used all over the world.

[7]Finally, as it is customary, each author is grateful to her/his coauthor for his/her wonderful typing of the manuscript.

Contents

List of Figures

List of Tables

Introduction

It is a fact that in mathematical sciences there has been a reluctance to deal with sequences of sets and set-valued maps. Despite the emergence of exciting new vistas for the applications of mathematics, our long familiarity with sequences (of elements) and with (single-valued) maps has perhaps been so deeply rooted in traditional mathematical conceptualizations that it has appeared easier to sacrifice the breadth of some problems or some simple underlying structure in order to avoid set-valued maps.

For this reason, we begin by providing examples of natural and/or general problems involving set-valued maps before giving a rough description of the results presented in the pages that follow.

Examples of Set-Valued Maps

1. First, we encounter set-valued maps each time we face *ill-posed problems* or *inverse problems*, i.e., problems for which either the existence of a solution or its uniqueness is not guaranteed for some data: Set-valued maps allow us to get away from the restriction that a map is bijective when we want to solve an equation.

   Indeed, the first natural instance when set-valued maps occur is the inverse f^{-1} of a single-valued map f from X to Y. We always can define f^{-1} as a set-valued map which associates with any data y the (possibly empty) set of solutions

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

   to the equation $f(x) = y$.

   Of the three commandments of Hadamard's tablets, *existence, uniqueness and stability*, we shall only retain the stability requirements, which can be encapsulated in adequate definitions of *continuity* of f^{-1}: This is one of the topics of the first chapter.

2. Taking into account uncertainties, disturbances, modeling errors, etc., leads naturally to set-valued maps and inclusions. They also arise when we wish to treat a problem *qualitatively*, by looking for solutions common to a set of data, sharing the same (qualitative) properties. Set-valued analysis should play an important role in the new field of *qualitative physics*, a rapidly growing branch of Artificial Intelligence.

3. Problems with constraints also yield specific set-valued maps: Solving the equation $f(x) = y$, where the solution x is required to belong to a subset K, amounts to solving the inclusion $f|_K(x) = y$ where $f|_K$, the restriction of f to K, is regarded as the set-valued map associating with x the point $f(x)$ when $x \in K$ and the empty set when x is not in K.

4. Unilateral problems in mechanics were formulated in the framework of variational inequalities (also called "generalized equations" by some authors), which are again inclusions in disguise. Their solution by Stampacchia and J.-L. Lions in the sixties gave a new impetus to set-valued maps, with a different vocabulary.

5. Set-valued maps provide a useful framework for control theory, since the early contributions of Ważewski and Filippov in the beginning of the sixties.

   Such set-valued maps, called *parametrized maps*, are associated with a family of maps $x \mapsto f(x,u)$ from X to Y when u ranges over a set $U(x)$ of parameters.

   The (single-valued) map f describes the dynamics of the system: It associates with the state x of the system and the control u the velocity $f(x,u)$ of the system. The set-valued map U describes a *feedback map* assigning to the state x the subset $U(x)$ of admissible controls.

   Hence the map F which associates with each state x the subset $F(x)$ of feasible velocities is defined by:

$$F(x) := f(x, U(x)) = \{f(x,u)\}_{u \in U(x)}$$

So, the control system governed by the family of parametrized differential equations

$$x'(t) = f(x(t), u(t)) \text{ where } u(t) \in U(x(t))$$

is actually governed by the *differential inclusion*

$$x'(t) \in F(x(t))$$

6. Optimization provides examples of problems where uniqueness of the solution is naturally lacking:

   Let W be a function from $X \times Y$ to $\mathbf{R}$. We consider the family of minimization problems

   $$\forall\, y \in Y,\ \ V(y) := \inf_{x \in X} W(x, y)$$

   parametrized by parameters y.

   The function V is called the *marginal* (or *performance* or *value*) function. For every $y \in Y$, let

   $$G(y) := \{x \in X \mid W(x, y) = V(y)\}$$

   be the subset of solutions to our minimization problem.

   One of the main issues of optimization theory is to study the set-valued map G (nonvacuity, continuity and differentiability in a suitable sense, and so on.) We shall call G the *marginal map*. It is no wonder that game theory and mathematical economics use set-valued maps in a natural way.

7. Another source of strong motivations came from optimization and mathematical programming, when necessary conditions (the Fermat rule[1], stating that the derivative of a function vanishes at points where it achieves an extremum) were needed to replace optimization problems by the resolution of equations.

[1] "Je désire seulement qu'il [Descartes] sache que nos questions *de Maximis et Minimis* et *de Tangentibus linearum curvarum* sont parfaites depuis huit ou dix ans et que plusieurs personnes qui les ont vues depuis cinq ou six ans le peuvent témoigner", Fermat wrote in 1638 when Descartes accused that the *Methodus de Maxima et Minima* was just due to luck and trial and error! ("à tâtons et par rencontre".)

The Fermat rule is indeed one idea which was revisited and enhanced again and again under different names:

— The Euler-Lagrange equations, when dealing with problems of the calculus of variations;

— Lagrange and Kuhn-Tucker multipliers, when *state constraints* were added to optimization problems;

— The Pontriagin principle when dealing with optimal control problems.

After the advances of Functional Analysis, it was time to uncover the common fact behind all these results. It is still and always the *Fermat rule*, provided we are able to "differentiate" larger and larger classes of functions beyond differentiable functions.

The crucial revolution in the history of the concept of gradients happened in the sixties when J.- J. Moreau and R. T. Rockafellar proposed in the framework of convex analysis the notion of *subdifferential* of a convex function, *which is no longer an element, but a set of "subgradients"*.

8. The use of set-valued maps in mathematical economics and game theory started when von Neumann asked for an extension of the Brouwer Fixed Point Theorem to set-valued maps, which was needed for finding noncooperative equilibria for n-person games, for instance. This was achieved with the famous Kakutani Fixed-Point Theorem, in the forties. It has been used by Arrow and Debreu in the early fifties to provide the long-expected proof of the existence of a Walrasian equilibrium price.

   While this achievement made set-valued maps popular among mathematical economists, it was not until the challenges raised by optimization, control theory and unilateral problems in mechanics at the beginning of the sixties that renewed motivations arose to study set-valued maps, as an important subject in its own rite.

   This was the time when Zarantonello introduced *monotone maps*, which cover many important nonlinear single-valued or set-valued maps of the Calculus of Variations.

Properties of Set-Valued Maps

Having briefly indicated the importance of set-valued maps in a wide spectrum of applications and of fundamental mathematics, we paint a broad picture of their properties. We follow, in doing so, the outline of the book.

- LIMITS AND CONTINUITY

  Limits of sets were introduced by Painlevé in the first years of this century, just after Fréchet axiomatized in 1906 the concept of $\mathcal{L}$-spaces (on which a notion of limit is defined[2].) Studying limits of sets together with limits of elements may have been very natural in this context.

  The topological ideas are, indeed, quite simple and straightforward. In the same way that topological concepts are based on the notions of limits and cluster points of sequences of elements, their set-valued analogues are rooted in the concepts of *lower and upper limits* of sequences of sets, which are, so to speak, *"thick" limits and cluster points respectively*: The lower limit of a sequence of subsets K_n is the set of limits of sequences of elements $x_n \in K_n$, and the upper limit is the set of cluster points of such sequences.

  We mentioned already that *stability* is the only requirement that we retain to study *ill posed or inverse problems*. *Stability* is a catch word which means that the set of solutions depends continuously upon the data.

  How can we proceed to define continuity of set-valued maps?

  If we try to adapt to the set-valued case the two equivalent definitions of continuity of single-valued maps, we obtain two notions which are no longer equivalent!

[2]In his famous thesis, judged at the time far too much abstract; it was published in the Rendiconti Cir. Mat. di Palermo!

Lebesgue wrote: "....set theory was placed outside the pale of mathematics by the high priests of analytic functions... Set theory, which developed from the theory of analytical functions, could prove useful to its elder sister and could show people of good will its qualities and richness".

This unfortunate situation led to two concepts of *semicontinuity* of set-valued maps, introduced at the beginning of the thirties by Bouligand and Kuratowski: Lower and upper semicontinuity. These issues were developed in the monographs of Hausdorff and Kuratowski and thoroughly investigated at this time by many (but not exclusively) Polish and French mathematicians.

For some reason, just a while later, set-valued maps yielded the way to single-valued maps: A set-valued map was viewed at the time as a single-valued map from a set to the power set of another set. However, as it turned out, the structures exported to power sets were too poor, and specific information was indeed wasted by doing so.

For instance, when we regard a set-valued map as a single-valued map from one set to the power set of the other (supplied with any one of the topologies we can think of), we arrive at continuity concepts which are stronger than both lower and upper semicontinuity, introducing parasitic artifacts. For example, using such topologies to differentiate set-valued maps, leads to such strong requirements, that most set-valued maps would become nondifferentiable.

This is the reason why we shall start this book with the study of limits and leave aside the examination of topologies on power sets.

Furthermore, we shall renew history, by regarding a map not ... as a map, but as a graph (a subset of the product of the departure and the arrival sets), reestablishing some symmetry by putting these two sets on the same footing. This brings us back to the source of analytical geometry, at the time of F. Viète, P. de Fermat and R. Descartes, before the concept of function and map evolved from the one of curves and graphs.

To regard a map as a graph is our constant and basic point of view throughout this book (which has been called the *graphical approach.*)

For instance, *closed maps*, that is maps with closed graph, shall play a starring role in this book. It is a weaker property than continuity or even, upper semicontinuity, very familiar and thus easy to check, common to both set-valued maps and their inverses.

- LINEAR FUNCTIONAL ANALYSIS

  What is the set-valued version of a continuous linear operator?

  Remembering that the graph of a continuous linear operator is a closed vector subspace, we are tempted to single out the maps whose graphs are closed linear subspaces (called *linear processes*.)

  This generalization is not bold enough, since dealing exclusively with closed vector subspaces is still too restrictive: We need to use the notion of closed convex cone, which is a kind of vector subspace in which it is forbidden to use subtraction. These cones enjoy many properties of the vector subspaces.

  For this reason, we select the *closed convex processes*, i.e., the maps whose graphs are closed convex cones, as the candidates to play the part of set-valued linear maps.

  We shall see later that derivatives of some set-valued maps are closed convex processes, which is a desirable property for a derivative[3]. Indeed, the two basic theorems on continuous linear operators due to Banach, the Closed Graph Theorem (equivalent to the Open Mapping Principle) and the Banach-Steinhaus Theorem, can be adapted to closed convex processes. The first one states that a closed convex process defined on the whole space is continuous, and the second states that pointwise bounded families of closed convex processes are bounded — a prerequisite for studying the convergence of closed convex processes. But most important of all, one can *transpose* closed convex processes and use the benefits of duality theory, based on the Bipolar Theorem.

  Closed convex processes also possess eigenvectors and invariant spaces and cones. They truly deserve the status of linear set-valued maps.

[3]keeping us in line with Hadamard's linearity bill enforced by Fréchet that a derivative should be linear with respect to the increment.

- NONLINEAR FUNCTIONAL ANALYSIS

  We are convinced that many problems can be regarded as *inclusions*

  $$\text{given } F : X \rightsquigarrow Y \text{ and } y \in Y, \text{ find } x \in X \text{ such that } F(x) \ni y$$

  Most theorems on existence of solutions to nonlinear equations can be extended to the case of inclusions.

  For example, this is the case for the Brouwer Fixed Point Theorem, whose generalization to set-valued maps is the famous Kakutani Fixed-Point Theorem. We shall prove an equivalent statement, called the *Equilibrium Theorem*, which provides the existence of an equilibrium of a set-valued map, a solution to the inclusion $F(x) \ni 0$.

  Of course, for applications, we need not only to solve such a problem, but also to approximate its solutions by solutions to approximate problems:

  $$\text{given } F_n : X_n \rightsquigarrow Y_n,\ y_n \in Y_n, \text{ find } x_n \in X_n \text{ with } F_n(x_n) \ni y_n$$

  where X_n and Y_n are subspaces of X and Y.

  The famous Lax's principle of numerical analysis states that *convergence* of the data y_n to y, *consistency* of F_n to F and *stability* of the F_n's imply the convergence of approximate solutions. As an important special case, when the spaces $X_n = X$, $Y_n = Y$ and the maps $F_n = F$ are constant, this principle boils down to the statement of the Inverse-Function Theorem.

  Convergence, consistency and stability, which were originally defined in the framework of linear equations, can be extended to this general case by introducing adequate notions of *convergence* and *derivatives* of set-valued maps. For instance, the concept of consistency is nothing other than the fact that the graph of F is the lower limit of the graphs of the approximate maps F_n, while *stability is the boundedness of the inverses of the derivatives of the maps F_n.*

  This provides a first motivation for devising a set-valued differential calculus. For that purpose, we need to begin with the study of tangents.

• Tangents and Normals

The concept of tangency has been overshadowed in some sense by the requirement that the space of tangent vectors must be a vector space, so that the original idea became concealed after its formal implementation in differential geometry.

If we come back to the idea underlying the notion of tangency to a subset K at some point $x \in K$, we are tempted to form *"thick" differential quotients*

$$\frac{K-x}{h}$$

and to take (in various ways) their limits when $h > 0$ goes to 0.

We obtain in this way a variety of closed cones made of what we call tangent vectors. The most popular of these tangent cones is for the time the *contingent cone* introduced in the thirties by Bouligand, (which is the upper limit of these differential quotients.) Some of these tangent cones are closed convex cones, and they enjoy a property which is the natural extension of linearity (without subtraction.)

These tangent cones possess a rich calculus which justifies their use in many questions, mainly in problems *with state constraints*: They are involved in the sufficient conditions for the existence of an equilibrium and for the stability of solutions to equations with constraints. They also appear in the formulation of necessary conditions in optimization problems with constraints and play a key role in viability theory.

In order to define space of normals, which in differential geometry consists of vectors othogonal to the tangent vector space, we are led to introduce the dual concept of *normal cones* to any subset.

• Differentiation of Maps

We already mentioned that the concept of stability in the Inverse-Function Theorem requires the notion of derivative of a set-valued map, leading to the question: *How do we formulate this concept?*

The idea is very simple and goes back to the prehistory of the differential calculus, when Pierre de Fermat introduced in the first half of the seventeenth century the concept of tangent to the graph of a function: *The tangent space to the graph of a function f at a point (x, y) of its graph is the line of slope $f'(x)$, i.e., the graph of the linear function $u \mapsto f'(x)u$.*

It is possible to implement this idea *for any set-valued map F*, since we have introduced a way to implement the tangency for any subset of a normed space. Therefore, in the framework of a given problem, *we can regard a tangent cone to the graph of the set-valued map F at some point (x, y) of its graph as the graph of the associated "derivative" of F at this point (x, y).*

Derivatives built in this way from the various choices of tangent cones are called *graphical derivatives* and the calculus of tangent cones can be transferred to a set-valued differential calculus, including chain rules.

With such derivatives of set-valued maps in our hands, we can *linearize* set-valued problems for approximating them by linearized ones. The latter involving closed convex processes, this strategy provides ways for transferring some properties of linear set-valued maps to nonlinear maps.

- GRADIENTS OF FUNCTIONS AND THE FERMAT RULE

The particular case of real-valued functions deserves a study by itself for taking into account the order relation of real numbers. We are led to do so whenever we look for a minimizer of a function or when we study the monotone behavior of a function along a solution to a differential equation or inclusion (Lyapunov property.)

The set-valued approach indicates the route: We associate with a function V the set-valued map $\mathbf{V}_\uparrow$ defined by

$$\mathbf{V}_\uparrow(x) \;:=\; [V(x), +\infty[$$

whose graph is the *epigraph of V*.

The graphs of the derivatives of such set-valued maps $\mathbf{V}_\uparrow$ are the epigraphs of functions which are called *epiderivatives*. We discover that they are close relatives of the directional derivatives introduced by Dini, who was among the first to revolt against the rigidity imposed by the heirs of Cauchy[4].

As far as optimization is concerned, the Fermat Rule can be extended to any function by using these epiderivatives. Since they enjoy a rich calculus, we obtain in this way many necessary conditions for a minimum. This can be done by transferring the set-valued differential calculus to what can be called an *epidifferential calculus*.

By duality, we associate with each of the epiderivatives a concept of *generalized gradient*: It is in general a subset of elements, reduced to the usual gradient whenever the function is differentiable in the usual way. In this framework, the Fermat Rule becomes: *If a point achieves the minimum of a function, then it is an equilibrium of the generalized gradient*, i.e., the generalized gradient at an optimal point contains 0.

- CONVERGENCE OF MAPS

  What about the convergence of a sequence of set-valued maps F_n?

  The first idea which comes to mind is to extend the various notions of uniform convergence of single-valued maps, regarded as a map from one space to another.

  Since we know how to deal with limits of sets, it is again natural to use the *graphical approach* and to study the upper and lower limits of graphs.

  We follow this hint and study *graphical upper and lower limits* of a sequence of set-valued maps as the maps whose graphs are the upper and lower limits of the graphs.

[4]Cauchy, however, had the merit to formalize the concept of limits, continuity and differentiability. His definitions have been canonized ever since: A function was allowed to be differentiated only if the differential quotients were converging to the derivative for the pointwise convergence topology. The need to use nondifferentiable functions has been felt several times. By Bouligand, with the notions of contingent and paratingent, by Dini who also broke Hadamard's linearity law, by L. Schwartz and S. Sobolev, with the discovery of weak derivatives of functions and distributions. But each of these extensions was devised for specific purposes (solving partial differential equations, for instance.)

When we deal with real-valued functions, we are led to use the set-valued maps $\mathbf{V}_{n\uparrow}$ associated with functions V_n, whose graphs are epigraphs of the functions V_n. The graphical limits of the set-valued maps $\mathbf{V}_{n\uparrow}$ induce what we call *epigraphical limits* of the functions V_n. This concept is closely related to *G*-convergence introduced by de Giorgi and has been extensively used in the study of stability of optimization problems[5].

An interesting question arises: What are the connections between the (epigraphical) convergence of a sequence of functions and the (graphical) convergence of their gradients? We shall answer such questions.

- MEASURES AND INTEGRATION

We encounter measurable maps whenever we deal with models of systems having measurable data, and in particular when we deal with random set-valued variables (an issue we shall not address in this book.)

Another important instance where measurable set-valued maps do arise is in the linearization of a control system (or a differential inclusion) along a solution.

Hence, we cannot escape the burden of studying measurable maps, which are the maps whose graphs are measurable, and checking in particular that all the standard operations preserve measurability.

We also need measurability for defining integrals of set-valued maps. Integrals of set-valued maps are involved in many convexification (also called relaxation) problems, since roughly speaking *the integral of a measurable set-valued map is always convex*.

This property was in fact the original motivation to introduce the integral of set-valued maps in mathematical economics and game theory (with a continuum of players.)

We shall also address the basic questions of ergodic theory, extending to set-valued maps the Poincaré Recurrence Theorem

[5]See the forthcoming book SET-VALUED ANALYSIS AND SUBDIFFERENTIAL CALCULUS [355] by Rockafellar and Wets.

and the existence of invariant measures.

- DIFFERENTIAL INCLUSIONS

  Control Theory on one hand, and the evolution of macro-systems under uncertainty on the other hand, constitute very strong motivations for extending differential and partial differential equations to differential and partial differential inclusions.

  We shall provide only an introductory survey of some basic results, since covering this material requires books by themselves[6].

  We shall state some existence theorems, show that the set of solutions depends continuously upon the initial data, describe some properties of Lyapunov functions (which, thanks to the concept of epiderivatives, can even be taken lower semicontinuous), relate the graphical derivatives of the solution map to the solution map of variational inclusions (which are linearizations of the differential inclusion along a solution) and state some applications of the Viability Theorem.

- SELECTIONS AND PARAMETRIZATION

  We cannot escape *in fine* answering two natural questions: Can we find selections of set-valued maps inheriting their regularity properties? Are set-valued maps *parametrizable*?

  We shall be able to show that measurable set-valued maps do have measurable selections and that continuous (Carathéodory, Lipschitz) maps do have continuous (Carathéodory, Lipschitz) selections under severe restrictions: The images of the set-valued map must be convex.

  Actually, a Lipschitz set-valued map F with closed convex images is parametrizable in the sense that there exists a "control space" U and a Lipschitz map $f : X \times U \mapsto X$ such that

$$\forall\, x, \;\; F(x) \;=\; \{f(x,u)\}_{u\in U}$$

[6] See for instances the books DIFFERENTIAL INCLUSIONS [33] by Aubin & Cellina, VIABILITY THEORY [50] by Aubin and SET-VALUED ANALYSIS AND CONTROL THEORY [195] by Frankowska.

Therefore, this limited class of set-valued maps is in essence made of families of single-valued maps.

Chapter 1

Continuity of Set-Valued Maps

Introduction

This chapter is devoted to the elementary topological properties of sequences of sets and set-valued maps.

We begin the first section by extending concepts of limits and cluster points of sequences of elements to sequences of sets. These set-valued analogues have been introduced by Painlevé in 1902 under the names of *upper* and *lower limits of sets*. They were then popularized by Kuratowski in his famous and influential book TOPOLOGIE and thus, often Christianized *Kuratowski lower and upper limits* of sequences of sets[1]. We shall call them simply *lower and upper limits*:

The lower limit *of a sequence of subsets K_n is the set of* limits *of sequences of elements $x_n \in K_n$ and the* upper limit *is the set of* cluster *points of such sequences.*

Some elementary properties of lower and upper limits are investigated in Section 1, whereas their calculus is presented in Section 2.

Set-valued maps can be regarded as "continuous sequences." But before adapting the above concepts to this case in Section 4, we gather in Section 3 the main definitions concerning set-valued maps.

[1]Had they been called *Painlevé* instead of *Kuratowski*, a easier name to pronounce indeed, these concepts could have been more popular.... This at least is the opinion of one of the authors.

Upper and lower limits of set-valued maps are naturally closely related to continuity concepts. Recall that for single-valued maps, continuity amounts to mapping converging sequences to converging sequences or, equivalently, to the celebrated incantation:

$\forall\, \varepsilon > 0,\ \ \exists\, \delta > 0, \ldots$.

Unfortunately, this characterization is no longer true for set-valued maps:

There are *two distinct ways* to extend the concept of continuity: The first one, encompassing the above "convergence" idea, is called *lower semicontinuity*. The second, extending the "$\forall\, \varepsilon$, $\exists\, \delta$ — definition," leads to the so-called *upper semicontinuity* of set-valued maps[2]. Kuratowski showed however that semicontinuous maps are generically continuous.

Lower semicontinuous maps are closely related to lower limits, but the situation is further complicated by the fact that upper limits characterize the maps with closed graph and are not always upper semicontinuous maps. These subtleties are explained in Section 4.

It happens that set-valued maps with closed graph taking their values in a compact set are upper semicontinuous. Hence we have an easy way to verify upper semicontinuity of a map.

Lower semicontinuity is not as simple to check: the last section is devoted to several useful lower semicontinuity criteria.

The definitions introduced in this chapter are basic and will be used throughout the book.

1.1 Limits of Sets

1.1.1 Definitions

Let X be a metric space supplied with a distance d. When K is a subset of X, we denote by

$$d_K(x) \ := \ d(x, K) \ := \ \inf_{y \in K} d(x, y)$$

[2]These concepts should not be confused with the upper and lower semicontinuity of real-valued functions.

the *distance from x to K,* where we set $d(x,\emptyset) := +\infty$. The *ball of radius $r>0$ around K in X* is denoted by

$$B_X(K,r) \ := \ \{x \in X \mid d(x,K) \le r\}$$

When there is no ambiguity, we set

$$B(K,r) \ := \ B_X(K,r)$$

When X is a Banach space whose *unit ball* is denoted by B (or B_X if the space must be mentioned), we observe that

$$B_X(K,r) \ = \ \overline{K + rB_X}$$

The balls $B(K,r)$ are neighborhoods of K. When K is compact, each neighborhood of K contains such a ball around K.

Limits of sets have been introduced by Painlevé[3] in 1902, as it is reported by his student Zoretti[4]. They have been popularized by Kuratowski in his famous book TOPOLOGIE and thus, often called *Kuratowski lower and upper limits* of sequences of sets.

Definition 1.1.1 *Let $(K_n)_{n\in\mathbf{N}}$ be a sequence of subsets of a metric space X. We say that the subset*

$$\mathrm{Limsup}_{n\to\infty} K_n \ := \ \left\{x \in X \mid \liminf_{n\to\infty} d(x,K_n) = 0\right\}$$

is the upper limit *of the sequence K_n and that the subset*

$$\mathrm{Liminf}_{n\to\infty} K_n \ := \ \{x \in X \mid \lim_{n\to\infty} d(x,K_n) = 0\}$$

is its lower limit. *A subset K is said to be the* limit *or the* set limit *of the sequence K_n if*

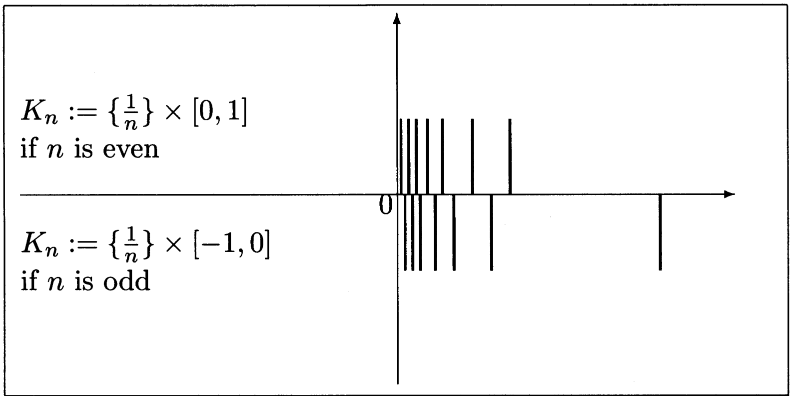
$$K \ = \ \mathrm{Liminf}_{n\to\infty} K_n \ = \ \mathrm{Limsup}_{n\to\infty} K_n \ =: \ \mathrm{Lim}_{n\to\infty} K_n$$

---

[3] "Je dirai qu'un point a appartient à *l'ensemble limite* de E_α si, quelque petits que soient les deux nombres r et ε, sous la condition $|\alpha - \alpha_0| \le \varepsilon$, le cercle de centre a et de rayon r renferme des points appartenant à certains des E_α."

[4] who wrote: "Je ne veux pas terminer ce travail sans adresser tous mes remerciements à tous mes maîtres de l'Ecole Normale dont la sollicitude à mon égard ne s'est jamais ralentie, et dont les encouragements m'ont été si précieux. D'ailleurs, dans ce milieu normalien, tout ce qui vous entoure est un enseignement, ..."

Figure 1.1: Example of Upper and Lower Limits of Sets



$\text{Liminf}_{n\to\infty} K_n = \{0\}$ & $\text{Limsup}_{n\to\infty} K_n = [-1,+1]$

Lower and upper limits are obviously *closed.* We also see at once that

$$\text{Liminf}_{n\to\infty} K_n \subset \text{Limsup}_{n\to\infty} K_n$$

and that the upper limits and lower limits of the subsets K_n and of their closures $\overline{K_n}$ do coincide, since $d(x, K_n) = d(x, \overline{K_n})$.

Any decreasing sequence of subsets K_n has a limit, which is the intersection of their closures:

$$\text{if } K_n \subset K_m \text{ when } n \geq m, \text{ then } \text{Lim}_{n\to\infty} K_n = \bigcap_{n\geq 0} \overline{K_n}$$

An upper limit may be empty (no subsequence of elements $x_n \in K_n$ has a cluster point.)

Concerning sequences of singleta $\{x_n\}$, the set limit, when it exists, is either empty (the sequence of elements x_n is not converging), or is a singleton made of the limit of the sequence.

It is easy to check that:

Proposition 1.1.2 *If* $(K_n)_{n\in\mathbf{N}}$ *is a sequence of subsets of a metric space, then* $\text{Liminf}_{n\to\infty} K_n$ *is the* set of limits of sequences $x_n \in K_n$ *and* $\text{Limsup}_{n\to\infty} K_n$ *is the set* of cluster points of sequences $x_n \in K_n$, *i.e., of limits of subsequences $x_{n'} \in K_{n'}$. The upper limit is also equal to the subset of cluster points of "approximate" sequences satisfying:*

$$\forall\, \varepsilon > 0, \;\; \exists\, N(\varepsilon) \quad \textit{such that}\ \forall\, n > N(\varepsilon), \;\; x_n \;\in\; B(K_n, \varepsilon)$$

Remark — Replacing the balls of a metric space by neighborhoods and the sequences of a metric space by generalized sequences, we can extend the concepts of upper and lower limits to generalized sequences of subsets of a topological space X.

We recall that a set $\mathcal{M}$ supplied with a preorder $\succeq$ is *directed* if every finite subset has an upper bound. A *generalized sequence* is a map

$$\mu \;\in\; \mathcal{M} \;\mapsto\; x_\mu \;\in\; X$$

An element $x \in X$ is the *limit* of $(x_\mu)_{\mu \in \mathcal{M}}$ if, for every neighborhood $\mathcal{V}$ of x, there exists $\mu_0 \in \mathcal{M}$ such that x_μ belongs to $\mathcal{V}$ for all $\mu \succeq \mu_0$. An element x is said to be a *cluster point* of this generalized sequence if, for every neighborhood $\mathcal{V}$ of x and every $\mu \in \mathcal{M}$, there exists $\nu \succeq \mu$ such that $x_\nu \in \mathcal{V}$.

We recall that a single-valued map f from a topological space X to another Y is continuous at x if it maps any generalized sequence converging to x to a generalized sequence converging to $f(x)$ and that a subset $K \subset X$ is compact if and only if every generalized sequence of elements of K has a cluster point. (See for instance [148, Section 1.7].) The main difference with sequences of a metric space is that a converging generalized sequence is not necessarily bounded. (We can also replace generalized sequences by "filters.")

The upper limit of a generalized sequence of subsets K_μ where μ ranges over a directed subset $\mathcal{M}$ is the set of cluster points of generalized sequences x_μ of elements of K_μ and the lower limit is the set of limits of such generalized sequences.

We may need the above extension when dealing with weak topologies of a Banach space X and of its dual denoted by $X^\star$. We say that the bilinear map $< \cdot, \cdot >$

$$(p, x) \in X^\star \times X \mapsto < p, x > := \; p(x)$$

is the *duality pairing*.

We recall that the *weakened topology* $\sigma(X, X^\star)$ of X is defined by the semi-norms

$$\mathbf{p}_M(x) \;:=\; \sup_{q \in M} | < q, x > |$$

when $M := \{q_1, \ldots, q_l\}$ ranges over the finite subsets of $X^\star$. The *weak $-\star$ topology* $\sigma(X^\star, X)$ of the dual $X^\star$ is defined by the semi-norms

$$\mathbf{p}_N(q) \;:=\; \sup_{y \in N} | < q, y > |$$

when $N := \{x_1, \ldots, x_n\}$ ranges over the finite subsets of X.

In other words, a generalized sequence of elements $x_\mu \in X$ converges weakly to $x \in X$ if and only if for any $q \in X^\star$, $< q, x_\mu >$ converges to $< q, x >$ and a sequence $p_\mu \in X^\star$ converges weakly to p if and only if for any $y \in X$, $< p_\mu, y >$ converges to $< p, y >$.

The key facts that can be recalled now are that $X^\star$ *is still the dual of* X *supplied with the weakened topology*, that X *is the dual of* $X^\star$ *supplied with the weak-*$\star$ *topology*, and that *the bounded subsets of the dual* $X^\star$ *are weakly relatively compact*.

However, in general X is a closed subspace of the *bidual* $X^{\star\star}$ of X, which is the dual of the Banach space $X^\star$, endowed with the *norm*

$$\|p\|_\star := \sup_{\|x\| \leq 1} | < p, x > |$$

The space X is called *reflexive* if $X = X^{\star\star}$. In this case, it enjoys both the properties of a Banach space and of the dual of a Banach space, including the weak compactness of the unit ball. □

We could for instance use the fact that *the weakly compact subsets of the dual* $X^\star$ *of a* separable[5] *Banach space are metrizable (for the weak-*$\star$ *topology) (see [148, Theorem 5.6.3])* to avoid generalized sequences.

Instead, for the sake of simplicity, we shall use the concept of *sequentially weak* upper limit:

Definition 1.1.3 *Let us consider a sequence of subsets* K_n *of the dual of a Banach space. We shall say that the subset*

$$\sigma - \mathrm{Limsup}_{n \to \infty} K_n$$

of weak-$\star$ *limits of subsequences of elements* $x_n \in K_n$ *is the* sequentially weak upper limit *of the subsets* K_n.

In this way, we can present lower and upper limits in the framework of metric spaces or of (countable) sequences, leaving to the interested reader the task of checking some extensions to the case of non metrizable topological spaces.

[5]This means that there is countable basis of open subsets, or, equivalently, that there exists a countable basis spanning a dense vector space. The separability concept goes back to Fréchet.

We also point out the quite impressive equivalent formulation of upper and lower limits which follows from Proposition 1.1.2:

$$\mathrm{Limsup}_{n\to\infty} K_n = \bigcap_{N>0} \overline{\bigcup_{n\geq N} K_n} = \bigcap_{\varepsilon>0} \bigcap_{N>0} \bigcup_{n\geq N} B(K_n, \varepsilon)$$

and

$$\mathrm{Liminf}_{n\to\infty} K_n = \bigcap_{\varepsilon>0} \bigcup_{N>0} \bigcap_{n\geq N} B(K_n, \varepsilon)$$

and single out the following property of upper limits:

Theorem 1.1.4 *Let K be a subset of a metric space X satisfying the following property:*

$$\text{for any neighborhood } \mathcal{U} \text{ of } K,\ \exists\, N \text{ such that } \forall\, n \geq N,\ \ K_n \subset \mathcal{U}$$

Then $\mathrm{Limsup}_{n\longrightarrow\infty} K_n \subset \overline{K}$.

Conversely, if X is compact*, then the upper limit* $\mathrm{Limsup}_{n\to\infty} K_n$ *enjoys the above property (and thus, is the* smallest closed subset satisfying it.*)*

Proof — The first statement is obvious. The second one is a consequence of the following more general result:

Proposition 1.1.5 *Let us consider sequences of subsets L_n and M_n of a metric space and assume that there exists a compact subset M satisfying the following property:*

$$\text{for any neighborhood } \mathcal{W} \text{ of } M,\ \exists\, N \text{ such that } \forall\, n \geq N,\ \ M_n \subset \mathcal{W}$$

Then, for any neighborhood $\mathcal{U}$ of $M \cap (\mathrm{Limsup}_{n\to\infty} L_n)$, there exists an integer N such that $L_n \cap M_n \subset \mathcal{U}$ whenever $n \geq N$.

Proof — If the neighborhood $\mathcal{U}$ contains M, the result follows from the assumption on M. Otherwise, by taking an open neighborhood $\mathcal{U}$, the subset $K := M \backslash \mathcal{U}$ is not empty, disjoint of $\mathrm{Limsup}_{n\to\infty} L_n$ and is compact by assumption.

Let y belong to K. Since y does not belong to $\mathrm{Limsup}_{n\to\infty} L_n$, there exist $\varepsilon_y > 0$ and N_y such that, for all $n \geq N_y$, y does not belong to $B(L_n, \varepsilon_y)$. The subset K being compact, it can be covered by p

balls $B(y_i, \varepsilon_{y_i})$. This implies that for all $n \geq N_0 := \max_{i=1,\ldots,p} N_{y_i}$ and

$$\mathcal{V} := \bigcup_{i=1}^{p} B(y_i, \varepsilon_{y_i})$$

the intersections $L_n \cap \mathcal{V}$ are empty.

On the other hand, $\mathcal{W} := \mathcal{U} \cup \mathcal{V}$ being a neighborhood of M, we deduce from the assumption that there exists N_1 such that

$$\forall n \geq N_1, \ \ M_n \ \subset \ \mathcal{U} \cup \mathcal{V}$$

Therefore $L_n \cap M_n \subset \mathcal{U}$ for all $n \geq \max(N_0, N_1)$. □

Remark — If M is not compact, but just closed, the conclusion of the proposition remains true for any neighborhood $\mathcal{U}$ of $M \cap (\mathrm{Limsup}_{n\to\infty} L_n)$ whose complement in M is compact. □

We also provide a useful technical lemma:

Lemma 1.1.6 *Let us consider a sequence of subsets $L_n \subset Z$ of a metric space Z and a sequence of subsets $M_n \subset Y$ of a compact metric space Y. Let $\varphi : Z \times Y \mapsto R$ be an upper semicontinuous function. We denote by $M^\sharp$ the upper limit of the M_n and $L^\flat$ the lower limit of the L_n. Then*

$$\mathrm{limsup}_{n\to\infty} \left(\sup_{y\in M_n} \inf_{z\in L_n} \varphi(z,y) \right) \ \leq \ \sup_{y\in M^\sharp} \inf_{z\in L^\flat} \varphi(z,y)$$

Proof — Let y belong to $M^\sharp$. Since φ is upper semicontinuous, we know that for any $\varepsilon > 0$ and any $z \in L^\flat$, there exist neighborhoods $N(z)$ of z and $N(y)$ of y such that

$$\forall\, z' \in N(z),\ \forall\, y' \in N(y),\ \ \varphi(z', y') \ \leq \ \varphi(z,y) + \varepsilon/2$$

In particular, by taking $z_y \in L^\flat$ such that

$$\varphi(z_y, y) \ \leq \ \inf_{z\in L^\flat} \varphi(z,y) + \varepsilon/2$$

and approximating z_y by elements $z_y^n \in L_n$, we infer that there exists $N_y > 0$ satisfying

$$\forall\, n \geq N_y,\ \forall\, y' \in N(y),\ \ \varphi(z_y^n, y') \ \leq \ \varphi(z_y, y) + \varepsilon/2$$

and thus, that for any $y \in M^\sharp$,

$$\forall\, n \geq N_y,\ \forall\, y' \in N(y),\ \ \inf_{z\in L_n} \varphi(z, y') \ \leq \ \inf_{z\in L^\flat} \varphi(z,y) + \varepsilon$$

On the other hand, the compact set $M^\sharp$ can be covered by n neighborhoods $N(y_i)$ so that, by Theorem 1.1.4, there exists an integer N_0 such that

$$\forall\, n \geq N_0,\;\; M_n \;\subset \bigcup_{i=1,\ldots,n} N(y_i)$$

Set $N := \max_{i=0,\ldots,n} N_{y_i}$. Then, for all $n \geq N$ and $y \in M_n$, y belongs to some $N(y_i)$, so that,

$$\inf_{z\in L_n} \varphi(z,y) \;\leq\; \inf_{z\in L^\flat} \varphi(z,y_i) + \varepsilon \;\leq\; \sup_{y\in M^\sharp} \inf_{z\in L^\flat} \varphi(z,y) + \varepsilon$$

Hence we have proved that for any $\varepsilon > 0$, there exists $N > 0$ such that

$$\forall\, n \geq N,\;\; \sup_{y\in M_n} \inf_{z\in L_n} \varphi(z,y) \;\leq\; \sup_{y\in M^\sharp} \inf_{z\in L^\flat} \varphi(z,y) + \varepsilon \;\; \square$$

1.1.2 The Compactness Theorem

The Bolzano-Weierstrass Compactness Theorem was adapted in 1927 to the set-valued framework:

Theorem 1.1.7 (Zarankiewicz) *Every sequence of subsets K_n of a separable metric space X contains a subsequence which has a (possibly empty) limit.*

Proof — Since X is separable, there exists a countable family of open subsets $\mathcal{U}_m$ satisfying the following property:

$$\forall \text{ open subset } \mathcal{U},\;\; \forall\, x \in \mathcal{U},\;\; \exists\, \mathcal{U}_m \;\;\text{ such that }\;\; x \in \mathcal{U}_m \subset \mathcal{U}$$

Let us consider a sequence of subsets K_n. We shall construct a sequence of subsequences $(K_n^{(m)})_{n>0}$ by induction.

For $m = 0$, we set $K_n^{(0)} := K_n$. Assume that the $m-1$ first subsequences $\left(K_n^{(p)}\right)_{n>0}$, $0 \leq p \leq m-1$ have been constructed.

Consider the m^{th} open subset $\mathcal{U}_m$. Then either for every subsequence n_j,

$$\mathcal{U}_m \cap \left(\mathrm{Limsup}_{j\to\infty} K_{n_j}^{(m-1)}\right) \;\neq\; \emptyset$$

in which case we set $K_j^{(m)} := K_j^{(m-1)}$, or there exists a subsequence n_j such that

$$\mathcal{U}_m \cap \left(\mathrm{Limsup}_{j\to\infty} K_{n_j}^{(m-1)}\right) \;=\; \emptyset$$

in which case we set $K_j^{(m)} := K_{n_j}^{(m-1)}$. (The choice of such a subsequence does not matter.)

These sequences $\left(K_n^{(m)}\right)_{n>0}$ being constructed, we extract the diagonal subsequence $D_n := K_n^{(n)}$. We claim that it has a set limit.

If not, there would exist

$$x_0 \in \mathrm{Limsup}_{n\to\infty} D_n \quad \& \quad x_0 \notin \mathrm{Liminf}_{n\to\infty} D_n$$

The latter condition means that there exists an open neighborhood $\mathcal{U}$ of x_0 and a subsequence D_{n_j} such that $\mathcal{U} \cap D_{n_j} = \emptyset$ for any j. Let us fix an open subset $\mathcal{U}_m$ such that $x_0 \in \mathcal{U}_m \subset \mathcal{U}$. We thus deduce that

$$\mathcal{U}_m \cap \left(\mathrm{Limsup}_{j\to\infty} D_{n_j}\right) = \emptyset$$

Since for $n_j \geq m$, $D_{n_j} := K_{n_j}^{(n_j)} = K_{p_j}^{(m-1)}$ for some p_j, we observe that D_{n_j} is a subsequence of the sequence $\left(K_n^{(m-1)}\right)_{n>0}$, the upper limit of which is disjoint from $\mathcal{U}_m$.

By the very construction of $\left(K_n^{(m)}\right)_{n>0}$, we infer that

$$K_j^{(m)} = K_{p_j}^{(m-1)}$$

and consequently, that

$$\mathcal{U}_m \cap \left(\mathrm{Limsup}_{j\to\infty} K_j^{(m)}\right) = \mathcal{U}_m \cap \left(\mathrm{Limsup}_{j\to\infty} K_{p_j}^{(m-1)}\right) = \emptyset$$

Since $D_n := K_n^{(n)} = K_{p_n}^{(m)}$ for some p_n, we deduce that the sequence $(D_n)_{n\geq m}$ is a subsequence of the sequence $\left(K_j^{(m)}\right)_{j>0}$. Thus

$$x_0 \in \mathrm{Limsup}_{n\to\infty} D_n \subset \mathrm{Limsup}_{j\to\infty} K_j^{(m)} \subset X \setminus \mathcal{U}_m$$

which contradicts the fact that x_0 belongs to $\mathcal{U}_m$. □

1.1.3 The Duality Theorem

For closed convex cones K_n, upper limits and lower limits can be exchanged by duality.

We introduce *the (negative) polar cones* to subsets $K \subset X$ and $L \subset X^\star$ defined by

$$K^- := \{p \in X^\star \mid \forall\, x \in K, \ < p, x > \leq 0\}$$

and

$$L^- := \{x \in X \mid \forall\, p \in L, \ < p, x > \leq 0\}$$

Let $\sigma - \mathrm{Limsup}_{n\to\infty} K_n^-$ denote the sequentially weak upper limit of the polar cones K_n^-.

Theorem 1.1.8 *Let $(K_n)_{n\in\mathbf{N}}$ be a sequence of closed convex cones of a Banach space X. Then*

$$\mathrm{Liminf}_{n\to\infty} K_n = \left(\sigma - \mathrm{Limsup}_{n\to\infty} K_n^-\right)^-$$

Proof — Inclusion

$$\mathrm{Liminf}_{n\to\infty} K_n \subset \left(\sigma - \mathrm{Limsup}_{n\to\infty} K_n^-\right)^-$$

is obvious: If $x \in \mathrm{Liminf}_{n\to\infty} K_n$ is the limit of a sequence of elements $x_n \in K_n$ and

$$p \in \sigma - \mathrm{Limsup}_{n\to\infty} K_n^-$$

is the weak-$\star$ limit of a subsequence $p_{n'} \in K_{n'}^-$, inequalities $< p_{n'}, x_{n'} > \leq 0$ imply that $< p, x > \leq 0$.

Conversely, assume that some

$$x \in \left(\sigma - \mathrm{Limsup}_{n\to\infty} K_n^-\right)^-$$

does not belong to the lower limit $\mathrm{Liminf}_{n\to\infty} K_n$. Then there exists $\varepsilon > 0$ and a subsequence (again denoted $(K_n)_{n\in\mathbf{N}}$) such that

$$\forall\, n \geq 0, \ (x + \varepsilon B) \cap K_n = \emptyset$$

The Separation Theorem implies the existence of elements $p_n \in X^\star$ of norm equal to one such that

$$\sigma_{K_n}(p_n) \leq < p_n, x > -\varepsilon \|p_n\| = < p_n, x > -\varepsilon$$

The subsets K_n being cones, we deduce that $p_n \in K_n^-$ and that $\sigma_{K_n}(p_n) = 0$. Then a subsequence (again denoted) p_n converges weakly$-\star$ to some p, which thus belongs to $\sigma - \mathrm{Limsup}_{n\to\infty} K_n^-$, so that $< p, x > \leq 0$. Inequalities

$$0 \leq \langle p_n, x_n \rangle - \varepsilon$$

imply that $\varepsilon \leq 0$, a contradiction. $\square$

1.1.4 Convex Hull of Limits

Since the distance function to a subset of a normed space is convex if and only if this subset is convex, we infer that *the lower limit of a sequence of convex subsets is closed and convex.*

It is useful to have a characterization of the closed convex hull of upper limit. We denote by $\mathrm{co}(K)$ the *convex hull* of K and by $\overline{\mathrm{co}}(K)$ its *closed convex hull.*

Lemma 1.1.9 *Let us consider a sequence of subsets K_n contained in a bounded subset of a finite dimensional vector space X. Then*

$$\overline{\mathrm{co}}\,(\mathrm{Limsup}_{n\to\infty} K_n) \;=\; \bigcap_{N>0} \overline{\mathrm{co}}\left(\bigcup_{n\geq N} K_n\right)$$

Proof — The closed convex hull of the upper limit is obviously contained in the closed convex subset

$$A \;:=\; \bigcap_{N>0} \overline{co}\left(\bigcup_{n\geq N} K_n\right)$$

We have to prove that it is equal to it when the dimension of X is finite and the subsets K_n are contained in a bounded set.

Since an element x of A is the limit of a subsequence of convex combinations v_N of elements of $\bigcup_{n\geq N} K_n$ and since the dimension of X is an integer p, Carathéodory's Theorem allows us to write that

$$v_N \;:=\; \sum_{j=0}^{p} a_j^N x_{N_j}$$

where

$$N_j \;\geq\; N,\;\; a_j^N \;\geq\; 0,\;\; \sum_{j=0}^{p} a_j^N \;=\; 1$$

and x_{N_j} belong to K_{N_j}. The vector a^N of $p+1$ components a_j^N contains a subsequence (again denoted by) a^N which converges to some nonnegative vector a of $p+1$ components a_j such that $\sum_{j=0}^{p} a_j = 1$.

The subsets K_n being contained in a given compact subset, we can extract successively subsequences (again denoted by) x_{N_j} converging to elements x_j, which belong to the upper limit of the subsets K_n. Hence x is equal to the convex combination $\sum_{j=0}^{p} a_j x_j$, and the lemma is proved. □

1.2 Calculus of Limits

We begin by pointing out the following obvious properties:

Proposition 1.2.1 *Let K_n, L_n, K_n^i, $(i = 1, \dots, p)$ be sequences of subsets of a metric space. Then*

$$\begin{cases} i) & \mathrm{Limsup}_{n\to\infty}(K_n \cap L_n) & \subset & \mathrm{Limsup}_{n\to\infty}K_n \cap \mathrm{Limsup}_{n\to\infty}L_n \\ ii) & \mathrm{Liminf}_{n\to\infty}(K_n \cap L_n) & \subset & \mathrm{Liminf}_{n\to\infty}K_n \cap \mathrm{Liminf}_{n\to\infty}L_n \\ iii) & \mathrm{Limsup}_{n\to\infty}(K_n \cup L_n) & = & \mathrm{Limsup}_{n\to\infty}K_n \cup \mathrm{Limsup}_{n\to\infty}L_n \\ iv) & \mathrm{Liminf}_{n\to\infty}(K_n \cup L_n) & \supset & \mathrm{Liminf}_{n\to\infty}K_n \cup \mathrm{Liminf}_{n\to\infty}L_n \\ v) & \mathrm{Limsup}_{n\to\infty}\prod_{i=1}^p K_n^i & \subset & \prod_{i=1}^p \mathrm{Limsup}_{n\to\infty}K_n^i \\ vi) & \mathrm{Liminf}_{n\to\infty}\prod_{i=1}^p K_n^i & = & \prod_{i=1}^p \mathrm{Liminf}_{n\to\infty}K_n^i \end{cases}$$

We need also to relate direct and inverse images of upper and lower limits of a sequence of subsets to the upper and lower limits of their direct and inverse images. We mention now the obvious relations and postpone the proofs of criteria which transform the following inclusions to equalities.

Proposition 1.2.2 *Let K_n be a sequence of subsets of a metric space X, M_n be a sequence of subsets of a metric space Y and $f : X \mapsto Y$ be a (single-valued) continuous map. Then*

$$\begin{cases} i) & f(\mathrm{Limsup}_{n\to\infty}K_n) & \subset & \mathrm{Limsup}_{n\to\infty}f(K_n) \\ ii) & f(\mathrm{Liminf}_{n\to\infty}K_n) & \subset & \mathrm{Liminf}_{n\to\infty}f(K_n) \\ iii) & \mathrm{Limsup}_{n\to\infty}f^{-1}(M_n) & \subset & f^{-1}(\mathrm{Limsup}_{n\to\infty}M_n) \\ iv) & \mathrm{Liminf}_{n\to\infty}f^{-1}(M_n) & \subset & f^{-1}(\mathrm{Liminf}_{n\to\infty}M_n) \end{cases}$$

The question arises of providing converse results under adequate assumptions.

1.2.1 Direct Images

We begin with the direct images of upper limits: We easily obtain equalities when f is *proper*. We recall that a continuous single-valued map from a metric space X to a metric space Y is *proper* if and only if one of the following equivalent statements

$$\text{If } f(x_n) \text{ converges in } Y, \text{ then } x_n \text{ has a cluster point}$$

or

$$\begin{cases} i) & f \text{ maps closed subsets to closed subsets} \\ ii) & \forall \text{ compact } M \subset Y,\ f^{-1}(M) \text{ is compact} \end{cases}$$

holds true.

Proposition 1.2.3 *We posit the assumptions of Proposition 1.2.2. Let us assume that f is proper, then*

$$f(\mathrm{Limsup}_{n\to\infty} K_n) = \mathrm{Limsup}_{n\to\infty} f(K_n)$$

Furthermore, if f is surjective, we obtain

$$f^{-1}(\mathrm{Limsup}_{n\to\infty} M_n) = \mathrm{Limsup}_{n\to\infty} f^{-1}(M_n)$$

The proof is a simple consequence of definitions and is omitted.

In the case of continuous linear operators, we obtain a more specific criterion using *the polar set*

$$K^\circ := \{p \in X^\star \mid \forall x \in K, \ < p, x > \leq 1\}$$

of a subset K of X.

When X and Y are Banach spaces and f is a continuous linear operator $A \in \mathcal{L}(X,Y)$, we obtain the following result:

Theorem 1.2.4 *Let X and Y be Banach spaces, $(K_n)_{n\in\mathbf{N}}$ be a sequence of subsets of X and $A \in \mathcal{L}(X,Y)$ be a continuous linear operator satisfying*

$$0 \in \mathrm{Int}\left(\mathrm{Im}(A^\star) + \bigcup_{N>0}\bigcap_{n>N} K_n^\circ\right) \tag{1.1}$$

— *If X is a finite dimensional space, then*

$$\mathrm{Limsup}_{n\to\infty} A(K_n) = A(\mathrm{Limsup}_{n\to\infty} K_n)$$

— *If X is reflexive, then*

$$\sigma - \mathrm{Limsup}_{n\to\infty} A(K_n) = A(\sigma - \mathrm{Limsup}_{n\to\infty} K_n)$$

Proof — Observe that the first statement follows from the second one and Proposition 1.2.2 *i*). Hence we have to prove only the second claim.

Let us consider a sequence $x_n \in K_n$ such that a subsequence of elements $A(x_n)$ (again denoted $A(x_n)$) converges weakly to some y in Y. We shall check that $(x_n)_{n\in\mathbf{N}}$ has a weak cluster point, by showing that it is weakly bounded, and thus, weakly relatively compact. Assumption (1.1) implies that the set

$$\mathrm{Im}(A^\star) + \bigcup_{N>0} \bigcap_{n>N} K_n^\circ$$

contains a ball around zero of radius $\gamma > 0$. Then any $p \in X^\star$, $\|p\|_\star \leq \gamma$ can be written,

$$p := A^\star q + r, \text{ where } q \in Y^\star, \ r \in \bigcup_{N>0} \bigcap_{n>N} K_n^\circ$$

Therefore, there exists $N > 0$ such that $r \in \bigcap_{n>N} K_n^\circ$ and consequently,

$$\begin{cases} \sup_{n>N} < p, x_n > = \sup_{n>N}(< q, Ax_n > + < r, x_n >) \\ \leq \sup_{n>N}(\|q\|\|Ax_n\| + \sup_{x\in K_n} < r, x >) \\ \leq \sup_{n>N} \|q\|\|Ax_n\| + 1 \leq +\infty \end{cases}$$

since the weakly converging sequence Ax_n is bounded.

Then x_n remains in a bounded subset, which is weakly relatively compact. Hence it has a cluster point x which belongs to $\sigma - \mathrm{Limsup}_{n\to\infty} K_n$. Since A is continuous from X to Y, when they are supplied with their weak topologies, Ax is a weak cluster point of the sequence of elements Ax_n, which converges weakly to $y = Ax$. Hence

$$\sigma - \mathrm{Limsup}_{n\to\infty} A(K_n) \subset A(\sigma - \mathrm{Limsup}_{n\to\infty} K_n)$$

The opposite inclusion being obvious, the proof ensues. □

1.2.2 Inverse Images

We study next the case of inverse images of limits.

We begin first with a simple criterion which holds true when the subsets M_n are convex and the map f is a continuous linear operator.

Proposition 1.2.5 *Let us consider two Banach spaces X and Y, a continuous linear operator $A \in \mathcal{L}(X,Y)$ and a sequence of convex subsets $M_n \subset Y$. We assume that the "constraint qualification assumption"*

$$\exists\, \gamma > 0,\ c > 0 \quad \textit{such that } \gamma B \subset cA(B_X) - M_n$$

holds true for large enough n's. Then

$$\begin{cases} \mathrm{Liminf}_{n\to\infty} A^{-1}(M_n) & = & A^{-1}(\mathrm{Liminf}_{n\to\infty} M_n) \\ \mathrm{Limsup}_{n\to\infty} A^{-1}(M_n) & = & A^{-1}(\mathrm{Limsup}_{n\to\infty} M_n) \end{cases}$$

This statement is a consequence of a more precise result:

Proposition 1.2.6 *Let us consider two Banach spaces X and Y, a continuous linear operator $A \in \mathcal{L}(X,Y)$ and two sequences of convex subsets $L_n \subset X$ and $M_n \subset Y$. Assume that there exist $\gamma > 0$, $c > 0$, $N > 0$ such that*

$$\forall\, n \geq N,\quad \gamma B \subset A(L_n \cap cB_X) - M_n$$

Then

$$\mathrm{Liminf}_{n\to\infty}(L_n \cap A^{-1}(M_n)) = (\mathrm{Liminf}_{n\to\infty} L_n) \cap A^{-1}(\mathrm{Liminf}_{n\to\infty} M_n)$$

$$\mathrm{Limsup}_{n\to\infty}(L_n \cap A^{-1}(M_n)) \supset (\mathrm{Limsup}_{n\to\infty} L_n) \cap A^{-1}(\mathrm{Liminf}_{n\to\infty} M_n)$$

and

$$\mathrm{Limsup}_{n\to\infty}(L_n \cap A^{-1}(M_n)) \supset (\mathrm{Liminf}_{n\to\infty} L_n) \cap A^{-1}(\mathrm{Limsup}_{n\to\infty} M_n)$$

Proof — The inclusion

$$\mathrm{Liminf}_{n\to\infty}(L_n \cap A^{-1}(M_n)) \subset (\mathrm{Liminf}_{n\to\infty} L_n) \cap A^{-1}(\mathrm{Liminf}_{n\to\infty} M_n)$$

being obvious, let us prove the other one, by checking that any x in the intersection

$$\mathrm{Liminf}_{n\to\infty} L_n \cap A^{-1}(\mathrm{Liminf}_{n\to\infty} M_n)$$

is the limit of a sequence of elements x_n belonging to L_n such that $A(x_n)$ belongs to M_n.

We know that x can be approximated by elements $u_n \in L_n$ and that $A(x)$ can be approximated by elements $v_n \in M_n$. Then

$$\varepsilon_n \ := \ \|A(u_n) - v_n\|$$

converges to 0 and

$$\theta_n \ := \ \gamma/(\gamma + \varepsilon_n) \ \in \left]0, 1\right[$$

converges to 1 and satisfies $\theta_n \varepsilon_n = (1 - \theta_n)\gamma$. Therefore,

$$\theta_n(v_n - A(u_n)) \in \theta_n\varepsilon_n B = (1-\theta_n)\gamma B \subset (1-\theta_n)(A(L_n \cap cB_X) - M_n)$$

Consequently, there exist elements $u'_n \in L_n \cap cB_X$ and $v'_n \in M_n$ such that

$$A\left(\theta_n u_n + (1 - \theta_n)u'_n\right) \ = \ \theta_n v_n + (1 - \theta_n)v'_n$$

If we set

$$x_n \ := \ \theta_n u_n + (1 - \theta_n)u'_n$$

we observe that x_n belongs to L_n and that $A(x_n)$ belongs to M_n, for these subsets are convex.

Furthermore,

$$\|x_n - u_n\| \ = \ (1 - \theta_n)\|u_n - u'_n\|$$

converges to 0 since the sequences u_n and u'_n are bounded, the first one because it is convergent to u and the second because it is bounded by c by assumption. Hence x_n converges to x. The proof of the second and third statements is analogous and left as an exercise. □

Remark — Actually, Proposition 1.2.5 is equivalent to Proposition 1.2.6: The former used with $\mathbf{1} \times A$ and $L_n \times M_n$ implies obviously the latter. □

We can prove a similar result for non convex subsets and nonlinear maps, which is based on extensions of Graves' Inverse Function Theorem 3.4.2 proved in Chapter 3. It states that *if f is a continuous map from an open subset of a Banach space X to a Banach space Y, continuously differentiable at x_0, the surjectivity of $f'(x_0)$ implies the existence of a constant $l > 0$ such that, whenever y lies in a neighborhood of $f(x_0)$, the equation $f(x) = y$ has a solution satisfying*

$$\|x - x_0\| \leq l\|f(x_0) - y\| \tag{1.2}$$

The simplest extension reads:

Proposition 1.2.7 *Let us assume that X and Y are Banach spaces, that the map $f : X \mapsto Y$ is continuously differentiable at some point x of $f^{-1}(\mathrm{Liminf}_{n\to\infty} M_n)$, and that $f'(x)$ is surjective.*

Then x belongs to the lower limit $\mathrm{Liminf}_{n\to\infty} f^{-1}(M_n)$, and a similar statement holds true for upper limits.

Proof — Set $y = f(x)$ and consider a sequence of elements $y_n \in M_n$ converging to y. Then (1.2) implies the existence of a solution x_n to the equation $f(x_n) = y_n$ satisfying inequality $\|x - x_n\| \leq l\|y - y_n\|$ for n large enough. Hence x being the limit of $x_n \in f^{-1}(M_n)$ belongs to the lower limit of the inverse images of the subsets M_n. The proof of the second statement is analogous. □

We shall generalize the above inverse function theorem in Chapter 3 and thus be able to weaken the surjectivity assumption of $f'(x)$ by assuming a *transversality condition* bearing on f and the subsets M_n. These conditions involve the concept of contingent cone $T_{M_n}(y_n)$ to a subset M_n at some point $y_n \in M_n$, which will be studied in Chapter 4. The *contingent cone* $T_M(y)$ is defined by

$$v \in T_M(y) \iff \liminf_{h\to 0+} \frac{d_M(y + hv)}{h} = 0$$

However, this is a natural place to state this important extension despite the technical complexity of this transversality assumption:

Theorem 1.2.8 *Let X and Y be two Banach spaces. We consider a sequence of closed subsets $M_n \subset Y$ and a map $f : X \mapsto Y$ continuously differentiable at some point*

$$x \in f^{-1}(\mathrm{Liminf}_{n\to\infty} M_n)$$

Let us assume that there exist constants $c > 0$ *and* $\eta > 0$ *such that*

$$\forall\, y_n \in B(f(x),\eta) \cap M_n, \;\; B_Y \subset cf'(x)(B_X) - T_{M_n}(y_n)$$

Then x *belongs to* $\mathrm{Liminf}_{n\to\infty} f^{-1}(M_n)$ *and a similar statement holds true for upper limits.*

This result is contained in the following, more general,

Theorem 1.2.9 *Let* X *and* Y *be two Banach spaces. We consider a sequence of closed subsets* $M_n \subset Y$ *and a map* $f : X \mapsto Y$ *differentiable on a neighborhood* $\mathcal{U}$ *of* $f^{-1}(\mathrm{Liminf}_{n\to\infty} M_n)$ *(respectively* $f^{-1}(\mathrm{Limsup}_{n\to\infty} M_n)$*) such that the derivatives* $f'(x)$ *are uniformly bounded on* $\mathcal{U}$*. Let us assume that there exist constants* $c > 0$ *and* $\eta > 0$ *such that for every* $x \in \mathcal{U}$*,*

$$\forall\, y_n \in B(f(x),\eta) \cap M_n, \;\; B_Y \subset cf'(x)(B_X) - T_{M_n}(y_n)$$

Then

$$\mathrm{Liminf}_{n\to\infty} f^{-1}(M_n) \;=\; f^{-1}(\mathrm{Liminf}_{n\to\infty} M_n)$$

respectively

$$\mathrm{Limsup}_{n\to\infty} f^{-1}(M_n) \;=\; f^{-1}(\mathrm{Limsup}_{n\to\infty} M_n)$$

We postpone the proof of the last theorem to Chapter 3.

1.3 Set-Valued Maps

Sequences of subsets can be regarded as set-valued maps defined on the set $\mathbf{N}$ of integers.

Naturally, we can replace $\mathbf{N}$ by a metric (or even, topological) space X, and sequences of subsets $n \rightsquigarrow K_n$ by set-valued maps $x \rightsquigarrow F(x)$ (which associates with a point x a subset $F(x)$) and adapt the definition of upper and lower limits to this case, called the *continuous case*.

Before proceeding further, we recall in this section the basic definitions dealing with set-valued maps, also called *multifunctions*, *multivalued functions*, *point to set maps* or *correspondences*.

Definition 1.3.1 *Let X and Y be metric spaces. A set-valued map F from X to Y is characterized by its* graph *Graph(F), the subset of the product space $X \times Y$ defined by*

$$\mathrm{Graph}(F) \ := \ \{(x,y) \in X \times Y \mid y \in F(x)\}$$

We shall say that $F(x)$ is the image *or the* value *of F at x.*

A set-valued map is said to be nontrivial *if its graph is not empty, i.e., if there exists at least an element $x \in X$ such that $F(x)$ is not empty.*

We say that F is strict *if all images $F(x)$ are not empty. The* domain *of F is the subset of elements $x \in X$ such that $F(x)$ is not empty:*

$$\mathrm{Dom}(F) \ := \ \{x \in X \mid F(x) \neq \emptyset\}$$

The image *of F is the union of the images (or values) $F(x)$, when x ranges over X:*

$$\mathrm{Im}(F) \ := \ \bigcup_{x \in X} F(x)$$

The inverse *F^{-1} of F is the set-valued map from Y to X defined by*

$$x \in F^{-1}(y) \iff y \in F(x) \iff (x,y) \in \mathrm{Graph}(F)$$

The domain of F is thus the image of F^{-1} and coincides with the projection of the graph onto the space X and, in a symmetric way, the image of F is equal to the domain of F^{-1} and to the projection of the graph of F onto the space Y.

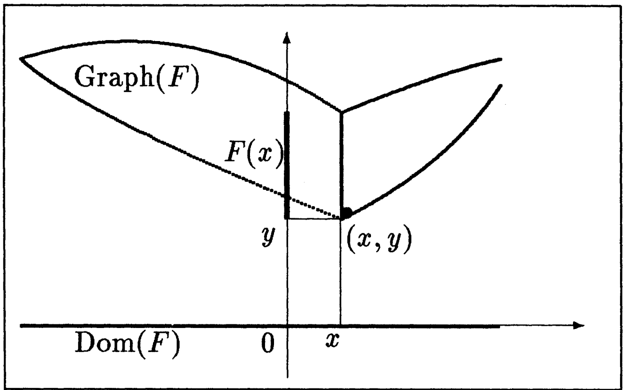
If K is a subset of X, we denote by $F|_K$ the *restriction* of F to K, defined by

$$F|_K(x) := \begin{cases} F(x) & \text{if } x \in K \\ \emptyset & \text{if } x \notin K \end{cases}$$

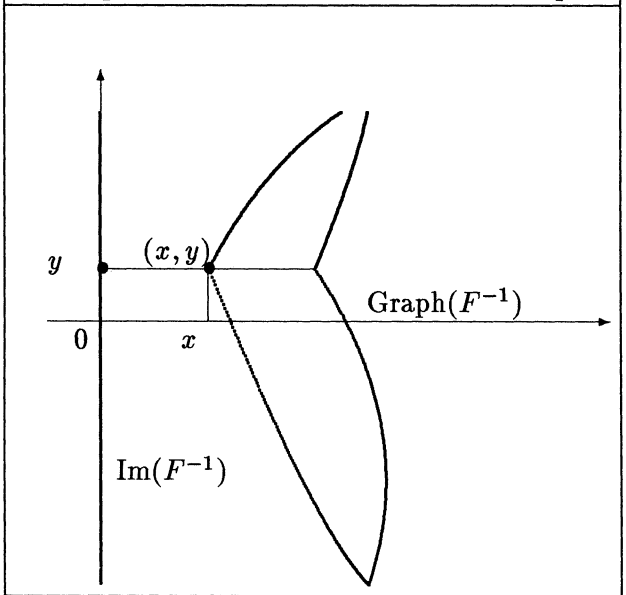
Let $\mathcal{P}$ be a property of a subset (for instance, closed, convex, etc..) Since we shall emphasize the symmetric interpretation of a set-valued map as a graph (instead of a map from a set to another one), we shall say as a general rule that *a set-valued map satisfies property $\mathcal{P}$ if and only if its graph satisfies it.*

For instance, a set-valued map is said to be closed (respectively convex, closed convex, measurable, monotone, maximal monotone)

Figure 1.2: Graph of a Set-Valued Map and of its Inverse



Graph and Domain of a Set-Valued Map



Graph and Image of its Inverse

if and only if its graph is closed (respectively convex, closed convex, measurable, monotone, maximal monotone.)

If the images of a set-valued map F are closed, convex, bounded, compact, and so on, we say that F is *closed-valued*, *convex-valued*, *bounded-valued*, *compact-valued*, and so on.

When $\star$ denotes an operation on subsets, we use the same notation for the operation on set-valued maps, which is defined by

$$F_1 \star F_2 : x \rightsquigarrow F_1(x) \star F_2(x)$$

We define in that way $F_1 \cap F_2$, $F_1 \cup F_2$, $F_1 + F_2$ (in vector spaces), etc. Similarly, if α is a map from the subsets of Y to the subsets of Y, we define

$$\alpha(F) : x \rightsquigarrow \alpha(F(x))$$

For instance, we shall use $\overline{F}$, $\text{co}(F)$, etc., to denote the set-valued maps $x \rightsquigarrow \overline{F}(x)$, $x \rightsquigarrow \text{co}(F(x))$, etc.

We shall write

$$F \subset G \iff \text{Graph}(F) \subset \text{Graph}(G)$$

and say that G is an *extension of* F.

Let us mention the following elementary properties:

$$\begin{cases} i) & F(K_1 \cup K_2) & = & F(K_1) \cup F(K_2) \\ ii) & F(K_1 \cap K_2) & \subset & F(K_1) \cap F(K_2) \\ iii) & F(X \backslash K) & \supset & \text{Im}(F) \setminus F(K) \\ iv) & K_1 \subset K_2 & \Longrightarrow & F(K_1) \subset F(K_2) \end{cases}$$

There are two manners to define the inverse image by a set-valued map F of a subset M:

$$\begin{cases} i) & F^{-1}(M) := \{x \mid F(x) \cap M \neq \emptyset\} \\ ii) & F^{+1}(M) := \{x \mid F(x) \subset M\} \end{cases}$$

The subset $F^{-1}(M)$ is called the *inverse image* of M by F and $F^{+1}(M)$ is called the *core* of M by F.

They naturally coincide when F is single-valued.

We observe at once that

$$F^{+1}(Y \setminus M) \;=\; X \setminus F^{-1}(M) \;\; \& \;\; F^{-1}(Y \setminus M) \;=\; X \setminus F^{+1}(M)$$

One can conceive as well two dual ways for defining composition products of set-valued maps (which coincide when the maps are single-valued):

Definition 1.3.2 *Let X, Y, Z be metric spaces and*

$$G : X \rightsquigarrow Y \;\; \& \;\; H : Y \rightsquigarrow Z$$

be set-valued maps.

1 — The usual composition product (called simply the product*) $H \circ G : X \rightsquigarrow Z$ of H and G at x is defined by*

$$(H \circ G)(x) \;:=\; \bigcup_{y \in G(x)} H(y)$$

2 — The square product *$H \square G : X \rightsquigarrow Z$ of H and G at x is defined by*

$$(H \square G)(x) \;:=\; \bigcap_{y \in G(x)} H(y)$$

Let $\mathbf{1}$ denote the identity map from one set to itself. We deduce the following formulas

$$\left\{ \begin{array}{lcl} \text{Graph}(H \circ G) & = & (G \times \mathbf{1})^{-1}(\text{Graph}(H)) \\ & = & (\mathbf{1} \times H)(\text{Graph}(G)) \\ & & \\ \text{Graph}(H \square G) & = & (G \times \mathbf{1})^{+1}(\text{Graph}(H)) \end{array} \right. \tag{1.3}$$

as well as formulas which state that the inverse of a product is the product of the inverses (in reverse order):

$$\left\{ \begin{array}{lll} i) & (H \circ G)^{-1}(z) & = \; G^{-1}(H^{-1}(z)) \; = \; \left(G^{-1} \circ H^{-1}\right)(z) \\ & & \\ ii) & (H \square G)^{-1}(z) & = \; G^{+1}(H^{-1}(z)) \end{array} \right.$$

We also observe that

$$\begin{cases} i) & x \in (H\square G)^{-1}(z) \iff G(x) \subset H^{-1}(z) \\ ii) & x \in (G^{-1}\square H^{-1})(z) \iff H^{-1}(z) \subset G(x) \end{cases}$$

and thus, that

$$G(x) = H^{-1}(z) \iff x \in (G^{-1}\square H^{-1})(z) \cap (H\square G)^{-1}(z)$$

Let us also point out the following relations: When M is a subset of Z, then

$$\begin{cases} i) & (H \circ G)^{-1}(M) = G^{-1}(H^{-1}(M)) \\ ii) & (H \circ G)^{+1}(M) = G^{+1}(H^{+1}(M)) \end{cases}$$

1.4 Continuity of Set-Valued maps

The concepts of semi-continuous maps have been introduced in 1932 by G. Bouligand[6] and K. Kuratowski[7]. We begin with the upper semicontinuous set-valued maps:

1.4.1 Definitions

In this section, X, Y, Z denote metric spaces. We describe the concepts of semicontinuous set-valued maps introduced by Bouligand, Kuratowski and Wilson in the early thirties.

Definition 1.4.1 *A set-valued map $F : X \rightsquigarrow Y$ is called* upper semicontinuous *at $x \in \mathrm{Dom}(F)$ if and only if for any neighborhood $\mathcal{U}$ of $F(x)$,*

$$\exists\, \eta > 0 \quad \textit{such that } \forall\, x' \in B_X(x,\eta), \;\; F(x') \subset \mathcal{U}.$$

[6]who wrote: "Peut-on rendre plus profond hommage à la mémoire de René Baire qu'en poursuivant les conséquences d'une idée dégagée par lui et dont l'importance se révèle chaque jour accrue, la semi-continuité? Elle échappa tout le XIX^e siècle aux adeptes de la théorie des fonctions, et à plus forte raison, aux purs géomètres, qui s'adonnaient à des occupations moins subtiles."

[7]who also wrote:
"D'après Monsieur Baire, une fontion est dite semi-continue supérieurement..... La notion de semi-continuité dont nous allons nous servir ici est tout à fait analogue à celle-ci, mais concerne le cas où la fonction $F(x)$ admet comme valeurs des sous-ensembles fermés...."

It is said to be upper semicontinuous *if and only if it is upper semicontinuous at any point of* $\mathrm{Dom}(F)$.

When $F(x)$ is compact, F is upper semicontinuous at x if and only if

$$\forall\, \varepsilon > 0,\ \exists\, \eta > 0 \text{ such that } \forall\, x' \in B_X(x,\eta),\ \ F(x') \subset B_Y(F(x),\varepsilon)$$

We observe that this definition is a natural adaptation of the definition of a *continuous* single-valued map. Why then do we use the adjective *upper semicontinuous* instead of continuous? One of the reasons is that the celebrated characterization of continuous maps — *a map f is continuous at x if and only if it maps sequences converging to x to sequences converging to $f(x)$* — does not hold true any longer in the set-valued case.

Indeed, the set-valued version of this characterization leads to the following definition.

Definition 1.4.2 *A set-valued map $F : X \rightsquigarrow Y$ is called* lower semicontinuous *at $x \in \mathrm{Dom}(F)$ if and only if for any $y \in F(x)$ and for any sequence of elements $x_n \in \mathrm{Dom}(F)$ converging to x, there exists a sequence of elements $y_n \in F(x_n)$ converging to y.*

It is said to be lower semicontinuous *if it is lower semicontinuous at every point $x \in \mathrm{Dom}(F)$.*

Actually, as in the single-valued case, this definition is equivalent to:

For any open subset $\mathcal{U} \subset Y$ such that $\mathcal{U} \cap F(x) \neq \emptyset$,

$$\exists\, \eta > 0 \text{ such that } \forall\, x' \in B_X(x,\eta),\ \ F(x') \cap \mathcal{U} \neq \emptyset$$

Unfortunately, there exist set-valued maps which enjoy one property without satisfying the other.

Examples — The set-valued map $F_1 : \mathbf{R} \rightsquigarrow \mathbf{R}$ defined by

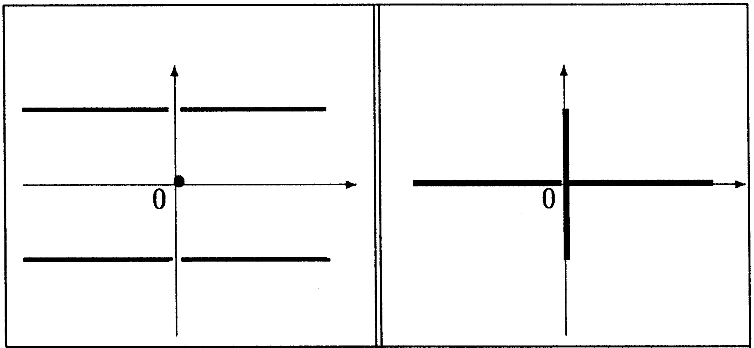
$$F_1(x) := \begin{cases} [-1,+1] & \text{if} \quad x \neq 0 \\ \{0\} & \text{if} \quad x = 0 \end{cases}$$

is lower semicontinuous at zero but not upper semicontinuous at zero.

The set-valued map $F_2 : \mathbf{R} \rightsquigarrow \mathbf{R}$ defined by

$$F_2(x) := \begin{cases} \{0\} & \text{if} \quad x \neq 0 \\ [-1,+1] & \text{if} \quad x = 0 \end{cases}$$

Figure 1.3: Semicontinuous and Noncontinuous Maps



lower s.c. & not upper s.c.	upper s.c. & not lower s.c.

is upper semicontinuous at zero but not lower semicontinuous at zero. □

We are therefore led to introduce still another

Definition 1.4.3 *We shall say that set-valued map F is* continuous *at x if it is* both *upper semicontinuous and lower semicontinuous at x, and that it is* continuous *if and only if it is continuous at every point of* $\mathrm{Dom}(F)$.

Remark — We can use the concepts of inverse images and cores to characterize upper and lower semicontinuous maps:

Proposition 1.4.4 *A set-valued map $F : X \rightsquigarrow Y$ is upper semicontinuous at x if the core of any neighborhood of $F(x)$ is a neighborhood of x and a set-valued map is lower semicontinuous at x if the inverse image of any open subset intersecting $F(x)$ is a neighborhood of x.*

Hence, F is upper semicontinuous if and only if the core of any open subset is open and it is lower semicontinuous if and only if the inverse image of any open subset is open.

If $\mathrm{Dom}(F)$ *is closed, then F is lower semicontinuous if and only if the core of any closed subset is closed and F is upper semicontinuous if and only if the inverse image of any closed subset is closed.*

We shall also need to adapt to the set-valued case the concept of Lipschitz applications.

Definition 1.4.5 *When X and Y are normed spaces, we shall say that $F : X \leadsto Y$ is* Lipschitz around $x \in X$ *if there exist a positive constant l and a neighborhood $\mathcal{U} \subset \mathrm{Dom}(F)$ of x such that*

$$\forall\, x_1, x_2 \in \mathcal{U}, \quad F(x_1) \subset F(x_2) + l\|x_1 - x_2\| B_Y$$

In this case F is also called Lipschitz (or $l-$Lipschitz) on $\mathcal{U}$.

If y is given in $F(x)$, F is said pseudo-Lipschitz *around $(x, y) \in$* Graph(F) *if there exist a positive constant l and neighborhoods $\mathcal{U} \subset$* Dom(F) *of x and $\mathcal{V}$ of y such that*

$$\forall\, x_1, x_2 \in \mathcal{U}, \quad F(x_1) \cap \mathcal{V} \subset F(x_2) + l\|x_1 - x_2\| B_Y$$

We can also adapt the definitions of upper and lower limits to the case of a set-valued map $F : X \leadsto Y$. For that purpose, the notation $x' \to_F x$ means that x' remains in Dom(F) and converges to x.

Definition 1.4.6 *When $F : X \leadsto Y$ is a set-valued map, we say that*

$$\mathrm{Limsup}_{x' \to x} F(x') \;:=\; \left\{ y \in Y \mid \liminf_{x' \to_F x} d\left(y, F(x')\right) = 0 \right\}$$

is the upper limit *of $F(x')$ when $x' \to x$ and that*

$$\mathrm{Liminf}_{x' \to x} F(x') \;:=\; \left\{ y \in Y \mid \lim_{x' \to_F x} d\left(y, F(x')\right) = 0 \right\}$$

is the lower limit *of $F(x')$ when $x' \to x$.*

They are obviously *closed.* We also see at once that

$$\mathrm{Liminf}_{x' \to x} F(x') \subset \overline{F(x)} \subset \mathrm{Limsup}_{x' \to x} F(x')$$

Remark — The use of the concept of *filter* would avoid duplicating these definitions in the discrete and continuous cases. We have preferred this longer, hopefully more pedagogical, presentation. □

Remark — As in the case of sequences of sets, we can easily obtain the following characterization of upper and lower limits. Set $K := \mathrm{Dom}(F)$. Then

$$\begin{cases} \mathrm{Limsup}_{x' \to x} F(x') & = \bigcap_{\eta > 0} \overline{\bigcup_{x' \in B_K(x, \eta)} F(x')} \\ & = \bigcap_{\varepsilon > 0} \bigcap_{\eta > 0} \bigcup_{x' \in B_K(x, \eta)} B(F(x'), \varepsilon) \end{cases}$$

and

$$\mathrm{Liminf}_{x'\to x}F(x') = \bigcap_{\varepsilon>0}\bigcup_{\eta>0}\bigcap_{x'\in B_K(x,\eta)} B(F(x'),\varepsilon) \ \square$$

The connections between semi-continuity of set-valued maps and semi limits are given by

Proposition 1.4.7 *A point (x, y) belongs to the* closure of the graph *of a set-valued map $F : X \rightsquigarrow Y$ if and only if*

$$y \in \mathrm{Limsup}_{x'\to x}F(x')$$

and F is lower semicontinuous *at $x \in \mathrm{Dom}(F)$ if and only if*

$$F(x) \subset \mathrm{Liminf}_{x'\to x}F(x')$$

Then we can measure the lack of closedness (of the graph) or the lack of lower semicontinuity by the discrepancy between the sets

$$F(x), \ \mathrm{Liminf}_{x'\to x}F(x') \ \text{and} \ \mathrm{Limsup}_{x'\to x}F(x')$$

Remark — This proposition led several authors to call upper semi-continuous maps the ones which are closed in our terminology. Naturally, these two concepts coincide for compact-valued maps, since Theorem 1.1.4 can be easily adapted to the case of set-valued maps:

Proposition 1.4.8 *The graph of an upper semicontinuous set-valued map $F : X \rightsquigarrow Y$ with closed domain and closed values is closed.*

The converse is true if we assume that Y is compact.

More generally, the "continuous version" of Proposition 1.1.5 reads:

Proposition 1.4.9 *Let F and G be two set-valued maps from X to Y. Assume that F is closed, that $G(x)$ is compact and that G is upper semicontinuous at $x \in \mathrm{Dom}(F \cap G)$. Then $F \cap G$ is upper semicontinuous at x.*

The above proposition provides an easy way to construct upper semi-continuous set-valued maps, by "cutting" set-valued maps with closed graph by set-valued maps with ball values, the radius of which are upper semi-continuous (real-valued) functions:

Corollary 1.4.10 *Let $F : X \leadsto Y$ be a* closed *set-valued map and*

$$r : X \mapsto \mathbf{R}_+$$

be an upper semicontinuous function. If the dimension of Y is finite , *then the* cut set-valued map $F_r : X \leadsto Y$ *defined by*

$$F_r(x) := F(x) \cap r(x)B$$

is upper semicontinuous.

This is due to the fact that $x \leadsto r(x)B$ *is upper semicontinuous (respectively lower semicontinuous, Lipschitz) whenever r is upper semicontinuous (respectively lower semicontinuous, Lipschitz.)*

Remark — We shall prove later (Theorem 2.2.5) that the set-valued analogues of continuous linear operators, called closed convex processes, from a Banach space to another are also upper semicontinuous (and even, Lipschitz): this is an extension of Banach's Closed Graph Theorem. □

We can extend the concept of proper maps to set-valued maps in the following way:

Definition 1.4.11 (Proper Set-Valued Map) *Let us consider a closed set-valued map $F : X \leadsto Y$. The two following properties are obviously equivalent:*

1. — The projection π_Y : Graph(F) $\mapsto Y$ is proper
2. — If a sequence $y_n \in F(x_n)$ converges in Y, then the sequence $(x_n)_{n\in\mathbf{N}}$ has a cluster point.

If they are satisfied, we shall say that the set-valued map $F : X \leadsto Y$ is proper.

When F is proper, *the images $F(K)$ of closed subsets $K \subset X$ are closed and the inverse images $F^{-1}(M)$ of compact subsets $M \subset Y$ are compact.*

Indeed, we can write

$$\begin{cases} i) & F(K) \;=\; \pi_Y\left(\mathrm{Graph}(F) \cap (K \times Y)\right) \\ ii) & F^{-1}(M) \;=\; \pi_X\left(\mathrm{Graph}(F) \cap \pi_Y^{-1}(M)\right) \end{cases}$$

In particular, the image of a proper set-valued map is closed.

Proposition 1.4.12 *Let us assume that X is locally compact and that for every compact subset $K \subset X$, the graph of the restriction $F|_K$ of $F : X \rightsquigarrow Y$ to K is compact. Then*

1. — F is upper semicontinuous
2. — F^{-1} is proper (and thus, its domain is closed)

Proof — Since for all $x \in \mathrm{Dom}(F)$, there exists a compact neighborhood K of x, the restriction $F|_K$ is upper semicontinuous on K, having a compact graph. Then F is upper semicontinuous at x.

Let a sequence

$$x_n \in F^{-1}(y_n) \subset \mathrm{Dom}(F)$$

converge to some x. We have to check that the sequence y_n has a cluster point. Since the convergent sequence $(x_n)_{n\in\mathbf{N}}$ remains in a compact subset K, and since the pair (x_n, y_n) belongs to the graph of $F|_K$, which is compact, the sequence (x_n, y_n) has a cluster point of the form (x, y), which belongs to the graph of F. Hence $y \in F^{-1}(x)$ is a cluster point of $(y_n)_{n\in\mathbf{N}}$. □

1.4.2 Generic Continuity

Although the notions of upper semicontinuity and lower semicontinuity are distinct, they *generically coincide*, i.e., a semicontinuous map is continuous on a *residual*.

Recall that a *residual* of a metric space X is a countable intersection of dense open subsets $A_n \subset X$. Countable intersections of residuals are residuals. *Baire's Theorem*[8] states that *a residual of a complete metric space is dense*. A property which is true for every element of a residual is called *generic*.

Theorem 1.4.13 (Generic Continuity) *Let F be a set-valued map from a complete metric space X to a complete separable metric space Y.*

1. — If F is upper semicontinuous, it is continuous on a residual of X.

2. — If F is lower semicontinuous with compact values, it is continuous on a residual of X.

3. — If F is lower semicontinuous with closed values, then there exists a residual R of X such that

$$\forall\, x \in R, \quad \mathrm{Limsup}_{x' \to x} F(x') \;=\; F(x)$$

[8]proved in 1899...

Proof — Since Y is a separable metric space, there exists a countable family of open subsets $\mathcal{V}_n \subset Y$, stable by finite union satisfying the following property:

$$\forall \text{ open subset } \mathcal{V} \subset Y, \ \forall y \in \mathcal{V}, \ \exists \mathcal{V}_n \text{ such that } y \in \overline{\mathcal{V}_n} \subset \mathcal{V}$$

1. — Assume first that F is upper semicontinuous. We associate with any open subset $\mathcal{V}_n$ the subset

$$L_n := F^{-1}(\overline{\mathcal{V}_n}) = \{x \in X \mid F(x) \cap \overline{\mathcal{V}_n} \neq \emptyset\}$$

which is closed by Proposition 1.4.4 since the $\overline{\mathcal{V}_n}$ are closed.

We then note that if $x \in X$ is such that

$$\forall n \in \mathbf{N}, \ x \in L_n \implies x \in \operatorname{Int}(L_n)$$

then F is lower semicontinuous at x. Indeed, let $\mathcal{V} \subset Y$ be any open subset such that $F(x) \cap \mathcal{V} \neq \emptyset$. Then there exists $n \in \mathbf{N}$ and

$$y \in F(x) \cap \overline{\mathcal{V}_n} \subset F(x) \cap \mathcal{V}$$

Thus $x \in L_n$ and therefore it belongs to $\operatorname{Int}(L_n.)$ Since

$$L_n = F^{-1}(\overline{\mathcal{V}_n}) \subset F^{-1}(\mathcal{V})$$

we deduce that F is lower semicontinuous at x.

Therefore the subset $D^\flat$ of points where F is not lower semicontinuous is contained in the countable union of the subsets ∂L_n, that are closed subsets with empty interior.

By taking the complements, we infer that the complement of $D^\flat$, which is the set of elements at which F is continuous, contains a residual.

2. — If F is lower semicontinuous, then the subsets

$$K_n := F^{+1}(\overline{\mathcal{V}_n}) = \{x \in X \mid F(x) \subset \overline{\mathcal{V}_n}\}$$

are closed by Proposition 1.4.4. If $x \in X$ is such that

$$\forall n \in \mathbf{N}, \ x \in K_n \implies x \in \operatorname{Int}(K_n)$$

F is upper semicontinuous at x. Indeed, let an open set $\mathcal{V} \subset Y$ be such that $F(x) \subset \mathcal{V}$. For every $y \in F(x)$, let $\mathcal{V}_{n_y}$ be such that $y \in \mathcal{V}_{n_y} \subset \mathcal{V}$. Since F has compact images, $F(x)$ can be covered by a finite number of sets $\mathcal{V}_{n_j} := \mathcal{V}_{n_{y_j}}$, $j = 1, ..., m$. Thus

$$F(x) \subset \bigcup_{j=1}^{m} \overline{\mathcal{V}_{n_j}} := \overline{\bigcup_{j=1}^{m} \mathcal{V}_{n_j}} \subset \mathcal{V}$$

The family of $\mathcal{V}_n$ being stable by finite union, we deduce that for some n,

$$x \in F^{+1}\left(\overline{\bigcup_{j=1}^{m} \mathcal{V}_{n_j}}\right) = K_n \subset F^{+1}(\mathcal{V})$$

Thus, $x \in \mathrm{Int}(F^{+1}(\mathcal{V})$ and we have proved that F is upper semicontinuous at x.

Therefore the subset $D^\sharp$ of points where F is not upper semicontinuous is contained in the countable union of the subsets $\partial K_n := K_n \backslash \mathrm{Int}(K_n)$. The interiors of these subsets are empty because they are closed.

By taking the complements, we deduce that the complement of $D^\sharp$, which is the set of elements at which F is continuous, contains a residual.

3. — We follow Choquet's proof to deduce the third statement. It is based on the fact that any complete separable metric space Y is homeomorphic to a subset Z_0 of a compact metric space Z.

Denote by φ the homeomorphism from Y onto Z_0, and set

$$G_0 := \varphi \circ F : X \rightsquigarrow Z_0 \ \& \ G(x) := \overline{G_0(x)}$$

the closure of $G_0(x)$ in Z.

Since F is lower semicontinuous, so is G_0 and G. By the preceding statement, there exists a residual R of X such that G is upper semicontinuous from R to Z. Therefore, for any $\varepsilon > 0$, there exists η such that, for any $x' \in B(x, \eta)$,

$$G_0(x') \subset G(x') \subset B(G(x), \varepsilon) = B(G_0(x), \varepsilon)$$

On the other hand, $G_0(x) = G(x) \cap Z_0$ since the images $G_0(x)$ are closed in Z_0 for the induced topology. Therefore

$$\forall\, x \in R, \ \ \mathrm{Limsup}_{x' \to x} G_0(x') = G_0(x)$$

Since φ is an homeomorphism, F enjoys the same property. □

Remark — It is easy to construct a lower semicontinuous set-valued map F (with closed but noncompact values) which is nowhere upper semicontinuous: Take $X = \mathbf{R}$, $Y = \mathbf{R}^2$ and $F(t) := \{(x, y) \mid y = tx\}$. □

1.4.3 Example: Parametrized Set-Valued Maps

Let us consider three metric spaces X, Y and Z, a set-valued map

$$U : X \rightsquigarrow Z$$

and a single-valued map

$$f : \mathrm{Graph}(U) \mapsto Y$$

We associate with these data the set-valued map $F : X \rightsquigarrow Y$ defined by

$$\forall\, x \in X, \quad F(x) := (f(x,u))_{u \in U(x)}$$

Proposition 1.4.14 *Assume that f is* continuous *from* Graph(U) *to Y.*

- *If U is lower semicontinuous, so is F.*
- *If U is upper semicontinuous with compact values, so is F.*

Proof — Let us consider a sequence $x_n \in \mathrm{Dom}(F)$ converging to $x \in \mathrm{Dom}(F)$ and take $y := f(x,u)$ belonging to $F(x)$, where $u \in U(x)$. Since U is lower semicontinuous, there exists a sequence $u_n \in U(x_n)$ converging to u. Then the sequence $y_n = f(x_n, u_n)$, where y_n belongs to $F(x_n)$, converges to y because f is continuous. Hence F is lower semicontinuous.

Let us fix $x \in \mathrm{Dom}(U)$, $\varepsilon > 0$ and consider the neighborhood $B(F(x), \varepsilon)$ of $F(x)$. Since it is a neighborhood of each $f(x,u)$ when $u \in U(x)$, the continuity of f implies the existence of $\eta_u > 0$ and $\delta_u > 0$ such that

$$\forall\, (x', v) \in \mathrm{Graph}(U) \cap (B(x, \eta_u) \times B(u, \delta_u)), \;\; f(x', v) \in B(F(x), \varepsilon)$$

The subset $U(x)$ being compact, it can be covered by p balls $B(u_i, \delta u_i)$ Since U is upper semicontinuous, there exists $\eta_0 > 0$ such that

$$\forall\, x' \in B(x, \eta_0), \;\; U(x') \subset \bigcup_{i=1}^{p} B(u_i, \delta_{u_i})$$

We take $\eta := \min(\eta_0, \min_{i=1,\cdots,p} \eta_{u_i}) > 0$. Then we infer that

$$\forall\, x' \in B(x, \eta), \;\; F(x') \subset B(F(x), \varepsilon)$$

i.e., that F is upper semicontinuous at x. $\square$

1.4.4 Marginal Maps

Definition 1.4.15 (Marginal Functions) *Let us consider a set-valued map* $F : X \rightsquigarrow Y$ *and a function* $f : \text{Graph}(F) \mapsto \mathrm{R}$. *We associate with them the* marginal function $g : X \mapsto \mathrm{R} \cup \{+\infty\}$ *defined by*

$$g(x) := \sup_{y \in F(x)} f(x, y)$$

Theorem 1.4.16 (Maximum Theorem) *Let metric spaces* X, Y, *a set-valued map* $F : X \rightsquigarrow Y$ *and a function* $f : \text{Graph}(F) \mapsto \mathrm{R}$ *be given.*

— *If* f *and* F *are* lower semicontinuous, *so is the marginal function.*

— *If* f *and* F *are* upper semicontinuous *and if the* values of F are compact, *so is the marginal function.*

Proof — Let us consider a sequence x_n converging to x, fix $\lambda < g(x)$ and choose $y \in F(x)$ such that $\lambda \le f(x, y)$. Then there exist elements $y_n \in F(x_n)$ converging to y (because F is lower semicontinuous) and we know that $f(x_n, y_n) \le g(x_n)$. Since f is lower semicontinuous, we infer that

$$\lambda \le f(x, y) \le \liminf_{n \to \infty} f(x_n, y_n) \le \liminf_{n \to \infty} g(x_n)$$

By letting λ converge to $g(x)$, the claim ensues.

For proving the second statement, pick $x \in X$ and fix $\varepsilon > 0$. Since f is upper semicontinuous, we can associate with any $y \in F(x)$ open neighborhoods $\mathcal{V}(y)$ of y and $\mathcal{U}_y(x)$ of x such that

$$\forall\, u \in \mathcal{U}_y(x) \text{ and } v \in \mathcal{V}(y), \;\; f(u, v) \le f(x, y) + \varepsilon \tag{1.4}$$

Since $F(x)$ is compact, it can be covered by n neighborhoods $\mathcal{V}(y_i)$, $i = 1, \cdots, p$, the union of which makes up a neighborhood of $F(x)$. Then there exists a neighborhood $\mathcal{U}_0(x)$ such that

$$\forall\, x' \in \mathcal{U}_0(x), \;\; F(x') \subset \bigcup_{i=1}^{p} \mathcal{V}(y_i)$$

because F is upper semicontinuous. By taking u in the neighborhood

$$\mathcal{U}(x) := \mathcal{U}_0(x) \cap \bigcap_{i=1}^{p} \mathcal{U}_{y_i}(x)$$

we observe that

$$\forall\, u \in \mathcal{U}(x), \quad \forall\, v \in F(u), \quad f(u,v) \leq \sup_{i=1,\cdots,p} f(x,y_i) + \varepsilon \;\leq\; g(x) + \varepsilon$$

(thanks to (1.4)) and we deduce that

$$\forall\, u \in \mathcal{U}(x), \quad g(u) \;\leq\; g(x) + \varepsilon \;\; \square$$

We will use quite often the following corollary:

Corollary 1.4.17 *If a set-valued map F is lower semicontinuous* (resp. *upper semicontinuous with compact values*)*, then the function* $(x,y) \mapsto d(y, F(x))$ *is upper semicontinuous* (resp. *lower semicontinuous*) *on* $\mathrm{Dom}(F)$.

1.5 Lower Semi-Continuity Criteria

We provide in this section lower semicontinuity criteria, some of them being adaptations of results on lower limits of sequences to the continuous case.

Proposition 1.5.1 *Consider a metric space X, two normed spaces Y and Z, two set-valued maps G and F from X to Y and Z respectively, and a (single-valued) map f from $X \times Z$ to Y satisfying the following assumptions:*

$$\begin{cases} i) & G \text{ and } F \text{ are lower semicontinuous with convex values} \\ ii) & f \text{ is continuous} \\ iii) & \forall\, x \in X, \;\; u \mapsto f(x,u) \text{ is affine} \end{cases}$$

We posit the following condition:
$\forall\, x \in X,\ \exists\, \gamma > 0,\ \delta > 0,\ c > 0,\ r > 0$ *such that* $\forall\, x' \in B(x,\delta)$ *we have*

$$\gamma B_Y \;\subset\; f(x', F(x') \cap rB_Z) - G(x')$$

Then the set-valued map $R : X \rightsquigarrow Z$ defined by

$$R(x) \;:=\; \{u \in F(x) \;\mid\; f(x,u) \in G(x)\} \tag{1.5}$$

is lower semicontinuous with nonempty convex values.

The proof is a trivial adaptation of the one of Proposition 1.2.6. We state now another condition which is less symmetric.

Proposition 1.5.2 *Consider a metric space X, two normed spaces Y and Z, two set-valued maps G and F from X to Y and Z respectively and a (single-valued) map f from $X \times Z$ to Y such that*

$$\begin{cases} i) & F \text{ is lower semicontinuous with convex values} \\ ii) & f \text{ is continuous} \\ iii) & \forall\, x \in X, \;\; u \mapsto f(x,u) \text{ is affine} \\ iv) & \forall\, x \in X, \;\; G(x) \text{ is convex and its interior is nonempty} \\ v) & \text{the graph of the map } X \ni x \rightsquigarrow \operatorname{Int}(G(x)) \text{ is open} \end{cases}$$

We posit the following condition:

$$\forall\, x \in X, \; \exists\, u \in F(x) \text{ such that } f(x,u) \in \operatorname{Int}(G(x)) \tag{1.6}$$

Then the set-valued map R defined by (1.5) is lower semicontinuous with convex values.

Proof

1. — We introduce the set-valued map $S : X \rightsquigarrow Z$ defined by

$$S(x) \;= \{u \in F(x) \;\mid\; f(x,u) \in \operatorname{Int}(G(x))\} \subset R(x)$$

Assumption (1.6) implies that $S(x)$ is not empty. *We claim that S is lower semicontinuous.* Indeed, if $x_n \to x$ and if u belongs to $S(x) \subset F(x)$, there exists a sequence of elements $u_n \in F(x_n)$ which converges to u because F is lower semicontinuous. Since

$$(x_n, f(x_n,u_n)) \text{ converges to } (x, f(x,u)) \in \operatorname{Graph}(\operatorname{Int}(G(\cdot)))$$

by continuity of f and since the graph of $\operatorname{Int}(G(\cdot))$ is open, the elements $f(x_n,u_n)$ belong to $\operatorname{Int}(G(x_n))$ for n large enough and thus, the elements u_n belong to $S(x_n)$ and converge to u.

2. — Convexity of $F(x)$ and $G(x)$ implies that $\overline{S(x)} = R(x)$. Indeed, let us fix $u \in R(x)$ and $u_0 \in S(x)$. Then $v_\theta := \theta u_0 + (1-\theta)u$ belongs to $S(x)$ when $\theta \in]0,1[$, because $G(x)$ is convex and $f(x,u_0)$ belongs to the interior of $G(x)$, so that for every $\theta \in]0,1[$,

$$f(x,u) + \theta(f(x,u_0) - f(x,u)) = f(x, y + \theta y_0 - \theta y) = f(x, v_\theta)$$

belongs to the interior of $G(x)$. Then u is the limit of v_θ when $\theta > 0$ converges to 0.

3. — The theorem follows because the closure of any lower semicontinuous set-valued map is still lower semicontinuous. □

We extend now the above lower semicontinuity criterion to infinite intersection of set-valued maps.

Theorem 1.5.3 *Let us consider a metric space X, normed vector-spaces Y and Z and set-valued maps $F : X \times Y \rightsquigarrow Z$ and $H : X \rightsquigarrow Y$. We assume that*

$$\begin{cases} i) & F \text{ is lower semicontinuous with convex values} \\ ii) & H \text{ is upper semicontinuous with compact values} \end{cases}$$

and that there exist positive constants γ, δ, c such that for every single-valued map $e : Y \mapsto \gamma B$ we have

$$\forall\, x' \in B(x,\delta), \;\; cB \cap \bigcap_{y \in H(x')} (F(x',y) - e(y)) \neq \emptyset \tag{1.7}$$

Then the set-valued map $G : X \rightsquigarrow Z$ defined by

$$\forall\, x \in X, \;\; G(x) := \bigcap_{y \in H(x)} F(x,y)$$

is lower semicontinuous (with nonempty convex images.)

Remark — When the set-valued map F is locally bounded (in the sense that it maps some neighborhood of each point to a bounded subset), we do not need the constant c and we can replace (1.7) by

$$\forall\, x' \in B(x,\delta), \;\; \bigcap_{y \in H(x')} (F(x',y) - e(y)) \neq \emptyset \;\; \square$$

Proof — Let us choose any sequence of elements $x_n \in \mathrm{Dom}(F)$ converging to x and $z \in G(x)$. We have to approximate z by elements $z_n \in G(x_n)$.

We introduce the following numbers:

$$e_n := \sup_{y \in H(x_n)} d(z, F(x_n, y))/2 \tag{1.8}$$

Now, let us choose for each $y \in H(x_n)$ an element $u_n(y) \in F(x_n, y)$ satisfying

$$\|z - u_n(y)\| \le 2d(z, F(x_n, y)) \le e_n$$

and set $\theta_n := \gamma/(\gamma + e_n)$. Consequently,

$$\theta_n(z - u_n(y)) \in \theta_n e_n B = (1 - \theta_n)\gamma B$$

So that there exists $a_n(y) \in \gamma B$ such that

$$\theta_n(z - u_n(y)) = (1 - \theta_n)a_n(y)$$

Therefore, assumption (1.7) implies the existence for all n large enough of elements $w_n \in cB$ and elements $v_n(y) \in F(x_n, y)$ such that $a_n(y) = v_n(y) - w_n$ for all $y \in H(x_n)$.

Hence we can write

$$\theta_n(z - u_n(y)) = (1 - \theta_n)(v_n(y) - w_n)$$

So that the common value:

$$z_n := \theta_n z + (1 - \theta_n)w_n = \theta_n u_n(y) + (1 - \theta_n)v_n(y)$$

does not depend on y, belongs to all $F(x_n, y)$ (by convexity), converges to z because

$$\|z - z_n\| = (1 - \theta_n)\|z - w_n\| \le (1 - \theta_n)(\|z\| + c)$$

and, because $1 - \theta_n = e_n/(\gamma + e_n)$ converges to 0 for e_n, converges to 0 thanks to the following lemma. □

Lemma 1.5.4 *Let us assume that F is lower semicontinuous and that H is upper semicontinuous with compact images. Then the numbers e_n defined by (1.8) converge to* 0.

Proof — Since F is lower semicontinuous, Corollary 1.4.17 to the Maximum Theorem implies that the function

$$(x, y, z) \mapsto d(z, F(x, y))$$

is upper semicontinuous. Therefore, for any $\varepsilon > 0$ and any $y \in H(x)$, there exist an integer N_y and a neighborhood $\mathcal{V}_y$ of y such that

$$\forall\, y' \in \mathcal{V}_y,\ \forall\, n \geq N_y,\ d(z, F(x_n, y')) \leq \varepsilon \tag{1.9}$$

because $d(z, F(x, y)) = 0$. Hence the compact set $H(x)$ can be covered by p neighborhoods $\mathcal{V}_{y_i}$. Furthermore, H being upper semicontinuous, there exists an integer N_0 such that,

$$\forall\, n \geq N_0,\ H(x_n) \subset \bigcup_{i=1,\ldots,p} \mathcal{V}_{y_i}$$

Set $N := \max_{i=0,\ldots,p} N_{y_i}$. Then, for all $n \geq N$ and $y \in H(x_n)$, y belongs to some $\mathcal{V}_{y_i}$, so that, by (1.9), $d(z, F(x_n, y)) \leq \varepsilon$. Thus,

$$\forall\, n \geq N,\ e_n := \sup_{y \in H(x_n)} d(z, F(x_n, y))/2 \leq \varepsilon/2$$

i.e., our lemma is proved. □

For set-valued maps with non convex images, we deduce from Theorem 1.2.9 its continuous version:

Theorem 1.5.5 *Let $G : X \rightsquigarrow Z$ be a closed lower semicontinuous set-valued map from a metric space X to a Banach space Z and $f : X \times Y \mapsto Z$ a continuous (single-valued) map, where Y is another Banach space. Let us assume that f is differentiable with respect to y and that there exist constants $c > 0$ and $\eta > 0$ such that*

$$\begin{cases} \forall\, x \in B(x_0, \eta),\ y \in B(y_0, \eta),\ z \in B(f(x_0, y_0), \eta) \cap G(x) \\ B_Z \subset c f'_y(x, y)(B_Y) - T_{G(x)}(z) \end{cases} \tag{1.10}$$

Then the set-valued map R defined by

$$R(x) := \{y \in Y \mid f(x, y) \in G(x)\}$$

is lower semicontinuous at x_0.

Chapter 2

Closed Convex Processes

Introduction

Naturally, the first question which arises is: *What are the set-valued analogues of continuous linear operators?*

Since the graph of a continuous linear operator $A \in \mathcal{L}(X, Y)$ is a (closed) vector subspace of $X \times Y$, it is quite natural to regard set-valued maps, with closed convex cones as their graphs, as these set-valued analogues. Such set-valued maps are called *closed convex processes*[1] and the maps the graph of which are vector subspaces are called *linear processes*.

The main class of examples of closed convex processes is provided by derivatives of set-valued maps which are introduced and studied exhaustively in Chapter 5.

We shall prove that closed convex processes enjoy (almost) all properties of continuous linear operators, including Banach's Open Mapping and Closed Graph Theorems (Section 2) and the Uniform Boundedness Theorem (Section 3.)

As continuous linear operators, closed convex processes can be *transposed* and the Bipolar Theorem can be adapted to closed convex processes. They thus enjoy the benefits of a duality theory exposed in Section 5. For instance, a duality criterion of invariance by a

[1]The term "process" has been coined by R.T. Rockafellar for denoting maps the graph of which are cones in a study of economic "processes" (with constant return to scale.)

closed convex process is given in this section for linear processes and in Chapter 4 (Section 2) in the general case.

We shall also provide in Chapter 3 a theorem on existence of eigenvectors of closed convex processes (Theorem 3.6.2.)

For proving these results, we recall in Section 4 the Bipolar Theorem for continuous linear operators, the Closed Range Theorem and the properties of *support functions*. The latter allow characterization of closed convex subsets by an (infinite) family of linear inequalities thanks to a version of the Hahn-Banach Separation Theorem. It also enables one to deal with the class of upper semicontinuous convex positively homogeneous functions instead of handling closed convex subsets.

We conclude the chapter with a short section on upper *hemicontinuous set-valued maps*, characterized by the upper semicontinuity of the support functions of their values, a more familiar property. This characterization is mainly useful for set-valued maps with closed convex images.

2.1 Definitions

Let us introduce the set-valued analogues of continuous linear operators, which are the closed convex processes.

Definition 2.1.1 (Closed Convex Process) *Let $F : X \rightsquigarrow Y$ be a set-valued map from a normed space X to a normed space Y. We shall say that F is*

— convex *if its graph is convex*

— closed *if its graph is closed*

— *a* process *(or* positively homogeneous*) if its graph is a cone*

— *a* linear process *if its graph is a vector subspace.*

Hence a closed convex process *is a set-valued map whose graph is a closed convex cone.*

We shall see that most of the properties of continuous linear operators are enjoyed by closed convex processes.

Let us begin with the following obvious statements.

Lemma 2.1.2 *A set-valued map F is convex if and only if*

$$\begin{cases} \forall\, x_1,\, x_2 \in \mathrm{Dom}(F),\ \forall\, \lambda \in [0,1], \\ \lambda F(x_1) + (1-\lambda)F(x_2) \subset F(\lambda x_1 + (1-\lambda)x_2) \end{cases}$$

It is a process if and only if

$$\forall\, x \in X,\, \forall\, \lambda > 0,\ \lambda F(x) = F(\lambda x) \ \text{and}\ 0 \in F(0)$$

and a convex process if and only if it is a process satisfying

$$\forall\, x_1,\, x_2 \in X,\ F(x_1) + F(x_2) \subset F(x_1 + x_2)$$

We observe that the domain and the image of a closed convex process are convex cones (not necessarily closed.)

The main examples of closed processes are provided by contingent derivatives of set-valued maps that we shall introduce in Chapter 5.

We associate with a closed convex process its *norm* defined in the following way.

Definition 2.1.3 (Norm of a Closed Convex Process) *Let $F : X \rightsquigarrow Y$ be a closed convex process. Its* norm $\|F\|$ *is equal to*

$$\begin{cases} \|F\| \ := \ \sup_{x\in\mathrm{Dom}(F)} d(0,F(x))/\|x\| \\ \qquad = \ \sup_{x\in\mathrm{Dom}(F)} \inf_{v\in F(x)} \|v\|/\|x\| \\ \qquad = \sup_{x\in\mathrm{Dom}(F)\cap B} \inf_{v\in F(x)} \|v\| \end{cases} \tag{2.1}$$

2.2 Open Mapping and Closed Graph Theorems

The Banach Open Mapping Theorem can be extended to closed convex processes:

Theorem 2.2.1 (Open Mapping) *Let X, Y be Banach spaces. Assume that a closed convex process $F : X \rightsquigarrow Y$ is* surjective *(in the sense that* $\mathrm{Im}(F) = Y$*). Then F^{-1} is Lipschitz:*

There exists a constant $l > 0$ such that, for all $x_1 \in F^{-1}(y_1)$ and for any $y_2 \in Y$, we can find a solution $x_2 \in F^{-1}(y_2)$ satisfying:

$$\|x_1 - x_2\| \;\leq\; l\|y_1 - y_2\|$$

Actually, this theorem holds true for closed convex maps:

Theorem 2.2.2 (Robinson-Ursescu) *Let X, Y be Banach spaces, $F : X \rightsquigarrow Y$ be a closed convex set-valued map. Suppose that y_0 belongs to the interior of the image of F and let $x_0 \in F^{-1}(y_0)$.*

Then there exist positive constants l and γ such that for any $y \in y_0 + \gamma B$, there exists a solution x to the inclusion $F(x) \ni y$ satisfying

$$\|x - x_0\| \;\leq\; l\|y - y_0\|$$

Proof — For simplicity, we prove this theorem only when the Banach space X is *reflexive*[2].

Let us introduce the function ρ defined by

$$\rho(y) \;:=\; \inf_{x \in F^{-1}(y)} \|x - x_0\| \;=\; d(x_0, F^{-1}(y)) \tag{2.2}$$

(It takes the value $+\infty$ when y does not belong to $\mathrm{Im}(F)$.)

Since the set-valued map F is convex, this function is obviously convex. Assume for a while it is lower semicontinuous.

Then, Baire's Theorem implies that ρ is continuous[3] on the interior of $\mathrm{Im}(F)$, which is not empty by assumption. Since continuous convex functions are locally Lipschitz, there exists a ball of radius $\gamma > 0$ centered at y_0 and a constant $l' > 0$ such that for all y in this ball,

$$\|\rho(y)\| \;=\; \|\rho(y) - \rho(y_0)\| \;\leq\; l'\|y - y_0\|$$

---

[2] See for instance [35, Theorem 3.3.1] for the nonreflexive case and the original papers of Robinson and Ursescu.

[3] Let us recall this result: The domain of ρ is the union of the sections

$$S_n \;:=\; \{y \mid \rho(y) \leq n\}$$

which are closed because ρ is lower semicontinuous. Since the interior of the domain of ρ is not empty, the interior of one of these sections is not empty. So, the convex function ρ being bounded on an open subset $\mathrm{Int}(S_n)$ for some n, it is continuous (and even, locally Lipschitz) on the interior of its domain.

because $\rho(y_0) = 0$. Therefore

$$d\left(x_0, F^{-1}(y)\right) \le l'\|y - y_0\|$$

By setting $l = 2l'$ we end the proof. It remains to check that the function ρ is lower semicontinuous. $\square$

Lemma 2.2.3 *Let us consider a reflexive Banach space X, a Banach space Y, a closed convex process $F : X \rightsquigarrow Y$.*

Then the function ρ defined by (2.2) is lower semicontinuous.

Proof — It is enough to prove that nonempty sections

$$\{y \mid \rho(y) \le \lambda\}$$

are closed. Let us consider a sequence of elements y_n of such a section converging to some y. Then, by reflexivity of X, there exists $x_n \in F^{-1}(y_n)$ satisfying $\|x_n - x_0\| = \rho(y_n) \le \lambda$. Hence the elements x_n remain in the ball $x_0 + \lambda B$. Since X is reflexive, there exists a (weak) cluster point x of the sequence x_n. Then (x, y) is a weak cluster point of the sequence (x_n, y_n), which thus belongs to the graph of F because, being closed and convex, it is closed in $X \times Y$ when X is supplied with the weak topology and Y with the norm topology. Since the elements x_n belong to the ball $x_0 + \lambda B$, which is weakly closed, the cluster point x belongs also to this ball, so that

$$\rho(y) \le \|x - x_0\| \le \lambda \ \square$$

Proof of Theorem 2.2.1 — We take $x_0 = 0$, $y_0 = 0$ in Theorem 2.2.2. To say that 0 belongs to the interior of the cone $\mathrm{Im}(F)$ amounts to saying that $\mathrm{Im}(F)$ is equal to Y, i.e., that F is surjective. By Theorem 2.2.2 for a constant $l > 0$ we have:

$$\forall\, y \in Y, \exists\, x \in F^{-1}(y) \text{ such that } \|x\| \le l\|y\|$$

Fix y_1, $y_2 \in Y$ and let us choose any solution $x_1 \in F^{-1}(y_1)$ and $e \in F^{-1}(y_2 - y_1)$ satisfying $\|e\| \le l\|y_1 - y_2\|$. Then $x_2 := x_1 + e$ belongs to $F^{-1}(y_2)$, since F is a convex process and satisfies the estimate

$$\|x_1 - x_2\| = \|e\| \le l\|y_1 - y_2\| \ \square$$

Corollary 2.2.4 *Let us consider Banach spaces X, Y, a continuous linear operator $A \in \mathcal{L}(X,Y)$ and a closed convex subset $K \subset X$. Let us assume that there exists $x_0 \in K$ such that $Ax_0 \in \operatorname{Int}(A(K))$.*

Then there exist positive constants l and γ such that for any $y \in A(x_0) + \gamma B$, there exists a solution $x \in K$ to the equation $Ax = y$ satisfying $\|x - x_0\| \leq l\|y - A(x_0)\|$.

Remark — Actually, Corollary 2.2.4 is equivalent to Theorem 2.2.2; we apply it with $K := \operatorname{Graph}(F)$ and $A := \pi_Y$. □

Since the restriction $F := A|_K$ of a continuous linear operator $A \in \mathcal{L}(X,Y)$ to a closed convex cone is a closed convex process, we obtain the following consequence:

Corollary 2.2.5 *Let us consider Banach spaces X, Y, a continuous linear operator $A \in \mathcal{L}(X,Y)$ and a closed convex cone $K \subset X$ such that $A(K) = Y$. Then the set-valued map $y \rightsquigarrow A^{-1}(y) \cap K$ is Lipschitz: there exists a positive constant l such that,*

$$\forall\, y_1,\, y_2 \,\in\, Y,\;\; A^{-1}(y_1) \cap K \;\subset\; A^{-1}(y_2) \cap K \;+\; l\|y_1 - y_2\|B$$

As in the case of continuous linear operators, the Open Mapping Theorem is equivalent to the Closed Graph Theorem, which can be stated as follows.

Theorem 2.2.6 (Closed Graph Theorem) *A closed convex process F from a Banach space X to another Y whose domain is the whole space is Lipschitz: there exists a (Lipschitz) constant $l > 0$ such that*

$$\forall\, x_1,\, x_2 \,\in X,\;\; F(x_1) \;\subset\; F(x_2) \;+\; l\|x_1 - x_2\|B \qquad (2.3)$$

Thus, the norm of F is finite whenever $\operatorname{Dom}(F) = X$.

Proof — It is sufficient to apply the Open Mapping Theorem 2.2.1 to the closed convex process F^{-1}. □

2.3 Uniform Boundedness Theorem

We can now adapt the Uniform Boundedness Theorem to the case of closed convex processes.

Theorem 2.3.1 (Uniform Boundedness) *Let X and Y be Banach spaces and F_h be a family of closed convex processes from X to Y, "pointwise bounded" in the sense that*

$$\forall\, x \in X,\ \exists\, y_h \in F_h(x) \text{ such that } \sup_h \|y_h\| < +\infty \qquad (2.4)$$

Then this family is "uniformly bounded" in the sense that

$$\sup_h \|F_h\| < +\infty$$

Hence we can speak of *bounded* families of closed convex processes, without specifying whether it is pointwise or uniform.

Proof — Let us consider the positively homogeneous convex lower semicontinuous functions ρ_h defined by

$$\rho_h(x) := \inf_{y \in F_h(x)} \|y\| = d(0, F_h(x))$$

(which are lower semicontinuous because the closed convex processes F_h are Lipschitz) and the function ρ defined by

$$\forall\, x \in X,\ \rho(x) := \sup_h \rho_h(x)$$

Assumption (2.4) implies that this function ρ is finite. Since it is also positively homogeneous, convex and lower semicontinuous (being the supremum of such functions), it is continuous at 0. Hence there exists a constant l such that $\sup_h d(0, F_h(x)) = \rho(x) \le l\|x\|$, i.e., $\|F_h\| \le l < \infty$. □

The following consequence of Theorem 2.3.1 extends to closed convex processes the following useful convergence result.

Theorem 2.3.2 (Crossed Convergence) *Consider a metric space U, Banach spaces X, Y and a set-valued map associating to each*

$u \in U$ *a* closed convex process $F(u) : X \rightsquigarrow Y$. *Let us assume that the family of closed convex processes* $F(u)$ *is pointwise bounded.*

Then the following conditions are equivalent:

$$\begin{cases} i) & \textit{the map } u \rightsquigarrow \mathrm{Graph}(F(u)) \textit{ is lower semicontinuous} \\ ii) & \textit{the map } (u,x) \rightsquigarrow F(u)(x) \textit{ is lower semicontinuous} \end{cases}$$

Proof — For proving that $i)$ implies $ii)$, let us consider a sequence of elements (u_n, x_n) converging to (u, x) and an element $y \in F(u)(x)$. We have to approximate it by elements $y_n \in F(u_n)(x_n)$.

Since $u \rightsquigarrow \mathrm{Graph}(F(u))$ is lower semicontinuous, we can approximate (x, y) by elements $(\widehat{x}_n, \widehat{y}_n) \in \mathrm{Graph}(F(u_n))$. By the pointwise boundedness assumption and Theorem 2.3.1, there exist $l > 0$ and solutions $f_n \in F(u_n)(x_n - \widehat{x}_n)$ satisfying

$$\|f_n\| \leq l\|x_n - \widehat{x}_n\|$$

The right hand side of the above inequality converges to zero when n goes to infinity. Because $F(u_n)$ is a convex process, the element $y_n := \widehat{y}_n + f_n$ does belong to $F(u_n)(x_n)$. Consequently, y_n converging to y, we deduce that the set-valued map $(u, x) \rightsquigarrow F(u)(x)$ is lower semicontinuous at (u, x).

The converse is obviously true even when the family $(F(u))_{u \in U}$ is unbounded. □

2.4 The Bipolar Theorem

Definition 2.4.1 *Let* K *be a nonempty subset of a Banach space* X. *We associate with any continuous linear form* $p \in X^\star$

$$\sigma_K(p) := \sigma(K, p) := \sup_{x \in K} < p, x > \in \mathbf{R} \cup \{+\infty\}$$

The function $\sigma_K : X^\star \mapsto \mathbf{R} \cup \{+\infty\}$ *is called the* support function of K. *Its domain is a convex cone called the* barrier cone *denoted by*

$$b(K) := \mathrm{Dom}(\sigma_K) := \{p \in X^\star \mid \sigma_K(p) < \infty\} \qquad (2.5)$$

We say that the subsets of $X^\star$ *defined by*

$$\begin{cases} i) & K^\circ & := & \{p \in X^\star \mid \sigma_K(p) \leq 1\} \\ ii) & K^- & := & \{p \in X^\star \mid \sigma_K(p) \leq 0\} \\ iii) & K^+ & := & -K^- \\ iv) & K^\perp & := & \{p \in X^\star \mid \forall x \in K, \ <p,x> = 0\} \end{cases}$$

are the polar set, (negative) polar cone, positive polar cone *and* orthogonal *of* K *respectively.*

When $L \subset X^\star$, we define the polar set $L^\circ \subset X$ as the subset of elements $x \in X$ (and not $X^{\star\star}$) satisfying $<p,x> \leq 1$ for all $p \in L$. The polar cone $L^- \subset X$ and the orthogonal $L^\perp \subset X$ of L are defined in the same way. The subsets

$$K^{\circ\circ} := (K^\circ)^\circ \subset X \quad \& \quad K^{--} := (K^-)^- \subset X$$

are called respectively the *bipolar set* and *bipolar cone* of a subset $K \subset X$ and the subspace $K^{\perp\perp} := (K^\perp)^\perp \subset X$ the *biorthogonal* of K.

It is clear that K° is a *closed convex subset containing* 0, that K^- is a *closed convex cone*, that $K^\perp$ is a *closed subspace* of $X^\star$ and that

$$K^\perp = K^- \cap K^+ \subset K^- \subset K^\circ \subset b(K)$$

Examples

- When $K = \{x\}$, then $\sigma_K(p) = <p,x>$
- When $K = B_X$, then $\sigma_{B_X}(p) = \|p\|_\star$
- If K is a *cone*, then

$$\sigma_K(p) = \begin{cases} 0 & \text{if} \quad p \in K^- \\ +\infty & \text{if} \quad p \notin K^- \end{cases} \square$$

When $K = \emptyset$, we set $\sigma_\emptyset(p) = -\infty$ for every $p \in X^\star$.
The Separation Theorem can be stated in the following way:

Theorem 2.4.2 (Separation theorem) *Let K be a nonempty subset of a Banach space X. Its* closed convex hull *is characterized by linear constraint inequalities in the following way:*

$$\overline{co}(K) = \{x \in X \mid \forall p \in X^\star, \ <p,x> \le \sigma_K(p)\}$$

Furthermore, there is a bijective correspondance between *nonempty closed convex subsets of X* and *nontrivial lower semicontinuous positively homogeneous convex functions on $X^\star$.*

Remark — The Separation Theorem holds true not only in Banach spaces, but in any Hausdorff locally convex topological vector-space. In particular, we can use it when X is supplied with the weakened topology. The geometrical interpretation can be stated as follows: the closed convex hull of a nonempty subset is the intersection of all closed half-spaces containing it. □

We observe that a subset K is *bounded if and only if its support function is finite.*

We mention the following consequence, known as the *Bipolar theorem.*

Theorem 2.4.3 (Bipolar Theorem) *Let X, Y be Banach spaces and $K \subset X$. The* bipolar *cone K^{--} is the closed convex cone spanned by K.*

If $A \in \mathcal{L}(X,Y)$ is a continuous linear operator from X to Y and K is a subset of X, then

$$(A(K))^- = A^{\star^{-1}}(K^-)$$

where $A^\star$ denotes the transpose of A.

Thus the closed cone spanned by $A(K)$ is equal to $\left(A^{\star^{-1}}(K^-)\right)^-$

We provide now a simple criterion which implies that the image of a closed subset is closed.

Theorem 2.4.4 (Closed Range Theorem) *Let X be a Banach space, Y be a reflexive space, $K \subset X$ be a* weakly closed *subset and $A \in \mathcal{L}(X,Y)$ a continuous linear operator satisfying*

$$\mathrm{Im}(A^\star) + b(K) = X^\star \tag{2.6}$$

Then the image $A(K)$ is closed. In particular, if K is a closed convex cone *and if*

$$\mathrm{Im}(A^\star) + K^- = X^\star$$

then

$$A(K) = \left(A^{\star^{-1}}(K^-)\right)^-$$

Proof — Let us consider a sequence of elements $x_n \in K$ such that $A(x_n)$ converges to some y in Y. We shall check that this sequence is weakly bounded, and thus, weakly relatively compact. Let us take for that purpose any $p \in X^\star$, which can, by assumption (2.6), be written $p := A^\star q + r$, where $q \in Y^\star$, $r \in b(K)$. Therefore,

$$\begin{cases} \sup_n < p, x_n > = \sup_n(< q, Ax_n > + < r, x_n >) \\ \leq \sup_n(\|q\| \|Ax_n\| + \sigma_K(r)) < +\infty \end{cases}$$

since the converging sequence $(Ax_n)_{n\in\mathbf{N}}$ is bounded.

Therefore the sequence $(x_n)_{n\in\mathbf{N}}$ has a weak cluster point x which belongs to K since it is weakly closed. □

Remark — Actually, arguments of the above proof imply that, under assumptions of the Closed Range Theorem, every sequence $x_n \in K$ such that $A(x_n)$ is weakly converging, has a weak cluster point □

For the convenience of the reader, we list below some useful calculus of support functions and barrier cones[4].

[4]See [35, Chapter 3] for instance.

Table 2.1: **Properties of Support Functions.**

(1)	▷	If $K \subset L \subset X$, then $b(L) \subset b(K)$ and $\sigma_K \leq \sigma_L$
(2)	▷	If $K_i \subset X$, $\quad i \in I$, then $b(\overline{co}(\bigcup_{i\in I} K_i)) \subset \bigcap_{i\in I} b(K_i)$ $\sigma(\overline{co}(\bigcup_{i\in I} K_i), p) = \sup_{i\in I} \sigma_{K_i}(p)$
(3)	▷	If $K_i \subset X_i$, $\quad (i = 1, \ldots, n)$, then $b(\prod_{i=1}^n K_i) = \prod_{i=1}^n b(K_i)$ $\sigma(\prod_{i=1}^n K_i, (p_1, \ldots, p_n)) = \sum_{i=1}^n \sigma_{K_i}(p_i)$
(4)a)	▷	If $A \in \mathcal{L}(X, Y)$, then $b(\overline{A(K)}) = A^{\star^{-1}} b(K)$ $\sigma_{\overline{A(K)}}(p) = \sigma_K(A^\star p)$
(4)b)	▷	If K_1 and K_2 are contained in X, then $b(K_1 + K_2) = b(K_1) \cap b(K_2)$ $\sigma_{K_1+K_2}(p) = \sigma_{K_1}(p) + \sigma_{K_2}(p)$ In particular, if $K \subset X$ and P is a cone, then $b(K + P) = b(K) \cap P^-$ and $\sigma_{K+P}(p) = \sigma_K(p)$ if $p \in P^-$ and $+\infty$ if not
(5)	▷	If $L \subset X$ and $M \subset Y$ are *closed convex* subsets and $A \in \mathcal{L}(X, Y)$ is a continuous linear operator such that the following *constraint qualification condition* $0 \in \mathrm{Int}(M - A(L))$ holds true, then $b(L \cap A^{-1}(M)) = b(L) + A^\star b(M)$ and $\forall\, p \in b(L \cap A^{-1}(M))$, $\exists\, \overline{q} \in Y^\star$ such that $\sigma_{L\cap A^{-1}(M)}(p) = \sigma_L(p - A^\star \overline{q}) + \sigma_M(\overline{q})$ $= \inf_{q\in Y^\star}(\sigma_L(p - A^\star q) + \sigma_M(q))$
(5)a)	▷	If $M \subset Y$ is a *closed convex* subset and if $A \in \mathcal{L}(X, Y)$ is a continuous linear operator such that $0 \in \mathrm{Int}(\mathrm{Im}(A) - M)$, then $b(A^{-1}(M)) = A^\star b(M)$ and, for every $p \in b(A^{-1}(M))$, there exists $\overline{q} \in b(M)$ such that $\sigma_{A^{-1}(M)}(p) = \sigma_M(\overline{q}) = \inf_{A^\star q = p}(\sigma_M(q))$
(5)b)	▷	If K_1 and K_2 are closed convex subsets of X such that $0 \in \mathrm{Int}(K_1 - K_2)$, then $b(K_1 \cap K_2) = b(K_1) + b(K_2)$ and $\forall\, p \in b(K_1 \cap K_2)$, $\exists\, \overline{q}_i \in X^\star, (i = 1, 2)$ such that $\sigma_{K_1\cap K_2}(p) = \sigma_{K_1}(\overline{q}_1) + \sigma_{K_2}(\overline{q}_2)$ $= \inf_{p=p_1+p_2}(\sigma_{K_1}(p_1) + \sigma_{K_2}(p_2))$

2.5 Transposition of Closed Convex Process

Definition 2.5.1 (Transpose of a Process) *Let X, Y be Banach spaces, $F : X \rightsquigarrow Y$ be a process. Its* left-transpose *(in short, its* transpose*) $F^\star$ is the closed convex process from $Y^\star$ to $X^\star$ defined by*

$$p \in F^\star(q) \iff \forall\, x \in X,\ \forall\, y \in F(x),\ \ <p, x> \ \leq \ <q, y> \tag{2.7}$$

In particular, the transpose $F^\star$ of a linear process F is defined by

$$p \in F^\star(q) \iff \forall\, x \in X, \forall\, y \in F(x),\ \ <p, x> = <q, y>$$

The graph of the transpose $F^\star$ of F is related to the polar cone of the graph of F in the following way:

Lemma 2.5.2 (Graph of the Transpose) *Let us consider Banach spaces X, Y and let $F : X \rightsquigarrow Y$ be a process. Then*

$$(q, p) \in \mathrm{Graph}(F^\star) \iff (p, -q) \in (\mathrm{Graph}(F))^-$$

In the case of linear processes, we observe that $p \in F^\star(q)$ if and only if $(p, -q)$ belongs to $\mathrm{Graph}(F)^\perp$ and we see at once that the *bitranspose* of a closed linear process F coincides with F.

The definition of a bitranspose of a convex process is not symmetric: If $G : Y^\star \rightsquigarrow X^\star$ is a convex process, we define its transpose $G^\star : X \rightsquigarrow Y$ by the formula

$$(-y, x) \in (\mathrm{Graph}(G))^-$$

(instead of the formula $(y, -x) \in (\mathrm{Graph}(G))^-$ obtained by exchanging the roles of X and $Y^\star$, Y and $X^\star$ respectively.)

With this definition, *the bitranspose of a closed convex process F coincides with F.*

We provide now a formula for transposing the product of closed convex processes.

Theorem 2.5.3 (Transpose of a Product) *Let W, X, Y, Z be Banach spaces, F be a closed convex process from X to Y, $A \in \mathcal{L}(W, X)$ and $B \in \mathcal{L}(Y, Z)$ be continuous linear operators. Assume that*

$$\mathrm{Im}(A) - \mathrm{Dom}(F) = X \tag{2.8}$$

Then the transpose of BFA is equal to:

$$(BFA)^\star = A^\star F^\star B^\star$$

Proof — First, we prove that the formula

$$(BF)^\star = F^\star B^\star$$

always holds true, since the graph of BF is equal to $(\mathbf{1}\times B)\mathrm{Graph}(F)$. Consequently, thanks to the Bipolar Theorem 2.4.3

$$((\mathbf{1} \times B)\mathrm{Graph}(F))^- = (\mathbf{1} \times B)^{\star^{-1}} (\mathrm{Graph}(F)^-)^-$$

so that $(p, -q)$ belongs to $\mathrm{Graph}(BF)^-$ if and only if $(p, -B^\star q)$ belongs to $\mathrm{Graph}(F)^-$, i.e., if and only if p belongs to $F^\star(B^\star q)$.

The graph of FA being equal to $(A \times \mathbf{1})^{-1}\mathrm{Graph}(F)$, we need to assume the constraint qualification property

$$\mathrm{Im}(A \times \mathbf{1}) - \mathrm{Graph}(F) = X \times Y \tag{2.9}$$

for deducing from the properties of polar cones that

$$(\mathrm{Graph}(FA))^- = (A^\star \times \mathbf{1})\,(\mathrm{Graph}(F)^-)$$

If this is the case, we infer that $r \in (FA)^\star(q)$ if and only if there exists $p \in F^\star(q)$ such that $r = A^\star p$.

It remains now to check that assumption (2.8) implies the constraint qualification property (2.9.)

Indeed, let (x, y) belong to $X \times Y$. Since x can be expressed as $x = Ax_0 - x_1$ where x_0 belongs to W and x_1 to the domain of F, we can write

$$(x, y) = (Ax_0, y_0) - (x_1, y_1)$$

where $y_1 \in F(x_1)$ and $y_0 = y + y_1$. Hence (x, y) belongs to $\mathrm{Im}(A \times \mathbf{1}) - \mathrm{Graph}(F)$.

We then deduce from the above proof the conclusion of the theorem. □

Corollary 2.5.4 *Let X, Y and Z be Banach spaces, F a closed convex process from X to Z, G a closed convex process from Y to Z and $A \in \mathcal{L}(X, Y)$. If*

$$A(\mathrm{Dom}(F)) - \mathrm{Dom}(G) = Y \tag{2.10}$$

then $(F + GA)^\star = F^\star + A^\star G^\star$.

Proof — We set

$$H(x) := F(x) + G(Ax), \quad B(y,z) := y + z$$

so that the set-valued map H can be written $H = B(F \times G)(\mathbf{1} \times A)$. Since $\mathrm{Dom}(F \times G)$ is equal to $\mathrm{Dom}(F) \times \mathrm{Dom}(G)$, assumption (2.10) implies that

$$\mathrm{Im}(\mathbf{1} \times A) - \mathrm{Dom}(F \times G) = X \times Y$$

Therefore from Theorem 2.5.3 follows that $H^\star = (\mathbf{1} \times A)^\star (F \times G)^\star B^\star$. Since

$$(F \times G)^\star = F^\star \times G^\star \quad \& \quad (\mathbf{1} \times A)^\star = \mathbf{1} + A^\star$$

we infer that $H^\star = F^\star + A^\star G^\star$. $\square$

Corollary 2.5.5 (Transpose of the Restriction) *Let X, Y be Banach spaces, $F : X \rightsquigarrow Y$ be a closed convex process and $K \subset X$ be a closed convex cone. Assume that*

$$K - \mathrm{Dom}(F) = X$$

Then the transpose of the restriction $F|_K$ of F to K is given by

$$(F|_K)^\star(q) = \begin{cases} F^\star(q) + K^- & \text{if } q \in \mathrm{Dom}(F^\star) \\ \emptyset & \text{otherwise} \end{cases}$$

Proof — We apply Corollary 2.5.4 with $A = \mathbf{1}$ and G defined by

$$G(x) = \begin{cases} \{0\} & \text{if } x \in K \\ \emptyset & \text{otherwise} \end{cases}$$

whose domain is K and whose transpose is the constant set-valued map defined by $G^\star(q) = K^-$. $\square$

We shall adapt to the case of closed convex processes the Bipolar Theorem 2.4.3. To this end, we begin by stating the following simple result.

Proposition 2.5.6 *Let X, Y be Banach spaces and $F : X \leadsto Y$ be a process. Then*

$$(\mathrm{Im}(F))^- = -F^{\star^{-1}}(0) \quad \& \quad F(0) = (\mathrm{Dom}(F^\star))^+$$

Therefore, if F is a convex process, the image of F is dense if and only if the kernel $F^{\star^{-1}}(0)$ of its transpose is equal to 0.

Furthermore a convex process F is surjective if $F^{\star^{-1}}(0) = \{0\}$ and either the dimension of Y is finite or the image of F is closed.

Proof — To say that q belongs to the polar cone of the image of F amounts to saying that the pair $(0, q)$ belongs to the polar cone of the graph of F, i.e., that $(-q, 0)$ belongs to the graph of the transpose $F^\star$, in other words, that 0 belongs to $F^\star(-q)$. The proof of the second statement is naturally analogous. □

The extension of the Bipolar Theorem 2.4.3 to closed convex processes is then a consequence of Corollary 2.5.5 and Proposition 2.5.6.

Theorem 2.5.7 (Bipolar Theorem) *Consider Banach spaces X, Y and let $F : X \leadsto Y$ be a closed convex process, and $K \subset X$ be a cone satisfying* $\mathrm{Dom}(F) - K = X$. *Then*

$$(F(K))^- = -F^{\star^{-1}}(K^+)$$

Proof — We apply Proposition 2.5.6 to the restriction $F|_K$ whose image is $F(K)$ and whose transpose is $F^\star(\cdot) + K^-$ thanks to Corollary 2.5.5. □

The above condition is obviously satisfied when the domain of F is the whole space. In this case, we obtain

Corollary 2.5.8 *Let $F : X \leadsto Y$ be a strict closed convex process. Then*

$$\mathrm{Dom}(F^\star) = F(0)^+$$

and $F^\star$ is upper hemicontinuous (see Definition 2.6.2 below) with bounded closed convex images, mapping the unit ball to the ball of radius $\|F\|$. In particular, $F^\star(0) = \{0\}$.

The restriction of $F^\star$ to the vector space

$$\mathrm{Dom}(F^\star) \cap (\mathrm{Dom}(F^\star))$$

is single-valued and linear.

Proof — We observe that

$$\forall\, q \in \mathrm{Dom}(F^\star), \quad \sup_{p\in F^\star(q)} \|p\| \;\le\; \|F\|\|q\|$$

because for all $x \in \mathrm{Dom}(F) = X$ and for all $p \in F^\star(q)$, we have, by definitions of the transpose and the norm of a closed convex process

$$\begin{cases} \|p\| & := \sup_{x\in B} < p, x > \le \sup_{x\in B} \inf_{y\in F(x)} < q, y > \\ & \le \sup_{x\in B} \inf_{y\in F(x)} \|q\|\|y\| \;=\; \|F\|\|q\| \end{cases}$$

Then $F^\star$ maps bounded sets to bounded sets. Therefore, the cone $F^\star(0)$ being bounded, is equal to $\{0\}$. Since the domain of F is the whole space, the assumptions of Proposition 2.6.4 below are met, so that for all $x \in X$, the function $q \mapsto \sigma(F^\star(q), x)$ is upper semicontinuous.

The domain of $F^\star$ is closed, thanks to the Closed Range Theorem 2.4.4 applied to the projection $\pi_{Y^\star}$ from $X^\star \times Y^\star$ and the cone $\mathrm{Graph}(F^\star)$. Then

$$\mathrm{Dom}(F^\star) \;=\; \pi_{Y^\star}\,(\mathrm{Graph}(F^\star))$$

is closed because, the domain of F being equal to X,

$$\mathrm{Im}\,((\pi_{Y^\star})^\star) \;-\; (\mathrm{Graph}(F^\star))^- \;=\; X \times Y$$

This and Proposition 2.5.6 imply that $\mathrm{Dom}(F^\star) = F(0)^+$.

If q belongs to both $\mathrm{Dom}(F^\star)$ and $-\mathrm{Dom}(F^\star)$, then

$$F^\star(q) + F^\star(-q) \;\subset\; F^\star(0) \;=\; \{0\}$$

Hence $F^\star(q)$ contains only one element and $F^\star(q) = -F^\star(-q)$. □

Example: Case of linear processes

When $F : X \leadsto Y$ is a linear process from a Banach space X to another Y, then, from the very definition of the adjoint, it follows that $F^\star$ is also a linear process from $Y^\star$ into $X^\star$. Hence its domain is a subspace of $Y^\star$ and, by Proposition 2.5.6, $\overline{\mathrm{Dom}(F^\star)} = F(0)^\perp$.

In general the space $\mathrm{Dom}(F^\star)$ is not closed. However this is always the case when Y has a finite dimension.

Furthermore Corollary 2.5.8 implies that when F is strict, then

$$\mathrm{Dom}(F^\star) \ = \ F(0)^\perp$$

and $F^\star$ is a linear operator from the subspace $F(0)^\perp$ into $X^\star$.

When $X = Y$, a closed subspace $P \subset \mathrm{Dom}(F)$ is called *invariant* under F if $F(P) \subset P$. We have the following

Proposition 2.5.9 *Assume that* $\mathrm{Dom}(F^\star)$ *is closed. If a closed subspace P is invariant under F, then its orthogonal space $P^\perp$ is invariant under $F^\star$.*

Consequently, if both $\mathrm{Dom}(F)$ *and* $\mathrm{Dom}(F^\star)$ *are closed, then a closed subspace P is invariant under F, if and only if its orthogonal space $P^\perp$ is invariant under $F^\star$.*

Proof — Let $P \subset \mathrm{Dom}(F)$ be an invariant closed subspace. Then $F(0) \subset P$ and therefore

$$P^\perp \ \subset \ F(0)^\perp \ = \ \mathrm{Dom}(F^\star)^{\perp\perp} \ = \ \mathrm{Dom}(F^\star)$$

Fixing $q \in P^\perp$ and $p \in F^\star(q)$, for every $x \in P,\ y \in F(x) \subset P$ we have

$$< p, x > = < q, y > = \ 0$$

Thus $p \in P^\perp$ and we have proved that if a subspace P is invariant under F, then its orthogonal $P^\perp$ is invariant under the adjoint process $F^\star$. Since $F = F^{\star\star}$ the last statement follows. □

For every $x \in X$ define recursively

$$F^1(x) \ = \ F(x) \ \text{ and for every integer } n \geq 1, \ \ F^{n+1}(x) \ = \ F(F^n(x))$$

Then for every n, $F^n(0)$ is a subspace of X and

$$F^n(0) \ \subset \ F^{n+1}(0)$$

Consider the subspace

$$Q \ = \ \bigcup_{n\geq 1} F^n(0)$$

Clearly $F(Q) \subset Q$. If F is strict, then it is Lipschitz on X. Thus $F(\overline{Q}) \subset \overline{Q}$ and therefore in this case $\overline{Q}$ is the smallest closed subspace of X invariant under F. We also have the following implication

$$F^k(0) = F^{k+1}(0) \implies \forall\, n \in \mathbf{N},\ F^k(0) = F^{k+n}(0) \qquad (2.11)$$

In particular this yields that when $X = \mathbf{R}^n$, then $Q = F^n(0)$. Indeed in this case $\{F^k(0)\}_{k\geq 1}$ is a nondecreasing sequence of subspaces of $\mathbf{R}^n$ satisfying (2.11.) The dimension of $\mathbf{R}^n$ being equal to n, the last claim follows.

Thus $F^n(0)$ is the smallest subspace of $\mathbf{R}^n$ invariant under F. □

The sum of two closed convex processes or the product BF is not necessarily closed. We have to provide sufficient conditions implying that they are still closed.

Proposition 2.5.10 (Closed Graph Criterion) *Let X, Y and Z be reflexive Banach spaces, $F : X \rightsquigarrow Y$ and $G : X \rightsquigarrow Z$ be closed convex processes and $B \in \mathcal{L}(Y, Z)$ be a continuous linear operator. If*

$$B^\star(\mathrm{Dom}(G^\star)) - \mathrm{Dom}(F^\star) = Y^\star$$

then the convex process $BF + G$ is closed.

Proof — This is a consequence of the Closed Range Theorem 2.4.4 with $K := \mathrm{Graph}(F \times G)$ and the continuous linear operator A defined by $A(x, y, z) := (x, By + z)$. □

Proposition 2.5.11 *Let X, Y be Banach spaces, $G : X \rightsquigarrow Y$ be a closed convex process and $P \subset X$ and $Q \subset Y$ be closed convex cones. Let us consider the convex process F defined by*

$$F(x) := \begin{cases} G(x) + Q & \textit{if } x \in P \\ \emptyset & \textit{if } x \notin P \end{cases}$$

It is closed when we suppose that

$$\mathrm{Dom}(G^\star) + Q^- = Y^\star$$

If we assume that $\mathrm{Dom}(G) - P = X$, *then its transpose is given by*

$$F^\star(q) := \begin{cases} G^\star(q) + P^- & \textit{if } q \in Q^+ \\ \emptyset & \textit{if } q \notin Q^+ \end{cases}$$

Proof — We observe that F is the sum of the closed convex process G and the closed convex process H defined by $H(x) := Q$ if $x \in P$ and $\emptyset$ if $x \notin P$, whose domain is P and whose transpose is defined by $H^\star(q) := P^-$ if $q \in Q^+$ and $\emptyset$ if not. Corollary 2.5.4 ends the proof. □

2.6 Upper Hemicontinuous Maps

We associate with a set-valued map F from a metric space X to a normed space Y the family of functions

$$x \mapsto \sigma(F(x), p) \ := \ \sup_{y \in F(x)} < p, y >$$

indexed by the continuous linear functionals $p \in Y^\star$.

We observe that

Corollary 2.6.1 *If a set-valued map F from a metric space X to a normed space Y is weakly upper semicontinuous and has compact values* (resp. *lower semicontinuous*), *then the function*

$$(x, q) \ \in \ X \times Y^\star \ \mapsto \ \sigma(F(x), q)$$

is upper semicontinuous (resp. *lower semicontinuous*).

Therefore, it is quite convenient to introduce the following definition.

Definition 2.6.2 (Upper Hemicontinuous Map) *We shall say that a set-valued map $F : X \rightsquigarrow Y$ is* upper hemicontinuous *at $x_0 \in$* $\mathrm{Dom}(F)$ *if and only if for any $p \in Y^\star$, the function $x \mapsto \sigma(F(x), p)$ is upper semicontinuous at x_0.*

It is said to be upper hemicontinuous if and only if it is upper hemicontinuous at every point of $\mathrm{Dom}(F)$.

Proposition 2.6.3 *The graph of an upper hemicontinuous set-valued map with* closed convex *values is* closed.

Proof — Let us consider elements $(x_n, y_n) \in \mathrm{Graph}(F)$ converging to a pair (x, y). Then, for every $p \in Y^\star$,

$$< p, y > = \lim_{n \to \infty} < p, y_n > \leq \ \limsup_{n \to \infty} \sigma(F(x_n), p) \ \leq \ \sigma(F(x), p)$$

by the upper semicontinuity of $x \mapsto \sigma(F(x), p)$. This inequality implies that $y \in F(x)$ since these subsets are closed and convex, thanks to the Separation Theorem 2.4.2. We have shown that (x, y) belongs to Graph(F), which ends the proof. □

It is useful to compare the support functions of $F(x)$ and $F^\star(q)$.

Proposition 2.6.4 (Support Function of the Transpose) *Let X, Y be Banach spaces and $F : X \rightsquigarrow Y$ be a closed convex process. Then for every $x \in X$ and $q \in Y^\star$*

$$\sigma(F^\star(q), x) \ + \ \sigma(F(x), -q) \ \leq \ 0$$

Furthermore for every x_0 in the interior of the domain of F there exists $p_0 \in F^\star(q_0)$ such that $< p_0, x_0 >$ is equal to the common value

$$\sigma(F^\star(q_0), x_0) \ = \ -\sigma(F(x_0), -q_0) \tag{2.12}$$

In the same way, if q_0 belongs to the interior of the domain of $F^\star$, there exists $y_0 \in F(x_0)$ such that $< q_0, y_0 >$ is equal to the common value (2.12.)

Therefore for every $x \in Dom(F)$, the function $q \mapsto \sigma(F^\star(q), x)$ is upper semicontinuous on the interior of the domain of $F^\star$ and for every $q \in Dom(F^\star)$ the function $x \mapsto \sigma(F(x), -q)$ is upper semicontinuous on the interior of the domain of F.

Proof — The first claim follows from the very definition of the support function.

Denoting by π_X the projector from $X \times Y$ to X, we can write

$$\sigma(F(x_0), q_0) \ = \ \sigma(\mathrm{Graph}(F) \cap \pi_X^{-1}(x_0), (0, q_0))$$

We apply the formula (5) of Table 2.1 on the support functions of an intersection and inverse image with $L := \mathrm{Graph}(F)$, $M := \{x_0\}$ and $A := \pi_X$. The constraint qualification assumption is satisfied because 0 belongs to the interior of

$$\pi_X(\mathrm{Graph}(F)) \ - \ x_0 \ = \ \mathrm{Dom}(F) \ - \ x_0$$

by assumption. Then

$$\sigma(F(x_0), q_0) \ = \ \inf_{p \in X^\star} (< p, x_0 > + \sigma(\mathrm{Graph}(F), (0, q_0) - \pi_X^\star(p)))$$

We observe that $\pi_X^\star(p) = (p, 0)$ and that

$$\sigma(\mathrm{Graph}(F), (-p, q_0)) \;=\; 0$$

if and only if $p \in F^\star(q_0)$, so that (2.12) ensues. Since the function $-\sigma(F(x_0), \cdot)$ is upper semicontinuous, the proof of the last claim ensues. $\square$

Chapter 3

Existence and Stability of an Equilibrium

Introduction

This chapter is devoted to the major questions of nonlinear functional analysis: provide criteria allowing to solve[1] a (nonlinear) equation or an inclusion (with or without constraints) and, also, approximate their solutions and study their stability.

We tried to present the shortest introduction to this so important field by choosing only the theorems we judge the most powerful, i.e., the theorems from which we obtain the most results with the least efforts.

Historically, nonlinear analysis is based on two fixed point theorems:

- Banach-Picard Successive Approximation Theorem for problems set on complete metric spaces
- Brouwer Fixed Point Theorem in a convex compact environment

Using these tools amounts to transforming the problem under investigation to a fixed point problem, for which one can apply one of

[1]Apparently Cauchy was the first to prove an existence theorem, instead of assuming a priori that the solution exists and compute it, risking in this way to obtain paradoxes.

these fixed point theorems or their numerous variations or extensions. The "fixed point" format being quite rigid, one encounters the risk of more or less considerable loss of information by doing so. It also often happens that this transformation may require additional assumptions and useless technical difficulties.

This is the reason why so many statements logically equivalent to the Brouwer Fixed Point Theorem — constituting the corpus of nonlinear analysis — have been designed to be readily adapted to classes of specific problems.

Still, among these equivalent results, one can look for those statements which, in some sense, incorporate more *labor-value*[2], and therefore, might be more useful.

This is the case of the Ky Fan and Ekeland Theorems, which provide less elegant but more efficient statements than the corresponding Fixed Point Theorems.

These two more flexible and powerful theorems, equivalent to each of the above Fixed Point Theorems, appeared at the beginning of the seventies:

- The *Ky Fan Inequality*, in 1972
- *Ekeland's Variational Principle*, in 1974.

We thus begin by deducing Ky Fan's Inequality from the Brouwer Fixed Point Theorem in the first section, and use it to derive in Section 2 not only the traditional fixed point theorems (including the Kakutani-Fan Fixed Point Theorem for set-valued maps), but also a Constrained Equilibrium Theorem, providing an equilibrium

[2]Although two statements $\mathcal{P}$ and $\mathcal{Q}$ may well be equivalent, it is common experience that the proof of one of the implications, say $\mathcal{P} \Longrightarrow \mathcal{Q}$, is more difficult or involves deeper results than the proof of the converse. In this case, one can say that $\mathcal{Q}$ *incorporates more labor value than* $\mathcal{P}$ and thus, expect as a general rule that the statement $\mathcal{Q}$ may be more useful than $\mathcal{P}$.

One could still follow Aristotle's observation that most theorems of geometry have valid converses, so that one can go leisurely from $\mathcal{P} \Longrightarrow \mathcal{Q}$ (called *analysis*) to $\mathcal{Q} \Longrightarrow \mathcal{P}$. However, more serious greek mathematicians, conscious of the pitfalls, used to add supplementary conditions (called *diorismoi*.) The reverse inference was called *synthesis*.

By the way, the most pleasant or intuitive statements are quite often the ones with the least labor value.

$\bar{x}$ of a set-valued map F (i.e., a solution to the inclusion $0 \in F(\bar{x})$ satisfying the constraints described by $\bar{x} \in K$.) Besides standard mild conditions on F, we shall assume that K is a *convex compact viability domain of* F; this means that we can find at every point x of the compact convex subset K an element $v \in F(x)$ tangent to K at x.

We then prove Ekeland's Variational Principle in Section 3, a result we shall use extensively in this book, above all for establishing the calculus of tangent cones, of derivatives of set-valued maps and epiderivatives of extended functions in the forthcoming chapters.

But we start to apply it in Section 4 to extend Graves' Inverse Function Theorem to *constrained problems* of the form:

$$\begin{cases} K \text{ being a closed subset describing the constraints,} \\ \text{look for a solution } x_0 \in K \text{ to the equation } f(x_0) = y_0. \end{cases}$$

In the unconstrained case, Graves' Theorem states that if f is $\mathcal{C}^1$ at x_0 and if $f'(x_0)$ is surjective, then the equation $f(x) = y$ still has a solution for any right-hand side y close enough to $f(x_0)$, and the set-valued map f^{-1} behaves in a Lipschitz way.

When constraints requiring that x belongs to K are present, we "differentiate" them by using the *contingent cone* $T_K(x)$ to K at x, which is a most reasonable way to translate the concept of tangency for an arbitrary subset. Roughly speaking, if $f'(\cdot)$ is continuous at x_0, if $T_K(\cdot)$ is lower semicontinuous at x_0 and if $f'(x_0)(T_K(x_0))$ is the whole space, then the equation $f(x) = y$ has a solution $x \in K$ for any y in a neighborhood of $f(x_0)$.

This stability property, often regarded as an existence theorem of a solution to nonlinear problems, is actually a particular case of an approximation procedure, and conceals the famous *Lax Principle*:

$$\begin{cases} \textit{Consistency and Stability imply Convergence} \\ \textit{for approximating solutions to given problems,} \end{cases}$$

which underlies numerical functional analysis. The constrained problem is approximated by the following constrained problems:

$$\text{find } x_n \in K_n \quad \text{satisfying} \quad f_n(x_n) = y_n$$

Consistency means that in some sense, f_n converges to f and K_n to K. *Stability* means that linear equations

$$\text{find } u_n \in T_{K_n}(x_n) \text{ satisfying } f'_n(x_n)u_n = v_n$$

have solutions satisfying the stability condition $\|u_n\| \leq c\|v_n\|$ and *Convergence* means that the convergence of the right-hand sides y_n to y_0 implies the convergence of approximate solutions x_n to x_0.

These theorems will be extended to inclusions as soon as we have defined the derivatives of set-valued maps in Chapter 5.

We describe in Section 5 the basic results on *monotone and maximal monotone* maps, which share many properties with continuous linear operators. Maximal monotone maps can be approximated by single-valued Lipschitz maps (*Yosida approximations*), a feature that is very useful[3].

We provide in Section 6 a result on existence of eigenvalues and eigenvectors of a closed convex process.

3.1 Ky Fan's Inequality

After what we said in the introduction, we state and prove the Ky Fan[4] inequality.

Theorem 3.1.1 (Ky Fan Inequality) *Let K be a compact convex subset of a Banach space[5] X and $\varphi : X \times X \mapsto \mathbf{R}$ be a function satisfying*

$$\left\{\begin{array}{lll} i) & \forall\, y \in K, & x \mapsto \varphi(x,y) \text{ is lower semicontinuous} \\ ii) & \forall\, x \in K, & y \mapsto \varphi(x,y) \text{ is concave} \\ iii) & \forall\, y \in K, & \varphi(y,y) \leq 0 \end{array}\right. \tag{3.1}$$

Then, there exists $\overline{x} \in K$, a solution to

$$\forall\, y \in K, \;\; \varphi(\overline{x},y) \leq 0 \tag{3.2}$$

[3] We shall use it in Chapter 9 to construct regular selections of set-valued maps.

[4] A student of Fréchet. He visited Paris for one year in 1939, with a small grant and the map of Paris' metro. He then had to survive there during the whole war.

[5] Actually, the proof we give shows that this theorem remains true for any Hausdorff locally convex topological vector space and in particular, when the Banach space X is endowed with the weak topology.

Proof — We shall prove it first in the finite-dimensional case, from which we shall derive the proof in Banach spaces.

— **The finite-dimensional case.** We shall derive a contradiction from the negation of the conclusion:

$$\forall\, x \in K, \ \exists\, y \in K \ \text{ such that } \ \varphi(x,y) > 0$$

so that K can be covered by the subsets

$$\mathcal{V}_y := \{x \in K \mid \varphi(x,y) > 0\}$$

which are open by assumption (3.1) *i*). Since K is compact, it can be covered by n such open subsets $\mathcal{V}_{y_i}$. Let us consider a *continuous partition of unity*[6] $(\alpha_i)_{i=1,\ldots,n}$ associated with this open covering of K and define the map $f : K \mapsto X$ by

$$\forall\, x \in K, \ f(x) := \sum_{i=1}^{n} \alpha_i(x) y_i$$

It maps K to itself because K is convex and the elements y_i belong to K. It is also continuous, so that Brouwer's Fixed Point Theorem implies the existence of a fixed point $\overline{y} = f(\overline{y}) \in K$ of f. Assumption (3.1) *ii*) imply that

$$\varphi(\overline{y},\overline{y}) = \varphi(\overline{y},\sum_{i=1}^{n}\alpha_i(\overline{y})y_i) \geq \sum_{i=1}^{n}\alpha_i(\overline{y})\varphi(\overline{y},y_i)$$

Let us introduce

$$I(\overline{y}) := \{i = 1,\ldots,n \mid \alpha_i(\overline{y}) > 0\}$$

It is not empty because $\sum_{i=1}^{n}\alpha_i(\overline{y}) = 1$. Furthermore

$$\sum_{i=1}^{n}\alpha_i(\overline{y})\varphi(\overline{y},y_i) = \sum_{i\in I(\overline{y})}\alpha_i(\overline{y})\varphi(\overline{y},y_i) > 0$$

[6] A *continuous partition of unity associated with a covering of K by n open subsets $\mathcal{V}_i$* is a sequence of n continuous maps $\alpha_i : K \mapsto \mathbf{R}$ such that,

$$\forall\, x \in K, \ \sum_{i=1}^{n}\alpha_i(x) = 1, \ \forall\, i = 1,\ldots,n, \ \alpha_i(x) \geq 0 \ \& \ \text{support}(\alpha_i) \subset \mathcal{V}_i$$

Such continuous partitions of unity do exist when K is a compact metric space.

because, whenever i belongs to $I(\overline{y})$, $\alpha_i(\overline{y}) > 0$, so that $\overline{y}$ belongs to $\mathcal{V}_{y_i}$, and thus, by the very definition of this subset, $\varphi(\overline{y}, y_i) > 0$. Hence, we have proved that $\varphi(\overline{y}, \overline{y})$ is strictly positive, a contradiction of assumption (3.1) *iii*).

— **The infinite-dimensional case.** Let us introduce the family $\mathcal{S}$ of finite subsets $M := \{y_1, \dots, y_m\} \subset K$ of K and the number

$$v := \sup_{M \in \mathcal{S}} \inf_{x \in K} \max_{y_i \in M} \varphi(x, y_i)$$

This number is nonpositive. Indeed, the function $\varphi_M(\lambda, \mu)$ defined on $S^m \times S^m$ (where the *simplex* S^m is a subset of elements $(\lambda_1, ..., \lambda_m) \in \mathbf{R}^m$ such that $\lambda_i \geq 0$ and $\sum_{i=1}^m \lambda_i = 1$) by

$$\varphi_M(\mu, \lambda) := \sum_{i=1}^{m} \mu_j \varphi(\sum_{i=1}^{n} \lambda_i y_i, y_j)$$

satisfies the assumption of Ky Fan's Theorem in finite dimension (it is obviously lower semicontinuous with respect to λ and linear with respect to μ, and $\varphi_M(\mu, \mu) \leq 0$ because φ is concave with respect to y and by assumption (3.1) *iii*). Hence there exists $\overline{\lambda} \in S^m$ such that

$$\forall \mu \in S^m, \quad \sum_{j=1}^{m} \mu_j \varphi(\overline{x}, y_j) = \varphi_M(\overline{\lambda}, \mu) \leq 0$$

where we set $\overline{x} := \sum_{i=1}^m \overline{\lambda}_i y_i \in co(M) \subset K$. We then deduce that

$$v_M := \inf_{x \in K} \max_{y_j \in M} \varphi(x, y_j) \leq \max_{y_j \in M} \varphi(\overline{x}, y_j) \leq \sup_{\mu \in S^m} \sum_{j=1}^{m} \mu_j \varphi(\overline{x}, y_j) \leq 0$$

and thus, that $v = \sup_{M \in \mathcal{S}} v_M$ is nonpositive.

It remains now to prove the existence of $\overline{x} \in K$ such that

$$\sup_{y \in K} \varphi(\overline{x}, y) \leq v$$

This results from the following

Lemma 3.1.2 *Let $K \subset X$ be a compact subset of a topological space X, L be any subset and $\varphi : K \times L \mapsto \mathbf{R}$ a function such that $x \mapsto \varphi(x, y)$ is lower semicontinuous for every $y \in L$. Let $\mathcal{S}$ be the family of finite subsets of L.*

Then there exists $\overline{x} \in K$ satisfying

$$\sup_{y \in L} \varphi(\overline{x}, y) \leq v := \sup_{M \in \mathcal{S}} \inf_{x \in K} \max_{y_j \in M} \varphi(x, y_j)$$

Proof — We introduce the subsets

$$S_y := \{x \in K \mid \varphi(x,y) \leq v\}$$

They are nonempty closed subsets of the compact K because $\varphi(\cdot, y)$ is lower semicontinuous. The very definition of v implies that these subsets enjoy the finite intersection property: indeed, consider n such subsets S_{y_i} and set $M := \{y_1, \ldots, y_n\}$. The compactness of K and the lower semicontinuity of $\varphi(\cdot, y)$ implies the existence of an element $x_M \in K$ minimizing the function $\max_{y_i \in M} \varphi(\cdot, y_i)$, which then satisfies

$$\max_{y_i \in M} \varphi(x_M, y_i) = \inf_{x \in K} \max_{y_i \in M} \varphi(x, y_i) \leq v$$

so that it belongs to $\bigcap_{i=1}^n S_{y_i}$.

Hence there exists some $\overline{x}$ in their nonempty intersection $\bigcap_{y \in K} S_y$, which is a solution to the desired inequality. □

3.2 Equilibrium and Fixed Point Theorems

3.2.1 The Equilibrium Theorem

We are ready to derive from the Ky Fan inequality the Constrained Equilibrium Theorem. It provides sufficient conditions for an upper semicontinuous set-valued map F from a Banach space X to closed convex subsets of X to have an equilibrium satisfying constraints, i.e., a solution $\overline{x}$ to the inclusion $0 \in F(\overline{x})$ required to belong to a given closed subset K.

Such a set K defining the constraints will be assumed to be at least *convex and compact*, properties which are usually demanded for any existence theorem derived from the Brouwer Fixed Point Theorem[7].

But we have to provide a further assumption which links the set K with the set-valued map F: we shall require that K be a *viability domain of the set-valued map F*. We recall that the *tangent cone* $T_K(x)$ to a convex subset K at $x \in K$ is the closed cone spanned by $K - x$, which is convex:

$$T_K(x) = \overline{\bigcup_{h>0} \frac{K-x}{h}}$$

[7] up to the reduction to this case, using *coerciveness* to trade-off with compactness, as we shall see at the end of this section, for instance.

(see Section 2 of Chapter 4 for a detailed exposition of tangent cones to convex subsets.)

A subset $K \subset \mathrm{Dom}(F)$ is said to be a *viability domain* of F if and only if

$$\forall\, x \in K, \quad F(x) \cap T_K(x) \neq \emptyset$$

This means that for any point $x \in K$, there exists at least a direction $v \in F(x)$ which is tangent to K at x. We shall see in Chapter 10 that this means also that from any initial state $x_0 \in K$, there exists a solution $x(\cdot)$ to the differential inclusion

$$x' \in F(x), \quad x(0) = x_0$$

which is *viable* in K in the sense that $x(t) \in K$ for any $t \geq 0$.

Theorem 3.2.1 (Equilibrium Theorem) *Assume that X is a Banach space*[8] *and that $F : X \rightsquigarrow X$ is an upper hemicontinuous set-valued map with closed convex images.*

If $K \subset X$ is a convex compact viability domain of F, then it contains an equilibrium of F.

Proof — We proceed by contradiction, assuming that the conclusion is false. Hence, for any $x \in K$, 0 does not belong to $F(x)$. Since the images of F are closed and convex, the Hahn-Banach Separation Theorem implies that there exists $p_x \in X^\star$ such that $\sigma(F(x), p_x) < 0$.

By setting

$$\mathcal{V}_p := \{x \in K \mid \sigma(F(x), p) < 0\}$$

the negation of the existence of an equilibrium of F in K implies that K can be covered by the subsets $\mathcal{V}_p$ when p ranges over the dual of X. These subsets are open by the very definition of upper hemicontinuity of F. So K can be covered by n such open subsets $\mathcal{V}_{p_i}$. Let us consider a continuous partition of unity $(\alpha_i)_{i=1,\ldots,n}$ associated with

[8]Actually, the proof we give shows that this fundamental theorem remains true for any Hausdorff locally convex topological vector space and in particular for a Banach space endowed with the weak topology.

this finite open covering and introduce the function $\varphi : K \times K \to \mathbf{R}$ defined by

$$\varphi(x,y) := \sum_{i=1}^{n} \alpha_i(x) < p_i, x - y >$$

Being continuous with respect to x and affine with respect to y, the assumptions of Ky Fan's Inequality (Theorem 3.1.1) are satisfied. Hence there exists $\overline{x} \in K$ such that for $\overline{p} := \sum_{i=1}^{n} \alpha_i(\overline{x})p_i$ we have

$$\forall\, y \in K, \quad \varphi(\overline{x}, y) \ = < \overline{p}, \overline{x} - y > \leq \ 0$$

The above inequality means that $-\overline{p}$ belongs to the polar cone $T_K(\overline{x})^-$ of the convex subset K at $\overline{x}$.

Since K is a viability domain of F, there exists $v \in F(\overline{x}) \cap T_K(\overline{x})$, and thus

$$\sigma(F(\overline{x}), \overline{p}) \ \geq < \overline{p}, v > \geq \ 0$$

We set

$$I(\overline{x}) \ := \ \{i = 1, \ldots, n \mid \alpha_i(\overline{x}) \ > \ 0\}$$

which is not empty. Hence

$$\sigma(F(\overline{x}), \overline{p}) \leq \sum_{i \in I(\overline{x})} \alpha_i(\overline{x}) \sigma(F(\overline{x}), p_i) \ < \ 0$$

because, for any $i \in I(\overline{x})$, $\alpha_i(\overline{x}) > 0$, and thus, $\overline{x}$ belongs to the subset $\mathcal{V}_{p_i}$, which means precisely that $\sigma(F(\overline{x}), p_i) < 0$. The latter inequality is then a contradiction of the previous one. □

By modifying slightly the proof of the Equilibrium Theorem, we can prove the existence of zeros of a set-valued map from a Banach space X to another Banach space Y.

Theorem 3.2.2 (Solvability Theorem) *Let K be a convex compact subset of a Banach space X and F be an upper hemicontinuous set-valued map with closed convex values from K to another Banach space Y.*

Let us consider also a continuous map $B : K \mapsto \mathcal{L}(X,Y)$. If K, F and B are related by the condition

$$\forall\, x \in K, \quad F(x) \cap \overline{B(x)T_K(x)} \ \neq \ \emptyset$$

then

$$\begin{cases} i) & \exists\, \overline{x} \ \in \ K \quad \textit{such that } 0 \ \in \ F(\overline{x}) \\ ii) & \forall\, y \ \in \ K, \ \exists\, \widehat{x} \ \in \ K \quad \textit{such that } B(\widehat{x})y \ \in \ B(\widehat{x})\widehat{x} + F(\widehat{x}) \end{cases}$$

Proof — The proof of the existence of an equilibrium $\overline{x} \in K$ of F is the same as the one of the Equilibrium Theorem, where we define the function φ by

$$\varphi(x,y) := \sum_{i=1}^{n} \alpha_i(x) < p_i, B(x)(x-y) >$$

Ky Fan's Inequality thus implies the existence of $\overline{x} \in K$ such that

$$-B(\overline{x})^\star \overline{p} \in T_K(\overline{x})^-$$

so that, taking a sequence $u_n \in T_K(\overline{x})$ such that $B(\overline{x})u_n$ converges to some

$$v \in F(\overline{x}) \cap \overline{B(\overline{x})T_K(\overline{x})}$$

(which exists by the tangential condition), we infer that

$$\left\{ \begin{array}{l} \sigma(F(\overline{x}), \overline{p}) \\ \geq < \overline{p}, v > = \lim_{n\to\infty} < \overline{p}, B(\overline{x})u_n > = \lim_{n\to\infty} < B(\overline{x})^\star \overline{p}, u_n > \geq 0 \end{array} \right.$$

This inequality is contradicted as in the proof of Theorem 3.2.1.

Take now $y \in K$ and introduce the set-valued map $G : X \leadsto Y$ defined by

$$G(x) := F(x) + B(x)(x-y)$$

which also satisfies the assumptions of our theorem. Then there exists a zero $\widehat{x} \in K$ of G, which is a solution to the inclusion $B(\widehat{x})y \in B(\widehat{x})\widehat{x} + F(\widehat{x})$. □

3.2.2 Fixed Point Theorems

These theorems are equivalent to the Brouwer Fixed Point Theorem; we shall prove here only this equivalence[9].

We begin by showing that Theorem 3.2.1 implies the Kakutani Fixed Point Theorem[10], which is the set-valued version of the Brouwer

[9]See [45, Appendix B] for a proof of the Brouwer Fixed Point Theorem based on Sperner's Lemma and [35, Chapter II] for a proof based on differential geometry.

[10]called Ky Fan's Fixed Point Theorem in infinite dimensional spaces.

The story began in 1910 with the Brouwer Fixed Point Theorem, which was proved later in 1926 via the *Three Polish Lemma*, the three Poles being Knaster, Kuratowski and Mazurkiewicz. Knaster saw the connection between Sperner's Lemma and the fixed point theorem, Mazurkiewicz provided a proof corrected by Kuratowski. The extension to Banach spaces was proved in 1930 by their colleague Schauder.

Von Neumann did need the set-valued version of this Fixed Point Theorem in game theory, which was proved by Kakutani in 1941.

Fixed Point Theorem.

Theorem 3.2.3 (Kakutani Fixed Point Theorem) *Let K be a convex compact subset of a Banach space X and $G : X \rightsquigarrow K$ be an upper hemicontinuous set-valued map with nonempty closed convex values. Then G has a fixed point*[11] $\overline{x} \in K \cap G(\overline{x})$.

Proof — We set $F(x) := G(x) - x$, which is also upper hemicontinuous with convex values. Since K is convex, then $K - x \subset T_K(x)$, and since $G(K) \subset K$, we deduce that K is a viability domain of F because

$$F(x) \subset K - x \subset T_K(x)$$

Hence there exists a viable equilibrium $\overline{x} \in K$ of F, which is a fixed point of G. □

Actually, we do not need to assume that G maps K to itself. It is enough to assume that K is a viability domain of $F := G - \mathbf{1}$, which can be written in the following form

$$\forall\, x \in K, \quad G(x) \cap (x + T_K(x)) \neq \emptyset \tag{3.3}$$

This leads to the following

Definition 3.2.4 (Inward & Outward Maps) *A map $G : K \rightsquigarrow X$ satisfying property (3.3) is said to be* inward. *It is called* outward *if*

$$\forall\, x \in K, \quad G(x) \cap (x - T_K(x)) \neq \emptyset$$

Since K is a viability domain of $F := G - \mathbf{1}$ when G is inward and of $F_- := \mathbf{1} - G$ when G is outward, and since the equilibria of F and F_- are fixed points of G, we obtain the useful

Theorem 3.2.5 *Let K be a convex compact subset of a Banach space X and $G : K \rightsquigarrow X$ be an upper hemicontinuous map with nonempty closed convex values. If G is either inward or outward, it has a fixed point*

$$\overline{x} \in K \cap G(\overline{x})$$

[11] which can be regarded as an equilibrium for the discrete set-valued dynamical system $x_{n+1} \in G(x_n)$.

Since the Kakutani Fixed Point Theorem implies obviously the Brouwer Fixed Point Theorem, we have proved that all these statements are equivalent. □

Remark — We can deduce from the Viability Theorem and other properties of differential inclusions two other results on the existence of an equilibrium of a set-valued map F.

We shall mention them in Chapter 10. □

We now provide a convenient sufficient condition implying that a set-valued map G is outward. We say that

$$\partial\sigma_K(p) := \{x \in K \mid < p, x > = \sigma_K(p)\}$$

is the *extremal face* of K at p. Observe that $x \in \partial\sigma_K(p)$ if and only if $p \in T_K(x)^-$.

Proposition 3.2.6 *Let K be a nonempty closed convex subset of X and let G be a set-valued map from K to K satisfying*

$$\forall\, p \in X^\star,\ \forall\, x \in \partial\sigma_K(p),\ \ G(x) \cap \partial\sigma_K(p) \neq \emptyset$$

Then G is outward. In particular, this is the case of a map G satisfying

$$\forall\, p \in X^\star,\ \ G(\partial\sigma_K(p)) \subset \partial\sigma_K(p)$$

Proof — Let $x \in \partial\sigma_K(p)$ and

$$y_0 \in G(x) \cap \partial\sigma_K(p)$$

be fixed. Hence, $v = y_0 - x$ belongs to $G(x) - x$ and $-T_K(x)$ because,

$$\forall\, p \in T_K(x)^-,\ < p, v > = < p, y_0 > - < p, x > = \sigma_K(p) - \sigma_K(p) = 0 \ \square$$

We can relax the assumption that K is compact, replacing it by a *coerciveness assumption.*

Theorem 3.2.7 *Let K be a closed convex subset of a finite dimensional space and let F be an upper hemicontinuous map with nonempty compact convex images satisfying the coerciveness assumption*

$$\limsup_{\substack{\|x\|\to\infty \\ x\in K}} \sigma(F(x), x) < 0 \tag{3.4}$$

We posit the tangential condition

$$\forall\, x \in K,\ F(x) \cap T_K(x) \neq \emptyset$$

Then there exists an equilibrium $\overline{x} \in K$ *of* F. *If we posit the stronger coerciveness assumption,*

$$\limsup_{\substack{\|x\| \to \infty \\ x \in K}} \frac{\sigma(F(x), x)}{\|x\|} \;=\; -\infty \tag{3.5}$$

Then for all $y \in K$, *there exists a solution* $\widehat{x} \in K$ *to the inclusion* $y \in \widehat{x} - F(\widehat{x})$.

Proof — Coerciveness assumption (3.4) implies that there exist $\varepsilon > 0$ and $a > 0$ such that

$$\sup_{\substack{\|x\| \geq a \\ x \in K}} \sigma(F(x), x) \;\leq\; -\varepsilon \;<\; 0$$

This implies that for all $x \in K$ with $\|x\| \leq \alpha$, $F(x) \subset T_{aB}(x)$. By taking the number a large enough so that $K \cap a \overset{o}{B} \neq \emptyset$, we know that

$$T_{K \cap aB}(x) \;=\; T_K(x) \cap T_{aB}(x)$$

(see Chapter 4, Section 2.) So the tangential condition implies that

$$\forall\, x \in K \cap aB,\; F(x) \cap T_{K \cap aB}(x) \;\neq\; \emptyset$$

It is sufficient to apply Theorem 3.2.1. To prove the second part of the theorem, we fix $y \in K$ and replace F by the map G defined by

$$G(x) \;=\; F(x) + y - x$$

which satisfies the coerciveness assumption (3.4) whenever F satisfies the stronger coerciveness assumption (3.5.) □

3.2.3 The Leray-Schauder Theorem

We can deduce the set-valued version of the Leray-Schauder Theorem[12] on the existence of stationary points from Theorem 3.2.1 by *Poincaré's continuation method*. We take $X = \mathbf{R}^n$ and K to be a compact convex subset with a nonempty interior, so that the boundary

$$\partial K \;=\; K \setminus \operatorname{Int}(K)$$

of K is distinct from K.

[12] proved in 1934.

Theorem 3.2.8 (Leray-Schauder) *Let us consider a compact convex subset $K \subset \mathbf{R}^n$ with nonempty interior and an upper hemicontinuous set-valued map F from $[0,1] \times K$ to $\mathbf{R}^n$, with nonempty closed convex values. Suppose the set-valued map $x \rightarrow F(0,x)$ satisfies the tangential condition*

$$\forall\, x \,\in\, \partial K, \;\; F(0,x) \cap T_K(x) \,\neq\, \emptyset \tag{3.6}$$

and

$$\forall\, x \,\in\, \partial K, \; \forall\, \lambda \,\in\, [0,1[, \;\; 0 \,\notin\, F(\lambda,x) \tag{3.7}$$

Then there exists $\overline{x} \in K$ such that $0 \in F(1,\overline{x})$.

Proof — We suppose that the conclusion is false and derive a contradiction. We set $N := \partial K$, a closed subset of K and introduce the subset

$$M \;:=\; \{x \in K \mid \exists\, \lambda \in [0,1] \text{ satisfying } 0 \,\in\, F(\lambda,x)\}$$

The subset M is nonempty because it contains a solution $\tilde{x} \in K$ to the inclusion $0 \in F(0,\tilde{x})$, which exists by Theorem 3.2.1 (thanks to assumption (3.6).) It is closed, since the graph of F is closed. The intersection $M \cap N$ is empty : If $x \in N$ and if $\lambda \in [0,1[$, assumption (3.7) implies that $x \notin M$.

Consider a continuous function φ from K to $[0,1]$

$$\varphi(x) \;:=\; \frac{d(x,N)}{d(x,M) + d(x,N)}$$

which is equal to zero on N and to one on M and define the set-valued map G on K by

$$G(x) \;:=\; F(\varphi(x),x)$$

Then G is clearly upper hemicontinuous with nonempty closed convex values. It coincides with $F(0,\cdot)$ on $N = \partial K$ and, consequently, satisfies the assumptions of Theorem 3.2.1. Hence, there exists a solution $\overline{x} \in K$ to the inclusion

$$0 \,\in\, G(\overline{x}) \;=\; F(\varphi(\overline{x}),\overline{x})$$

But this implies that $\overline{x} \in M$ and, therefore, $\varphi(\overline{x}) = 1$, so that $0 \in F(1,\overline{x})$ which is a contradiction. □

3.3 Ekeland's Variational Principle

Ekeland's Theorem, we are about to prove, provides an approximate minimizer of a bounded from below lower semicontinuous function in a given neighborhood of a point. This localization property is very useful and explains the importance of this result which has been used extensively since its discovery in 1974.

Theorem 3.3.1 (Ekeland's Variational Principle) *Let*

$$V : X \mapsto \mathbf{R} \cup \{+\infty\}$$

be a lower semicontinuous nontrivial extended bounded from below function defined on a complete metric space X. Let $x_0 \in \mathrm{Dom}(V)$ and $\varepsilon > 0$ be fixed. Then there exists $\overline{x} \in X$, a solution to

$$\left\{ \begin{array}{ll} i) & V(\overline{x}) + \varepsilon d(x_0, \overline{x}) \ \leq \ V(x_0) \\ & \\ ii) & \forall\, x \neq \overline{x}, \ \ V(\overline{x}) \ < \ V(x) + \varepsilon d(x, \overline{x}) \end{array} \right. \tag{3.8}$$

Proof — We can take $\varepsilon := 1$ and assume that V is non negative by replacing it by $V(\cdot) - \inf V$ if needed. We associate with V the set-valued map $F : X \rightsquigarrow X$ defined by

$$F(x) \ := \ \{y \in X \mid V(y) + d(x, y) \leq V(x)\}$$

The subsets $F(x)$ are closed and F is a preorder (given by $y \succ x \iff y \in F(x)$) in the sense that

$$\left\{ \begin{array}{llr} i) & \forall\, x \ \in \ \mathrm{Dom}(V), \ \ x \ \in \ F(x) & \text{(reflexivity)} \\ ii) & y \ \in \ F(x) \ \Longrightarrow \ F(y) \ \subset \ F(x) & \text{(transitivity)} \end{array} \right.$$

Indeed, the second condition is obvious when $x \notin \mathrm{Dom}(V)$, since in this case $F(x) = X$.

Assume now that $V(x)$ is finite. By taking $y \in F(x)$, $z \in F(y)$, adding inequalities

$$V(z) + d(y, z) \ \leq \ V(y) \quad \& \quad V(y) + d(x, y) \ \leq \ V(x)$$

and using the triangular inequality, we obtain $V(z) + d(z, x) \leq V(x)$, i.e., $z \in F(x)$.

Let us define the function g on $\mathrm{Dom}(V)$ by

$$\forall\, y \in \mathrm{Dom}(V), \;\; g(y) := \inf_{z \in F(y)} V(z)$$

It is clear that

$$\forall\, y \in F(x), \;\; d(x,y) \le V(x) - g(x)$$

so that the diameter $\sup_{y,z \in F(x)} \|y - z\|$ of $F(x)$ is not greater than $2(V(x) - g(x))$.

Define now the following sequence starting at x_0: we take $x_{n+1} \in F(x_n)$ such that

$$V(x_{n+1}) \le g(x_n) + 2^{-n}$$

Since $F(x_{n+1}) \subset F(x_n)$ by transitivity, we deduce that

$$g(x_n) \le g(x_{n+1})$$

On the other hand, inequality $g(y) \le V(y)$ implies that

$$g(x_{n+1}) \le V(x_{n+1}) \le g(x_n) + 2^{-n} \le g(x_{n+1}) + 2^{-n}$$

and thus, that

$$0 \le V(x_{n+1}) - g(x_{n+1}) \le 2^{-n}$$

Therefore, the diameters of the closed subsets $F(x_n)$ converge to 0. Since the sequence of these closed subsets is nonincreasing and since the metric space is complete, we infer that

$$\bigcap_{n \ge 0} F(x_n) = \{\overline{x}\}$$

This implies that $\overline{x} \in F(x_0)$, i.e., inequality (3.8) i). We also observe that $F(\overline{x}) = \{\overline{x}\}$. Indeed, $\overline{x}$ belonging to all the subsets $F(x_n)$, then, by transitivity, $F(\overline{x})$ is contained in all the $F(x_n)$'s, and thus, in $\{\overline{x}\}$.

Consequently, when $x \ne \overline{x}$, then x does not belong to $F(\overline{x})$, i.e.,

$$V(x) + d(x, \overline{x}) > V(\overline{x})$$

This is inequality (3.8) ii). □

3.4 Constrained Inverse Function Theorem

3.4.1 Derivatives of Single-Valued Maps

We recall first some classical definitions and notations.

Definition 3.4.1 *Let Ω be an open subset of a normed space X and*

$$f : \Omega \mapsto Y$$

be a single-valued map from Ω to a normed space[13] *Y. We denote by*

$$\nabla_h f(x) : v \mapsto (f(x+hv) - f(x))/h$$

---

[13]The concepts of functionals and of derivatives of functionals are due to Volterra in 1887. Then Gâteaux, in a note written in 1913 and published in 1919 after his death during the First World War, introduced the concept of *first variation*

$$\delta f(x) := \left[\frac{d}{d\lambda} f(x + \lambda \delta x)\right]_{\lambda=0}$$

As we can see, Gâteaux's definition did not involve linearity with respect to the increment — long before our time when we are at last ready not to require linearity anymore in order to go forward — and for that, was criticized by Hadamard and his former student, Fréchet. Mathematicians of this period still asked many properties for the derivatives of functionals and were not ready to give them away. But Gâteaux was forgiven, and even *rewarded* posthumously since linearity was introduced in the definition of the Gâteaux derivative!

Fréchet proposed his concept of derivative (with the mandatory linearity) for a function as early as in 1912 in the case of functions and in 1925 for maps from a normed space to another one.

However, he did not choose the stronger concept (called *strong* or *strict* differentiability) derived from a definition of the derivative of a function that Peano suggested in 1892

$$f'(x) := \lim_{(x_1,x_2)\to(x,x)} \frac{f(x_1) - f(x_2)}{x_1 - x_2}$$

on the ground of convincing physical motivations. This track was not used until 1961 (by S. Leach to give an improved statement of the Inverse Function Theorem.) However, Peano's point of view on how to take limits is very much the one which underlies the concepts of Clarke tangent cones and paratingent cones which will be studied in Chapter 4.

Since then, the history of derivatives in Banach spaces is a kind of *mathematical striptease*, where one by one, and very shyly, the required properties of the derivative of a functional are taken away. We shall go quite far in Chapter 6 to leave the derivatives with the bare minimum.

This led to a ménagerie of concepts: strong or weak Fréchet and Gâteaux derivatives, Hadamard, bounded (Suchomlinov), locally uniform (Vainberg) derivatives, directional semiderivatives or derivatives from the right, Sindalowsky *the* differential quotient of f at x in the direction v.

If the differential quotients $\nabla_h f(x)(v)$ *converge to a limit denoted by* $df(x)(v)$, *then* $df(x)(v)$ *is called* the Dini directional derivative *of* f *at* x *in the direction* v.

If the map $v \mapsto df(x)(v)$ *is linear and continuous, we say that* f *is* Gâteaux differentiable *at* x *and the continuous linear operator*

$$f'(x) : v \mapsto df(x)(v) =: f'(x)v$$

is called the derivative *of* f *at* x. *When* f *is real-valued, we say that* $f'(x) \in X^\star$ *is the* gradient *of* f *at* x.

If furthermore

$$\lim_{y \to x} \frac{f(y) - f(x) - f'(x)(y-x)}{\|y - x\|} = 0$$

then f *is said to be* Fréchet differentiable *at* x. *When the derivative* $y \mapsto f'(y)$ *is continuous at* x, f *is called* continuously differentiable (*or* of class C^1) *at* x.

Recall that if f is continuously differentiable, it is Fréchet differentiable and when f is Fréchet differentiable, it is continuous.

Naturally, there are many other ways to take limits[14] of the differential quotients, using the weak topologies of either X or Y or both, taking limsup's or liminf's, etc.

3.4.2 Constrained Inverse Function Theorems

Let us consider now a Banach space X, a normed space Y, a closed subset $K \subset X$ and a continuous (single-valued) map $f : K \mapsto Y$.

---

or Muravev variations, and the ones we shall study in Chapter 6 and quote in the bibliographical comments of this chapter.

We shall face a similar situation in the set-valued case. This is quite natural, though, because each problem demands its own amount of properties that the derivative should enjoy, i.e., its own degree of regularity. Without going too far by always requiring minimal assumptions, some problems could not be solved by sticking to the richest structure. The right balance between generality and readability is naturally a subjective choice.

[14] too many of them having names of French mathematicians..., according at least to some American ones.

We shall deduce from Ekeland's Theorem a "nonlinear extension" of Graves' Theorem 3.4.2, which allows one to solve (locally) a "constrained problem" of the form

$$\text{find } x \in K \quad \text{such that } f(x) = y$$

for any right-hand side y in a neighborhood of some

$$y_0 := f(x_0) \in f(K)$$

When K is equal to the whole space X, a solution to this unconstrained problem is given by Graves' version of the Inverse-Function Theorem[15]:

Theorem 3.4.2 (Graves' Theorem) *Let X, Y be Banach spaces and $f : X \mapsto Y$ a continuous (single-valued) map. We assume that f is continuously differentiable on a neighborhood of x_0 and that*

$$f'(x_0) \text{ is surjective}$$

Then the set-valued map $y \rightsquigarrow f^{-1}(y)$ is pseudo-Lipschitz *around $(f(x_0), x_0)$.*

Proof — It is an obvious consequence of Theorem 3.4.10 below with $K := X$. □

For constrained problems, i.e., when K is no longer the whole space, we need to introduce right now the concept of contingent cone, the exhaustive study of which is the purpose of Chapter 4.

The contingent cone $T_K(x)$ to $K \subset X$ at x is the upper limit of the subsets $(K - x)/h$ when $h \to 0+$:

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

Graves' Inverse-Function Theorem can be extended to our constrained problem provided that an adequate transversality condition holds true:

[15]This result is sometimes called in the literature Ljusternik's theorem. However Graves was the first to state it in 1950, while Ljusternik investigated in 1934 tangent spaces to level sets of a function. Both Graves and Ljusternik used in their proofs an approximation procedure which is a standard technique of functional analysis.

Theorem 3.4.3 (Constrained Inverse Function Theorem) *Let X be a Banach space and Y be a normed space. We consider a (single-valued) continuous map $f : X \mapsto Y$, a closed subset $K \subset X$ and an element x_0 of K.*

We assume that f is Fréchet differentiable on a neighborhood of x_0 and we posit the following transversality assumption*:*

there exist constants $c > 0$, $\alpha \in [0, 1[$ and $\eta > 0$ such that

$$\forall\, x \in K \cap B(x_0, \eta), \quad B_Y \subset f'(x)(T_K(x) \cap cB_X) + \alpha B_Y$$

Then $f(x_0)$ belongs to the interior of $f(K)$ and the set-valued map $y \rightsquigarrow f^{-1}(y) \cap K$ is pseudo-Lipschitz *around $(f(x_0), x_0)$.*

This theorem is a consequence of the still more general Theorem 3.4.5 below:

Indeed, not only are we interested in knowing whether a solution to the constrained problem does exist, but we wish to approximate it by solutions to the approximate problems

$$\text{find } x_n \in K_n \quad \text{such that } f_n(x_n) = y_n$$

where y_n converges to y_0, f_n converges to f and K_n to K in the following sense:

Definition 3.4.4 *We say that the maps f_n are* consistent *with f on K_n at $x_0 \in K$ if there exist elements $x_{0n} \in K_n$ converging to x_0 such that $f_n(x_{0n})$ converges to $f(x_0)$.*

We have to add to this consistency assumption a stability assumption to be able to approximate the solution of the constrained problem:

Theorem 3.4.5 *Let X be a Banach space and Y a normed space. Consider a sequence of continuous single-valued maps f_n from X to Y, a sequence of closed subsets $K_n \subset X$ and an element x_0 in the lower limit of the subsets K_n such that the maps f_n are* consistent *with f on K_n at x_0.*

We assume that f_n are Fréchet differentiable on a neighborhood of x_0 and verify the following stability assumption*:*

there exist constants $c > 0$, $\alpha \in [0, 1[$ *and* $\eta > 0$ *such that*

$$\forall\, x \in K_n \cap B(x_0, \eta), \quad B_Y \subset f_n'(x)(T_{K_n}(x) \cap cB_X) + \alpha B_Y \quad (3.9)$$

Then there exist $l > 0$ *and* $\gamma > 0$ *such that for any* $x_{0n} \in B_{K_n}(x_0, \gamma)$ *satisfying* $f_n(x_{0n}) \in B(f(x_0), \gamma)$, *we have*

$$\forall\, y_n \in B(f(x_0), \gamma), \quad d\left(x_{0n}, f_n^{-1}(y_n) \cap K_n\right) \leq l\|y_n - f_n(x_{0n})\|$$

When the functions $f_n \equiv f$ and the subsets $K_n \equiv K$ are constant, the conclusion of the theorem means simply that f^{-1} is pseudo-Lipschitz at $(f(x_0), x_0)$, so that Theorem 3.4.3 follows obviously from it.

Before proving the above theorem we show that it implies Theorem 1.2.9 from Chapter 1.

Proof of Theorem 1.2.9 — Let $u_0 \in f^{-1}(\text{Liminf}_{n\to\infty} M_n)$: There exists a sequence of elements $v_{0n} \in M_n$ converging to $f(u_0)$. We shall construct a sequence of elements $u_n \in f^{-1}(M_n)$ converging to u_0 thanks to Theorem 3.4.5.

Indeed, we apply it with $X \times Y$ playing the role of X, the map

$$f \ominus \mathbf{1} : (u, v) \mapsto f(u) - v$$

playing the role of f and f_n, the sets $K_n := X \times M_n$, $x_{0n} = (u_0, v_{0n})$ and $y_n = 0$. The transversality assumption of Theorem 1.2.9 implies that for every $x \in \mathcal{U}$

$$B_Y \subset f'(x)(dB_X) - T_{M_n}(y) \cap dB_Y$$

(with $d = c\sup_{x\in\mathcal{U}} \|f'(x)\| + 1$) which can be written

$$B_Y \subset (f \ominus \mathbf{1})'(x)\,(T_{K_n}(x) \cap 2dB_X)$$

Hence the transversality condition is satisfied with $\alpha = 0$.

Therefore, there exist a constant $l > 0$ and solutions

$$x_n = (u_n, v_n) \in X \times M_n \text{ to } (f \ominus \mathbf{1})(x_n) = f(u_n) - v_n = 0$$

satisfying

$$\begin{cases} \|x_n - x_{0n}\| & = \|u_0 - u_n\| + \|v_{0n} - v_n\| \\ & \leq l\|y_n - (f \ominus \mathbf{1})(x_{0n})\| = l\|f(u_0) - v_{0n}\| \end{cases}$$

Hence $u_n \in f^{-1}(M_n)$ and converges to u_0, so that u_0 belongs to the lower limit of the inverse images $f^{-1}(M_n)$. □

Proof of Theorem 3.4.5 — We choose $\rho > 0$, $\varepsilon > 0$ such that

$$\frac{3\rho}{\eta} < \varepsilon < \frac{1-\alpha}{c}$$

and consider elements x_{0n} satisfying

$$\|x_{0n} - x_0\| \leq \eta/3 \quad \& \quad \|f_n(x_{0n}) - y_0\| \leq \eta/3$$

and elements $y_n \in B(f_n(x_{0n}), \rho)$.

By Ekeland's Variational Principle 3.3.1 applied to the function

$$V(x) := \|y_n - f_n(x)\|$$

on the complete metric space K_n, we know that there exists a solution $\widehat{x}_n \in K_n$ to

$$\begin{cases} i) & \|y_n - f_n(\widehat{x}_n)\| + \varepsilon\|\widehat{x}_n - x_{0n}\| \leq \|y_n - f_n(x_{0n})\| \\ ii) & \forall\, x_n \in K_n,\ \|y_n - f_n(\widehat{x}_n)\| \leq \|y_n - f_n(x_n)\| + \varepsilon\|x_n - \widehat{x}_n\| \end{cases} \tag{3.10}$$

We deduce from inequality (3.10) $i)$ that

$$\|\widehat{x}_n - x_{0n}\| \leq \frac{1}{\varepsilon}\|y_n - f_n(x_{0n})\| \leq \frac{\rho}{\varepsilon} < \frac{\eta}{3}$$

so that

$$\|\widehat{x}_n - x_0\| \leq \eta/3 + \|x_{0n} - x_0\| \leq 2\eta/3$$

Stability assumption (3.9) implies that there exist $u_n \in T_{K_n}(\widehat{x}_n)$ and $w_n \in Y$ satisfying

$$\begin{cases} i) & y_n - f_n(\widehat{x}_n) = f_n'(\widehat{x}_n)u_n + w_n \\ ii) & \|u_n\| \leq c\|y_n - f_n(\widehat{x}_n)\| \ \& \ \|w_n\| \leq \alpha\|y_n - f_n(\widehat{x}_n)\| \end{cases}$$

By definition of the contingent cone, there exist elements $h_p > 0$ and $e_p \in X$ converging to 0+ and 0 respectively such that, f_n being Fréchet differentiable,

$$\begin{cases} x_n := \widehat{x}_n + h_p u_n + h_p e_p \in K_n \\ f_n(x_n) = f_n(\widehat{x}_n) + h_p(f_n'(\widehat{x}_n)(u_n + e_p) + \varepsilon(h_p)) \end{cases}$$

where $\varepsilon(h_p)$ converges to 0 with h_p.

By taking in inequality (3.10) *ii)* such an x_n, by observing that

$$y_n - f_n(x_n) = (1-h_p)(y_n - f_n(\widehat{x}_n)) + h_p w_n - h_p(f_n'(\widehat{x}_n)e_p + \varepsilon(h_p))$$

we deduce that

$$h_p\|y_n - f_n(\widehat{x}_n)\| \leq h_p\|w_n\| + h_p\|f_n'(\widehat{x}_n)e_p + \varepsilon(h_p)\| + \varepsilon h_p\|u_n + e_p\|$$

Dividing by $h_p > 0$ and letting $p \to +\infty$, we get:

$$\|y_n - f_n(\widehat{x}_n)\| \leq \|w_n\| + \varepsilon\|u_n\| \leq (\alpha + \varepsilon c)\|y_n - f_n(\widehat{x}_n)\|$$

Since we have chosen ε such that $\alpha + \varepsilon c < 1$, we infer that $\widehat{x}_n$ is a solution to

$$\widehat{x}_n \in K_n \quad \& \quad f_n(\widehat{x}_n) = y_n$$

satisfying

$$d\left(x_{0n}, f_n^{-1}(y_n) \cap K_n\right) \leq \|\widehat{x}_n - x_{0n}\| \leq \frac{1}{\varepsilon}\|y_n - f_n(x_{0n})\|$$

from which the error estimate follows.

Observe also that by letting ε converge to $c/(1-\alpha)$, we obtain the estimate

$$d(x_{0n}, f_n^{-1}(y_n) \cap K_n) \leq c\|y_n - f_n(x_{0n})\|/(1-\alpha) \quad \square$$

Remark — This theorem extends to nonlinear constrained problems Lax's principle that *consistency and stability imply convergence*. Due to the importance of this principle in numerical analysis, we reformulate it in this framework:

Theorem 3.4.6 *Let us consider a sequence of continuous single-valued maps f_n from a Banach space X to a normed space Y, a sequence of closed subsets $K_n \subset X$ and an element*

$$x_0 \in \mathrm{Liminf}_{n\to\infty} K_n$$

such that the maps f_n are consistent *with f on K_n at x_0.*

Assume that f_n are Fréchet differentiable on a neighborhood of x_0 and that stability assumption *(3.9) holds true.*

Then there exists a constant $l > 0$ such that:

if y_n converges to $f(x_0)$, then x_0 can be approximated by solutions $x_n \in K_n$ to the equation $f_n(x_n) = y_n$ satisfying the following error estimate:

Given any "reference " sequence of elements $x_{0n} \in K_n$ converging to x_0 such that $f_n(x_{0n})$ converges to $f(x_0)$, for all n large enough

$$\|x_n - x_0\| \ \leq \ l(\|x_0 - x_{0n}\| + \|f(x_0) - f_n(x_{0n}\| + \|y_n - y_0\|)$$

Remark — Case of Linear Operators

When we consider the case of continuous linear operators $f_n := A_n \in \mathcal{L}(X,Y)$, we can restrict y_n to belong to the subset $A_n(K_n)$ in the proof of Theorem 3.4.5 and thus, to relax somewhat the stability assumption. Indeed, in this case,

$$y_n - A_n\widehat{x}_n \ \in \ A_n(K_n - \widehat{x}_n)$$

When $x_n \in K_n$, we denote by

$$S_{K_n}(x_n) \ := \ \bigcup_{h>0} \frac{K_n - x_n}{h}$$

the cone spanned by $K_n - x_n$:

Theorem 3.4.7 (Inverse Stability Theorem) *Let X be a Banach space and Y a normed space. Consider a sequence of continuous linear operators $A_n \in \mathcal{L}(X,Y)$ and a sequence of closed subsets $K_n \subset X$ such that A_n are consistent with $A \in \mathcal{L}(X,Y)$ on K_n at $x_0 \in \mathrm{Liminf}_{n\to\infty}K_n$.*

We posit the following stability assumption*:*

there exist constants $c > 0$, $\alpha \in [0,1[$ and $\eta > 0$ such that

$$\begin{cases} \forall\, x_n \ \in \ K_n \cap B(x_0,\eta), \\ \\ A_nS_{K_n}(x_n) \cap B_Y \ \subset \ A_n\left(T_{K_n}(x_n) \cap cB_X\right) + \alpha B_Y \end{cases} \tag{3.11}$$

Then there exist $l > 0$ and $\gamma > 0$ such that for any $x_{0n} \in B_{K_n}(x_0,\gamma)$ satisfying $A_nx_{0n} \in B(Ax_0,\gamma)$ we have

$$\forall\, y_n \in A_n(K_n) \cap B(Ax_0,\gamma), \ \ d\left(x_{0n}, A_n^{-1}(y_n) \cap K_n\right) \ \leq \ l\|y_n - A_nx_{0n}\|$$

In particular, for fixed K_n and A_n, we obtain:

Corollary 3.4.8 (Criterion of Pseudo-Lipschitzeanity) *Consider any closed subset K of a Banach space X and let $A \in \mathcal{L}(X,Y)$ be a continuous linear operator from X to a normed space Y. Assume that for some $x_0 \in K$, there exist constants $c > 0$, $\alpha \in [0,1[$ and $\eta > 0$ such that*

$$\forall\, x \in K \cap B(x_0,\eta), \;\; AS_K(x) \cap B_Y \subset A(T_K(x) \cap cB_X) + \alpha B_Y \quad (3.12)$$

Then the set-valued map

$$A(K) \ni y \rightsquigarrow A^{-1}(y) \cap K$$

is pseudo-Lipschitz *around (Ax_0, x_0).*

3.4.3 Pointwise Stability Conditions

The presence of the constant $\alpha \in [0,1[$ in the transversality conditions of the Inverse Function Theorem 3.4.3 is useful because it enables its deduction from a weaker pointwise condition and a kind of "continuity." Namely, we introduce the following definition:

Definition 3.4.9 (Sleek and Uniformly Sleek Sets) *A closed subset K is called* sleek *at $x_0 \in K$ if the cone-valued map*

$$K \ni x \rightsquigarrow T_K(x)$$

is lower semicontinuous at x_0. It is called uniformly sleek *at $x_0 \in K$ if*

$$\lim_{x \to x_0} \left(\sup_{u \in T_K(x_0) \cap B_X} d(u, T_K(x)) \right) = 0$$

We shall prove in Theorem 4.1.8 that *the contingent cone $T_K(x_0)$ to a subset K sleek at x_0 is convex.*

Theorem 3.4.10 (Pointwise Inverse Function Theorem) *Let X and Y be Banach spaces, K be a closed subset of X and $f : X \mapsto Y$ be a continuously differentiable at $x_0 \in K$ map. Assume that*

$$f'(x_0) T_K(x_0) = Y$$

1. — If K is uniformly sleek at x_0, then $f(x_0)$ belongs to the interior of $f(K)$ and the set-valued map

$$y \rightsquigarrow f^{-1}(y) \cap K$$

is pseudo- Lipschitz *around* $(f(x_0), x_0)$.

2. — If the dimension of Y is finite, it is sufficient to assume that K is sleek at x_0 to obtain the same conclusion.

Proof — Replacing K by a closed neighborhood of x_0 in K, we may assume that f is continuous on K. We have to prove that in both cases, the stability assumption is satisfied.

1. — The proof of the first case is easy. There exists a constant $c > 0$ such that, for all v in the unit sphere S_Y, there exists a solution u_0 to the equation $f'(x_0)u = v$ such that $\|u_0\| \leq c\|v\|$, thanks to Robinson-Ursescu's Theorem (because $T_K(x_0)$ is a closed convex cone.)

Since K is uniformly sleek at x_0 and $f'(\cdot)$ is continuous at x_0, we can associate with any $\varepsilon > 0$ an $\eta > 0$ such that, for all $u_0 \in X$ and all $x \in B_K(x_0, \eta)$, there exists $u \in T_K(x)$ satisfying

$$\|u - u_0\| \leq \varepsilon\|u_0\| \quad \& \quad \|f'(x) - f'(x_0)\| \leq \varepsilon$$

Hence any $v \in S_Y$ can be written $v = f'(x)u + w$ where $u \in T_K(x)$, $\|u\| \leq (1 + \varepsilon)c\|v\|$ and $w = f'(x_0)u_0 - f'(x)u$. Thus

$$\|w\| \leq \varepsilon\|u_0\|(\|f'(x_0)\| + 1 + \varepsilon) \leq \varepsilon c\|v\|(\|f'(x_0)\| + 1 + \varepsilon)$$

Taking $\varepsilon > 0$ small enough and using Theorem 3.4.3 we complete the proof of the first statement.

2. — When the dimension of Y is finite, the unit sphere S_Y is compact. We know from the Robinson-Ursescu Theorem that for any $v_i \in S_Y$, there exits a solution $u_{0i} \in T_K(x_0)$ to the equation $f'(x_0)u_{0i} = v_i$ such that $\|u_{0i}\| \leq c\|v_i\| = c$. Fix $\varepsilon > 0$.

Since K is sleek at x_0, for any v_i, there exists $\eta_i > 0$ such that, for all $x \in B_K(x_0, \eta)$, we can find $u_i \in T_K(x)$ satisfying

$$\|u_i - u_{0i}\| \leq \varepsilon\|u_{0i}\|$$

The map $f'(\cdot)$ being continuous at x_0, there exists $\eta_0 > 0$ such that $\|f'(x) - f'(x_0)\| \leq \varepsilon$.

We can cover S_Y by p balls $B(v_i, \varepsilon)$ so that, by taking

$$\eta := \min_{0 \leq i \leq p} \eta_i$$

we obtain that for any $v \in S_Y$ and any $x \in B_K(x_0, \eta)$, there exist $u_i \in T_K(x)$ and $w_i \in Y$ satisfying

$$v = f'(x)u_i + w_i, \quad \|u_i\| \leq c(1+\varepsilon), \quad w_i = v - v_i + f'(x_0)u_{0i} - f'(x)u_i$$

As above, we can show that

$$\|w_i\| \leq \varepsilon(1 + c\|f'(x_0)\| + (1+\varepsilon)c)$$

and use Theorem 3.4.3 with ε small enough. □

3.4.4 Local Uniqueness

We conclude this section with a local uniqueness criterion of a solution to a constrained nonlinear problem:

Proposition 3.4.11 *Let K be a subset of a finite dimensional vector-space X and $f : X \mapsto Y$ a single-valued map to a normed space Y. Let $x_0 \in K$ and assume that f is Fréchet differentiable at x_0.*

If

$$T_K(x_0) \cap \ker(f'(x_0)) = \{0\}$$

then x_0 is the only local solution $x \in K$ to the equation $f(x) = f(x_0)$, in the sense that there exists a neighborhood $\mathcal{N}(x_0)$ of x_0 such that

$$\forall\, x \in \mathcal{N}(x_0) \cap K, \quad x \neq x_0, \quad f(x) \neq f(x_0)$$

Proof — Assume the contrary: for all $n > 0$, there exists

$$x_n \in K \cap B(x_0, 1/n), \quad x_n \neq x_0$$

such that $f(x_n) = f(x_0)$. Let us set

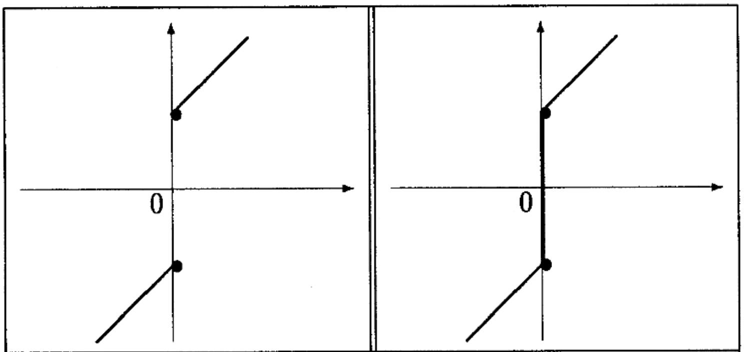
$$h_n := \|x_n - x_0\| > 0$$

which converges to 0 and

$$v_n := \frac{x_n - x_0}{h_n}$$

Since v_n belongs to the unit sphere, which is compact, a subsequence (again denoted) v_n converges to some element v of the unit sphere.

Figure 3.1: Example of Monotone and Maximal Monotone Maps



Monotone Map	Maximal Monotone Map

This limit v belongs also to the contingent cone $T_K(x_0)$ because, for all $n > 0$, $x_0 + h_n v_n = x_n$ belongs to K. Furthermore, f being Fréchet differentiable at x_0, we obtain the equation $f'(x_0)v_n = \varepsilon(h_n)$ where $\varepsilon(h_n)$ converges to 0 with h_n. We thus infer that v belongs also to the kernel of $f'(x_0)$. Since $v \neq 0$, we derive a contradiction of the assumption. □

3.5 Monotone and Maximal Monotone Maps

In this section, *we assume once and for all that X is a Hilbert space identified with its dual.*

3.5.1 Monotone Maps

Definition 3.5.1 *We shall say that a set-valued map A from X to X is* monotone *if its graph is monotone in the sense that*

$$\forall\,(x,p) \in \mathrm{Graph}(A),\ \forall\,(y,q) \in \mathrm{Graph}(A),\ \ < p-q, x-y > \;\geq\; 0$$

Example The map

$$A(x) = \begin{cases} \{x-1\} & \text{if } \ x < 0 \\ \{-1,+1\} & \text{if } \ x = 0 \\ \{x+1\} & \text{if } \ x > 0 \end{cases}$$

is a monotone map as well as the map

$$A(x) = \begin{cases} \{x-1\} & \text{if } x < 0 \\ [-1,1] & \text{if } x = 0 \\ \{x+1\} & \text{if } x > 0 \end{cases}$$

The terminology comes from the monotone maps from $\mathbf{R}$ to $\mathbf{R}$. For instance, a nondecreasing map F from $\mathbf{R}$ to $\mathbf{R}$ is monotone.

More generally, if f is a nondecreasing map from $\mathbf{R}$ to $\mathbf{R}$, the map $x \mapsto A(x) := [f(x-), f(x+)] \cap \mathbf{R}$ is monotone.

Let us give another example of monotone map. Let f be a nondecreasing function from an interval $\mathrm{Dom}(f) \subset \mathbf{R}$ to $\mathbf{R}$. Let $\Omega \subset \mathbf{R}^n$ be open and $X := L^2(\Omega)$. Define the set-valued map A from X to X by

$$\text{for almost all } \omega \in \Omega, \quad A(x(\cdot))(\omega) := f(x(\omega))$$

It is clear that A is monotone.

We shall see in Chapter 6 that the subdifferential map $x \mapsto \partial V(x)$ of a nontrivial convex function $V : X \to \mathbf{R} \cup \{+\infty\}$ is monotone. □

Since monotonicity is a property bearing on the graph of A,

$$A \textit{ is monotone if and only if } A^{-1} \textit{ is monotone}$$

If A and B are monotone and $\lambda > 0$, $\mu > 0$, then $\lambda A + \mu B$ is monotone. If we supply $X \times X$ with the product of the strong topology and the weak topology, we can see that the closure of a monotone graph is still the graph of a monotone map.

We begin by mentioning the following characterization:

Proposition 3.5.2 *A set-valued map A from X to X is monotone if and only if for every $\lambda > 0$*

$$\forall\ (x,p),\ (y,q) \in \mathrm{Graph}(A), \quad \|x-y\| \ \leq\ \|x-y+\lambda(p-q)\| \qquad (3.13)$$

Proof — We compute

$$\|x-y+\lambda(p-q)\|^2 \ =\ \|x-y\|^2 + \lambda^2\|p-q\|^2 + 2\lambda < p-q, x-y >$$

Hence, if A is monotone, inequality (3.13) ensues; conversely, if (3.13) holds true, we deduce that

$$\lambda^2\|p-q\|^2 + 2\lambda < p-q, x-y > \geq 0$$

and obtain monotonicity by dividing by $\lambda > 0$ and letting λ converge to zero. □

Remark — We observe that condition (3.13) does not involve the scalar product; therefore, it can be used on Banach spaces. In this case, maps A satisfying (3.13) are called *accretive maps*. □

We associate with the map A its *resolvents*

$$\forall\, \lambda > 0, \;\; J_\lambda \; := \; (1+\lambda A)^{-1}$$

An important consequence of property (3.13) is given in

Proposition 3.5.3 *The resolvent J_λ of a monotone map A is a single-valued nonexpansive map from* $\mathrm{Im}(1+\lambda A)$ *to X.*

Proof — Let $x \in J_\lambda(u)$ and $y \in J_\lambda(v)$. We can write

$$u \;\in\; x + \lambda A(x), \quad v \;\in\; y + \lambda A(y)$$

Property (3.13) implies that

$$\|x-y\| \;\leq\; \left\|(x-y) + \lambda\left(\frac{u-x}{\lambda} - \frac{v-y}{\lambda}\right)\right\| \;=\; \|u-v\|$$

This shows that J_λ is nonexpansive and, by taking $u = v$, that J_λ contains a unique point. □

3.5.2 Maximal Monotone Maps

We shall characterize monotone maps A such that

$$\forall\, \lambda \; > \; 0, \quad \mathrm{Im}(I+\lambda A) \; = \; X$$

They are the *maximal monotone maps.*

Definition 3.5.4 *A monotone set-valued map A is* maximal *if there is no other monotone set-valued map whose graph strictly contains the graph of A.*

We begin by pointing out the following:

A set-valued map is maximal monotone if and only if its inverse A^{-1} is maximal monotone.

Also, the graph of any monotone set-valued map is contained in the graph of a maximal monotone set-valued map by Zorn's lemma, because the union of an increasing family of graphs of monotone set-valued maps is the graph of a set-valued monotone map.

The following characterization of maximal monotone maps provides a useful and manageable way for recognizing that an element u belongs to $A(x)$.

Proposition 3.5.5 *A necessary and sufficient condition for a set-valued map A to be maximal monotone is that the property*

$$\forall\ (y, v) \in \text{Graph}(A), \quad < u - v, x - y > \ \geq\ 0$$

is equivalent to

$$u\ \in\ A(x)$$

Proposition 3.5.6 *Let A be maximal monotone. Then*

1. — Its images $A(x)$ are closed and convex

2. — Its graph is strongly-weakly closed in the sense that if x_n converges to x and if $u_n \in A(x_n)$ converges weakly to u, then $u \in A(x)$.

Proof —

1. — By the preceding proposition, $A(x)$ is the intersection of the closed half-spaces

$$\{u \in X \mid \ < u - v, x - y > \ \geq\ 0\}$$

when (y, v) ranges over the graph of A. Hence, $A(x)$ is closed and convex.

2. — Let x_n converge to x and let $u_n \in A(x_n)$ converge weakly to u. Let us choose (y, v) in the graph of A. Inequalities

$$< u_n - v, x_n - y > \ \geq\ 0$$

imply, by going to the limit, inequalities

$$< u - v,\ x - y > \geq 0$$

Hence, $u \in A(x)$ by Proposition 3.5.5. $\square$

Proposition 3.5.7 *If the restriction of a monotone single-valued map $A : X \mapsto X$ to finite dimensional vector subspaces is continuous, then A is maximal monotone.*

Proof — Let $x \in X$ and $u \in X$ such that

$$\langle u - A(y),\ x - y\rangle \geq 0 \text{ for all } y \in X$$

To show that A is maximal monotone, we have to check that $u = A(x)$ (Proposition 3.5.5.) For that purpose, we take[16] $y = x - \lambda(z - x)$ where $\lambda \in]0, 1[$ and $z \in X$. The above inequality becomes

$$\langle u - A(x - \lambda(z - x)),\ z - x\rangle \geq 0 \text{ for all } z \in X$$

By letting λ converge to zero and by the continuity of A on finite dimensional vector subspaces, we deduce that

$$\forall\, z \in X,\ \langle u - A(x), z - x\rangle \geq 0$$

that is $u = A(x)$. $\square$

The Minty Theorem we are about to prove provides a very important characterization of maximal monotone maps.

Theorem 3.5.8 (Minty) *A monotone map A is maximal if and only if the map $\mathbf{1} + A$ is surjective.*

Proof —

1. — Assume that $\mathbf{1} + A$ is surjective. Let $x \in X$ and $u \in X$ satisfy

$$\forall\, (y, v) \in \text{Graph}(A),\ < u - v, x - y > \geq 0 \qquad (3.14)$$

[16] This method is known under the name of *Minty's trick*.

By Proposition 3.5.5, we have to prove that $u \in A(x)$. Since $\mathbf{1} + A$ is surjective, we can choose y in (3.14) to be a solution y_0 of the inclusion

$$u + x \in y_0 + A(y_0)$$

Let $v_0 \in A(y_0)$ such that $u + x = y_0 + v_0$. Then, by (3.14),

$$\|x - y_0\|^2 = < x - y_0, x - y_0 > = - < u - v_0, x - y_0 > \leq 0$$

Hence, $x = y_0$, and thus

$$u = v_0 \in A(y_0) = A(x)$$

2. — Assume now that A is maximal monotone. Let $y \in X$. We have to prove that there exists x such that $y \in x + A(x)$. It is sufficient to choose $y = 0$, since this amounts to replacing A by $x \to -y + A(x)$, which is also maximal monotone. So we have to show that there exists $\overline{x}$ such that $-\overline{x} \in A(\overline{x})$. By Proposition 3.5.5, we must prove that there exists $\overline{x}$ satisfying

$$\forall\, (y, v) \in \mathrm{Graph}(A), \quad < -\overline{x} - v, \overline{x} - y > \geq 0$$

For convenience, set $\varphi(x; (y, v)) :=< x + v, x - y >$, or, equivalently,

$$\varphi(x; (y, v)) := \|x\|^2 + < x, v - y > - < v, y >$$

We have to check that there exists $\overline{x}$ such that

$$\forall\, (y, v) \in \mathrm{Graph}(A), \quad \varphi(\overline{x}; (y, v)) \leq 0 \tag{3.15}$$

Fix (y_0, v_0) to be any point in the graph of A. If a solution to (3.15) exists, it certainly belongs to

$$L := \{x \in \mathrm{Dom}(A) \mid \varphi(x; (y_0, v_0)) \leq 0\}$$

The map $x \to \varphi(x; (y_0, v_0))$ is quadratic. Consequently, the set

$$\{x \mid \varphi(x; (y_0, v_0)) \leq 0\}$$

is convex, closed, and bounded, hence, weakly compact. Let L denote its intersection with $\mathrm{Dom}(A)$. Then $\overline{co}(L)$ is weakly compact.

Also, the maps $x \mapsto \varphi(x;(y,v))$ are convex and (strongly) continuous. Therefore, their lower sections are convex and closed, hence weakly closed. It follows that the above maps are weakly lower semicontinuous.

We use Lemma 3.1.2: Set $\mathcal{S}$ to be the family of all finite subsets $M := \{(y_1,v_1),\dots,(y_n,v_n)\}$ of Graph(A). Then there exists $\overline{x} \in \overline{co}(L)$ such that

$$\begin{aligned}
& \sup_{(y,v)\in \mathrm{Graph}(A)} \varphi(\overline{x};(y,v)) \\
\leq\ & \sup_{M\in\mathcal{S}} \inf_{x\in\overline{co}(L)} \max_{i=1,\dots,n} \varphi(x;(y_i,v_i)) \\
\leq\ & \sup_{M\in\mathcal{S}} \inf_{x\in co(y_1,\dots,y_n)} \max_{i=1,\dots,n} \varphi(x;(y_i,v_i)) \\
\leq\ & \sup_{M\in\mathcal{S}} \inf_{x\in co(y_1,\dots,y_n)} \sup_{\mu\in S^n} \textstyle\sum_{j=1}^n \mu_j \varphi(x;(y_j,v_j)) \\
\leq\ & \sup_{M\in\mathcal{S}} \inf_{\lambda\in S^n} \sup_{\mu\in S^n} \varphi_M(\lambda,\mu)
\end{aligned}$$

where for all λ, μ in the simplex S^n

$$\varphi_M(\lambda,\mu) := \sum_{j=1}^n \mu_j \varphi(\beta(\lambda);(y_j,v_j)), \quad \text{and} \quad \beta(\lambda) := \sum_{i=1}^n \lambda_i y_i$$

Every function φ_M is continuous with respect to λ and

$$\begin{aligned}
& \varphi_M(\mu,\mu) = \textstyle\sum_{i,k=1}^n \mu^i\mu^k < \beta(\mu) + v_i, y_k - y_i > \\
& = \textstyle\sum_{i,k=1}^n \mu^i\mu^k < \beta(\mu), y_k - y_i > + \sum_{i,k=1}^n \mu^i\mu^k < v_i, y_k - y_i >
\end{aligned}$$

The first term is zero for reasons of symmetry, while the second can be written

$$\begin{aligned}
& \textstyle\sum_{i,k=1}^n \mu^i\mu^k < v_i, y_k - y_i > + \sum_{j,l=1}^n \mu^j\mu^l < -v_j, y_j - y_l > \\
= \ & \textstyle\sum_{i,k=1}^n \mu^i\mu^k < v_i - v_k, y_k - y_i > \leq 0
\end{aligned}$$

since A is monotone. Hence, assumptions of the Ky Fan inequality (Theorem 3.1.1) are satisfied. Therefore, for all M,

$$\inf_{\lambda\in S^n} \sup_{\mu\in S^n} \varphi_M(\lambda,\mu) \leq 0$$

so that inequality (3.15) holds true. We have shown that $\overline{x}$ is a solution to $0 \in \overline{x} + A(\overline{x})$. $\square$

3.5.3 Yosida Approximations

Suppose that A is maximal monotone. We show that A can be approximated in some sense by single-valued maps A_λ that are also maximal monotone. These maps, called *Yosida approximations*, play an important role, thanks to:

Theorem 3.5.9 *Let A be maximal monotone. Then for every $\lambda > 0$, the resolvent $J_\lambda = (\mathbf{1}+\lambda A)^{-1}$ is a nonexpansive single-valued map from X to X and the map $A_\lambda := (\mathbf{1} - J_\lambda)/\lambda$ satisfies*

$$\begin{cases} i) & \forall\, x \in X, \ \ A_\lambda(x) \in A(J_\lambda(x)) \\ ii) & A_\lambda \text{ is Lipschitz with constant } 1/\lambda \text{ and maximal monotone} \end{cases}$$

Let $m(A(x))$ denote the element of $A(x)$ with the smallest norm. We also have

$$\forall\, x \in \mathrm{Dom}(A), \ \ \|A_\lambda(x) - m(A(x))\|^2 \ \leq \ \|m(A(x))\|^2 - \|A_\lambda(x)\|^2$$

and for all $x \in \mathrm{Dom}(A)$,

$$\begin{cases} i) & J_\lambda(x) \text{ converges to } x \text{ when } \lambda \to 0+ \\ ii) & A_\lambda(x) \text{ converges to } m(A(x)) \text{ when } \lambda \to 0+ \end{cases}$$

Definition 3.5.10 *The maps A_λ are called the* Yosida Approximations *of A.*

Proof

1. — Let x_i $(i = 1, 2)$ be solutions to the inclusions

$$y_i \in x_i + \lambda A(x_i) \ \ (i = 1, 2)$$

So $y_i = x_i + \lambda v_i$ where $v_i \in A(x_i)$. We obtain

$$\begin{cases} \|y_1 - y_2\|^2 \ = \ \|x_1 - x_2 + \lambda(v_1 - v_2)\|^2 \\ = \ \|x_1 - x_2\|^2 + \lambda^2\|v_1 - v_2\|^2 + 2\lambda < v_1 - v_2, x_1 - x_2 > \\ \geq \ \|x_1 - x_2\|^2 + \lambda^2\|v_1 - v_2\|^2 \end{cases}$$

Hence,

$$\|x_1 - x_2\| \leq \|y_1 - y_2\| \quad \& \quad \|v_1 - v_2\| \leq \|y_1 - y_2\|/\lambda \tag{3.16}$$

By taking $y_1 = y_2$, (3.16) proves the uniqueness of the solution. We note that

$$x_i = J_\lambda y_i \text{ and } v_i = A_\lambda(y_i)$$

Hence, inequalities (3.16) prove that J_λ and A_λ are Lipschitz with constants 1 and $1/\lambda$, respectively.

2. — By the very definitions of J_λ and A_λ, we have

$$A_\lambda(y) = \frac{1}{\lambda}(y - J_\lambda(y)) \in A(J_\lambda(y)) \text{ for all } y \in X$$

Therefore, since $y_i = J_\lambda(y_i) + \lambda A_\lambda(y_i)$, we obtain

$$\begin{cases} < A_\lambda(y_1) - A_\lambda(y_2), y_1 - y_2 > \\ = < A_\lambda(y_1) - A_\lambda(y_2), J_\lambda(y_1) - J_\lambda(y_2) > + \lambda\|A_\lambda(y_1) - A_\lambda(y_2)\|^2 \\ \geq \lambda\|A_\lambda(y_1) - A_\lambda(y_2)\|^2 \geq 0 \end{cases}$$

Hence, A_λ is monotone (and by Proposition 3.5.7, maximal monotone.)

3. — Let $x \in \text{Dom}(A.)$ Then

$$\begin{aligned} & \|A_\lambda(x) - m(A(x))\|^2 \\ = \ & \|A_\lambda(x)\|^2 + \|m(A(x))\|^2 - 2 < A_\lambda(x), m(A(x)) > \\ = \ & \|m(A(x))\|^2 - \|A_\lambda(x)\|^2 - 2 < A_\lambda(x), m(A(x)) - A_\lambda(x) > \end{aligned}$$

Using that A is monotone, that $m(A(x)) \in A(x)$ and $A_\lambda(x) \in A(J_\lambda(x))$, we obtain

$$< A_\lambda(x), m(A(x)) - A_\lambda(x) >= \frac{1}{\lambda} < m(A(x)) - A_\lambda(x), x - J_\lambda(x) > \geq 0$$

Therefore, we have proved inequality

$$\|A_\lambda(x) - m(A(x))\|^2 \leq \|m(A(x))\|^2 - \|A_\lambda(x)\|^2 \tag{3.17}$$

4. — Then when $x \in \mathrm{Dom}(A)$, we have

$$\|x - J_\lambda(x)\| = \lambda\|A_\lambda(x)\| \leq \lambda\|m(A(x))\|$$

therefore, $J_\lambda(x)$ converges to x when λ converges to zero.

5. — We deduce that $y = A_\lambda(x)$ is a solution to the equation $y \in A(x - \lambda y)$. Indeed, setting $z = x - \lambda y$, this equation becomes $x \in z + \lambda A(z)$. Hence,

$$z = J_\lambda(x) \quad \& \quad y = (x - J_\lambda(x))/\lambda = A_\lambda(x)$$

This remark implies that

$$A_{\mu+\lambda}(x) = (A_\mu)_\lambda(x)$$

Indeed, $y = A_{\mu+\lambda}(x)$ is a solution to the equation $y \in A(x - \lambda y - \mu y)$; then $y \in A_\mu(x - \lambda y)$. Applying again the preceding remark to the Yosida approximation A_μ, which is maximal monotone, we deduce that $y = (A_\mu)_\lambda(x)$.

6. — Now we use inequality (3.17), replacing A by A_μ. Since $m(A_\mu(x)) = A_\mu(x)$, we obtain

$$\|A_{\mu+\lambda}(x) - A_\mu(x)\|^2 \leq \|A_\mu(x)\|^2 - \|A_{\lambda+\mu}(x)\|^2$$

Then the sequence $\|A_\mu(x)\|^2$ is monotone and bounded from above by $\|m(A(x))\|^2$, so that it converges to some real number α when $\lambda \to 0+$. This implies that

$$\lim_{\lambda,\mu \to 0} \|A_{\mu+\lambda}(x) - A_\mu(x)\|^2 \leq \alpha - \alpha = 0$$

Hence, $A_\lambda(x)$ satisfies the Cauchy criterion and converges to some element v in X. Since $A_\lambda(x) \in A(J_\lambda(x))$ and the graph of A is closed, we deduce that $v \in A(x)$. Also

$$\|v\| = \lim_{\lambda \to 0+} \|A_\lambda(x)\| \leq \|m(A(x))\|$$

Since $A(x)$ is closed and convex, the projection of zero onto $A(x)$ is unique and consequently, $v = m(A(x))$. Therefore, $A_\lambda(x)$ converges to $m(A(x))$ for all $x \in \mathrm{Dom}(A)$. □

We shall provide an existence and uniqueness theorem for differential inclusions the right-hand side of which is minus a maximal monotone in Chapter 10.

Remark — We shall study in Chapter 5, Section 2 the derivatives of monotone maps and in Chapter 7, Section 1, the graphical convergence of maximal monotone maps. □

3.6 Eigenvectors of Closed Convex Processes

We shall prove now that a closed convex process F does have an eigenvector in cones with compact soles.

Let K be a closed convex cone in $\mathbf{R}^n$. We recall that the following conditions are equivalent:

$$\begin{cases} i) & \text{K is spanned by a convex compact set disjoint from } 0 \\ ii) & \text{the interior of the polar cone } K^+ \text{ is not empty} \\ iii) & S := \{x \in K \mid < p_0, x >= 1\} \text{ where } p_0 \in \mathrm{Int}(K^+) \text{ spans } K \end{cases}$$

The compact convex subset S is called the *sole*, and such closed convex cone are called *cones with compact sole.*

Definition 3.6.1 *We say that a solution $\overline{x}$ to the problem*

$$\overline{x} \in K, \ \overline{x} \neq 0, \ \overline{\lambda} \in \mathbf{R} \ \ \& \ \ \overline{\lambda}\overline{x} \in F(\overline{x}) \tag{3.18}$$

is an eigenvector $\overline{x} \in K$ *of the closed convex process F associated with an* eigenvalue $\overline{\lambda}$.

Theorem 3.6.2 *Let X be a finite dimensional vector-space and $F : X \leadsto X$ be a closed convex process. Assume that a closed convex cone $K \subset X$ enjoys the following properties:*

$$\begin{cases} i) & K \text{ has a compact sole} \\ ii) & K \text{ is a viability domain of } F \\ iii) & \text{the norm } \|R\| \text{ of the map } R \text{ defined by:} \\ & \forall\, x \in K, \ \ R(x) := F(x) \cap T_K(x) \neq \emptyset \text{ is finite} \end{cases}$$

Then there exists a nonzero eigenvector $\overline{x} \in K$ *of the closed convex process* F *associated with an* eigenvalue $\overline{\lambda}$.

Proof — We associate with the closed convex cone K and an element $p_0 \in \mathrm{Int}(K^+)$ the "compact sole"

$$S := \{x \in K \mid \; < p_0, x > = \; 1\}$$

By Lemma 4.2.5 below, its tangent cone is equal to

$$T_S(x) \;\; = \;\; \{v \in T_K(x) \mid \; < p_0, v > = \; 0\}$$

For an element $y \in S$ the projection $\varpi(y)$ onto the orthogonal hyperplane to p_0 is defined by

$$\forall \, z \in X, \;\; \varpi(y)z \; := \; z- < p_0, z > y$$

We then remark that:

$$\forall \, y \in S, \;\; \varpi(y)T_K(y) \; \subset \; T_S(y) \tag{3.19}$$

Indeed, by Lemma 4.2.5 below, if $u \in T_K(y)$, then

$$\varpi(y)u := u- < p_0, u > y \; \in \; \overline{(K + \mathbf{R}y)} + \mathbf{R}y \; \subset \; \overline{K + \mathbf{R}y} \; = \; T_K(y)$$

because K is a closed convex cone. Furthermore

$$< p_0, u > = < p_0, u > - < p_0, u >< p_0, y > = \; 0$$

because $< p_0, y >= 1$. We deduce that $\varpi(y)u$ belongs to $T_S(y)$.

Let us associate with the closed convex process F the set-valued map G defined on the compact sole S by

$$G(y) \; := \; \varpi(y)(F(y) \cap \|R\|B)$$

It is obviously a set-valued map with closed convex images contained in the ball $\|R\|B$, which is compact.

Since its graph is closed, we deduce that G is upper semicontinuous from S to X.

By (3.19), S is a viability domain of the set-valued map G since the cone K is a viability domain of the closed convex process F.

Since the sole S is a convex compact subset, we can apply Equilibrium Theorem 3.2.1 to the set-valued map G. There exists an equilibrium $\overline{x} \in S$ of G, i.e., a solution to $0 \in G(\overline{x})$, in other words, a solution to

$$\overline{x} \in S, \quad 0 = \bar{y} - < p_0, \bar{y} > \overline{x} \text{ where } \bar{y} \in F(\overline{x})$$

By setting $\overline{\lambda} := < p_0, \bar{y} >$, we see that the pair $(\overline{\lambda}, \overline{x})$ is a solution to inclusion (3.18.) □

Remark — Other theorems dealing with positive eigenvalues and an extension of Perron-Frobenius's Theorem to closed convex processes can be found in [35, Chapter 3]. □

Chapter 4

Tangent Cones

Introduction

We already had to use tangent cones in the preceding chapter, both for solving inclusions under constraints and studying stability and convergence of approximations of their solutions.

They also play an important role in viability theory and in optimization. It is time now to conduct a thorough investigation of their properties in order to apply theorems involving tangent cones.

There are many ways to describe technically the concept of tangency to a given subset K at a point $x \in K$. The idea can well be expressed in terms of constraints: We regard a (closed) subset K as the subset of elements of the whole space X obeying given constraints.

Pick any direction $v \in X$ and start from x in the direction v, ranging over the line $x + hv$ when $h > 0$.

We would like to distinguish those directions which, for small h, do not lead us too far away from K. Such directions encapsulate the idea of tangency.

But how can we translate rigorously this idea in mathematical terms? It certainly involves the distance $d_K(x + hv)$ from the points $x + hv$ to K, which should tend to 0 faster than h, in the sense that $d_K(x + hv)/h$ should converge to 0 with h.

Naturally, but unfortunately, there are many ways of taking such limits; we can take either the liminf or the limsup, we can require a little more uniformity in both cases. We feel like opening the door of a *ménagerie of tangents*, and facing the choice of a favorite pet! We are in the situation of the particle physicists when so many unsuspected particles popped up, destroying the harmonious pre-particle description of matter. We also face a whole collection of more or less colored "quark cones", each one having a charm of its own.

Even though the subjective choice of a given category of tangents can be justified by accumulated experience, the reasons to emphasize one concept at a given time can very well disappear later on. This is why we shall *use the concept of "tangent cone" as a generic term to designate any of the particular implementations inherited from history* (as the contingent and paratingent cones since Bouligand.) Fortunately, most of these different implementations coincide for familiar classes of subsets (differential manifolds, convex subsets.)

An analogous situation happened with the concept of derivative of a function, rigorously defined after Cauchy and Weierstrass, and severely taught to generations of young, properly brainwashed mathematicians. The acquired dogma were difficult to overrun when Sobolev (very timidly) and then L. Schwartz (much more boldly) introduced the weak derivatives leading on to derivatives in the sense of distributions, when Dini involved liminf and limsup to let differential quotients converge (more on this in next chapters) and when Bouligand introduced the concepts of contingent and paratingent sets. Furthermore, even for the usual derivatives, different ways of taking the limits and various topologies were used, giving rise to a manifold of concepts (Gâteaux, Hadamard, Fréchet, $\mathcal{C}^1$, etc., derivatives) which were adapted more or less reluctantly, each problem requiring its particular brand of derivative.

Geometers became eagerly "differential geometers" one century ago, and would hardly accept leaving the luxurious environment where the sets of tangents are comfortable vector spaces, with smoothness all around them.

But non smooth sets do exist, unfortunately, as soon as objects of investigation are no longer freely given; they are encountered in numerous problems as the result of some operations. This may destroy smoothness, so that we lose the possibility of imposing *a priori* regularity conditions[1].

Many of us did actually encounter them. The pressure coming from many problems[2] in control theory and differential games, in economy theory, in biological regulation, in systems theory, etc., led mathematicians from many countries and schools, motivated by different problems, to investigate various ways of overcoming regularity requirements.

The first step away from differential geometry was to adapt many of the results to the case of *convex subsets*, after Fenchel, J.-J. Moreau and Rockafellar. It was a surprise to see that tangent cones and normal cones to convex subsets shared many properties of tangent and normal spaces to differentiable manifolds.

After being born with Bouligand in the thirties, and then revived during the early days of control theory, the problem of defining and using tangent cones to arbitrary subsets got a new start in 1975, when Clarke introduced a tangent cone $C_K(x)$ to a set K at $x \in K$ which he proved to be always a closed convex cone. The price of this nice property, however, was quite high since this tangent cone may often be too small or even reduced to the singleton 0. At that time, notwithstanding this cost, it was a good candidate to be "the" tangent cone that was waited for as a messiah!

Further investigations both in viability theory and control theory suggested however that, for the time being at least, the tangent cone which was the most often adequate was the *contingent cone* $T_K(x)$ introduced by Bouligand: it is the upper limit of the *differential quotients*[3] $(K - x)/h$ when $h \to 0+$.

It happens however that if the set-valued map $y \rightsquigarrow T_K(y)$ is lower semicontinuous at x, then the contingent cone $T_K(x)$ coincides

[1]However Differential Geometers can find solace in the 2,600 year old Xenophanes of Colophon's statement that *God was spherical*, unique, motionless and "shaking" all things by the power of thought. He made fun of the anthropomorphic character of the gods of the Olympian theology, arguing that cows would make their gods in their own image.

Of greater philosophic interest, he stated that men can have *no certain knowledge, only opinions*.

[2]Just remember to begin with that the intersection of two smooth sets may no longer be smooth.

[3]whereas the tangent cone $C_K(x)$ is the lower limit of the differential quotients $(K - x')/h$, when $h \to 0+$ *and* $x' \to_K x$.

with $C_K(x)$ and is thus convex. This fortunate situation deserves a name: we say that *K is sleek at x*. Both smooth manifolds and closed convex subsets are sleek.

This, and elementary properties of the contingent cones, are presented in Section 1. Section 2 is devoted to the tangent cones to *convex subsets*, where we recall (most often, without proof) the properties of the tangent cones from convex analysis.

The main motivation for introducing *sleek* subsets (and/or Clarke tangent cones) is that they play a crucial role in the *transversality assumptions* that we need for stating that a tangent cone to an inverse image (respectively an intersection) is equal to the inverse image (respectively the intersection) of the tangent cones. These crucial formulas are indeed a consequence of the Constrained Inverse Function Theorem from Chapter 3, and are proved in Section 3.

We then devote Section 4 to *dual concepts*: the normal space to a differentiable manifold being defined as the orthogonal space to the tangent space, the normal cone to a convex subset as the polar cone to the tangent cone, we consider the polar cones to the tangent cones and translate by duality the properties of the corresponding tangent cones to normal cones. We shall not dwell on this, since many other monographs emphasize this view point, couched in the traditional Lagrange formulation of necessary conditions for optimality. In this sense, one can say that this book is "orthogonal" to most of the texts devoted to optimization and nonsmooth analysis!

Section 5, which can be omitted in a first reading, deals with other tangent cones.

When convexity is required, we can use what we call the *convex kernel* of the contingent cone, which is a convex cone lying between the Clarke tangent cone and the contingent cone, and which is large enough: *it is the intersection of the maximal closed convex cones contained in the contingent cone.*

We then examine the *paratingent cones* introduced also by Bouligand, which contain the contingent cones. The *graph of the paratingent map is closed* (an important property indeed.) We prove Choquet's Theorem which states that the contingent and paratingent cones are generically equal.

We study in Section 6 the limits of tangent cones $T_{K_n}(x_n)$ to a sequence of subsets K_n and relate them to the upper and lower limits of differential quotients $(K_n - x_n)/h_n$.

Since we can regard tangent cones to a subset K as first-order approximations of K, we conclude this chapter with a short introduction to higher order approximations, which are various limits when $h \mapsto 0+$ of "differential quotients"

$$\frac{K - x - hu_1 - \cdots - h^{m-1}v_{m-1}}{h^m}$$

They enjoy many of the properties of first-order approximations.

4.1 Tangent Cones to a Subset

4.1.1 Contingent Cones

We begin with a presentation of the contingent cones:

Definition 4.1.1 (Contingent Cones) *Let $K \subset X$ be a subset of a normed vector space X and $x \in \overline{K}$ belong to the closure of K. The* contingent[4] cone *$T_K(x)$ is defined by*

$$T_K(x) \ := \ \{v \mid \liminf_{h \to 0+} d_K(x + hv)/h = 0\}$$

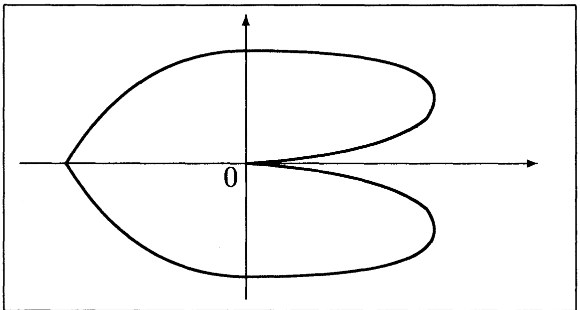
We see at once that *the contingent cone $T_K(x)$ is the upper limit of the subsets $(K - x)/h$*, so that *$T_K(x)$ is a closed cone.*

[4]from the Latin *contingere*, to touch on all sides, introduced by G. Bouligand in the 30's. This term was already used by R. Descartes, in a 1638 letter to Mersenne criticizing P. de Fermat's method on tangents:

"Puis, outre cela, sa règle prétendue n'est pas universelle comme il lui semble, et elle ne peut s'étendre à aucune des questions qui sont un peu difficiles, mais seulement aux plus aisées, ainsi qu'il pourra éprouver si, après l'avoir mieux digérée, il tâche de s'en servir pour trouver les *contingentes*, par exemple, de la ligne courbe BDN, que je suppose être telle qu'en quelque lieu de sa circonférence qu'on prenne le point B, ayant tiré la perpendiculaire BC, les deux cubes des deux lignes BC et CD soient ensemble égaux au parallélépipède des deux mêmes lignes BC, CD et de la ligne donnée P."

This is quite a long and nasty sentence indeed.

Figure 4.1: Contingent Cone at a Boundary Point may be the Whole Space



Subset K such that $T_K(0) = X$

Denote by

$$S_K(x) \ := \ \bigcup_{h>0} \frac{K-x}{h}$$

the cone spanned by $K - x$. Hence the contingent cone $T_K(x)$ is contained in $\overline{S_K(x)}$.

It is very convenient to have the following characterization of this cone in terms of sequences:

$$\left\{ \begin{array}{l} v \ \in \ T_K(x) \ \text{ if and only if } \exists\, h_n \to 0+ \ \text{ and } \exists\, v_n \to v \\ \text{such that } \forall\, n, \ \ x + h_n v_n \in K \end{array} \right.$$

We also observe that

$$\text{if} \ \ x \in \ \mathrm{Int}(K), \ \ \text{then} \ T_K(x) \ = \ X$$

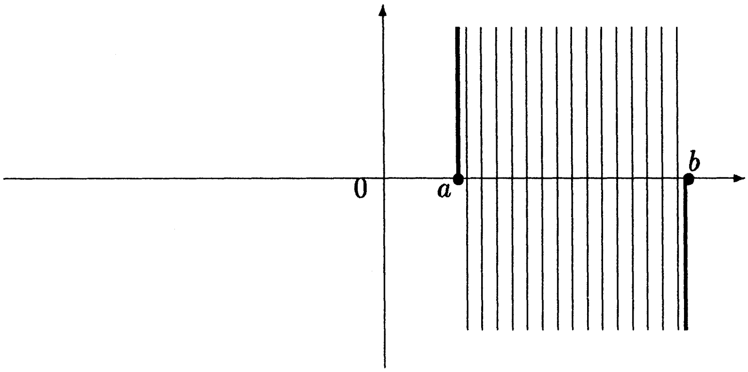
This situation may also happen when x does not belong to the interior of K (see Figure 4.1.)

Unfortunately, the *graph of the set-valued map*

$$K \ \ni \ x \ \rightsquigarrow \ T_K(x) \ \subset \ X$$

is not necessarily closed. This important property is lacking whenever inequality constraints are involved in the definition of K, as the

Figure 4.2: The Graph of $T_{[a,b]}(\cdot)$



following simple example, where $K := [a,b]$, shows (see Figure 4.2.) We shall see in a while that this set-valued map is closed when K is described by equality constraints and that it is lower semicontinuous when K is convex.

Example Let us consider a normed space X, a continuous map

$$g = (g_1, \ldots, g_p) : X \mapsto \mathbf{R}^p$$

and the subset K of X defined by inequality constraints

$$K := \{x \in X \mid g_i(x) \geq 0, \; i = 1, \ldots, p\}$$

Fix x in K. We denote by

$$I(x) := \{i = 1, \ldots, p \mid g_i(x) = 0\}$$

the subset of *active constraints*. We observe that $T_K(x) = X$ whenever $I(x) = \emptyset$ and that, otherwise, inclusion

$$T_K(x) \subset \{u \in X \mid \forall\, i \in I(x), \; < g_i'(x), u > \geq 0\}$$

holds true when g is Fréchet differentiable at x. If we posit moreover the *constraint qualification assumption*:

$$\exists\, v_0 \in X \quad \text{such that } \forall\, i \in I(x), \; < g_i'(x), v_0 > > 0$$

Table 4.1: **Properties of Contingent Cones.**

(1)	▷	If $K \subset L$ and $x \in \overline{K}$, then $T_K(x) \subset T_L(x)$
(2)	▷	If $K_i \subset X$, $(i = 1, \dots, n)$ and $x \in \overline{\bigcup_i K_i}$, then $T_{\bigcup_{i=1}^n K_i}(x) \; = \; \bigcup_{i \in I(x)} T_{K_i}(x)$ where $I(x) := \{i \mid x \in \overline{K_i}\}$
(3)	▷	If $K_i \subset X_i$, $(i = 1, \dots, n)$ and $x_i \in \overline{K_i}$, then $T_{\prod_{i=1}^n K_i}(x_1, \dots, x_n) \; \subset \; \prod_{i=1}^n T_{K_i}(x_i)$
(4)	▷	If $g \in \mathcal{C}^1(X, Y)$, if $K \subset X$, $x \in \overline{K}$ and $M \subset Y$, then $\overline{g'(x)(T_K(x))} \; \subset \; T_{g(K)}(g(x))$; $T_{g^{-1}(M)}(x) \; \subset \; g'(x)^{-1} T_M(g(x))$
(5)	▷	If $K_i \subset X$, $(i = 1, \dots, n)$ and $x \in \overline{\bigcap_i K_i}$, then $T_{\bigcap_{i=1}^n K_i}(x) \; \subset \; \bigcap_{i=1}^n T_{K_i}(x)$

then the contingent cone

$$T_K(x) \; = \; \{u \in X \mid \forall\, i \in I(x), \; < g_i'(x), u > \geq \; 0\}$$

Indeed, let u satisfy $< g_i'(x), u > \geq 0$ for any $i \in I(x)$. For $i \notin I(x)$, strict inequalities $g_i(x) > 0$ imply that for some $\alpha > 0$, we have

$$\forall\, h \in [0, \alpha], \; \forall\, i \notin I(x), \; g_i(x + hu) \; \geq \; 0$$

Consider first the case when $\langle g_i'(x), u \rangle > 0$ for any $i \in I(x)$. Then

$$\forall\, i \in I(x), \;\; g_i(x + hu) \; = \; g_i(x + hu) - g_i(x) \; = < g_i'(x), u > + \, h\varepsilon_i(h)$$

where $\varepsilon_i(h)$ converges to 0 with h. This implies that $g_i(x + hu) \geq 0$ for h small enough and all $i \in I(x)$, and thus, for all $i = 1, \dots, p$. Then such an element u belongs to the contingent cone $T_K(x)$.

Consider now the general case. By assumption, we deduce that for any $\beta \in]0, 1[$, $u_\beta := (1 - \beta)u + \beta v_0$ satisfies strict inequalities $\langle g_i'(x), u_\beta \rangle > 0$ for any $i \in I(x)$ and, by what precedes, belongs also to the contingent cone $T_K(x)$. Letting β converge to 0, we infer that the limit u of the u_β's belongs also to the contingent cone $T_K(x)$. □

4.1.2 Elementary Properties of Contingent Cones

Proposition 4.1.2 *Assume that the Banach space X is* smooth, *i.e., that the norm of X is Gâteaux differentiable off the origin, and denote by $J(x)$ its gradient*[5] *at x. Let K be a closed subset of X.*

We set

$$\Pi_K(y) := \{z \in K \mid \|z-y\| = d_K(y)\}$$

Let $y \notin K$ be such that $\Pi_K(y)$ is not empty. Then

$$\forall\, z \in \Pi_K(y), \quad \forall\, v \in \overline{co}(T_K(z)), \quad \langle J(y-z), v\rangle \;\leq\; 0$$

Proof — Fix $v \in T_K(z)$: There exist $h_n > 0$ and $v_n \in X$ converging to 0 and v respectively such that $z + h_n v_n$ belongs to K for all $n \geq 0$. Since the function $\|\cdot\|$ is convex and Gâteaux differentiable off the origin, the following inequality holds true: for every $y_1 \neq 0$ and y_2,

$$< J(y_1), y_2 - y_1 > \leq \|y_2\| - \|y_1\|$$

We thus deduce that

$$\begin{cases} < J(y-z-h_n v_n), v_n > \\ = < J(y-z-h_n v_n), y-z-(y-z-h_n v_n) > /h_n \\ \leq (\|y-z\| - \|y-z-h_n v_n\|)/h_n \;\leq\; 0 \end{cases}$$

Since J is continuous from X to $X^\star$ supplied with the weak-$\star$ topology, we infer that $J(y-z-h_n v_n)$ converges weakly to $J(y-z)$. Hence $< J(y-z-h_n v_n), v_n >$ converges to $< J(y-z), v >$ because v_n converges strongly to v, so that

$$\forall\, v \in T_K(z), \quad < J(y-z), v > \leq 0$$

and thus the same inequality holds true for any v in the closed convex hull of $T_K(z)$. □

We introduce the Dubovitskij-Miljutin cone defined by

[5] which is a map from X to the unit sphere of $X^\star$, called the *duality map*, satisfying in particular $< J(x), x >= \|x\|$. The duality map is continuous from X to $X^\star$ supplied with the weak-$\star$ topology.

Definition 4.1.3 *The "Dubovitskij-Miljutin tangent cone" $D_K(x)$ to K at $x \in \overline{K}$ is defined by:*

$v \in D_K(x)$ if and only if $\exists\ \varepsilon > 0$ such that $x +]0, \varepsilon](v + \varepsilon B) \subset K$

This cone is the complement to the contingent cone to the complement:

Lemma 4.1.4 *Let x belong to the boundary of K and $\widehat{K}$ denote the complement of K. Then the complement of the contingent cone $T_{\widehat{K}}(x)$ to $\widehat{K}$ at x is the "Dubovitskij-Miljutin cone" $D_K(x)$ to K at x.*

Remark — The last example shows that when

$$K := \{x \mid g_j(x) \geq 0,\ j = 1, \ldots, p\}$$

then the open subset

$$\{v \mid < g_j'(x), v > > 0,\ j \in I(x)\}$$

is contained in $D_K(x)$. □

4.1.3 Adjacent and Clarke Tangent Cones

To go further by having equalities in some of the formulas of Table 4.1, we need more assumptions in order to use the Constrained Inverse Function Theorems of the preceding chapter.

These further assumptions are regularity assumptions, which require that the contingent cone $T_K(x)$, which is the upper limit of the subsets $(K - x)/h$ when $h \to 0+$, coincides with the lower limit of these subsets when $h \to 0+$ and even, of the subsets $(K - y)/h$ when both $h \to 0+$ and $y \in K$ converge to x.

This is the reason why we have to introduce two more tangent cones:

Definition 4.1.5 *Let $K \subset X$ be a subset of a normed vector space X and $x \in \overline{K}$ belong to the closure of K.*

1. — *the* intermediate or adjacent[6] cone $T^\flat_K(x)$ *is defined by*

$$T^\flat_K(x) \;:=\; \{v \mid \lim_{h\to 0+} d_K(x+hv)/h = 0\}$$

2. — *the* Clarke[7] tangent cone *or* circatangent cone $C_K(x)$ *is defined by*

$$C_K(x) \;:=\; \{v \mid \lim_{h\to 0+, x'\to_K x} d_K(x'+hv)/h = 0\}$$

where $\to_K$ *denotes the convergence in* K*.*

We shall say that a subset $K \subset X$ *is* derivable *at* $x \in \overline{K}$ *if and only if* $T^\flat_K(x) = T_K(x)$ *and* tangentially regular *at* x *if* $T_K(x) = C_K(x)$*.*

We see at once that these tangent cones are lower limits

$$T^\flat_K(x) \;=\; \mathrm{Liminf}_{h\to 0+}\frac{K-x}{h}$$

and

$$C_K(x) \;=\; \mathrm{Liminf}_{h\to 0+,\, K\ni x'\to x}\frac{K-x'}{h}$$

so that they are closed cones, that

$$C_K(x) \;\subset\; T^\flat_K(x) \;\subset\; T_K(x) \;\subset\; \overline{S_K(x)}$$

that these tangent cones to K and the closure $\overline{K}$ of K do coincide and that

$$\text{if } x\in \mathrm{Int}(K), \text{ then } C_K(x) \;=\; X$$

We shall see later that the converse is true when the dimension of X is finite.

These tangent cones may all be different: see Figure 4.5.2 of Section 4.5.2 for an example when these tangent cones as well as other ones are distinct. In the example of the set of Figure 4.1, the

[6] from the Latin *adjacere*, to lie near, also used under the name of *derivable* cone by R. T. Rockafellar.

[7] from the Canadian *Frank H. Clarke*; we shall use the adjective *circatangent* to mention properties derived from this tangent cone, for instance, circatangent derivatives and epiderivatives.

adjacent and the contingent cones at zero coincide (with the whole space) and the cone $C_K(0)$ is equal to the horizontal axis $\mathbf{R} \times \{0\}$.

It is very convenient to use the following characterization of these cones in terms of sequences.

$$\left\{ \begin{array}{l} v \in T_K^\flat(x) \text{ if and only if } \forall \;\; h_n \to 0+, \\ \exists\, v_n \to v \;\text{ such that }\; \forall\, n, \;\; x + h_n v_n \in K \end{array} \right.$$

and

$$\left\{ \begin{array}{l} v \in C_K(x) \text{ if and only if } \forall \;\; h_n \to 0+, \;\; \forall\, x_n \to_K x, \\ \exists\, v_n \to v \;\text{ such that }\; \forall\, n, \;\; x_n + h_n v_n \in K \end{array} \right.$$

They are not necessarily equal (see Figure 4.1) for instance.

Let us single out an astonishing fact: the tangent cone $C_K(x)$ is always a closed convex cone.

Proposition 4.1.6 *The tangent cone $C_K(x)$ is a closed convex cone satisfying the following properties*

$$C_K(x) + T_K(x) \;\subset\; T_K(x) \;\;\&\;\; C_K(x) + T_K^\flat(x) \;\subset\; T_K^\flat(x)$$

Proof — Let v_1 and v_2 belong to $C_K(x)$. To prove that $v_1 + v_2$ belongs to this cone, let us choose any sequence $h_n > 0$ converging to 0 and any sequence of elements $x_n \in K$ converging to x. There exists a sequence of elements v_{1n} converging to v_1 such that the elements $x_{1n} := x_n + h_n v_{1n}$ do belong to K for all n. But since the sequence x_{1n} does also converge to x in K, there exists a sequence of elements v_{2n} converging to v_2 such that

$$\forall\, n, \;\; x_{1n} + h_n v_{2n} \;=\; x_n + h_n(v_{1n} + v_{2n}) \;\in\; K$$

This implies that $v_1 + v_2$ belongs to $C_K(x)$ because the sequence of elements $v_{1n} + v_{2n}$ converges to $v_1 + v_2$.

The proof of the two other inclusions is analogous and left as an exercise. □

Unfortunately, the price to pay for enjoying this convexity property of the Clarke tangent cones is that they may often be reduced to the trivial cone $\{0\}$.

Table 4.2: **Properties of Adjacent Tangent Cones.**

(1)	$\triangleright$	If $K \subset L$ and $x \in \overline{K}$, then $T^\flat_K(x) \subset T^\flat_L(x)$
(2)	$\triangleright$	If $K_i \subset X$, $(i = 1, \ldots, n)$ and $x \in \overline{\bigcup_i K_i}$, then $T^\flat_{\bigcup_{i=1}^n K_i}(x) \supset \bigcup_{i \in I(x)} T^\flat_{K_i}(x)$ where $I(x) := \{i \mid x \in \overline{K_i}\}$
(3)	$\triangleright$	If $K_i \subset X_i$, $(i = 1, \ldots, n)$ and $x_i \in \overline{K}_i$, then $T^\flat_{\prod_{i=1}^n K_i}(x_1, \ldots, x_n) = \prod_{i=1}^n T^\flat_{K_i}(x_i)$
(4)	$\triangleright$	If $g \in \mathcal{C}^1(X, Y)$, if $K \subset X$, $x \in \overline{K}$ and $M \subset Y$, then $\overline{g'(x)(T^\flat_K(x))} \subset T^\flat_{g(K)}(g(x))$, $T^\flat_{g^{-1}(M)}(x) \subset g'(x)^{-1} T^\flat_M(g(x))$
(5)	$\triangleright$	If $K_i \subset X$ and $x \in \bigcap_i K_i$ $(i = 1, \ldots, n)$, then $T^\flat_{\bigcap_{i=1}^n K_i}(x) \subset \bigcap_{i=1}^n T^\flat_{K_i}(x)$

As an example, consider the subset

$$K := \{(x, y) \in \mathbf{R}^2 \mid |x| = |y|\}$$

We see that

$$T_K(0) = T^\flat_K(0) = K \text{ and } C_K(0) = \{0\}$$

But we shall show in just a moment that the Clarke tangent cone and the contingent cone do coincide at those points x where K is sleek, i.e., where the set-valued map $x \rightsquigarrow T_K(x)$ is *lower semicontinuous*. Hence the cone $C_K(x)$ can be seen as a "regularization" of the contingent cone $T_K(x)$.

The elementary properties of the adjacent cones are the same than the ones of contingent cones, except for the formula of the intermediate tangent cone to the product, which is the product of the intermediate tangent cones, and the formula for the union. They are summarized in Table 4.2.

4.1.4 Sleek Subsets

We recall the definition of sleek subsets:

Definition 4.1.7 *We shall say that a closed subset K is* sleek *at $x_0 \in K$ if the cone-valued map*

$$K \ni x \rightsquigarrow T_K(x)$$

is lower semicontinuous at x_0 and that it is sleek *if it is sleek at every point of K.*

Theorem 4.1.8 (Tangent Cones to Sleek Subsets) *Let K be a closed subset of a Banach space. If K is sleek at $x \in K$, then the contingent and Clarke tangent cones to K at x do coincide, and consequently, are convex.*

It follows from the "lower semicontinuous version" of the next result:

Theorem 4.1.9 *Let K be a nonempty closed subset of a Banach space X and x_0 belong to K. Then*

$$\mathrm{Liminf}_{x \to_K x_0} T_K(x) \subset C_K(x_0)$$

because if K is sleek at x_0, then

$$T_K(x_0) \subset \mathrm{Liminf}_{x \to_K x_0} T_K(x) \subset C_K(x_0) \subset T_K(x_0)$$

Before proving Theorem 4.1.9 in full generality, we shall prove stronger versions first in the case of finite dimensional vector-spaces, and then, in the case of Hilbert spaces and more generally, in the case of uniformly smooth Banach spaces X such that the norm of its dual $X^\star$ is Fréchet differentiable away from zero.

4.1.5 Limits of Contingent Cones; Finite Dimensional Case

Theorem 4.1.10 *Let X be a finite dimensional vector-space and K be a closed subset of X. Then for every $x \in K$*

$$\mathrm{Liminf}_{y \to_K x} T_K(y) = \mathrm{Liminf}_{y \to_K x} \overline{co}(T_K(y)) = C_K(x)$$

Consequently, K is sleek at x if and only if it is tangentially regular at x.

Proof

1. — Let us take $v \neq 0$ in the lower limit of the closed convex hull of the contingent cones $T_K(z)$ when z converges to $x \in K$: This means that for any $\varepsilon > 0$, there exists $\eta > 0$ such that for

$$\forall\, z \in B(x,\eta) \cap K, \quad B(v,\varepsilon) \cap \overline{co} T_K(z) \neq \emptyset$$

For proving that $v \in C_K(x)$, we take any

$$y \in B(x,\eta/4) \cap K, \quad t \leq \eta/4\|v\|, \quad z \in \Pi_K(y+tv)$$

and we introduce the function g defined by $g(t) := d_K(y+tv)$, which, being Lipschitz, is almost everywhere differentiable. Let

$$t \in [0, \eta/4\|v\|]$$

be any point where g is differentiable.

We claim that $g'(t) \leq \varepsilon$ whenever $g(t) > 0$, i.e., whenever $y+tv$ does not belong to K.

Let $h > 0$ be small enough. Since

$$d_K(y+(t+h)v) - d_K(y+tv) \leq \|y-z+hv\| - \|y-z\|$$

and since the euclidian norm is convex and differentiable off 0, we infer that

$$\begin{cases} g(t+h) - g(t) \leq \|y-z+hv\| - \|y-z\| \leq h\left\langle \frac{y-z+hv}{\|y-z+hv\|}, v\right\rangle \\ \leq h\left\langle \frac{y-z+hv}{\|y-z+hv\|} - \frac{y-z}{\|y-z\|}, v\right\rangle + h\left\langle \frac{y-z}{\|y-z\|}, v\right\rangle \end{cases}$$

We check next that $z \in B(x,\eta) \cap K$ because

$$\|z-x\| \leq \|z-y-tv\| + \|y-x+tv\| \leq 2\|y-x+tv\| \leq \eta$$

whenever $y \in B(x,\eta/4)$ and $t \leq \eta/4\|v\|$. Therefore, there exists $w \in \overline{co}(T_K(z))$ such that $\|v-w\| \leq \varepsilon$, so that, by Proposition 4.1.2,

$$\begin{cases} (g(t+h) - g(t))/h \\ \leq \left\langle \frac{y-z}{\|y-z\|}, w\right\rangle + \|v-w\| + \left\langle \frac{y-z+hv}{\|y-z+hv\|} - \frac{y-z}{\|y-z\|}, v\right\rangle \\ \leq \varepsilon + \left\langle \frac{y-z+hv}{\|y-z+hv\|} - \frac{y-z}{\|y-z\|}, v\right\rangle \end{cases}$$

Taking the limit when $h \to 0+$ we deduce that $g'(t) \leq \varepsilon$ whenever $g(t) > 0$. Hence the claim is proved.

We prove next that $g(t) \leq \varepsilon t$. This being obviously true when $g(t) = 0$, assume that $g(t) > 0$. Let

$$t_0 \ := \ \sup\{\tau \in [0,t] \mid g(\tau) \ = \ 0\}$$

Since g is continuous, $g(t_0)$ is equal to 0, so that we can write

$$g(t) = \int_{t_0}^{t} g'(\tau)d\tau \ \leq \ \varepsilon(t-t_0) \ \leq \ \varepsilon t$$

because $g'(\tau) \leq \varepsilon$ for any $\tau \in]t_0, t]$. This shows that $v \in C_K(x)$.

2. — Let us prove now that $C_K(x)$ is contained in the lower limit of the contingent cones $T_K(y)$ when y converges to x in K.

Indeed, let v belong to $C_K(x)$ and be a sequence of elements $x_n \in K$ converging to x. Then, for all $\varepsilon > 0$, there exist N and $\beta > 0$ such that, for all $0 < h \leq \beta$, $n \geq N$

$$d_K(x_n + hv) \ \leq \ h\varepsilon$$

Let us associate with such x_n elements $y_n^h \in K$ satisfying

$$\|y_n^h - x_n - hv\| \ \leq \ h\varepsilon$$

We set $v_n^h := (y_n^h - x_n)/h$. Since $\|v_n^h - v\| \leq \varepsilon$ and since the dimension of the space X is finite, there exists a cluster point $v_n \in v + \varepsilon B$ of the sequence $(v_n^h)_{h>0}$. Such a v_n belongs to the contingent cone $T_K(x_n)$. Hence v is the limit of the elements v_n. □

4.1.6 Limits of Contingent Cones; Infinite Dimensional Case

The proof of the first part of the above theorem can readily be extended to the case of Hilbert spaces when K is weakly closed. This is too strong an assumption, since we need to use such a result for closed subsets. We can overcome this difficulty and extend Theorem 4.1.10

to uniformly smooth[8] Banach spaces by using the following Edelstein Theorem[9] (see [35, Theorem 5.7.13. p.294] for instance), which states that in Hilbert spaces and some Banach spaces, we can approximate any point by another which has a unique projection of best approximation on a closed subset K.

Theorem 4.1.11 *Assume that X is reflexive and that the norms of X and $X^\star$ are Fréchet differentiable off the origin. Let K be a closed subset of X. Then there exists a dense subset D of X of points y with a unique projection on K.*

In infinite dimensional spaces, we have to replace the contingent cones $T_K(x)$ by the *weak contingent cones* $T_K^\sigma(x)$ defined by

$$v \in T_K^\sigma(x) \iff v \text{ is a weak cluster point of } \frac{y_h - x}{h} \text{ when } h \to 0+$$

where y_h belongs to K.

We observe that Proposition 4.1.2 can be extended to the case when the contingent cones $T_K(z)$ are replaced by the weak contingent cones $T_K^\sigma(z)$.

Proposition 4.1.12 *Assume that the norm of the Banach space X is Fréchet differentiable off the origin and denote by $J(x)$ its gradient at x. Let K be a closed subset of X and $y \notin K$ be such that $\Pi_K(y)$ is not empty. Then*

$$\forall\, z \in \Pi_K(y),\ \ \forall\, v \in \overline{co}(T_K^\sigma(z)),\ \ < J(y-z), v > \ \leq\ 0$$

Proof — Since the norm is Fréchet differentiable at $y - z \neq 0$, we know that for any $\varepsilon > 0$, there exists $\eta > 0$ such that whenever $\|w\| \leq \eta$,

$$|\|y - z + w\| - \|y - z\| - < J(y-z), w > | \ \leq\ \varepsilon \|w\|$$

[8] A Banach space is *uniformly smooth* if and only if its norm is uniformly Fréchet differentiable in the sense that there exists $\varepsilon(h)$ converging to 0 with $h > 0$ such that

$$\forall\, x \neq 0,\ x \in B,\ v \in B,\ \ |\|x + hv\| - \|x\| - h < J(x), v >| \ \leq\ h\varepsilon(h)$$

In this case, J is uniformly continuous from X to $X^\star$ supplied with the norm topology. Uniformly smooth Banach spaces are reflexive. Lebesgue spaces $L^p(\Omega)$ and Sobolev spaces $W^{m,p}(\Omega)$ are uniformly smooth for $1 < p < +\infty$.

[9] which can be derived from Ekeland's Variational Principle.

Let v belong to the contingent cone $T_K^\sigma(z)$. Then there exist sequences $h_n > 0$ converging to 0 and $v_n \in X$ converging weakly to v such that $z + h_n v_n$ belongs to K for any $n \geq 0$, so that

$$\|y - z\| \;\leq\; \|y - z - h_n v_n\|$$

For n large enough, the norm of $w := h_n v_n$ is smaller than or equal to η since v_n is bounded by a constant c and h_n converges to 0. The above inequalities imply

$$\begin{cases} < J(y-z), v_n > \\ \leq \; (\|y - z - h_n v_n\| - \|y - z\| + h_n < J(y-z), v_n >)/h_n \\ \leq \; \varepsilon\|v_n\| \;\leq\; \varepsilon c \end{cases}$$

and thus, by letting v_n converge weakly to v, that $< J(y-z), v > \leq \varepsilon c$ for any $\varepsilon > 0$. Hence $< J(y-z), v > \leq 0$ for every $v \in T_K^\sigma(z)$, and thus, for any v in its closed convex hull. □

Theorem 4.1.13 *Assume that X is uniformly smooth*[10] *and that the norm of $X^\star$ is Fréchet differentiable off the origin. Let K be a closed subset of X. Then*

$$\mathrm{Liminf}_{y \to_K x}\, T_K^\sigma(y) \;=\; \mathrm{Liminf}_{y \to_K x}\, \overline{co}(T_K^\sigma(y)) \;=\; C_K(x)$$

Proof — The proof is similar to the one of Theorem 4.1.10

1. — To prove the first inclusion, let us take $v \neq 0$ in the lower limit of the closed convex hull of the weak contingent cones $T_K^\sigma(z)$ when $z \in K$ converges to $x \in K$: This means that for any $\varepsilon > 0$, there exists $\eta > 0$ such that

$$\forall\, z \in B(x,\eta) \cap K, \;\; B(v,\varepsilon) \cap \overline{co} T_K(z) \;\neq\; \emptyset$$

We have to show that it belongs to $C_K(x)$. For that purpose, we take $y \in B(x, \eta/6) \cap K$, $\;t \leq \eta/6\|v\|$ and we introduce the function

$$g(t) \;:=\; d_K(y + tv)$$

[10] Actually, we only need that the norm of X is Fréchet differentiable off the origin and that J is uniformly continuous from the unit sphere of X to $X^\star$ supplied with the weak-$\star$ topology.

which, being Lipschitz, is almost everywhere differentiable. Let

$$t \in [0, \eta/6\|v\|]$$

be any point where g is differentiable. We claim that $g'(t) \leq \varepsilon$ whenever $g(t) > 0$.

Let $h > 0$ be small enough. By Edelstein's Theorem 4.1.11, we can associate with any $h > 0$ an element $y_h \in B(y + tv, h^2)$ which has a unique projection $z_h \in K$. Therefore,

$$\begin{cases} d_K(y + (t+h)v) - d_K(y + tv) \leq d_K(y_h + hv) - d_K(y_h) + 2h^2 \\ \leq \|y_h - z_h + hv\| - \|y_h - z_h\| + 2h^2 \end{cases}$$

Let $J : X \mapsto B_\star$ be the gradient of the norm of X. Since X is uniformly smooth, we infer that there exists $\varepsilon(h)$ converging to 0 with h such that

$$\|y_h - z_h + hv\| - \|y_h - z_h\| \leq h < J(y_h - z_h), v > +h\varepsilon(h)$$

Finally, we check that $z_h \in B(x, \eta)$ because

$$\begin{cases} \|z_h - x\| \leq \|z_h - y_h\| + \|y_h - y - tv\| + \|y - x + tv\| \\ \\ \leq d_K(y + tv) + 2h^2 + \|y - x + tv\| \leq 2(\|y - x + tv\| + h^2) \leq \eta \end{cases}$$

whenever $h \leq \sqrt{\eta/6}$, $y \in B(x, \eta/6)$ and $t \leq \eta/6\|v\|$. Therefore, there exists $v_h \in \overline{co}(T_K^\sigma(z_h))$ such that $\|v - v_h\| \leq \varepsilon$, so that, by Proposition 4.1.12,

$$\begin{cases} g(t+h) - g(t) = d_K(y + tv + hv) - d_K(y + tv) \\ \\ \leq h < J(y_h - z_h), v_h > +h(\|v - v_h\| + \varepsilon(h) + 2h) \\ \\ \leq h(\varepsilon + \varepsilon(h) + 2h) \end{cases}$$

so that $g'(t) \leq \varepsilon$ whenever $g(t) > 0$.

This yields that $g(t) \leq \varepsilon t$ in the same way as in the proof of Theorem 4.1.10 and shows that $v \in C_K(x)$.

2. — The second inclusion is a consequence of

Lemma 4.1.14 *Let X be a reflexive Banach space, $K \subset X$ a closed subset and x_0 belong to K. Then*

$$C_K(x_0) \subset \operatorname{Liminf}_{x \to_K x_0} T_K^\sigma(x)$$

Proof — Let v belong to $C_K(x_0)$ and be a sequence of elements $x_n \in K$ converging to x_0. Fix $\varepsilon > 0$ and let N and $\beta > 0$ be such that, for all $0 < h \leq \beta$, $n \geq N$

$$d_K(x_n + hv) \leq h\varepsilon$$

Let us associate with such x_n elements $y_n^h \in K$ satisfying

$$\|y_n^h - x_n - hv\| \leq 2h\varepsilon$$

We set $v_n^h := (y_n^h - x_n)/h$. Since $\|v_n^h - v\| \leq 2\varepsilon$ and since the space is reflexive, there exists a weak cluster point $v_n \in v + 2\varepsilon B$ of the sequence $(v_n^h)_{h>0}$. Such a v_n belongs to the contingent cone $T_K^\sigma(x_n)$. Hence v is the (strong) limit of the elements v_n. □

Proof of Theorem 4.1.9 — Since the lower limit of the contingent cones $T_K(x)$ when $x \in K$ converges to x_0 in K is a cone, it is enough to show that any $v \in \operatorname{Liminf}_{x \to_K x_0} T_K(x)$ with norm equal to 1 belongs to $C_K(x_0)$. Fix such a v. Then for any $\varepsilon > 0$, there exists $N \geq 1$ such that

$$\forall\, n \geq N, \ \forall\, z \in B_K(x_0, 1/n), \ B(v, \varepsilon/2) \cap T_K(z) \neq \emptyset \tag{4.1}$$

Assume for a while that v does not belong to $C_K(x_0)$. This means that there exist $\varepsilon \in]0, 1/4[$ and sequences $x_n \in B_K(x_0, 1/2n)$, $h_n \in]0, 1/4n[$ with

$$(x_n + h_n B(v, \varepsilon)) \cap K = \emptyset \tag{4.2}$$

Fix N such that (4.1) holds true, $n \geq N$ and set $\delta := \varepsilon/(1 + \varepsilon)$. We use Ekeland's Variational Principle (Theorem 3.3.1) with the closed subset

$$Q := K \cap (x_n + [0, h_n] B(v, \varepsilon)) \ni x_n$$

and the continuous function V defined on Q by $V(x) := -\|x - x_n\|$, which is bounded from below on Q: There exists then a solution $z \in Q$ to the δ-minimization problem

$$\forall\, x \in Q, \;\; \|x - x_n\| \;\leq\; \|z - x_n\| + \delta\|z - x\| \tag{4.3}$$

We shall exhibit an element $x \in Q$ which, plugged in the above inequality, yields a contradiction.

For that purpose, we observe that z can be written

$$z \;=\; x_n + (1-\alpha)h_n w_n$$

with $\alpha \in [0,1]$ and $\|w_n - v\| \leq \varepsilon$. We first infer from (4.2) that $\alpha \neq 0$. We check next that

$$z + [0, \alpha h_n]B(v,\varepsilon) \;\subset\; x_n + [0, h_n]B(v,\varepsilon) \tag{4.4}$$

Indeed, the subset $x_n + [0, h_n]B(v,\varepsilon)$ being convex,

$$\begin{cases} z + [0, \alpha h_n]B(v,\varepsilon) = (1-\alpha)(x_n + h_n w_n) + \alpha(x_n + [0, h_n]B(v,\varepsilon)) \\ \subset x_n + [0, h_n]B(v,\varepsilon) \end{cases}$$

Finally, z belongs to the ball $B_K(x_0, 1/n)$. Indeed,

$$\|z - x_n\| \;\leq\; (1-\alpha)h_n(\|v\| + \varepsilon)$$

and thus, the norm of v being equal to one,

$$\|z - x_0\| \;\leq\; \|z - x_n\| + \|x_n - x_0\| \;\leq\; h_n(1+\varepsilon) + 1/2n < 1/n$$

Property (4.1) implies the existence of $u \in T_K(z) \cap B(v, \varepsilon/2)$. Therefore there exist sequences u_p converging to u and $k_p > 0$ converging to 0 such that $z + k_p u_p$ belongs to K. By (4.4), it also belongs to $x_n + [0, h_n]B(v,\varepsilon)$ for p large enough. Hence there exists p such that $z + k_p u_p \in Q$ and $u_p \in B(v,\varepsilon)$. Taking $x := z + k_p u_p$ in inequality (4.3), we get

$$\begin{cases} ((1-\alpha)h_n + k_p)\|w_n\| \;-\; k_p\|u_p - w_n\| \\ \leq\; \|((1-\alpha)h_n + k_p)w_n + k_p(u_p - w_n)\| \\ = \|(1-\alpha)h_n w_n + k_p u_p)\| = \|x - x_n\| \leq (1-\alpha)h_n\|w_n\| + \delta k_p\|u_p\| \end{cases}$$

and thus, using the fact that $\|u_p - w_n\| \leq 2\varepsilon$,

$$k_p\|w_n\| \;\leq\; \delta k_p\|u_p\| + k_p\|u_p - w_n\| \;\leq\; k_p(\delta(1+\varepsilon) + 2\varepsilon)$$

Dividing by $k_p > 0$, we obtain

$$1 - \varepsilon \;\leq\; \|w_n\| \;\leq\; \delta(1+\varepsilon) + 2\varepsilon \;\leq\; 3\varepsilon$$

Hence $\varepsilon \geq 1/4$, the looked for contradiction of the choice of ε. □

4.2 Tangent Cones to Convex Sets

For convex subsets K, the situation is dramatically simplified by the fact that the Clarke tangent cones and the contingent cones coincide with the closed cone spanned by $K - x$:

Proposition 4.2.1 (Tangent Cones to Convex Sets) *Let us assume that K is convex. Then the contingent cone $T_K(x)$ to K at x is convex and*

$$C_K(x) \;=\; T_K^\flat(x) \;=\; T_K(x) \;=\; \overline{S_K(x)}$$

We shall denote by $T_K(x)$ the common value of these cones, and call it the tangent cone *to the convex subset K at x.*

Proof — We begin by stating the following consequence of convexity:

$$\forall\, v \in S_K(x),\;\; \exists\, h > 0,\;\; \text{such that}\;\; \forall\, t \in [0,h],\; x + tv \in K$$

since we can write that for any $t \in [0,h]$

$$x + tv \;=\; \left(1 - \frac{t}{h}\right)x + \frac{t}{h}(x + hv)$$

is a convex combination of elements of K.

It is enough to prove that $S_K(x)$ is contained in $C_K(x)$. Let $v := (y-x)/h$ belong to $S_K(x)$ (where $y \in K$ and $h > 0$) and let us consider sequences of elements $h_n > 0$ and $x_n \in K$ converging to 0 and x respectively. We see that $v_n := (y - x_n)/h$ converges to v and that

$$x_n + h_n v_n \;=\; \left(1 - \frac{h_n}{h}\right)x_n + \frac{h_n}{h}y \;\in\; K$$

since it is a convex combination of elements of K. $\square$

The negative polar cone of the tangent cone $T_K(x)$ to a convex subset, is called the *normal cone to* K at x and is denoted by

$$N_K(x) := T_K(x)^- = S_K(x)^-$$

It can be easily characterized by:

$$N_K(x) = \{p \in X^\star \mid \max_{y \in K} < p, y > = < p, x >\}$$

We observe that

$$\begin{cases} i) & N_K(x) \subset b(K) \\ ii) & \text{the "asymptotic cone" } b(K)^- \subset T_K(x) \end{cases}$$

Theorem 4.2.2 *Any closed convex subset of a Banach space is sleek.*

Proof — Let K be a closed subset of a Banach space X. We begin by proving that the graph of the set-valued map

$$K \ni x \rightsquigarrow N_K(x)$$

is closed in $X \times X^\star$, when X is supplied with the norm topology and $X^\star$ with the weak-$\star$ topology:

Let us consider sequences of elements $x_n \in K$ converging to x and $p_n \in N_K(x_n)$ converging weakly to p. Then inequalities

$$\forall\, y \in K, \ < p_n, y > \leq < p_n, x_n >$$

imply by passing to the limit inequalities

$$\forall\, y \in K, \ < p, y > \leq < p, x >$$

which state that p belongs to $N_K(x)$. Hence the graph is closed, so that the set-valued map $T_K(\cdot)$ is lower semicontinuous thanks to Proposition 1.1.8. $\square$

It may be useful to characterize the interior of the tangent cone to a convex subset.

Proposition 4.2.3 (Interior of a Tangent Cone) *Assume that the interior of a convex subset $K \subset X$ is not empty. Then*

$$\forall\, x \in K, \quad \mathrm{Int}(T_K(x)) \;=\; \bigcup_{h>0} \left(\frac{\mathrm{Int}(K) - x}{h} \right)$$

Furthermore, the graph of the set-valued map

$$K \;\ni\; x \;\rightsquigarrow\; \mathrm{Int}(T_K(x))$$

is open.

For the convenience of the reader, we list in Table 4.3 some useful formulas of the calculus of tangent cones to convex subsets derived from convex analysis (see [35, Section 4.1.].) The subsets K, K_i, L, $M, \ldots$ are assumed to be convex, the spaces X, Y are Banach spaces and $\mathcal{L}(X, Y)$ denotes the space of continuous linear operators.

Remark — Property

$$T_{K_1 \cap K_2}(x) \;=\; T_{K_1}(x) \cap T_{K_2}(x)$$

is false when assumption $0 \in \mathrm{Int}(K_1 - K_2)$ is not satisfied. Take for instance two balls K_1 and K_2 tangent at a point x. The tangent cone to the intersection $\{x\}$ is reduced to $\{0\}$, whereas the intersection of the tangent cones is a hyperplane. This shows that we cannot dispense with the *constraint qualification* assumptions in the calculus of tangent cones to inverse images and intersections. □

We provide examples of tangent cones to some specific convex subsets:

Examples

— 1. We observe first that an element v belongs to $T_{\mathbf{R}^n_+}(x)$ if and only if $v_i \geq 0$ whenever $x_i = 0$.

— 2. We denote by

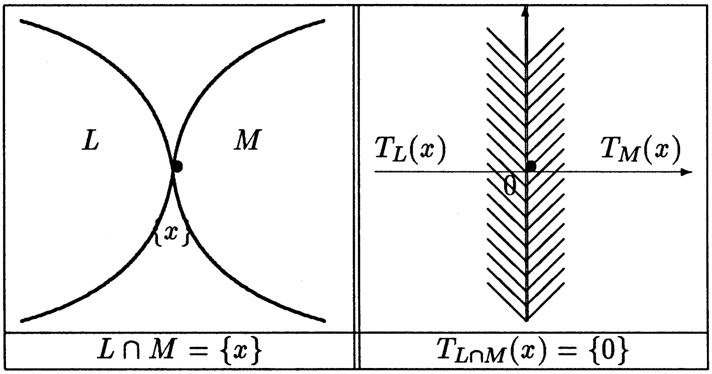
$$S^n \;:=\; \left\{ x \in \mathbf{R}^n_+ \mid \sum_{i=1}^{n} x_i \;=\; 1 \right\}$$

the probability simplex.

Table 4.3: **Properties of Tangent Cones to Convex Sets.**

(1)	$\triangleright$	If $x \in K \subset L \subset X$, then $T_K(x) \subset T_L(x)$ & $N_L(x) \subset N_K(x)$
(3)	$\triangleright$	If $x_i \in K_i \subset X_i, \quad (i = 1, \dots, n)$, then $T_{\prod_{i=1}^n K_i}(x_1, \dots, x_n) \ = \ \prod_{i=1}^n T_{K_i}(x_i)$ $N_{\prod_{i=1}^n K_i}(x_1, \dots, x_n) \ = \ \prod_{i=1}^n N_{K_i}(x_i)$
(4)a)	$\triangleright$	If $A \in \mathcal{L}(X, Y)$ and $x \in K \subset X$, then $T_{A(K)}(Ax) \ = \ \overline{A(T_K(x))}$ $N_{A(K)}(Ax) \ = \ {A^{\star}}^{-1} N_K(x)$
(4)b)	$\triangleright$	If $K_1, \ K_2 \subset X, \ x_i \in K_i, \ i = 1, 2$, then $T_{K_1+K_2}(x_1 + x_2) \ = \ \overline{T_{K_1}(x_1) + T_{K_2}(x_2)}$ $N_{K_1+K_2}(x_1 + x_2) \ = \ N_{K_1}(x_1) \cap N_{K_2}(x_2)$ In particular, if $x_1 \in K$ and x_2 belongs to a subspace P of X, then $T_{K+P}(x_1 + x_2) \ = \ \overline{T_{K_1}(x_1) + P}$ $N_{K+P}(x_1 + x_2) \ = \ N_K(x_1) \cap P^{\perp}$
(5)	$\triangleright$	If $L \subset X$ and $M \subset Y$ are *closed convex* subsets and $A \in \mathcal{L}(X, Y)$ satisfies the *constraint qualification assumption* $0 \ \in \ \mathrm{Int}(M - A(L))$, then, for every $x \in L \cap A^{-1}(M)$, $T_{L \cap A^{-1}(M)} \ = \ T_L(x) \cap A^{-1} T_M(Ax)$ $N_{L \cap A^{-1}(M)} \ = \ N_L(x) + A^{\star} N_M(Ax)$
(5)a)	$\triangleright$	If $M \subset Y$ is *closed convex* and if $A \in \mathcal{L}(X, Y)$ satisfies $0 \ \in \ \mathrm{Int}(\mathrm{Im}(A) - M)$, then for $x \in A^{-1}(M)$ $T_{A^{-1}(M)}(x) \ = \ A^{-1} T_M(Ax)$ $N_{A^{-1}(M)}(x) \ = \ A^{\star} N_M(Ax)$
(5)b)	$\triangleright$	If $K_1, \ K_2 \subset X$ are closed convex and satisfy $0 \ \in \ \mathrm{Int}(K_1 - K_2)$, then for any $x \in K_1 \cap K_2$ $T_{K_1 \cap K_2}(x) \ = \ T_{K_1}(x) \cap T_{K_2}(x)$ $N_{K_1 \cap K_2}(x) \ = \ N_{K_1}(x) + N_{K_2}(x)$
(5)c)	$\triangleright$	If $K_i \subset X, \ (i = 1, \dots, n)$, are closed and convex, $x \in \bigcap_{i=1}^n K_i$ and if there exists $\gamma > 0$ satisfying $\forall x_i$ such that $\|x_i\| \leq \gamma, \ \bigcap_{i=1}^n (K_i - x_i) \neq \emptyset$, then $T_{\bigcap_{i=1}^n K_i}(x) \ = \ \bigcap_{i=1}^n T_{K_i}(x)$ $N_{\bigcap_{i=1}^n K_i}(x) \ = \ \sum_{i=1}^n N_{K_i}(x)$

Figure 4.3: Counterexample: Tangent Cone to the Intersection



Lemma 4.2.4 *The contingent cone $T_{S^n}(x)$ to S^n at $x \in S^n$ is the cone of elements $v \in \mathbf{R}^n$ satisfying*

$$\sum_{i=1}^{n} v_i \;=\; 0 \;\;\&\;\; v_i \;\geq\; 0 \;\; \text{whenever} \;\; x_i \;=\; 0 \tag{4.5}$$

Proof — Let us take $v \in T_{S^n}(x)$. There exist sequences $h_p > 0$ converging to 0 and v_p converging to v such that $y_p := x + h_p v_p$ belongs to S^n for any $p \geq 0$. Then

$$\sum_{i=1}^{n} v_{p_i} \;=\; \frac{1}{h_p}\left(\sum_{i=1}^{n} y_{p_i} - \sum_{i=1}^{n} x_{p_i}\right) \;=\; 0$$

so that $\sum_{i=1}^{n} v_i = 0$. On the other hand, if $x_i = 0$, then

$$v_{p_i} \;=\; y_{p_i}/h_p \;\geq\; 0$$

so that $v_i \geq 0$.

Conversely, let us take v satisfying (4.5) and deduce that

$$y \;:=\; x + hv$$

belongs to the simplex for h small enough. First, the sum of the y_i is obviously equal to 1. Second, $y_i \geq 0$ when $x_i = 0$ because in this case v_i is nonnegative, and, when $x_i > 0$, it is sufficient to take $h < x_i/|v_i|$ for having $y_i \geq 0$. Hence y does belong to the simplex. $\square$

In the case of general convex cones, we have:

Lemma 4.2.5 *Let $K \subset X$ be a convex cone of a normed space X and $x \in K$. Then $T_K(x) = \overline{K + \mathbf{R}x}$.*

Furthermore,

$$p \in N_K(x) \iff x \in K,\ p \in K^- \ \&\ < p, x > = 0 \iff x \in N_{K^-}(p)$$

where $N_{K^-}(p) := \{x \in K \mid \forall\, q \in K^-,\ \ < q - p, x > \leq 0\}$.

Assume that $p_0 \in$ Int(K^+) and set

$$S \ := \ \{x \in K \mid \ < p_0, x > = \ 1\}$$

Then

$$T_S(x) \ = \ \{v \in T_K(x) \mid \ < p_0, v > = \ 0\} \tag{4.6}$$

Proof — Indeed,

$$K + \mathbf{R}x \ = \ K - \mathbf{R}_+ x \ \subset \ T_K(x)$$

because, for any $z \in K$, $\lambda > 0$,

$$\forall\, h < 1/\lambda, \quad x + h(z - \lambda x) \ = \ (1 - h\lambda)x + hz \ \in \ K$$

Conversely, $v \in T_K(x)$, being the limit of v_n such that $x + h_n v_n \in K$, belongs to $K + \mathbf{R}x$ for all n, because

$$v_n \ \in \ K - x/h_n \ \subset \ K - \mathbf{R}_+ x$$

To say that $p \in N_K(x)$ means that $< p, x >= \sigma_K(p)$, which is equal to 0 if and only if $p \in K^-$, so that the second statement of the lemma ensues.

Observe that S can be written in the form $K \cap p_0^{-1}(1)$ and the constraint qualification assumption $0 \in$ Int$(p_0(K) - 1)$ is satisfied because $p_0(K)$ is a cone of $\mathbf{R}$ containing 1. We then deduce that

$$T_S(x) \ = \ T_K(x) \cap p_0^{-1} T_{\{1\}}(1)$$

i.e., formula (4.6.) □

Remark — If we assume that $K^- + \{x\}^- = X^\star$ and that X is reflexive, then $T_K(x) = K + \mathbf{R}x$ thanks to the Closed Range Theorem 2.4.4. □

The next result plays an important role in control theory for proving that controllability and observability for closed convex process are dual concepts:

Theorem 4.2.6 (Polar of a Viability Domain) *Let X be a reflexive Banach space, $F : X \rightsquigarrow X$ be a closed convex process whose domain is the whole space and K be a closed convex cone. The following statements are equivalent:*

$$\begin{cases} i) & \forall\, x \in K, & F(x) \subset T_K(x) \\ ii) & \forall\, q \in K^+, & F^\star(q) \cap T_{K^+}(q) \neq \emptyset \end{cases}$$

where $K^+ := -K^- = \{p \in X^\star \mid \forall\, x \in K,\ < p, x > \geq 0\}$.

Proof — The first condition is equivalent to

$$\forall\, x \in K,\ \forall\, q \in -(T_K(x))^-,\ \sup_{y \in F(x)} < -q, y > \leq 0 \tag{4.7}$$

But we know from the latter lemma that $q \in -(T_K(x))^- = -N_K(x)$ if and only if

$$x \in N_{K^-}(-q) = (K^- + \mathbf{R}q)^- = (K^+ + \mathbf{R}q)^+ = (T_{K^+}(q))^+$$

On the other hand, since the domain of F is the whole space X, Proposition 2.6.4 implies that

$$\sup_{y \in F(x)} < -q, y > + \sup_{p \in F^\star(q)} < p, x > = 0$$

Therefore, condition (4.7) is equivalent to

$$\forall\, q \in K^+,\ \forall\, x \in (T_{K^+}(q))^+,\ \sup_{p \in F^\star(q)} < p, x > \geq 0 \tag{4.8}$$

Now, since the images of $F^\star(q)$ are weakly compact, the Separation Theorem implies that the intersection of $F^\star(q)$ and $T_{K^+}(q)$ is not empty, (i.e., that 0 belongs to the closed convex subset $F^\star(q) - T_{K^+}(q)$), if and only if

$$\forall\, x \in X,\ 0 \ \leq\ \sigma(F^\star(q) - T_{K^+}(q), x)$$

It is enough to observe that this statement is equivalent to (4.8), because

$$\sigma(F^\star(q) - T_{K^+}(q), x) \ = \ \begin{cases} \sigma(F^\star(q), x) & \text{if } \ x \ \in \ (T_{K^+}(q))^+ \\ +\infty & \text{if } \ x \ \notin \ (T_{K^+}(q))^+ \end{cases} \square$$

Definition 4.2.7 (Pseudo-Convex & Star-Shaped Sets) *A set K of a normed space X is said to be* star-shaped around $x \in K$ *if*

$$\forall\, y \in K,\ \ \forall\, \lambda \in [0,1],\ \ x + \lambda(y - x) \ \in \ K$$

and pseudo-convex *at $x \in K$ if and only if either one of the equivalent properties*

$$\begin{cases} i) & T_K(x) \ = \ \overline{S_K(x)} \ =: \ \overline{\bigcup_{h>0} \frac{K-x}{h}} \\ & \\ ii) & K \ \subset \ x + T_K(x) \end{cases}$$

holds true.

We observe the following

Lemma 4.2.8 *If $K \subset X$ is star-shaped around $x \in K$, then it is pseudo-convex at this point.*

Convex subsets are star-shaped around each of their elements and thus, share with them some properties. For instance:

Proposition 4.2.9 *If X, Y are normed spaces, $A \in \mathcal{L}(X, Y)$ is a continuous linear operator and if $K \subset X$ is pseudo-convex at some point x of K, then*

$$\overline{AT_K(x)} \ = \ T_{A(K)}(Ax)$$

In particular, when K is convex, we have for every $y \in A(K)$

$$\bigcap_{x \in K \cap A^{-1}(y)} \overline{A(T_K(x))} \ = \ T_{A(K)}(y)$$

Proof — Take $v \in T_{A(K)}(Ax)$ and let $h_n > 0$ and $v_n \in Y$ be sequences converging to 0 and v respectively such that

$$Ax + h_n v_n \in A(K) \subset Ax + A(T_K(x))$$

Therefore the v_n's belong to $A(T_K(x))$ and so their limit belongs to the closure of this cone. $\square$

4.3 Calculus of Tangent Cones

By using the tangent cone $C_K(x)$ instead of sleekness, we can state the following version of the "pointwise inverse function theorem", which follows from Theorem 4.1.10 and the proof of Theorem 3.4.10.

Theorem 4.3.1 (Pointwise Inverse Mapping Theorem) *Let X be a Banach space, K be a* closed *subset of X, Y be a finite dimensional vector-space $f : X \mapsto Y$ be continuously differentiable around an element $x_0 \in K$. If*

$$f'(x_0)(C_K(x_0)) = Y$$

then the set-valued map $y \rightsquigarrow f^{-1}(y) \cap K$ is pseudo-Lipschitz around $(f(x_0), x_0)$.

In particular, we single out the following consequence:

Corollary 4.3.2 *If K is a closed subset of a finite dimensional vector-space X, then $x_0 \in \mathrm{Int}(K)$ if and only if $C_K(x_0) = X$.*

If X is a Banach space, $x_0 \in \mathrm{Int}(K)$ if and only if there exist constants $\alpha \in [0, 1[$, $\eta > 0$ such that

$$\forall\, x \in B_K(x_0, \eta), \;\; B_X \subset T_K(x) + \alpha B_X \quad \square$$

4.3.1 Intersection and Inverse Image

We are ready to derive the basic results of the calculus of tangent cones.

Theorem 4.3.3 *Let X and Y be Banach spaces, $L \subset X$ and $M \subset Y$ be* closed *subsets, $f : X \mapsto Y$ be continuously differentiable around an element $x_0 \in L \cap f^{-1}(M)$.*

If Y is finite dimensional, we posit the pointwise transversality condition

$$f'(x_0)(C_L(x_0)) \; - \; C_M(f(x_0)) \; = \; Y$$

Then

$$T_L^\flat(x_0) \cap f'(x_0)^{-1} \left(T_M(f(x_0))\right) \; \subset \; T_{L\cap f^{-1}(M)}(x_0)$$

$$T_L^\flat(x_0) \cap f'(x_0)^{-1} \left(T_M^\flat(f(x_0))\right) \; = \; T^\flat_{L\cap f^{-1}(M)}(x_0)$$

and

$$C_L(x_0) \cap f'(x_0)^{-1} \left(C_M(f(x_0))\right) \; \subset \; C_{L\cap f^{-1}(M)}(x_0)$$

Otherwise, to obtain the same results in the case of any Banach space, we have to replace the pointwise transversality condition by the local transversality condition*:*

there exist constants $c > 0$, $\alpha \in [0, 1[$ and $\eta > 0$ such that

$$\begin{cases} \forall\, x \in L \cap B(x_0, \eta), \;\; \forall\, y \in M \cap B(f(x_0), \eta) \\ \\ B_Y \; \subset \; f'(x) \left(T_L^\flat(x) \cap cB_X\right) - T_M(y) + \alpha B_Y \end{cases}$$

Furthermore the above assumption yields that the set $L \cap f^{-1}(M)$ is sleek (respectively derivable) whenever L and M are sleek (respectively derivable.)

Proof — Let us prove for instance the inclusion for the Clarke tangent cones. Consider the closed subset

$$K \; := \; L \cap f^{-1}(M)$$

and take any sequence of elements $x_n \in K$ which converges to x. Let us pick any $u \in C_L(x_0)$ such that $f'(x_0)u \in C_M(f(x_0))$. Hence for any sequences $h_n > 0$ and $x_n \in K$ converging to 0 and x_0 respectively, there exist sequences u_n and v_n converging to u and $f'(x_0)u$ respectively such that, for all $n \geq 0$,

$$x_n + h_n u_n \in L \;\; \& \;\; f(x_n) + h_n v_n \in M$$

We now apply Theorem 4.3.1 to the subset $L \times M$ of $X \times Y$ and the continuous map $f \ominus \mathbf{1}$ associating to any (x, y) the element $f(x) - y$. It is obvious that the transversality condition

$$f'(x_0)(C_L(x_0)) - C_M(f(x_0)) \;=\; Y$$

implies the surjectivity assumption of Theorem 4.3.1.

The pair $(x_n + h_n u_n, f(x_n) + h_n v_n)$ belongs to $L \times M$ and

$$(f \ominus \mathbf{1})(x_n + h_n u_n, f(x_n) + h_n v_n) \;\text{ converges to } 0$$

because f is continuous at x.

Therefore, by Theorem 4.3.1, there exist $l > 0$ and a solution $(\widehat{x}_n, \widehat{y}_n) \in L \times M$ to the equation

$$(f \ominus \mathbf{1})(\widehat{x}_n, \widehat{y}_n) \;=\; 0$$

(i.e., $\widehat{y}_n = f(\widehat{x}_n)$) such that

$$\left\{ \begin{array}{l} \|x_n + h_n u_n - \widehat{x}_n\| + \|f(x_n) + h_n v_n - \widehat{y}_n\| \\ \leq l h_n \|(f(x_n + h_n u_n) - f(x_n))/h_n - v_n\| \end{array} \right.$$

Hence $\widehat{u}_n := (\widehat{x}_n - x_n)/h_n$ converges to u, and for all $n \geq 0$, we know that $x_n + h_n \widehat{u}_n$ belongs to $L \cap f^{-1}(M)$ because $x_n + h_n \widehat{u}_n = \widehat{x}_n$ and $f(x_n + h_n \widehat{u}_n) = \widehat{y}_n$. When Y is an arbitrary Banach space we apply Theorem 3.4.3 in exactly the same way. □

We list now three useful corollaries of this theorem:

Corollary 4.3.4 *Assume that X, Y are Banach spaces, $M \subset Y$ is a closed subset and that $f : X \mapsto Y$ is a continuously differentiable map around an element $x_0 \in f^{-1}(M)$.*

When the dimension of Y is finite, we suppose that

$$\mathrm{Im}(f'(x_0)) \;+\; C_M(f(x_0)) \;=\; Y$$

Then

$$T_{f^{-1}(M)}(x_0) \;=\; f'(x_0)^{-1}\left(T_M(f(x_0))\right)$$

Furthermore

$$T^{\flat}_{f^{-1}(M)}(x_0) \;=\; f'(x_0)^{-1}\left(T^{\flat}_M(f(x_0))\right)$$

and

$$C_{f^{-1}(M)}(x_0) \supset f'(x_0)^{-1}\left(C_M(f(x_0))\right)$$

Otherwise, to obtain the same result in the case of a Banach space Y, we have to assume that there exist constants $c > 0$, $\alpha \in [0,1[$ and $\eta > 0$ such that

$$\begin{cases} \forall\, x \in B(x_0,\eta), \quad \forall\, y \in B(f(x_0),\eta) \cap M \\ B_Y \subset cf'(x)(B_X) + T_M(y) + \alpha B_Y \end{cases}$$

Corollary 4.3.5 *Let K_1 and K_2 be closed subsets of a Banach space X and $x \in K_1 \cap K_2$. If the dimension of X is finite, we assume that*

$$C_{K_1}(x) - C_{K_2}(x) = X$$

Then

$$T^\flat_{K_1}(x) \cap T_{K_2}(x) \subset T_{K_1 \cap K_2}(x) \quad \& \quad T^\flat_{K_1}(x) \cap T^\flat_{K_2}(x) = T^\flat_{K_1 \cap K_2}(x)$$

and

$$C_{K_1}(x) \cap C_{K_2}(x) \subset C_{K_1 \cap K_2}(x)$$

Otherwise, the same result holds true when X is a Banach space if we suppose that there exist constants $c > 0$, $\alpha \in [0,1[$ and $\eta > 0$ such that

$$\begin{cases} \forall\, x \in K_1 \cap B(x_0,\eta), \quad \forall\, y \in K_2 \cap B(x_0,\eta) \\ B_Y \subset \left(T^\flat_{K_1}(x) \cap cB_X\right) - T_{K_2}(y) + \alpha B_Y \end{cases}$$

Finally, for a finite intersection, we can state:

Corollary 4.3.6 *Let us consider n closed subsets K_i of a Banach space X and $x_0 \in \bigcap_{i=1}^n K_i$. When the dimension of X is finite, we assume that*

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

Then

$$\bigcap_{i=1}^n T^\flat_{K_i}(x_0) = T^\flat_{\bigcap_{i=1}^n K_i}(x_0)$$

and

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

If the dimension of X is infinite, we assume that there exist constants $c > 0$, $\alpha \in [0,1[$ and $\eta > 0$ such that

$$\begin{cases} \forall\, x_i \in K_i \cap B(x_0,\eta),\ \forall\, v_i \in X, \\ \exists\, w_i \in X,\ \ \exists\, u \in \bigcap_{i=1}^{n}(T_{K_i}(x_i) - v_i - w_i) \text{ such that} \\ \|u\| \le c\max_{i=1,\ldots,n}\|v_i\| \ \&\ \|w_i\| \le \alpha\max_{i=1,\ldots,n}\|v_i\| \end{cases}$$

to get the same conclusions.

Proof — It is enough to apply Theorem 4.3.3 to $M = \{0\}$, the subset $\prod_{i=1}^{n} K_i \times D$, where $D := \{(x,\ldots,x)\}_{x\in X}$, and the map f defined on this set by $f(x,y) = x - y$. $\square$

4.3.2 Example: Tangent cones to subsets defined by equality and inequality constraints

Consider a closed subset L of a Banach space X and two continuously differentiable maps

$$g := (g_1,\ldots,g_p) : X \mapsto \mathbf{R}^p \ \&\ h := (h_1,\ldots,h_q) : X \mapsto \mathbf{R}^q$$

defined on an open neighborhood of L.

Let K be the subset of L defined by the constraints

$$K := \{x \in L \mid g_i(x) \ge 0,\ i = 1,\ldots,p \ \&\ h_j(x) = 0,\ j = 1,\ldots,q\}$$

We denote by $I(x) := \{i = 1,\ldots,p \mid g_i(x) = 0\}$ the subset of active constraints.

Proposition 4.3.7 *Let us posit the following transversality condition at a given $x \in K$:*

$$\begin{cases} i) & h'(x)C_L(x) = \mathbf{R}^q \\ ii) & \exists\, v_0 \in C_L(x) \text{ such that } h'(x)v_0 = 0 \text{ and} \\ & \forall\, i \in I(x),\ < g_i'(x), v_0 > > 0 \end{cases}$$

Then u belongs to the contingent cone to K at x if and only if u belongs to the contingent cone to L at x and satisfies the constraints

$$\forall\, i \in I(x),\ \ < g_i'(x), u > \geq\ 0\ \ \&\ \ \forall\, j = 1, \ldots, q,\ \ h_j'(x)u\ =\ 0$$

Proof — It is clear that for every $v \in T_K(x) \subset T_L(x)$, equations $h_j'(x)v = 0,\ j = 1, \ldots, q$ are satisfied and inequalities $< g_i'(x), v > \geq 0$ are satisfied for all active constraints $i \in I(x)$.

The converse inclusion is a direct consequence of Theorem 4.3.3 with

$$Y\ :=\ \mathbf{R}^p \times \mathbf{R}^q,\ \ f\ =\ g \times h,\ \ M\ :=\ \mathbf{R}_+^p \times \{0\}$$

It is sufficient to check that the assumption implies the transversality condition

$$(g'(x) \times h'(x))\,(C_L(x)) - T_{\mathbf{R}_+^p \times \{0\}}(g(x), 0)\ =\ \mathbf{R}^p \times \mathbf{R}^q$$

Indeed, take $(y, z) \in \mathbf{R}^p \times \mathbf{R}^q$. There exists a solution $v \in C_L(x)$ to the equation $h'(x)v = z$ by the first assumption. Let

$$\alpha\ :=\ \min_{i \in I(x)} < g_i'(x), v_0 >$$

and

$$\lambda\ :=\ \max(0, y_1 - < g_1'(x), v >, \ldots, y_p - < g_p'(x), v >)/\alpha$$

We set

$$a_i\ := < g_i'(x), v + \lambda v_0 > -\ y_i$$

By construction, $a_i \geq 0$ for any $i \in I(x)$, so that $a := (a_1, \ldots, a_p)$ belongs to the contingent cone to $\mathbf{R}_+^p$ to $g(x)$. Hence $w := \lambda v_0 + v$ belongs to the convex cone $C_L(x)$ and is a solution to the equations $g'(x)w - a = y$ and $h'(x)w = z$, so that

$$(y, z)\ \in\ \Big(g'(x) \times h'(x))u - T_{\mathbf{R}_+^p \times \{0\}}(g(x), h(x))\Big)\quad \square$$

Remark — Observe that the above proof implies that if $L = X$ in Proposition 4.3.7, then an element $u \in C_K(x)$ if and only if

$$\forall\, i \in I(x),\ \ < g_i'(x), u > \geq\ 0\ \ \&\ \ \forall\, j = 1, \ldots, q,\ \ h_j'(x)u\ =\ 0\ \square$$

Table 4.4: **Properties of Contingent Cones to Derivable Sets in Finite Dimensional Spaces.**

(5)	▷	If $L \subset X$ and $M \subset Y$ are *closed derivable* subsets, $f \in \mathcal{C}^1(X,Y)$ is a continuously differentiable map and $x \in L \cap f^{-1}(M)$ satisfies the *transversality condition* $f'(x)C_L(x) - C_M(f(x)) = Y$, then $T_{L\cap f^{-1}(M)}(x) \ = \ T_L(x) \cap f'(x)^{-1}T_M(f(x))$
(5)a)	▷	If $M \subset Y$ is a *closed derivable* subset, $f \in \mathcal{C}^1(X,Y)$ is a continuously differentiable map and $x \in f^{-1}(M)$ satisfies $\mathrm{Im}(f'(x)) - C_M(f(x)) = Y$, then $T_{f^{-1}(M)}(x) \ = \ f'(x)^{-1}T_M(f(x))$
(5)b)	▷	If K_1 and K_2 are *closed derivable* subsets contained in X, $x \in K_1 \cap K_2$ satisfies $C_{K_1}(x) - C_{K_2}(x) = X$, then $T_{K_1\cap K_2}(x) \ = \ T_{K_1}(x) \cap T_{K_2}(x)$
(5)c)	▷	If $K_i \subset X$, $(i = 1, \ldots, n)$ are *closed derivable* subsets and $x \in \bigcap_i K_i$ satisfies $\forall v_i = 1, \ldots, n, \ \ \bigcap_{i=1}^n (C_{K_i}(x) - v_i) \neq \emptyset$, then $T_{\bigcap_{i=1}^n K_i}(x) \ = \ \bigcap_{i=1}^n T_{K_i}(x)$

4.3.3 Direct Image

Let us consider now two normed vector spaces X and Y, a subset $K \subset X$ and a differentiable single-valued map f from X to Y.

We saw that for any $y \in f(K)$, we have

$$\overline{\bigcup_{x \in K \cap f^{-1}(y)} f'(x)(T_K(x))} \subset T_{f(K)}(y)$$

and

$$\overline{\bigcup_{x \in K \cap f^{-1}(y)} f'(x)(T_K^\flat(x))} \subset T_{f(K)}^\flat(y)$$

It is not that easy to find elegant sufficient conditions implying the equality

$$\overline{\bigcup_{x \in K \cap f^{-1}(y)} f'(x)(T_K(x))} = T_{f(K)}(y)$$

Theorem 4.3.8 *Let X be a finite dimensional vector-space, Y a normed space, $\Omega \subset X$ be an open subset containing K and $f : \Omega \mapsto Y$ be a single-valued map satisfying*

$$f(K) \ni y \leadsto f^{-1}(y) \cap K$$

is pseudo-Lipschitz *at $(f(x_0), x_0)$: there exists a constant $l > 0$ such that, for any $y \in f(K)$ close to $f(x_0)$,*

$$d\left(x_0, f^{-1}(y) \cap K\right) \leq l\|y - f(x_0)\|$$

Then, if f is Fréchet differentiable at x_0, we obtain the equality

$$\overline{f'(x_0)T_K(x_0)} = T_{f(K)}(f(x_0))$$

Proof — Let v belong to $T_{f(K)}(f(x_0))$. Then there exist sequences of elements $h_n > 0$ and v_n converging to 0 and v respectively such that

$$f(x_0) + h_n v_n = f(x_n) \in f(K) \tag{4.9}$$

The point is to choose solutions $x_n \in K$ to the above equation (4.9) such that a subsequence of $u_n := (x_n - x_0)/h_n$ converges to some u. Such an element u belongs to the contingent cone $T_K(x_0)$ and is a solution to the equation $f'(x_0)u = v$.

Since the set-valued map

$$f(K) \ni y \rightsquigarrow K \cap f^{-1}(y)$$

is pseudo-Lipschitz at $(f(x_0), x_0)$ by assumption, there exist a constant l and solutions $x_n \in K$ to the equation (4.9) such that

$$\|x_0 - x_n\| \leq l\|f(x_0) - f(x_0) - h_n v_n\| = l h_n \|v_n\|$$

Therefore, the sequence of elements u_n is bounded, so that a subsequence (again denoted) u_n converges to some u. □

Remark — Therefore, any sufficient condition implying that for some $x_0 \in K$, the set-valued map

$$f(K) \ni y \rightsquigarrow K \cap f^{-1}(y)$$

is pseudo-Lipschitz at $(f(x_0), x_0)$ will automatically imply the above equality between the contingent cone to the image and the closure of the image of the contingent cone.

One of them is given by the Criterion of Pseudo-Lipschitzeanity (See Corollary 3.4.8.) □

Actually, when the dimension of X is finite, we have another criterion:

Proposition 4.3.9 *Let X, Y be normed spaces, $K \subset X$, $A \in \mathcal{L}(X,Y)$ be a continuous linear operator and $y_0 \in Y$. If the dimension of X is finite and if for some $x_0 \in K \cap A^{-1}(y_0)$*

$$\ker(A) \cap T_K(x_0) = \{0\}$$

then $\overline{A(T_K(x_0))} = T_{A(K)}(A(x_0))$.

Proof — Let v belong to $T_{A(K)}(A(x_0))$. There exist $h_n > 0$ converging to 0, $v_n \in Y$ converging to v and $x_n \in K$ such that

$$A(x_0) + h_n v_n = A(x_n) = A(x_0 + h_n u_n)$$

where we set $u_n := (x_n - x_0)/h_n$.

If the sequence u_n is bounded, being in a finite dimensional vector-space, a subsequence converges to some u, which belongs to

$T_K(x_0)$, and satisfies $Au = v$. If not, a subsequence (again denoted) $\|u_n\|$ would go to ∞. Then a subsequence of elements $\widehat{u}_n = u_n/\|u_n\|$, satisfying equations $A\widehat{u}_n = v_n/\|u_n\|$, converges to some u in the unit sphere, which is a solution to the equation $Au = 0$ and which belongs to the contingent cone $T_K(x_0)$. This is impossible by assumption. □

When the map is no longer linear and the dimension of X still finite, we can prove:

Proposition 4.3.10 *Let K be a subset of a finite dimensional vector-space X, Y be a normed space, $f : X \to Y$ a Fréchet differentiable function and $y_0 \in f(K)$, $x_0 \in K \cap f^{-1}(y_0)$.*

Assume that for some $\varepsilon > 0$, the subset $f^{-1}(B(y_0, \varepsilon) \cap K)$ is bounded. Then condition

$$\forall\, x \in f^{-1}(y_0) \cap \overline{K}, \;\; \ker(f'(x)) \cap T_K(x) \;=\; \{0\}$$

implies

$$\bigcup_{x \in \overline{K} \cap f^{-1}(y_0)} f'(x)(T_K(x)) \;=\; T_{f(K)}(y_0)$$

Proof — Let v belong to $T_{f(K)}(y_0)$ be different from 0. There exist $h_n > 0$ converging to 0, $v_n \in Y$ converging to v and $x_n \in K$ such that $y_0 + h_n v_n = f(x_n)$, which are then bounded by assumption. Then a subsequence (again denoted) x_n converges to some $x \in \overline{K}$ satisfying $f(x) = y_0$.

We set $\alpha_n := \|x_n - x\|$ and $u_n := (x_n - x)/\alpha_n$. Then α_n converges to 0 and the sequence u_n being bounded in a finite dimensional vector-space, has a subsequence converging to some u, which belongs to $T_K(x)$.

On the other hand,

$$f(x + \alpha_n u_n) \;=\; f(x) + \alpha_n f'(x)u + \alpha_n \varepsilon(\alpha_n) \;=\; y_0 + h_n v_n$$

Hence

$$f'(x)u + \varepsilon(\alpha_n) \;=\; h_n v_n/\alpha_n$$

The left hand side being bounded and v_n converging to $v \neq 0$, we deduce that the sequence h_n/α_n is bounded, so that a subsequence converges to some λ. Therefore $f'(x)u = \lambda v$. By assumption, λ must be different from 0. Consequently,

$$v \;=\; f'(x)u/\lambda \;\in\; f'(x)(T_K(x)) \;\square$$

4.4 Normal Cones

We introduce in this section the dual concept of tangent cones: the normal cones.

We begin with the following characterization of the polar of the contingent cone:

Proposition 4.4.1 *Let K be a subset of a finite dimensional vector-space X. Then $p \in (T_K(x))^-$ if and only*

$$\left\{ \begin{array}{l} \forall\, \varepsilon > 0,\ \exists\, \eta > 0 \text{ such that } \forall\, y \in K \cap B(x, \eta), \\ \\ < p, y - x > \leq \ \varepsilon \|y - x\| \end{array} \right. \tag{4.10}$$

Proof — Let p satisfy above property (4.10) and $v \in T_K(x)$. Then there exist h_n converging to 0 and v_n converging to v such that

$$y \ := \ x + h_n v_n \ \in \ K \cap B(x, \eta)$$

for n large enough. Consequently, inequalities $< p, v_n > \leq \varepsilon$ imply by taking the limit that $< p, v > \leq \varepsilon$ for all $\varepsilon > 0$. Hence $< p, v > \leq 0$, so that any element p satisfying the above property belongs to the polar cone of $T_K(x)$, even in the case of normed spaces.

Conversely, assume that p violates property (4.10): There exist $\varepsilon > 0$ and a sequence of elements $x_n \in K$ converging to x such that

$$\langle p, x_n - x \rangle \ > \ \varepsilon \|x_n - x\|$$

We set $h_n := \|x_n - x\|$, which converges to 0, and $v_n := (x_n - x)/h_n$. These elements belonging to the unit sphere, a subsequence (again denoted) v_n converges to some v. By definition, this limit belongs to $T_K(x)$, so that $< p, v > \leq 0$. But our choice implies that $< p, v_n >>$ ε, so that $< p, v > \geq \varepsilon$, a contradiction. □

Notice that Proposition 4.1.2 implies that when the Banach space X is *smooth*, then for any $y \notin K$ such that $\Pi_K(y)$ is not empty and for any $z \in \Pi_K(y)$, $J(y - z)$ belongs to the $T_K(z)^-$.

The most important justification for dealing with dual concepts, is the availability of the one to one correspondences between closed convex cones and their polar cones, continuous linear operators and their transposes, lower semicontinuous convex functions and their conjugates. This requires one to deal with convex notions.

Since the Clarke tangent cone is convex, it can then be characterized by its polar cone, which, by analogy with the case of smooth manifolds, can be regarded as the normal cone.

Definition 4.4.2 *Let x belong to $K \subset X$. We shall say that the (negative) polar cone*

$$N_K(x) := C_K(x)^- = \{p \in X^\star \mid \forall\, v \in C_K(x),\ < p, v > \leq 0\}$$

is the (Clarke) normal cone *to K at x.*

The normal cone $N_K(x)$ is equal to the whole space whenever $C_K(x)$ is reduced to zero.

The normal cone is related to the closed convex hull of the upper limit of the polar cones to the contingent cones:

Theorem 4.4.3 *Let X be a finite dimensional vector-space and K be a closed subset of X. Then*

$$N_K(x) = \overline{co}\left(\text{Limsup}_{y \to_K x} \left(T_K(y)\right)^-\right)$$

This a consequence of Theorem 4.1.10 and the Duality Theorem 1.1.8. In infinite dimensional spaces, we deduce from Theorem 4.1.13 its dual version

Theorem 4.4.4 *Assume that X is uniformly smooth and that the norm of $X^\star$ is Fréchet differentiable off the origin. Let K be a closed subset of X. Then*

$$N_K(x) = \overline{co}\left(\sigma - \text{Limsup}_{y \to_K x} \left(T_K^\sigma(y)\right)^-\right)$$

Unfortunately, since the contingent and intermediate cones are not convex, duality may just be a blissful wish, since many problems, which are not smooth by nature, force us to use naturally these larger tangent cones.

The price to pay in terms of loss of information for playing with duality, just to be able to conserve some familiar classical formulation, is indeed too high in many situations.

Therefore, these dual concepts we have presented are recommended for convex, or more generally, *sleek* subsets.

Table 4.5: **Properties of Normal Cones In Finite Dimensional Vector Spaces.**

(3)	▷	If $x_i \in K_i \subset X_i$, $(i = 1, \ldots, n)$, then $N_{\prod_{i=1}^n K_i}(x_1, \ldots, x_n) \;=\; \prod_{i=1}^n N_{K_i}(x_i)$
(5)	▷	If $L \subset X$ and $M \subset Y$ are *closed* , $x \in L \cap f^{-1}(M)$, $f \in \mathcal{C}^1(X, Y)$ is continuously differentiable and the *transversality condition* $f'(x)C_L(x) - C_M(f(x)) = Y$ holds true, then $N_{L \cap f^{-1}(M)}(x) \;\subset\; N_L(x) + f'(x)^\star N_M(f(x))$
(5)a)	▷	If $M \subset Y$ is a *closed* subset and if $f \in \mathcal{C}^1(X, Y)$ is a continuously differentiable map such that $\mathrm{Im}(f'(x)) - C_M(f(x)) = Y$, then $N_{f^{-1}(M)}(x) \;\subset\; f'(x)^\star N_M(f(x))$
(5)b)	▷	If K_1 and K_2 are closed subsets contained in X and $x \in K_1 \cap K_2$ satisfies $C_{K_1}(x) - C_{K_2}(x) = X$, then $N_{K_1 \cap K_2}(x) \;\subset\; N_{K_1}(x) + N_{K_2}(x)$
(5)c)	▷	If $K_i \subset X$, $(i = 1, \ldots, n)$ are closed, $x \in \bigcap_i K_i$ and $\forall\, v_i = 1, \ldots, n,\;\; \bigcap_{i=1}^n (C_{K_i}(x) - v_i) \neq \emptyset$, then $N_{\bigcap_{i=1}^n K_i}(x) \;\subset\; \sum_{i=1}^n N_{K_i}(x)$

4.5 Other Tangent Cones

4.5.1 Convex Kernel of a Cone

Let us recall that the *Minkowski difference* $K \ominus L$ of two subsets K and L of a vector space is the subset $\bigcap_{x \in L}(K - x)$ of elements z such that $L + z \subset K$.

Let T be a closed cone. We observe that *the Minkowski difference*

$$T \ominus T \ := \ \{y \in X \mid y + T \subset T\}$$

is a closed convex cone contained in T.

Indeed, it is obviously a cone, contained in T (because $0 \in T$) and is convex: If y and z belong to $T \ominus T$, then

$$y + (z + T) \ \subset \ y + T \ \subset \ T$$

so that $y + z$ belongs also to $T \ominus T$.

Definition 4.5.1 *When T is a closed cone, the Minkowski difference $T \ominus T$ is called the* convex kernel (or the asymptotic cone) *of T and is denoted by T^∞.*

The convex kernel provides an explicit convex subcone of T, which may not be the largest one, but is the intersection of the maximal[11] closed convex subcones of T:

Proposition 4.5.2 *The convex kernel of a closed cone T is the intersection of the maximal closed convex cones contained in T.*

Proof — Let $\mathcal{P}$ denote the family of the maximal closed convex cones contained in T. Then

$$T^\infty \subset \bigcap_{P \in \mathcal{P}} P$$

Indeed, if x belongs to T^∞ and $P \in \mathcal{P}$, then the closed convex cone $\overline{P + \mathbf{R}_+ x}$ is a closed convex cone containing P and contained in T because x is in the Minkowski difference. The maximal character of P implies that it is equal to $\overline{P + \mathbf{R}_+ x}$, and thus, that $x \in P$.

On the other hand,

$$\bigcup_{P \in \mathcal{P}} P \ = \ T$$

[11]We say that a closed convex cone $K \subset T$ is *maximal* if for any other closed convex cone P satisfying $K \subset P \subset T$, we have $K = P$.

because for any $x \in T$, the closed convex cone $\mathbf{R}_+ x$ is contained in a cone $P \in \mathcal{P}$, thanks to Zorn's Lemma: Indeed, for any family of closed convex cones $K_i \subset T$ totally ordered for inclusion, $\overline{\bigcup_i K_i}$ is a closed convex subcone of T which is an upper bound of the K_i's.

To prove that $\bigcap_{P \in \mathcal{P}} P \subset T^\infty$, pick any $x \in \bigcap_{P \in \mathcal{P}} P$ and any $y \in T$. Since y belongs to some $P \in \mathcal{P}$, then

$$x + y \in P + P \subset P \subset T \quad \square$$

Proposition 4.5.3 *Let K be a closed cone in a normed space and K^∞ its convex kernel. Then*

$$C_K(0) = K^\infty \ \& \ \forall\, x \in K, \ C_K(x) \supset \overline{K^\infty + \mathbf{R}x}$$

Proof — Fix $x \in K$. Take any element $\lambda x + y \in \mathbf{R}x + K^\infty$ where $y \in K^\infty$ and consider any sequence of elements $h_n > 0$ converging to 0 and $x_n \in K$ converging to x. Since

$$x_n + h_n(\lambda x_n + y) = (1 + \lambda h_n)x_n + h_n y \in K$$

we infer that $\lambda x + y$ belongs to $C_K(x)$.

Consider now $v \in C_K(0)$ and assume for a moment that it does not belong to K^∞. Then there exists $x \in K$ such that $x + v \notin K$, and thus, $\varepsilon > 0$ with

$$(x + v + \varepsilon B) \cap K = \emptyset$$

But for $h > 0$ small enough, we know that there exists $e \in \varepsilon B$ with

$$hx + h(v + e) = h(x + v + e) \in K$$

Since K is a cone, $x + v + e$ belongs to K, a contradiction. $\square$

We can interpret the second statement of Proposition 4.1.6 by saying that the *Clarke tangent cone is contained in the convex kernels of the contingent cone and adjacent cone.*

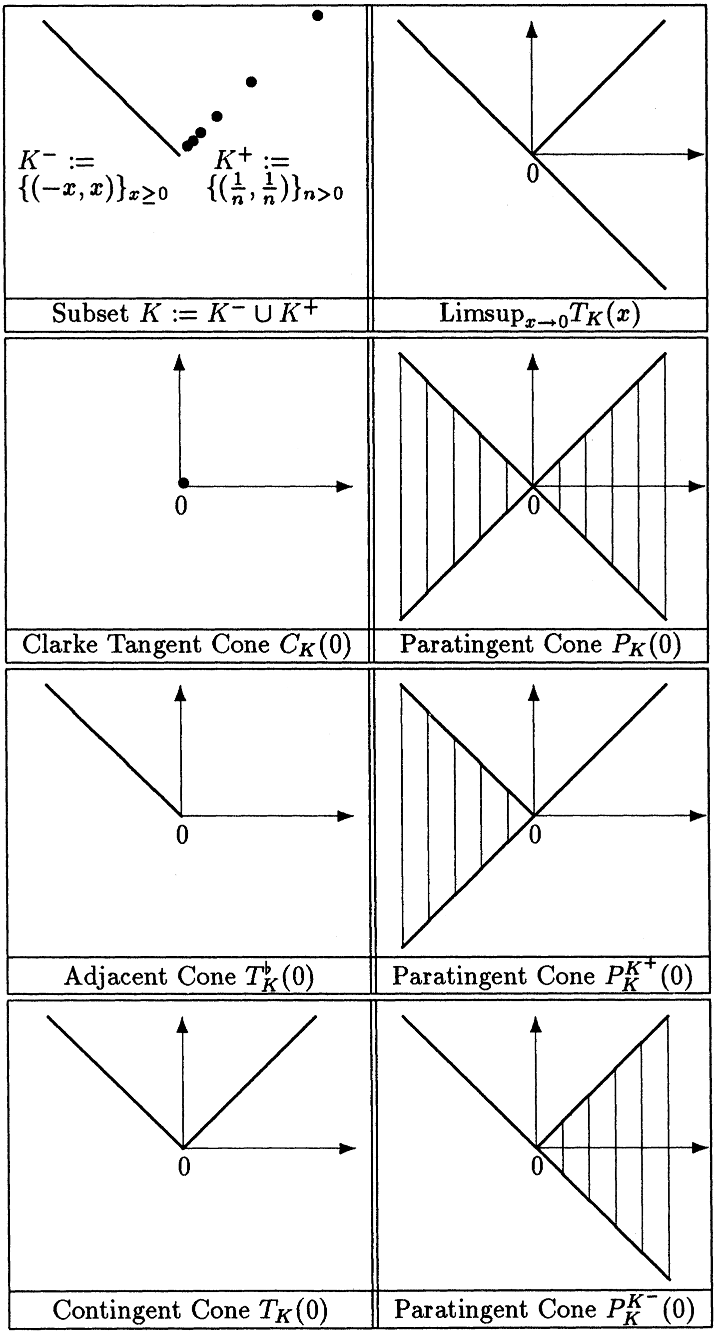
If one is interested in dealing with convex subcones of the contingent cone without requiring regularity, it may be advantageous to use the convex kernel of the contingent cone.

4.5.2 Paratingent Cones

Bouligand had also introduced in the thirties the concept of *paratingent cone* $P_K(x)$ to K at $x \in K$, which can be regarded as a symmetric concept to the one of Clarke tangent cone, in the sense that it is defined as

$$P_K(x) := \operatorname{Limsup}_{h \to 0+,\, x' \to_K x} \frac{K - x'}{h}$$

Figure 4.4: The Ménagerie of Tangent Cones: they may be all different



Actually, this cone may be too large to be useful:

For instance, take $X := \mathbf{R}^2$ and $K := \mathbf{R} \times \mathbf{R}_+$. Then

$$T_K(x,0) = C_K(x,0) = K \text{ and } P_K(x,0) = X$$

Therefore, we are led to introduce a more precise concept.

Definition 4.5.4 (Paratingent Cone) *Let $K \subset X$ be a subset of a normed space X, L a subset of X and $x \in \overline{L} \cap \overline{K}$. The Bouligand* paratingent cone *$P_K^L(x)$ to K relative to L at x is defined by*

$$P_K^L(x) := \left\{ v \mid \liminf_{h \to 0+,\ x' \to_L x} d_K(x' + hv)/h = 0 \right\}$$

When $L = K$, we set $P_K(x) := P_K^K(x)$.

In other words,

$$\forall\, x \in \overline{L} \cap \overline{K},\ P_K^L(x) = \operatorname{Limsup}_{h \to 0+, y \to_L x} \frac{K - y}{h}$$

One of the current choices of L is the boundary ∂K of K or the complement $\widehat{K}$ of K.

Lemma 4.5.5 *If K is open, then*

$$\forall\, x \in \partial K,\ P_K^{\partial K}(x) = P_K^{\widehat{K}}(x)$$

Proof — Since $P_K^{\partial K}(x) \subset P_K^{\widehat{K}}(x)$, we have to show that any $v \in P_K^{\widehat{K}}(x)$, i.e., satisfying

$$\exists\, x_n \to_{\widehat{K}} x,\ \exists\, h_n \to 0+,\ \exists\, v_n \to v \text{ such that } x_n + h_n v_n \in K$$

belongs also to $P_K^{\partial K}(x)$. Since K is open and $x \in \widehat{K}$, there exists $k_n \in]0, h_n[$ such that $y_n := x_n + k_n v_n$ belongs to ∂K and converges to x in ∂K. Take $l_n := h_n - k_n > 0$, which converges to 0 and note that

$$y_n + l_n v_n = x_n + h_n v_n$$

belongs to K for any n. Hence v does belongs to $P_K^{\partial K}(x)$. □

We observe at once that

$$\forall\, x \in \overline{K} \cap \overline{L},\ P_K^L(x) = -P_L^K(x)$$

and in particular, that $P_K(x) = -P_K(x)$ is symmetric.

Proposition 4.5.6 *Let L and K be closed subsets of a normed space X. The graph of the set-valued map $P_K^L : K \cap L \rightsquigarrow X$ is closed:*

$$\forall\, x \in K \cap L, \ \ \mathrm{Limsup}_{y \to_L x} P_K^L(y) \ \subset \ P_K^L(x)$$

Proof — Consider a sequence of elements (x_n, v_n) of the graph of the restriction of $P_K^L(\cdot)$ to $K \cap L$ converging to (x, v). There are sequences $x_{mn} \in L$ converging to x_n in L, h_{mn} converging to $0+$ and v_{mn} converging to v_n such that $x_{mn} + h_{mn} v_{mn}$ belongs to K for every n. We then can associate with any $k > 0$ integers m_k and n_k such that

$$\max\left(\|x_{n_k} - x\|, \|x_{m_k n_k} - x_{n_k}\|, h_{m_k n_k}, \|v_{n_k} - v\|, \|v_{m_k n_k} - v_{n_k}\|\right) \le 1/k$$

By taking

$$y_k \ := \ x_{m_k n_k}, \ \ h_k \ := \ h_{m_k n_k}, \ \ v_k \ := \ v_{m_k n_k}$$

we observe that $y_k \in L$ converges to x, $h_k > 0$ converges to 0, v_k converges to v and that $y_k + h_k v_k$ belongs to K for every $k > 0$. Hence v belongs to $P_K^L(x)$. □

A celebrated theorem of G. Choquet states that the contingent cone $T_K(x)$ and the paratingent cone $P_K^{\partial K}(x)$ coincide on a residual[12] of the boundary ∂K.

Theorem 4.5.7 (Choquet's Theorem) *Let us consider a separable Banach space X, $K \subset X$ and let L be a closed subset of X. Then there exists a residual $R \subset L$ of L such that*

$$\forall\, x \in \overline{K} \cap R, \ \ T_K(x) \ = \ P_K^L(x)$$

Consequently, the graph of the restriction of $T_K(\cdot)$ to $\overline{K} \cap R$ is closed.

Proof — By definition of the upper limit, we have:

$$\left\{ \begin{array}{ll} P_K^L(x) & = \bigcap_{\varepsilon,\delta,\eta>0} \bigcup_{y \in B_L(x,\eta)} \bigcup_{h \in]0,\delta]} \left((K - y)/h + \varepsilon B\right) \\ & = \bigcap_{\delta>0} \bigcap_{\varepsilon,\eta>0} \left(\bigcup_{y \in B_L(x,\eta)} \left[\bigcup_{h \in]0,\delta]} (K - y)/h \right] + \varepsilon B \right) \\ & = \bigcap_{n>0} \mathrm{Limsup}_{y \to_L x} \left[\bigcup_{h \in]0,1/n]} (K - y)/h \right] \end{array} \right.$$

[12]Recall that a *residual* is a countable intersection of dense open subsets A_n. Countable intersections of residuals are residuals. *Baire's Theorem* states that *a residual of a complete metric space is dense.* A property which is true for every element of a residual is called *generic.*

Set

$$S_n(y) := \overline{\bigcup_{h\in]0,1/n]} (K-y)/h}$$

The set-valued map

$$S_n : L \ni y \rightsquigarrow S_n(y)$$

is obviously lower semicontinuous with closed values. By Theorem 1.4.13, there exists a residual $A_n \subset L$ such that

$$\forall\, x \in A_n,\ \ \mathrm{Limsup}_{y\to_L x} S_n(y) \;=\; S_n(x)$$

Take $R := \bigcap_{n>0} A_n$, which is still a residual. Hence

$$\forall\, x \in R,\ \ \mathrm{Limsup}_{y\to_L x} S_n(y) \;=\; S_n(x)$$

so that

$$\forall\, x \in R,\ \ P_K^L(x) \;=\; \bigcap_{n>0} S_n(x) \;=\; \mathrm{Limsup}_{h\to 0+}(K-x)/h \;=\; T_K(x) \ \square$$

The paratingent cones $P_K^{\widehat{K}}(x)$ are *complementary convex*, i.e., their complements are convex, and the contingent cones are generically complementary convex. This follows from the next subsection.

4.5.3 Hypertangent Cones

We introduce the *hypertangent* cone to K at x defined as follows:

Definition 4.5.8 (Hypertangent Cone) *Let x belong to a subset K of a normed space X. The* hypertangent cone $H_K(x)$ *to K at $x \in K$ is the set of elements $v \in X$ such that there exist $\varepsilon > 0$, $\delta > 0$ and $\eta > 0$ for which*

$$B_K(x,\eta) + [0,\delta](v+\varepsilon B) \;\subset\; K$$

In the same way that the Dubovitskij-Miljutin cone (see Definition 4.1.3) to the complement of K is the complement of the contingent cone to K, it is easy to check the following

Proposition 4.5.9 *The hypertangent cone is an* open convex cone *contained in the Dubovitskij-Miljutin cone and satisfying*

$$H_K(x) + T_K^\flat(x) \;\subset\; D_K(x)$$

Furthermore,

$$\forall\, x \in \partial K,\ \ \text{the complement of } P_K^{\widehat{K}}(x) \text{ is equal to } H_{\widehat{K}}(x) \;=\; -H_K(x)$$

This and Choquet's Theorem 4.5.7 imply the following generic regularity result:

Theorem 4.5.10 (Shi Shuzhong) *Let X be a separable Banach space, $K \subset X$ be an open subset. Then there exists a residual $R \subset \partial K$ of the boundary of K such that*

$$\forall\, x \in \partial K,\ \ T_K(x) \ = \ P_K^{\partial K}(x) \ = \ -(X \setminus H_K(x))$$

In particular, $T_K(\cdot)$ is generically complementary convex on ∂K.

Proof — Indeed, Choquet's Theorem 4.5.7 implies that there exists a residual of ∂K on which $T_K(\cdot)$ and $P_K^{\partial K}(\cdot)$ coincide. Furthermore, K being open by the last lemma, $P_K^{\partial K}(x)$ is equal to $P_K^{\widehat{K}}(x) = -P_{\widehat{K}}^{K}(x)$. We then use the fact that $P_{\widehat{K}}^{K}(x)$ is the complement of $H_K(x)$. □

4.5.4 A Ménagerie of Tangent Cones

A. D. Ioffe has introduced a systematic way to classify the four main tangent cones which we have introduced, i.e., the paratingent, contingent, adjacent and circatangent cones, as well as the Dubovitskij-Miljutin and hypertangent cones.

Let K and L be subsets of a normed space X and $x \in \overline{K} \cap \overline{L}$. We denote by the notation $\dagger(a_n \to a)$ the fact that $a_n = a$ for all n. The subset $T_{QRS}^{L}(K,x)$ denotes the cone of elements $v \in X$ satisfying

$$Q(x_n \to_L x),\ R(h_n \to 0+),\ S(v_n \to v),\ \ x_n + h_n v_n \in K$$

where $Q(\cdot), S(\cdot) \in \{\forall,\ \exists,\ \dagger\}$ and $R(\cdot) \in \{\forall,\ \exists\}$.

Hence the cones we have introduced can be written

$$\begin{cases} T_K(x) & = & T_{\dagger,\exists,\exists}^{K}(K,x) \\ T_K^{\flat}(x) & = & T_{\dagger,\forall,\exists}^{K}(K,x) \\ C_K(x) & = & T_{\forall,\forall,\exists}^{K}(K,x) \\ P_K^{L}(x) & = & T_{\exists,\exists,\exists}^{L}(K,x) \\ D_K(x) & = & T_{\dagger,\forall,\forall}^{K}(K,x) \\ H_K(x) & = & T_{\forall,\forall,\forall}^{K}(K,x) \end{cases}$$

The largest of these cones (when $L = K$) is the paratingent cone $T^K_{\exists,\exists,\exists}(K,x)$ and the smallest is the hypertangent cone $T^K_{\forall,\forall,\forall}(K,x)$.

The proof of Proposition 4.1.6 implies that the cones $T^K_{\forall,\forall,S}(K,x)$ are convex and are contained in the convex kernel of $T^K_{QRS}(K,x)$:

$$T^K_{\forall,\forall,S}(K,x) + T^K_{QRS}(K,x) \subset T^K_{QRS}(K,x)$$

We also observe the following complementary property: recall that $\widehat{K}$ denotes the complement of K and set $\widehat{\forall} := \exists$, $\widehat{\exists} = \forall$ and $\widehat{\dagger} = \dagger$. Then

$$\text{the complement of } \; T^L_{QRS}(K,x) \; \text{ is equal to } \; T^L_{\widehat{Q}\widehat{R}\widehat{S}}(\widehat{K},x)$$

In infinite dimensional space, the complexity is compounded by the fact that (at least for the properties $S(\cdot)$) we can use either the strong or weak convergence....

It is clear that we can meet in the proofs of theorems various instances of these tangent cones, but few of them enjoy useful properties. As far as the experience of the authors and some of their colleagues is concerned, the contingent cone plays a major role both in optimization theory and viability theory, the use of the Clarke tangent cone is necessary when lower semicontinuity is required or convexity is needed (in this case, it may be still better to use convex kernels whenever the Clarke tangent cone is degenerate.) The paratingent cones are the only ones whose graphs are closed, a property which is quite important indeed, and which can be transmitted generically thanks to Choquet's Theorem.

The introduction of the cones $T^K_{\forall,R,S}(K,x)$ can also be justified as particular cases of asymptotic tangent cones to (constant) sequences of subsets. This is the topic of the next section.

4.6 Tangent Cones to Sequences of Sets

Let X be a normed vector space. We consider a sequence of subsets $K_n \subset X$ of X and their upper and lower limits denoted by $K^\sharp$ and $K^\flat$ respectively.

We shall adapt the concepts of contingent cones and intermediate cones to *sequences* of subsets K_n in the following way:

Definition 4.6.1 *We shall say that the* circatangent cone *to the sequence $(K_n)_n$ at $x \in K^\flat$ is the set $C^\flat_{(K_n)}(x)$ defined by*

$$C^\flat_{(K_n)}(x) \;:=\; \mathrm{Liminf}_{n\to\infty, K_n \ni x_n \to x, h_n \to 0+} \frac{K_n - x_n}{h_n}$$

If $x \in K^\sharp$, we shall say that

$$P^\sharp_{(K_n)}(x) \ := \ \mathrm{Limsup}_{n\to\infty, K_n \ni x_n \to x, h_n \to 0+} \frac{K_n - x_n}{h_n}$$

is the paratingent cone *to the sequence $(K_n)_n$ at x.*

Remark — When we consider a *constant sequence $K_n := K$*, we see that circatangent cone $C^\flat_{(K_n)}(x)$ coincides with the Clarke tangent cone $C_K(x)$.

In the case of the constant sequence $K_n := K$, the paratingent cone $P^\sharp_{(K_n)}(x)$ coincides with the *paratingent cone $P_K(x)$.* □

The upper limit of the paratingent cones to a sequence of K_n is contained in the paratingent cone to the sequence of the K_n's:

Proposition 4.6.2 *Let us consider a sequence of subsets $K_n \subset X$ of X and their upper limit $K^\sharp$. Then*

$$\mathrm{Limsup}_{K_n \ni x_n \to x} T_{K_n}(x_n) \ \subset \ \mathrm{Limsup}_{K_n \ni x_n \to x} P_{K_n}(x_n) \ \subset \ P^\sharp_{(K_n)}(x)$$

Proof — It is the same as the one of Proposition 4.5.6. □

We provide a characterization of the polar cone to the paratingent cone to a sequence of subsets:

Proposition 4.6.3 *Let X be a finite dimensional vector-space, K_n be a sequence of closed subsets of X. Then $p \in \left(P^\sharp_{(K_n)}(x)\right)^-$ if and only if there exists a sequence $\widehat{x}_n \in K_n$ converging to x such that*

$$\begin{cases} \forall\, \varepsilon > 0, \ \exists\, N, \ \eta > 0 \ \text{ such that } \forall\, n \geq N, \\ \forall\, x_n \in K_n \cap B(\widehat{x}_n, \eta), \ \ < p, x_n - \widehat{x}_n > \ \leq \ \varepsilon \|x_n - \widehat{x}_n\| \end{cases} \tag{4.11}$$

Proof — Let p satisfy above property (4.11) and $v \in P^\sharp_{(K_n)}(x)$. Then there exist h_n converging to 0 and v_n converging to v such that for n large enough

$$x_n \ := \ \widehat{x}_n + h_n v_n \ \in \ K_n \cap B(\widehat{x}_n, \eta)$$

From inequalities $< p, v_n > \leq \varepsilon$ we deduce that $< p, v > \leq \varepsilon$ for all $\varepsilon > 0$. Hence $< p, v > \leq 0$, so that

$$p \ \in \ \left(P^\sharp_{(K_n)}(x)\right)^-$$

even in the case of normed spaces.

Conversely, assume that $p \in P^\sharp_{(K_n)}(x)$ violates property (4.11): There exist $\varepsilon > 0$ and a sequence $x_n \in K_n$ converging to x such that

$$\langle p, x_n - \widehat{x}_n \rangle \; > \; \varepsilon \|x_n - \widehat{x}_n\|$$

Set $h_n := \|x_n - \widehat{x}_n\|$ (which converges to 0) and $v_n := (x_n - \widehat{x}_n)/h_n$. These elements belonging to the unit sphere, a subsequence (again denoted) v_n converges to some $v \in P^\sharp_{(K_n)}(x)$, so that $< p, v > \leq 0$. But our choice implies that $< p, v_n >> \varepsilon$, so that $< p, v > \geq \varepsilon$, a contradiction. □

The circatangent cone to a sequence of subsets enjoys the same properties as the Clarke tangent cones: it is always a closed convex cone and is actually equal to the lower limit of the contingent cones in finite dimensional spaces:

Theorem 4.6.4 *Let us consider a sequence of closed subsets $K_n \subset X$ of a Banach space X and an element $x_0 = \lim_{n\to\infty} x_n$ (where $x_n \in K_n$) of their lower limit $K^\flat$. Then*

1. — The circatangent cone $C^\flat_{(K_n)}(x_0)$ is always a closed convex cone

2. — The lower limit of the contingent cones $T_{K_n}(x_n)$ is contained in the circatangent cone:

$$\mathrm{Liminf}_{n\to\infty, K_n \ni x_n \to x} T_{K_n}(x_n) \;\subset\; C^\flat_{(K_n)}(x_0)$$

and equality holds when the dimension of X is finite.

Proof of the first statement — It is similar to the one of Proposition 4.1.6.

Proof of the second statement — Pick an element

$$v \in \mathrm{Liminf}_{K_n \ni x_n \to x_0} T_{K_n}(x_n)$$

with norm equal to one. This means that for any $\varepsilon > 0$, there exists $N \geq 1$ such that

$$\forall\, n \;\geq\; N,\;\; \forall\, z \;\in\; B_{K_n}(x_0, 1/n),\;\; B(v, \varepsilon/2) \cap T_{K_n}(z) \;\neq\; \emptyset \qquad (4.12)$$

We shall derive a contradiction by assuming that v does not belong to $C^\flat_{(K_n)}(x_0)$. This means that there exist $\varepsilon \in]0, 1/4[$ and sequences

$$x_n \;\in\; B_{K_n}(x_0, 1/2n), \quad h_n \in]0, 1/4n[$$

with

$$(x_n + h_n B(v,\varepsilon)) \cap K_n = \emptyset \tag{4.13}$$

Fix N such that (4.12) holds true, $n \geq N$ and set $\delta := \varepsilon/(1+\varepsilon)$.

We use Ekeland's Variational Principle (Theorem 3.3.1) in exactly the same way as in the proof of Theorem 4.1.9 to verify that (4.13) cannot hold.

Proof of the third statement — As for Lemma 4.1.14, the third statement follows from the second one and:

Lemma 4.6.5 *Let K_n be nonempty closed subsets of a reflexive Banach space and $x_0 = \lim_{n\to\infty} x_n$ (where $x_n \in K_n$). Then*

$$C^{\flat}_{(K_n)}(x_0) \subset \operatorname{Liminf}_{K_n \ni x_n \to x_0} T^{\sigma}_{K_n}(x_n)$$

Proof — The proof is analogous to the one of Lemma 4.1.14. □

When the subsets K_n are convex, we obtain the following relations:

Proposition 4.6.6 *Let us consider a sequence of closed convex subsets K_n of a reflexive Banach space X and its upper and lower limits $K^{\sharp}$ and $K^{\flat}$. Then for every $x \in K^{\flat}$:*

$$T_{K^{\flat}}(x) \subset \operatorname{Liminf}_{K_n \ni x_n \to x} T_{K_n}(x_n)$$

and

$$T_{K^{\sharp}}(x) \subset \operatorname{Limsup}_{K_n \ni x_n \to x} T_{K_n}(x_n)$$

Assume now that X is a Hilbert space and that the lower limit is equal to the sequentially weak upper limit $\sigma - \operatorname{Limsup}_{n\to\infty} K_n$ and denote by K the common value. Then

$$\operatorname{Liminf}_{n\to\infty, K_n \ni x_n \to x} T_{K_n}(x_n) = T_K(x)$$

Proof

— 1. Let us take $x \in K^{\flat}$ and $v \in S_{K^{\flat}}(x)$, the cone spanned by $K^{\flat} - x$. Hence there exists $h > 0$ such that $x + hv \in K^{\flat}$.

Consider any sequence of elements x_n belonging to K_n and converging to x, and let y_n denote the projection of $x+hv$ onto K_n. Since y_n converges to $x + hv$, the sequence of elements $v_n := (y_n - x_n)/h$ converges to v. Furthermore, for $h_n \in]0, h[$

$$x_n + h_n v_n = \left(1 - \frac{h_n}{h}\right) x_n + \frac{h_n}{h} y_n \in K_n$$

and we infer that v_n belongs to the contingent cone $T_{K_n}(x_n)$.

Hence v, the limit of the v_n's, belongs to the lower limit of the tangent cones $T_{K_n}(x_n)$.

— 2. Let x belong to $K^\sharp$ and $v := (y-x)/h$ be chosen arbitrarily in the cone spanned by $K^\sharp - x$. Denoting by x_n and y_n the projections onto K_n of x and $x+hv$ respectively, we infer that $v_n := (y_n - x_n)/h$ belongs to the tangent cone $T_{K_n}(x_n)$ and converges to v.

— 3. We deduce from Corollary 7.6.5 below that

$$N_K(x) \;=\; \sigma - \mathrm{Limsup}_{n\to\infty, K_n \ni x_n \to x} N_{K_n}(x_n)$$

By transposition, Theorem 1.1.8 implies that

$$\mathrm{Liminf}_{n\to\infty, K_n\ni x_n \to x} T_{K_n}(x_n) \;=\; T_K(x) \quad \square$$

We shall derive from the Stability Theorem a formula for the circatangent cone of an inverse image.

Theorem 4.6.7 *Let X and Y be two Banach spaces and $A \in \mathcal{L}(X,Y)$ be a continuous linear operator. Consider sequences of closed subsets $L_n \subset X$ and $M_n \subset Y$. Let us assume that there exist constants $c > 0$, $\alpha \in [0,1[$ and $\eta > 0$ such that*

$$\begin{cases} \forall\, x_n \;\in\; K_n \cap B(x_0,\eta),\;\; y_n \;\in\; M_n \cap B(Ax_0,\eta) \\ \\ B_Y \;\subset\; A\left(T^\flat_{L_n}(x_n) \cap cB_X\right) - T_{M_n}(y_n) + \alpha B_Y \end{cases}$$

Let x belong to the lower limit of the sequence $L_n \cap A^{-1}(M_n)$. Then

$$C^\flat_{(L_n)}(x) \cap A^{-1} C^\flat_{(M_n)}(Ax) \;\subset\; C^\flat_{L^\flat \cap A^{-1}(M^\flat)}(x)$$

Proof — Take any sequence of elements $x_n \in L_n \cap A^{-1}(M_n)$ which converges to x and any $u \in C^\flat_{(L_n)}(x)$ such that $Au \in C^\flat_{(M_n)}(Ax)$. Hence for any sequence $h_n > 0$, there exist sequences u_n and v_n converging to u and Au respectively such that, for all $n \geq 0$,

$$x_n + h_n u_n \in L_n \;\;\&\;\; Ax_n + h_n v_n \in M_n$$

We now apply Theorem 3.4.5 to the subsets $L_n \times M_n$ of $X \times Y$ and the continuous linear operators $A \ominus \mathbf{1}$ associating to any (x,y) the element $Ax - y$.

The stability assumptions of this Theorem are obviously satisfied. The pair $(x_n + h_n u_n, Ax_n + h_n v_n)$ belongs to $L_n \times M_n$ and

$$(A \ominus \mathbf{1})(x_n + h_n u_n, Ax_n + h_n v_n) \;\text{ converges to }\; 0$$

Therefore, by Theorem 3.4.5, there exist $l > 0$ and $(\widehat{x}_n, \widehat{y}_n) \in L_n \times M_n$ such that, for n large enough, $(A \ominus \mathbf{1})(\widehat{x}_n, \widehat{y}_n) = 0$ and

$$\|x_n + h_n u_n - \widehat{x}_n\| + \|Ax_n + h_n v_n - \widehat{y}_n\| \leq l h_n \|Au_n - v_n - 0\|$$

Hence $\widehat{u}_n := (\widehat{x}_n - x_n)/h_n$ converges to u, and we observe that for all $n \geq 0$, $x_n + h_n u_n$ belongs to $L_n \cap A^{-1}(M_n)$. $\square$

4.7 Higher Order Tangent Sets

The contingent cones are upper limits of the sets $(K - x)/h$ when $h > 0$ converges to 0.

We may need more accurate approximations, and consider the sets of cluster points of m^{th}-order differential quotients

$$\frac{K - x - hv_1 - \cdots - h^{m-1} v_{m-1}}{h^m}$$

when $h > 0$ converges to 0.

Definition 4.7.1 *Let x belong to a subset K of a normed space X and $v_1, \ldots, v_{m-1}$ be elements of X.*

We say that the subset

$$T_K^{(m)}(x, v_1, \ldots v_{m-1}) \ := \ \text{Limsup}_{h \to 0+} \frac{K - x - hv_1 - \cdots - h^{m-1} v_{m-1}}{h^m}$$

is the m^{th}-order contingent subset *of K at $(x, v_1, \ldots, v_{m-1})$.*

We see at once that $T_K^{(1)}(x) = T_K(x)$ is the contingent cone to K at x and that if $T^{(m)}(x, v_1, \ldots, v_{m-1})$ is not empty, then necessarily,

$$v_1 \in T_K^{(1)}(x), \ v_2 \in T_K^{(2)}(x, v_1), \ \ldots \ , v_{m-1} \in T_K^{(m-1)}(x, v_1, \ldots, v_{m-2})$$

Naturally, we may need more regularity and ask for stronger limits of the subsets $(K - x - hv_1 - \cdots - h^{m-1} v_{m-1})/h^m$, as in the case of first order approximations.

Definition 4.7.2 *Let x belong to a subset K of a normed space X and $v_1, \ldots, v_{m-1}$ be elements of X. We say that the subset*

$$T_K^{\flat(m)}(x, v_1, \ldots v_{m-1}) \ := \ \text{Liminf}_{h \to 0+} \frac{K - x - hv_1 - \cdots - h^{m-1} v_{m-1}}{h^m}$$

Table 4.6: **Properties of m^{th}-order Contingent Sets.**

(1)	▷	If $K \subset L$ and $x \in \overline{K}$, then $T_K^{(m)}(x, v_1, \dots, v_{m-1}) \subset T_L^{(m)}(x, v_1, \dots, v_{m-1})$
(2)	▷	If $K_i \subset X$, $(i = 1, \dots, n)$ and $x \in \overline{\bigcup_i K_i}$, then $T^{(m)}_{\cup_{i=1}^n K_i}(x, v_1, \dots, v_{m-1}) = \cup_{i \in I(x)} T^{(m)}_{K_i}(x, v_1, \dots, v_{m-1})$, where $I(x) := \{i \mid x \in \overline{K_i}\}$
(3)	▷	If $K_i \subset X_i$, $(i = 1, \dots, n)$ and $x_i \in \overline{K_i}$, then $T^{(m)}_{\prod_{i=1}^n K_i}(x_1, \dots, x_n, v_{1,1}, \dots, v_{1,n}, \dots, v_{m-1,1}, \dots, v_{m-1,n}) \subset \prod_{i=1}^n T^{(m)}_{K_i}(x_i, v_{1,i}, \dots, v_{m-1,i})$
(5)	▷	If $K_i \subset X$, $(i = 1, \dots, n)$ and $x \in \bigcap_i \overline{K_i}$, then $T^{(m)}_{\bigcap_{i=1}^n K_i}(x, v_1, \dots, v_{m-1}) \subset \bigcap_{i=1}^n T^{(m)}_{K_i}(x, v_1, \dots, v_{m-1})$

is the m^{th}-order adjacent subset *of* K *at* $(x, v_1, \dots, v_{m-1})$ *and that*

$$C_K^{(m)}(x, v_1, \dots v_{m-1}) := \operatorname{Liminf}_{h \to 0+, y \to_K x} \frac{K - y - hv_1 - \dots - h^{m-1} v_{m-1}}{h^m}$$

is the m^{th}-order circatangent subset *of* K *at* $(x, v_1, \dots, v_{m-1})$.

We shall say that K *is* m^{th}-derivable at $x, v_1, \dots, v_{m-1}$ *if the* m^{th}*-order contingent and adjacent sets coincide.*

They are closed subsets satisfying, for any $\lambda > 0$,

$$\begin{cases} T_K^{(m)}(x, \lambda v_1, \lambda^2 v_2, \dots, \lambda^{m-1} v_{m-1}) = \lambda^m T_K^{(m)}(x, v_1, v_2, \dots, v_{m-1}) \\ \\ T_K^{\flat(m)}(x, \lambda v_1, \lambda^2 v_2, \dots, \lambda^{m-1} v_{m-1}) = \lambda^m T_K^{\flat(m)}(x, v_1, v_2, \dots, v_{m-1}) \end{cases} \tag{4.14}$$

As for the m^{th}-order contingent sets, the adjacent set

$$T_K^{\flat(m)}(x, v_1, \dots, v_{m-1})$$

(respectively the circatangent set $C_K^{(m)}(x, v_1, \dots, v_{m-1})$) is empty if one of the v_j's does not belong to $T_K^{\flat(j)}(x, v_1, \dots, v_{j-1})$ (resp. $C_K^{(j)}(x, v_1, \dots, v_{j-1})$) for $j = 1, \dots, m-1$.)

We observe that

$$\begin{array}{l} C_K^{(m)}(x, u_1, \ldots, u_{m-1}) \;+\; T_K^{(m)}(x, v_1, \ldots, v_{m-1}) \\ \quad \subset\; T_K^{(m)}(x, u_1 + v_1, \ldots, u_{m-1} + v_{m-1}) \end{array}$$

and

$$\begin{array}{l} C_K^{(m)}(x, u_1, \ldots, u_{m-1}) \;+\; C_K^{(m)}(x, v_1, \ldots, v_{m-1}) \\ \quad \subset\; C_K^{(m)}(x, u_1 + v_1, \ldots, u_{m-1} + v_{m-1}) \end{array}$$

So that, taking into account (4.14), we obtain the following property of m^{th}-order circatangent sets:

$$\forall\, \lambda,\ \mu > 0, \quad \frac{\lambda^m C_K^{(m)}(x, u_1, \ldots, u_{m-1}) + \mu^m C_K^{(m)}(x, v_1, \ldots, v_{m-1})}{(\lambda + \mu)^m}$$

$$\subset\; C_K^{(m)}\left(x, \frac{\lambda u_1 + \mu v_1}{\lambda + \mu}, \ldots, \frac{\lambda^{m-1} u_{m-1} + \mu^{m-1} v_{m-1}}{(\lambda + \mu)^{m-1}}\right)$$

We also observe that

$$\left\{ \begin{array}{l} C_K^{(m)}(x, v_1, \ldots v_{m-1}) \\ := \;\mathrm{Liminf}_{h \to 0+, y \to_K x, v_1' \to v_1, \ldots, v_{m-1}' \to v_{m-1}} \frac{K - y - h v_1' - \cdots - h^{m-1} v_{m-1}'}{h^m} \end{array} \right.$$

so that if $C_K^{(m-1)}(x, v_1, \ldots, v_{j-1}) \neq \emptyset$, then $0 \in C_K^{(m)}(x, v_1, \ldots, v_{m-1})$.

Example — Consider for instance the m^{th}-order contingent set $T_{\mathbf{R}_+^n}^{(m)}(x)$. We obtain that for every $x \in \mathrm{Int}\,\mathbf{R}_+^n$ and all $v_i \in \mathbf{R}^n$

$$T_{\mathbf{R}_+^n}^{(m)}(x, v_1, \ldots, v_{m-1}) \;=\; \mathbf{R}^n$$

For every x from the boundary of $\mathbf{R}_+^n$ we have

$$v_m \;\in\; T_{\mathbf{R}_+^n}^{(m)}(x, v_1, \ldots, v_{m-1}) \;=\; T_{\mathbf{R}_+^n}^{\flat(m)}(x, v_1, \ldots, v_{m-1})$$

if and only if

$$\left\{ \begin{array}{l} v_{m_i} \geq 0 \;\text{ whenever }\; x_i \;=\; v_{1_i} \;=\; \cdots \;=\; v_{(m-1)_i} = 0 \\ \\ \text{and} \\ \\ v_{m_i} \in \mathbf{R} \text{ if } x_i = v_{1_i} = \cdots = v_{(k-1)_i} = 0,\ v_{k_i} > 0 \text{ for } 0 \leq k \leq m-1 \;\square \end{array} \right.$$

Table 4.7: **Properties of m^{th}-order Adjacent Sets.**

(1)	$\triangleright$	If $K \subset L$ and $x \in \overline{K}$, then $T_K^{\flat(m)}(x, v_1, \ldots, v_{m-1}) \subset T_L^{\flat(m)}(x, v_1, \ldots, v_{m-1})$
(2)	$\triangleright$	If $K_i \subset X$, $(i = 1, \ldots, n)$ and $x \in \overline{\bigcup_i K_i}$, then $T_{\cup_{i=1}^n K_i}^{\flat(m)}(x, v_1, \ldots, v_{m-1}) \supset \cup_{i \in I(x)} T_{K_i}^{\flat(m)}(x, v_1, \ldots, v_{m-1})$, where $I(x) := \{i \mid x \in \overline{K_i}\}$
(3)	$\triangleright$	If $K_i \subset X_i$, $(i = 1, \ldots, n)$ and $x_i \in \overline{K_i}$, then $T_{\prod_{i=1}^n K_i}^{\flat(m)}(x_1, \ldots, x_n, v_{1,1}, \ldots, v_{1,n}, \ldots, v_{m-1,1}, \ldots, v_{m-1,n})$ $= \prod_{i=1}^n T_{K_i}^{\flat(m)}(x_i, v_{1,i}, \ldots, v_{m-1,i})$
(5)	$\triangleright$	If $K_i \subset X$, $(i = 1, \ldots, n)$ and $x \in \bigcap_i \overline{K_i}$, then $T_{\bigcap_{i=1}^n K_i}^{\flat(m)}(x, v_1, \ldots, v_{m-1}) \subset \bigcap_{i=1}^n T_{K_i}^{\flat(m)}(x, v_1, \ldots, v_{m-1})$

The formulas enjoyed by the contingent and adjacent cones can be easily extended to the m^{th}-order contingent and adjacent sets. They are presented in Table 4.6 and Table 4.7. However, the formulas for the direct and inverse images involve higher order derivatives and are presented in the case of second order sets for simplicity.

Proposition 4.7.3 *Let X, Y be normed spaces and let $g \in \mathcal{C}^2(X,Y)$, $K \subset X$, $x \in \overline{K}$ and $M \subset Y$ be given. Then, for any $v_1 \in T_K(x)$*

$$\forall\, v_2 \in T_K^{(2)}(x, v_1),\;\; g'(x)T_K^{(2)}(x, v_1) \subset T_{g(K)}^{(2)}(g(x), g'(x)v_1) - \frac{1}{2}g''(x)(v_1, v_1)$$

and thus

$$T_{g^{-1}(M)}^{(2)}(x, v_1) \subset g'(x)^{-1}\left(T_M^{(2)}(g(x), g'(x)v_1) - \frac{1}{2}g''(x)(v_1, v_1)\right)$$

The same results hold true for the second order adjacent sets.

Remark — When $g = A \in \mathcal{L}(X,Y)$ is a continuous linear operator, the higher order derivatives vanish and we obtain the formulas

$$\begin{cases} A\left(T_K^{(m)}(x, v_1, \ldots, v_{m-1})\right) \subset T_{A(K)}^{(m)}(Ax, Av_1, \ldots, Av_{m-1}) \\ A\left(T_K^{\flat(m)}(x, v_1, \ldots, v_{m-1})\right) \subset T_{A(K)}^{\flat(m)}(Ax, Av_1, \ldots, Av_{m-1}) \end{cases} \square$$

In order to obtain the converse inclusions, we need to assume a transversality assumption.

Theorem 4.7.4 *Let X and Y be Banach spaces, $L \subset X$ and $M \subset Y$ be* closed *subsets, $f : X \mapsto Y$ be twice continuously differentiable around an element $x_0 \in L \cap f^{-1}(M)$.*

If Y is a finite dimensional vector-space, we posit the pointwise transversality condition

$$f'(x_0)(C_L(x_0)) \; - \; C_M(f(x_0)) \; = \; Y$$

Then

$$\left\{ \begin{array}{l} T_L^{\flat(2)}(x_0, u_1) \cap f'(x_0)^{-1} \left(T_M^{(2)}(f(x_0), f'(x_0)u_1) - \frac{1}{2} f''(x_0)(u_1, u_1)\right) \\ \subset \; T^{(2)}_{L \cap f^{-1}(M)}(x_0, u_1) \end{array} \right.$$

and, when M is twice derivable, we have the equality instead of the inclusion. Furthermore

$$\left\{ \begin{array}{l} T_L^{\flat(2)}(x_0, u_1) \cap f'(x_0)^{-1} \left(T_M^{\flat(2)}(f(x_0), f'(x_0)u_1) - \frac{1}{2} f''(x_0)(u_1, u_1)\right) \\ = \; T^{\flat(2)}_{L \cap f^{-1}(M)}(x_0, u_1) \end{array} \right.$$

Otherwise, to obtain the same results in the case of any Banach space, we have to replace the pointwise transversality condition by the local transversality condition*:*

there exist constants $c > 0$, $\alpha \in [0, 1[$ and $\eta > 0$ such that

$$\left\{ \begin{array}{l} \forall \, x \in L \cap B(x_0, \eta), \;\; \forall \, y \in M \cap B(f(x_0), \eta) \\ \\ B_Y \; \subset \; f'(x) \left(T_L^{\flat}(x) \cap cB_X\right) - T_M(y) + \alpha B_Y \end{array} \right.$$

Proof — Let us prove for instance the equality for the second order adjacent sets. Consider the closed subset $K := L \cap f^{-1}(M)$. We already know from Theorem 4.3.3 that every $u_1 \in T_L^{\flat}(x_0) \cap f'(x_0)^{-1}\left(T_M^{\flat}(f(x_0))\right)$ belongs to $T^{\flat}_{L \cap f^{-1}(M)}(x_0)$. Let us choose such an element u_1 and pick any $u_2 \in T_L^{\flat(2)}(x_0, u_1)$ such that

$$f'(x_0)u_2 \in T_M^{\flat(2)}(f(x_0), f'(x_0)u_1) - \frac{1}{2} f''(x_0)(u_1, u_1)$$

Hence for any sequence $h_n > 0$ converging to 0, there exist sequences u_{2n} and v_{2n} converging to u_2 and $f'(x_0)u_2 + \frac{1}{2} f''(x_0)(u_1, u_1)$ respectively such that, for all $n \geq 0$,

$$x_0 + h_n u_1 + h_n^2 u_{2n} \; \in \; L \;\; \& \;\; f(x_0) + h_n f'(x_0) u_1 + h_n^2 v_{2n} \; \in \; M$$

We apply now Theorem 4.3.1 to the subset $L \times M$ of $X \times Y$ and the continuous map $f \ominus \mathbf{1}$ associating to any (x, y) the element $f(x) - y$. It is obvious that the transversality condition

$$f'(x_0)(C_L(x_0)) - C_M(f(x_0)) \;=\; Y$$

implies the assumption of Theorem 4.3.1. On the other hand

$$\frac{1}{h_n^2}(f \ominus \mathbf{1})(x_0 + h_n u_1 + h_n^2 u_{2n}, f(x_0) + h_n f'(x_0)u_1 + h_n^2 v_{2n}) \text{ converges to } 0$$

because f is continuously differentiable at x_0.

Therefore, by Theorem 4.3.1, there exist $l > 0$ and a solution $(\widehat{x}_n, \widehat{y}_n) \in L \times M$ to the equation

$$(f \ominus \mathbf{1})(\widehat{x}_n, \widehat{y}_n) \;=\; 0$$

(i.e., $\widehat{y}_n = f(\widehat{x}_n)$) such that

$$\begin{cases} \|x_0 + h_n u_1 + h_n^2 u_{2n} - \widehat{x}_n\| + \|f(x_0) + h_n f'(x_0)u_1 + h_n^2 v_{2n} - \widehat{y}_n\| \\ \leq \; l h_n^2 \;\|(f(x_0 + h_n u_1 + h_n^2 u_{2n}) - f(x_0) - h_n f'(x_0)u_1)/h_n^2 - v_{2n}\| \end{cases}$$

Hence $\widehat{u}_{2n} := (\widehat{x}_n - x_0 - h_n u_1)/h_n^2$ converges to u_2, and for all $n \geq 0$, we know that $x_0 + h_n u_1 + h_n^2 \widehat{u}_{2n}$ belongs to $L \cap f^{-1}(M)$ because

$$x_0 + h_n u_1 + h_n^2 \widehat{u}_n \;=\; \widehat{x}_n$$

and $f(x_0 + h_n u_1 + h_n^2 \widehat{u}_n) = \widehat{y}_n$.

To prove the converse inclusion, fix $u_2 \in T^{\flat(2)}_{L \cap f^{-1}(M)}(x_0, u_1)$. Then for every sequence $h_n \to 0+$, there exists a sequence $u_{2n} \to u_2$ such that

$$x_0 + h_n u_1 + h_n^2 u_{2n} \;\in\; L, \quad f(x_0 + h_n u_1 + h_n^2 u_{2n}) \;\in\; M$$

Thus $u_2 \in T^{\flat(2)}_{L\cap}(x_0, u_1)$ and since

$$\begin{cases} f(x_0 + h_n u_1 + h_n^2 u_{2n}) \\ = \; f(x_0) + h_n f'(x_0)u_1 + \frac{1}{2}h_n^2 f''(x_0)(u_1, u_1) + h_n^2 f'(x_0)u_{2n} + \varepsilon(h_n)h_n^2 \end{cases}$$

where $\varepsilon(h_n)$ converges to zero with h_n, we deduce that

$$f'(x_0)u_2 + \frac{1}{2}f''(x_0)(u_1, u_1) \;\in\; T^{\flat(2)}_M(f(x_0), f'(x_0)u_1) \quad \square$$

Example: Second order contingent sets to subsets defined by equality and inequality constraints

Consider a closed subset L of a Banach space X and two twice continuously differentiable maps

$$g := (g_1, \ldots, g_p) : X \mapsto \mathbf{R}^p \ \ \& \ \ h := (h_1, \ldots, h_q) : X \mapsto \mathbf{R}^q$$

defined on an open neighborhood of L.

Let K be the subset of L defined by the constraints

$$K \ := \ \{x \in L \mid g_i(x) \geq 0, \ i = 1, \ldots, p \ \ \& \ h_j(x) = 0, \ j = 1, \ldots, q\}$$

and let

$$I(x) \ := \ \{i = 1, \ldots, p \mid g_i(x) = 0\}$$

denote the subset of active constraints.

We posit the following transversality condition at a given $x \in K$:

$$\begin{cases} i) & h'(x)C_L(x) \ = \ \mathbf{R}^q \\ & \\ ii) & \exists\, v_0 \in C_L(x) \ \text{ such that } \ h'(x)v_0 = 0 \\ & \& \ \forall\, i \in I(x), \ \ < g_i'(x), v_0 > \ > \ 0 \end{cases}$$

By Proposition 4.3.7, u_1 belongs to the contingent cone to K at x if and only if u_1 belongs to the contingent cone to L at x and satisfies the constraints

$$\forall\, i \in I(x), \ \ < g_i'(x), u_1 > \ \geq \ 0 \ \ \& \ \ \forall\, j = 1, \ldots, q, \ \ h_j'(x)u_1 = 0$$

Proposition 4.7.5 *Let us set*

$$I^{(2)}(x, u_1) \ := \ \{i \in I(x) \mid \ \ < g_i'(x), u_1 > \ = \ 0\}$$

Then u_2 belongs to the second order contingent set to K at (x, u_1) if and only if u_2 belongs to the second order contingent set to L at (x, u_1) and satisfies the constraints

$$\begin{cases} \forall\, i \in I^{(2)}(x, u_1), \ \ < g_i'(x), u_2 > + \frac{1}{2} g_i''(x)(u_1, u_1) \ \geq \ 0 \\ \\ \text{and } \ \forall\, j = 1, \ldots, q, \ \ h_j'(x)u_2 + \frac{1}{2} h_j''(x)(u_1, u_1) = 0 \end{cases}$$

Proof — Define

$$Y \ = \ \mathbf{R}^p \times \mathbf{R}^q, \ \ f \ = \ (g, h), \ \ M \ = \ \mathbf{R}^p_+ \times \mathbf{R}^q$$

The proof follows by exactly the same arguments as the ones of Proposition 4.3.7, after observing that the previous example yields that M is twice derivable and using Theorem 4.7.4. □

Chapter 5

Derivatives of Set-Valued Maps

Introduction

We need to differentiate set-valued maps as much as we need to differentiate single-valued maps, for extending the Inverse Function Theorem to set-valued maps for instance, and for many other reasons.

How should we go about it? The idea is very simple, and goes back to the prehistory of the differential calculus, when Pierre de Fermat introduced in the first half of the seventeenth century the concept of a tangent to the graph of a function:

the tangent space to the graph of a function f at a point (x, y) of its graph is the line of slope $f'(x)$, i.e., the graph of the linear function $u \mapsto f'(x)u$.

It is possible to implement this idea for any set-valued map F since we have introduced concepts of tangency for any subset of a normed space. For instance, *we can regard the contingent cone to the graph of the set-valued map F at some point (x, y) of its graph as the graph of the associated "contingent derivative" of F at this point (x, y).*

Since the contingent cone is at least... a cone, the contingent derivative is at least *positively homogeneous (or a process).* This is what remains of the familiar, but luxurious, requirement of linearity.

However, when the contingent cone happens to be convex (this is the case when the graph is sleek), *the contingent derivative is a closed convex process,* i.e., a set-valued analogue of a continuous linear operator.

The graphical derivatives associated to the Clarke tangent cone, which we call here the *circatangent derivatives*, are also closed convex processes. They are presented in the second section together with the *adjacent derivatives*, the graphs of which are the adjacent (or intermediate) tangent cones. Hence the *graphical derivatives* which are attached to each such tangent cone correspond to different regularity requirements.

It is also possible to define them as adequate limits of differential quotients. The limits, involving "limsupinf's", that would not have been proposed by any sensible or even sober person in a right frame of mind had they not been derived from the "graphical approach." They are still pointwise limits, which puts them in a different class than the *distributional derivatives*[1] introduced by Laurent Schwartz in the early fifties. We shall see in Chapter 7 that they are also adequate "graphical limits" of differential quotients.

Chain rules are provided in the third section.

With this differential calculus at hand, we are able to approximate solutions to an inclusion

$$F(x) \ni y$$

by solutions to approximate inclusions

$$F_n(x_n) \ni y_n$$

by adapting Lax's principle:

Convergence of the data y_n to y, *consistency* of F_n to F and *stability* of the F_n's imply the convergence of some solutions x_n's to the solution x.

Consistency means that (x, y) belongs to the lower limit of the graphs of the maps F_n. Stability means here that the norms of the

[1] which were designed to make sense out of Dirac's operational calculus and solve partial differential equations.

In order to keep the linearity of the differential operators, the limit of differential quotients is taken in spaces of distributions endowed with topologies much weaker than the pointwise convergence topology, used ever since Cauchy's rigorous definition.

inverses $DF_n(x_n, y_n)^{-1}$ are bounded when the pairs (x_n, y_n) range over a neighborhood of (x, y). We shall deduce that, if the right-hand side y_n converges to y, there exist approximate solutions $x_n \in F^{-1}(y_n)$ which converge to x, and we shall provide an error estimate.

When the maps $F_n = F$ are constant, this principle boils down to the statement of the Inverse-Function Theorem.

We consider also an application to *Qualitative Analysis*, a field started in mathematical economics under the name of "comparative statics" and taken over by computer scientists as a domain of Artificial Intelligence. For instance, if the vector-spaces X and Y are finite dimensional, we are just interested in the possibility of solving an equation or an inclusion $F(x) \ni y$ *for prescribed sign vectors*. We provide a dual criterion allowing us to answer this question.

Finally, by using high-order contingent sets to the graph of a set-valued map F, we define in Section 6 high-order derivatives of a set-valued map.

We derive in Chapter 10 *variational inclusions* by *linearizing* a differential inclusion, the solution map of which contains the contingent derivative of the solution map of the original differential inclusion.

5.1 Contingent Derivatives

By coming back to the original point of view proposed by Fermat, we are able to define geometrically derivatives of set-valued maps from the choice of tangent cones to the graphs, even though they yield *very strange* limits of differential quotients.

Definition 5.1.1 *Let X, Y be normed spaces and $F : X \rightsquigarrow Y$ be a set-valued map.*

The contingent derivative *$DF(x, y)$ of F at $(x, y) \in$ Graph(F) is the set-valued map from X to Y defined by*

$$\mathrm{Graph}(DF(x,y)) \;:=\; T_{\mathrm{Graph}(F)}(x,y)$$

When $F := f$ is single-valued, we set $Df(x) := Df(x, f(x))$.

We shall say that F is sleek *at $(x, y) \in \mathrm{Graph}(F)$ if the map*

$$\mathrm{Graph}(F) \ni (x', y') \rightsquigarrow \mathrm{Graph}(DF(x', y'))$$

is lower semicontinuous at (x, y) (i.e., if the graph of F is sleek at (x, y)).

It is said to be derivable at (x, y) *if* $\mathrm{Graph}(F)$ *is derivable at that point.*

The set-valued map F is sleek (*respectively* derivable) *if it is sleek (respectively derivable) at every point of its graph.*

Finally, we shall say that F is pseudo-convex *at $(x, y) \in \mathrm{Graph}(F)$ if and only if*

$$\forall\, x' \in \mathrm{Dom}(F), \quad F(x') \subset DF(x, y)(x' - x) + y$$

(i.e., if its graph is pseudo-convex at (x, y)).

Naturally, *the contingent derivative is a closed convex process whenever F is sleek at (x, y).*

We can easily compute the derivative of the inverse of a set-valued map F (or even of a non injective single-valued map): *The contingent derivative of the inverse of a set-valued map F is the inverse of the contingent derivative*:

$$D(F^{-1})(y, x) = DF(x, y)^{-1}$$

The restriction $F := f|_K$ of a single-valued map f to a subset $K \subset X$ provides an example of a set-valued map defined by

$$f|_K(x) := \begin{cases} f(x) & \text{if } x \in K \\ \emptyset & \text{if } x \notin K \end{cases}$$

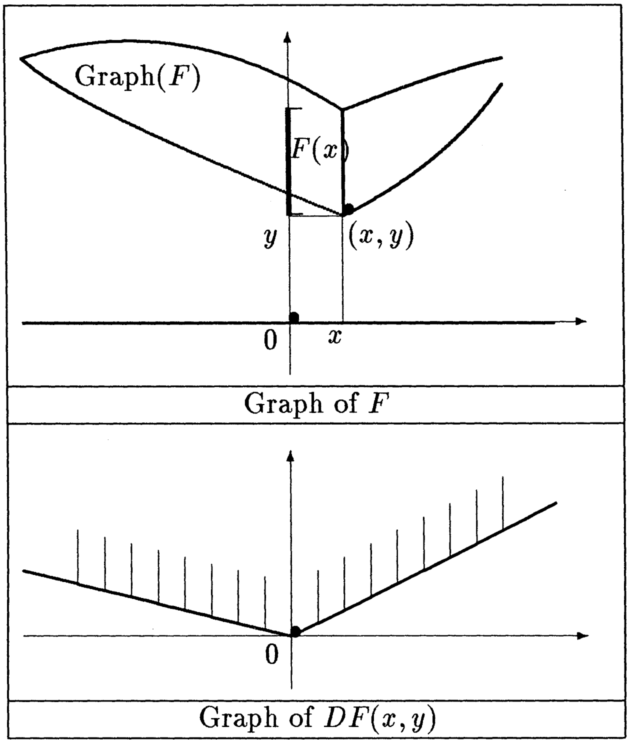
for which we obtain the following formula:

If f is Fréchet differentiable around a point $x \in K$, then *the contingent derivative of the restriction of f to K is the restriction of the derivative to the contingent cone*:

$$D(f|_K)(x) = D(f|_K)(x, f(x)) = f'(x)|_{T_K(x)}$$

Actually, this follows from the useful

Figure 5.1: Graph of the Contingent Derivative



Proposition 5.1.2 *Let X and Y be normed spaces, f be a single-valued map from an open subset $\Omega \subset X$ to Y, $M : X \rightsquigarrow Y$ be a set-valued map. Define the set-valued map $F : X \rightsquigarrow Y$ by:*

$$\forall\, x \in X, \quad F(x) := f(x) - M(x)$$

If f is Fréchet differentiable at $x \in \Omega \cap \mathrm{Dom}(M)$, then for every $y \in F(x)$

$$DF(x,y)(u) = f'(x)u - DM(x, f(x) - y)(u)$$

Proof — Let v belong to $DF(x,y)(u)$. Then there exist $h_n > 0$ converging to 0 and sequences u_n and v_n converging to u and v respectively such that for every n

$$y + h_n v_n \in f(x + h_n u_n) - M(x + h_n u_n)$$

Since

$$f(x + h_n u_n) = f(x) + h_n(f'(x)u + \varepsilon(h_n))$$

where $\varepsilon(h_n)$ converges to 0 with h_n we get

$$f(x) - y + h_n(f'(x)u - v_n + \varepsilon(h_n)) \in M(x + h_n u_n)$$

and thus $f'(x)u - v \in DM(x, f(x) - y)(u)$.

Conversely, assume that $f'(x)u - v$ belongs to $DM(x, f(x) - y)(u)$. Hence, there exist a sequence $h_n > 0$ converging to 0 and sequences u_n and w_n converging to u and $f'(x)u - v$ such that

$$f(x) - y + h_n w_n \in M(x + h_n u_n)$$

Then for some $\varepsilon(h_n) \to 0$, the sequence $v_n := f'(x)u + \varepsilon(h_n) - w_n$ converges to v and satisfies

$$y + h_n v_n \in f(x + h_n u_n) - M(x + h_n u_n), \quad \square$$

When we add "state constraints" L, this result becomes:

Proposition 5.1.3 *Let X, Y be normed spaces, $f : \Omega \mapsto Y$ be a single-valued map, where $\Omega \subset X$ is open, $M : X \rightsquigarrow Y$ be a set-valued map and $L \subset X$. Let $F : X \rightsquigarrow Y$ be the set-valued map defined by:*

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

If f is Fréchet differentiable at $x \in \Omega \cap \mathrm{Dom}(F)$, then for every $y \in F(x)$,

$$DF(x,y)(u) \subset \begin{cases} f'(x)u - DM(x, f(x)-y)(u) & \text{when} \quad u \in T_L(x) \\ \emptyset & \text{when} \quad u \notin T_L(x) \end{cases}$$

Equality holds true when we assume that either L or M is derivable at x and M is Lipschitz at x.

In particular if M is constant, then

$$\forall\, u \in T_L(x), \quad DF(x,y)(u) = f'(x)u - T_M(f(x)-y)$$

Remark — The above two Propositions remain true if f is locally Lipschitz and Gâteaux differentiable. □

Proof — Let v belong to $DF(x,y)(u)$. Then there exist $h_n > 0$ converging to 0 and sequences u_n and v_n converging to u and v respectively such that for all n

$$x + h_n u_n \in L \quad \& \quad y + h_n v_n \in f(x+h_n u_n) - M(x+h_n u_n)$$

This implies that u belongs to $T_L(x)$. Since we can write

$$f(x+h_n u_n) = f(x) + h_n(f'(x)u + \varepsilon(h_n))$$

where $\varepsilon(h_n)$ converges to 0 with h_n, we get

$$f(x) - y + h_n(f'(x)u - v_n + \varepsilon(h_n)) \in M(x+h_n u_n)$$

and thus $f'(x)u - v$ does belong to $DM(x, f(x)-y)(u)$.

Conversely, assume for instance that M is derivable, that u belongs to $T_L(x)$ and that $f'(x)u - v$ belongs to $DM(x, f(x)-y)(u)$. Hence, there exist a sequence $h_n > 0$ converging to 0 and sequences u_n, $\overline{u}_n$ and w_n converging to u, u and $f'(x)u - v$ respectively, such that $x + h_n u_n$ belongs to L and $f(x) - y + h_n w_n$ to $M(x + h_n \overline{u}_n)$. Since M is Lipschitz at x there exists a sequence $\overline{w}_n$ converging to $f'(x)u - v$ satisfying

$$\forall\, n \geq 1, \quad f(x) - y + h_n \overline{w}_n \in M(x+h_n u_n)$$

Then for some $\varepsilon(h_n) \to 0$, the sequence $v_n := f'(x)u + \varepsilon(h_n) - \overline{w}_n$ converges to v and satisfies

$$y + h_n v_n \in f(x + h_n u_n) - M(x + h_n u_n) \quad \square$$

Our first task is to characterize these contingent derivatives by adequate limits of differential quotients. We derive that the contingent derivative in the direction u is the upper limit of the differential quotients

$$DF(x,y)(u) = \mathrm{Limsup}_{h\to 0+, u'\to u} \frac{F(x+hu') - y}{h}$$

from

Proposition 5.1.4 *Let X, Y be normed spaces, $F : X \rightsquigarrow Y$ be a set-valued map and let $(x,y) \in \mathrm{Graph}(F)$. Then*

$$v \in DF(x,y)(u) \iff \mathrm{liminf}_{h\to 0+, u'\to u} d\left(v, \frac{F(x+hu') - y}{h}\right) = 0$$

If $x \in \mathrm{Int}(\mathrm{Dom}(F))$ and F is Lipschitz around x, then

$$v \in DF(x,y)(u) \iff \liminf_{h\to 0+} d\left(v, \frac{F(x+hu) - y}{h}\right) = 0$$

If moreover the dimension of Y is finite and l denotes the Lipschitz constant of F at x, then

$$\mathrm{Dom}(DF(x,y)) = X \text{ and } DF(x,y) \text{ is } l - \text{Lipschitz}$$

Proof — The first two statements being obvious, let us check the last one. Let u belong to X and l denote the Lipschitz constant of F on a neighborhood of x. Then, for all $h > 0$ small enough and $y \in F(x)$,

$$y \in F(x) \subset F(x+hu) + lh\|u\|B$$

Hence there exists $y_h \in F(x+hu)$ such that $v_h := (y_h - y)/h$ belongs to $l\|u\|B$, which is compact. Therefore the sequence v_h has a cluster point v, which belongs to $DF(x,y)(u)$.

Fix any u_1, $u_2 \in X$ and $v_1 \in DF(x,y)(u_1)$. To end the proof we have to show that there exists $v_2 \in DF(x,y)(u_2)$ such that

$\|v_1 - v_2\| \le l\,\|u_1 - u_2\|$. Consider $h_n \to 0+$ and $v_{1n} \to v_1$ satisfying

$$y + h_n v_{1n} \in F(x + h_n u_1)$$

and let $z_n \in F(x + h_n u_2)$ be such that for all large n,

$$\|z_n - y - h_n v_{1n}\| \le l h_n \|u_1 - u_2\|$$

(they exist by the Lipschitz continuity of F at x.) Taking a subsequence and keeping the same notations, we may assume that the bounded sequence $(z_n - y)/h_n$ converge to some $v_2 \in DF(x,y)(u_2)$. Clearly v_2 satisfies the required estimates. □

Remark — Domain of the Derivative It is quite useful to relate the tangent cones to the domain of a set-valued map to the domain of its derivative.

We always have

$$\overline{\mathrm{Dom}(DF(x,y))} \subset T_{\mathrm{Dom}(F)}(x)$$

If F is pseudo-convex at $(x,y) \in \mathrm{Graph}(F)$, (and in particular, if F is convex), then equality

$$\overline{\mathrm{Dom}(DF(x,y))} = T_{\mathrm{Dom}(F)}(x)$$

holds true, because

$$\mathrm{Dom}(F) = \pi_X \mathrm{Graph}(F)$$

where π_X is the projection from $X \times Y$ onto X. We thus deduce from Proposition 4.2.9 that

$$\left\{ \begin{array}{l} \overline{\mathrm{Dom}(DF(x,y))} = \overline{\pi_X \mathrm{Graph}(DF(x,y))} \\ = \overline{\pi_X T_{\mathrm{Graph}(F)}(x,y)} = T_{\pi_X(\mathrm{Graph}(F))}(x) = T_{\mathrm{Dom}(F)}(x) \end{array} \right. \square$$

Remark — Kernel of the Contingent Derivative The *kernel* of the contingent derivative is related to the contingent cone to the inverse image:

$$T_{F^{-1}(y)}(x) \subset \ker DF(x,y) := DF(x,y)^{-1}(0)$$

Equality holds true in this formula whenever F^{-1} is pseudo-Lipschitz around (y,x).

To prove the converse inclusion, let u belong to the kernel of $DF(x,y)$. Then there exist sequences $h_n > 0$ converging to 0, u_n and v_n converging respectively to u and 0 satisfying

$$\forall\, n, \;\; y + h_n v_n \;\in\; F(x + h_n u_n)$$

Since F^{-1} is pseudo-Lipschitz around (y,x), there exist $l > 0$ and an element $x_n^1 \in F^{-1}(y)$ such that

$$\|x_n^1 - (x + h_n u_n)\| \;\le\; l\|y - (y + h_n v_n)\| \;=\; l h_n \,\|v_n\|$$

Hence, by setting $u_n^1 := (x_n^1 - x)/h_n$, we see that

$$x + h_n u_n^1 \;=\; x_n^1 \;\in\; F^{-1}(y)$$

and that u_n^1 converges to u because $\|u_n^1 - u_n\| \;\le\; l\|v_n\|$ and because v_n converges to 0. Therefore we have proved that u belongs to the contingent cone to $F^{-1}(y)$ at x. □

Remark — Lower Semicontinuously Differentiable Maps

Definition 5.1.5 (Lower Semicontinuous Differentiability) *Let X, Y be normed spaces and $F : X \rightsquigarrow Y$ be a set-valued map. We say that F is* lower semicontinuously differentiable *at $(x,y) \in$ Graph(F) if the set-valued map*

$$(x,y,u) \;\in\; \mathrm{Graph}(F) \times X \;\rightsquigarrow\; DF(x,y)(u)$$

is lower semicontinuous.

Observe that it implies that F is sleek at (x_0, y_0). The converse needs further assumptions. We derive for instance from the "Cross-Convergence" Theorem 2.3.2 the following criterion:

Proposition 5.1.6 *Assume that X and Y are Banach spaces and that F is sleek on some neighborhood $\mathcal{U}$ of $(x_0, y_0) \in \mathrm{Graph}(F)$. If the boundedness property*

$$\forall\, u \;\in\; X, \qquad \sup_{(x,y)\,\in\,\mathcal{U}\cap\mathrm{Graph}(F)} \; \inf_{v \in DF(x,y)(u)} \|v\| \;<\; +\infty$$

holds true, then the set-valued map

$$(x,y,u) \;\in\; \mathrm{Graph}(F) \times X \rightsquigarrow DF(x,y)(u)$$

is lower semicontinuous on $(\mathcal{U} \cap \mathrm{Graph}(F)) \times X$.

5.2 Adjacent and Circatangent Derivatives

5.2.1 Definitions and Elementary Properties

Naturally, we can also associate with any other concept of tangent cone a concept of derivative. Since we were led to introduce Clarke and adjacent (or intermediate) tangent cones, we can introduce two more graphical derivatives:

Definition 5.2.1 *Let X, Y be normed spaces, $F : X \rightsquigarrow Y$ be a set-valued map and $y \in F(x)$.*

1. — the adjacent derivative $D^\flat F(x,y)$ *is the set-valued map from X to Y defined by*

$$\text{Graph}\left(D^\flat F(x,y)\right) := T^\flat_{\text{Graph}(F)}(x,y)$$

2. — the circatangent derivative $CF(x,y)$ *is the set-valued map from X to Y defined by*

$$\text{Graph}(CF(x,y)) := C_{\text{Graph}(F)}(x,y)$$

When $F := f$ is single-valued, we set

$$D^\flat f(x) := D^\flat f(x, f(x)), \ \ Cf(x) := Cf(x, f(x))$$

We see at once that these graphical derivatives are closed processes, that

$$\forall\, u, \ \ CF(x,y)(u) \subset D^\flat F(x,y)(u) \subset DF(x,y)(u)$$

When f is single-valued, we observe that the circatangent and adjacent derivatives $Cf(x)(u)$ and $D^\flat f(x)(u)$ are either empty or contain only one element. In particular,

$$\begin{cases} i) & D^\flat f(x)(u) = f'(x)u \ \text{ if } f \text{ is Fréchet differentiable at } x \\ ii) & Cf(x)(u) = f'(x)u \ \text{ if } f \text{ is continuously differentiable at } x \end{cases}$$

Naturally, the circatangent derivative is always a closed convex process.

Remark — Since $CF(x,y)$ is a closed convex process, it is equal to its bitranspose. The transpose

$$CF(x,y)^\star : Y^\star \leadsto X^\star$$

has been called the *codifferential* of F at (x,y). It is involved whenever dual conditions are used.

Instead of founding a "direct" differential calculus on tangent cones, as we did, it is possible to start with *any concept of normal cone* $N_{\mathrm{Graph}(F)}(x,y)$ to the graph of F at a point (x,y). We can associate with it the corresponding "codifferential" $\partial F(x,y)^\star$ defined by

$$p \in \partial F(x,y)^\star(q) \text{ if and only if } (p,-q) \in N_{\mathrm{Graph}(F)}(x,y)$$

When such a normal cone is not convex, it can no longer be defined as polar of a tangent cone, so that such codifferential, in general, is not the transpose of a graphical derivative. □

As for the contingent derivative, *a graphical derivative of the inverse of a set-valued map F is the inverse of the derivative*:

$$\begin{cases} i) & D^\flat(F^{-1})(y,x) = D^\flat F(x,y)^{-1} \\ ii) & C(F^{-1})(y,x) = CF(x,y)^{-1} \end{cases}$$

If K is a subset of X and $f : X \mapsto Y$ is a single-valued continuously differentiable around a point $x \in K$ map, then *the derivative of the restriction is the restriction of the derivative to the corresponding tangent cone*:

$$\begin{cases} i) & D^\flat(f|_K)(x) := D^\flat(f|_K)(x,f(x)) = f'(x)|_{T^\flat_K(x)} \\ ii) & C(f|_K)(x) := C(f|_K)(x,f(x)) = f'(x)|_{C_K(x)} \end{cases}$$

Actually, these formulas follow from:

Proposition 5.2.2 *Consider normed spaces X, Y, a single-valued map f from an open subset $\Omega \subset X$ to Y and a set-valued map $M : X \leadsto Y$. Define the set-valued map $F : X \leadsto Y$ by:*

$$\forall x \in X, \quad F(x) := f(x) - M(x)$$

If f is Fréchet differentiable at $x \in \Omega \cap \mathrm{Dom}(M)$, then for every $y \in F(x)$

$$D^\flat F(x,y)(u) \;=\; f'(x)u - D^\flat M(x, f(x)-y)(u)$$

and if f is continuously differentiable at x, then

$$CF(x,y)(u) \;=\; f'(x)u - CM(x, f(x)-y)(u)$$

The proof is analogous to the one of Proposition 5.1.2 and is left as an exercise.

Proposition 5.2.3 *Let X and Y be normed spaces, f a single-valued map from an open subset $\Omega \subset X$ to Y, $M : X \rightsquigarrow Y$ be a set-valued map and $L \subset X$. Consider the set-valued map $F : X \rightsquigarrow Y$ defined by:*

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

If f is Fréchet differentiable at $x \in \Omega \cap \mathrm{Dom}(F)$ and M is Lipschitz at x, then for every $y \in F(x)$ the adjacent derivative of F at (x,y) is equal to

$$D^\flat F(x,y)(u) = \begin{cases} f'(x)u - D^\flat M(x, f(x)-y)(u) & \text{when } \; u \in T_L^\flat(x) \\ \emptyset & \text{when } \; u \notin T_L^\flat(x) \end{cases}$$

The same formula holds true for a map f continuously differentible at x, the circatangent derivatives of F and M and the Clarke tangent cone to L.

The proof is the same as the one of Proposition 5.1.3 and these statements remain true when f is Lipschitz at x and Gâteaux differentiable. □

5.2.2 Limits of Differential Quotients

In order to characterize adjacent and circatangent derivatives in terms of limits of differential quotients, we need the concept of "lim sup inf" (or Γ-convergence) of functions of two variables introduced independently by several authors (de Giorgi, Rockafellar, ...):

Definition 5.2.4 (Lim sup inf) *Let L and M be two metric spaces and $\phi : L \times M \mapsto \mathbf{R}$ be a function. We set*

$$\limsup_{x' \to x} \inf_{y' \to y} \phi(x', y') := \sup_{\varepsilon > 0} \inf_{\eta > 0} \sup_{x' \in B(x, \eta)} \inf_{y' \in B(y, \varepsilon)} \phi(x', y')$$

We write $(x', y') \to_F (x, y)$ to denote that (x', y') converges to (x, y) while remaining in the graph of F.

Hence, by translating the definition of the intermediate and the Clarke tangent cones, we obtain the following characterizations:

Proposition 5.2.5 *Let (x, y) belong to the graph of a set-valued map $F : X \rightsquigarrow Y$ from a normed space X to another Y. Then*

$$\begin{cases} v \text{ belongs to } D^\flat F(x, y)(u) \text{ if and only if} \\ \limsup_{h \to 0+} \inf_{u' \to u} d\left(v, \frac{F(x+hu')-y}{h}\right) = 0 \end{cases}$$

and

$$\begin{cases} v \text{ belongs to } CF(x, y)(u) \text{ if and only if} \\ \limsup_{h \to 0+,\ (x', y') \to_F (x, y)} \inf_{u' \to u} d\left(v, \frac{F(x'+hu')-y'}{h}\right) = 0 \end{cases}$$

If F is Lipschitz around $x \in \mathrm{Int}(\mathrm{Dom}(F))$, then the formulas become much simpler:

$$\begin{cases} v \text{ belongs to } D^\flat F(x, y)(u) \text{ if and only if} \\ \lim_{h \to 0+} d\left(v, \frac{F(x+hu)-y}{h}\right) = 0 \end{cases}$$

and

$$\begin{cases} v \text{ belongs to } CF(x, y)(u) \text{ if and only if} \\ \lim_{h \to 0+,\ (x', y') \to_F (x, y)} d\left(v, \frac{F(x'+hu)-y'}{h}\right) = 0 \end{cases}$$

Remark: Kernel of the Derivative As for the contingent derivative, the kernels of other derivatives are linked to the associated tangent cones to the inverse image:

$$T^\flat_{F^{-1}(y)}(x) \subset \ker D^\flat F(x, y) := D^\flat F(x, y)^{-1}(0)$$

and equality holds true in this formula whenever F^{-1} is pseudo-Lipschitz around (y, x). In this case, we also have:

$$\ker CF(x,y) \subset C_{F^{-1}(y)}(x) \quad \square$$

The proofs are left as an exercise.

Proposition 5.2.6 *Let us assume that the images of F are convex and that F is Lipschitz around x. Then for any $(x, y) \in \text{Graph}(F)$ the images of the adjacent derivative $D^\flat F(x, y)$ are convex and*

$$\begin{cases} D^\flat F(x,y)(0) & = & T_{F(x)}(y) \\ \\ D^\flat F(x,y)(u) + D^\flat F(x,y)(0) & = & D^\flat F(x,y)(u) \end{cases}$$

Proof — Let v_1 and v_2 belong to $DF(x, y)(u)$. Then, for any sequence $h_n > 0$ converging to 0, there exist sequences u_{1n} and u_{2n} converging to u and sequences v_{1n} and v_{2n} converging to v_1 and v_2 respectively such that

$$\forall\, n, \quad y + h_n v_{in} \in F(x + h_n u_{in}) \quad (i = 1, 2)$$

Since F is Lipschitz around x, there exists $l > 0$ such that for all n large enough,

$$y + h_n v_{2n} \in F(x + h_n u_{1n}) + l h_n \|u_{2n} - u_{1n}\|$$

so that we can find another sequence v_{3n} converging to v_2 such that

$$y + h_n v_{3n} \subset F(x + h_n u_{1n})$$

Now, $F(x + h_n u_{1n})$ being convex, we deduce that for all $\lambda \in [0, 1]$,

$$y + h_n(\lambda v_{1n} + (1 - \lambda) v_{3n}) \in F(x + h_n u_{1n})$$

Since $\lambda v_{1n} + (1 - \lambda) v_{3n}$ converges to $\lambda v_1 + (1 - \lambda) v_2$, this element belongs to $D^\flat F(x, y)(u)$.

By Proposition 5.2.5, we see that $v \in D^\flat F(x, y)(0)$ if and only if $d(v, (F(x) - y)/h)$ converges to 0, i.e., v belongs to the adjacent tangent cone to $F(x)$ at y. Since $F(x)$ is convex, it coincides with the tangent cone.

Since $0 \in D^\flat F(x,y)(0)$ we obtain that

$$\forall\, u,\ D^\flat F(x,y)(u) \ \subset\ D^\flat F(x,y)(u) \ +\ D^\flat F(x,y)(0)$$

To prove the opposite inclusion fix

$$v \ \in\ D^\flat F(x,y)(u) \quad \& \quad w \ \in\ D^\flat F(x,y)(0)$$

Let $h_n \to 0+$, $v_n \to v$ be such that

$$\forall\, n,\ \ y + h_n v_n \ \in\ F(x + h_n u)$$

By convexity of $F(x)$, there exist $w_n \to w$ such that for n large enough, $y + \sqrt{h_n} w_n \in F(x)$. Then, by the Lipschitz continuity of F, for all large n and for some w'_n, we have

$$y \ +\ \sqrt{h_n} w'_n \ \in\ F(x + h_n u)\ ;\ \|w'_n - w_n\| \ \leq\ l\sqrt{h_n}\,\|u\|$$

Thus

$$\begin{cases} (1 - \sqrt{h_n})(y + h_n v_n) \ +\ \sqrt{h_n}(y + \sqrt{h_n} w'_n) \\ \\ =\ y + h_n(v_n + w'_n)\ -\ \sqrt{h_n}\, h_n v_n \ =\ y + h_n(v + w) + h_n \varepsilon(h_n) \\ \\ \in\ F(x + h_n u) \end{cases}$$

where $\varepsilon(h_n)$ converges to 0. Hence

$$\lim_{n \to \infty} \operatorname{dist}\left(v + w,\ \frac{F(x + h_n u) - y}{h_n}\right) \ =\ 0$$

This ends the proof. $\square$

5.2.3 Derivatives of monotone operators

Let X be a Hilbert space (identified with its dual.) We recall that a set-valued map $F : X \rightsquigarrow X$ is monotone if and only if

$$\forall\, (x,p),\ (y,q)\ \in\ \mathrm{Graph}(F),\ \ < p - q\ ,\ x - y > \geq\ 0$$

(see Chapter 3 for an introduction to monotone and maximal monotone maps.) We also recall that when F is monotone, its *resolvent* $J := (\mathbf{1} + F)^{-1}$ is single-valued and Lipschitz (with constant equal to 1) on its domain. Therefore, we can easily compute derivatives of F in terms of the derivatives of its resolvent.

Proposition 5.2.7 *Let X be a Hilbert space identified with its dual $X^\star$. We supply $X^\star$ with the weak-$\star$ topology. If $F : X \leadsto X^\star$ is monotone, then the contingent derivative $DF(x,p)$ at every pair $(x,p) \in Graph(F)$ is semi positive-definite in the sense that*

$$\forall \, (u,r) \in \text{Graph}(DF(x,p)), \quad < r, u > \geq 0$$

Furthermore, the following statements are equivalent:

$$\begin{cases} a) & r \in DF(x,p)(u) \\ \\ b) & u \in DJ(x+p)(r+u) \end{cases}$$

All the above statements remain true for the adjacent and circatangent derivatives.

Proof — The first claim is obvious, since $< r, u >$ is the limit of a sequence $< r_n, u_n >$, where $p + h_n r_n \in F(x + h_n u_n)$, (because u_n converges to u strongly and r_n converges to r weakly) and since

$$h_n^2 < r_n, u_n > = < x + h_n u_n - x, p + h_n r_n - p > \geq 0$$

For proving the second statement, we observe that p belongs to $F(x)$ if and only if $x = J(x+p)$, because $x + p \in (\mathbf{1} + F)(x)$ and because J is the inverse of $(\mathbf{1} + F)$. We also deduce from the latter that the two following statements are equivalent:

$$\begin{cases} i) & r + u \in D(\mathbf{1}+F)(x, x+p)(u) \\ \\ ii) & u \in DJ(x+p, x)(r+u) \; = \; DJ(x+p)(r+u) \end{cases}$$

On the other hand, Proposition 5.1.3 implies that

$$D(\mathbf{1}+F)(x, x+p)(u) \; = \; u + DF(x,p)(u)$$

We have obtained the formula we were looking for. □

Since the cone-valued map N_K associating with any $x \in K$ the normal cone $N_K(x)$ to a closed convex subset is maximal monotone (because the normal cone is the subdifferential of the indicator of K), and since its resolvent is the projector of best approximation onto K, we deduce the following corollary:

Corollary 5.2.8 *Let K be a closed convex subset of a Hilbert space and p belong to the normal cone $N_K(x)$ to K at some $x \in K$. Denote by Π_K the best approximation projector onto K. Then,*

$$q \in DN_K(x,p)(u) \iff u \in D\Pi_K(x+p)(u+q)$$

5.3 Chain Rules

We derive from the calculus of tangent cones the associated calculus of derivatives of set-valued maps. We begin naturally by the chain rule for computing the composition product of a set-valued map $G : X \rightsquigarrow Y$ and a set-valued map $H : Y \rightsquigarrow Z$.

We shall need the following result:

Proposition 5.3.1 *Let X, Y be normed spaces, $F : X \rightsquigarrow Y$ be a set-valued map and K be a subset of X. Assume that F is Lipschitz around some $x \in K$. Then, for any $y \in F(x)$, we have*

$$D^\flat F(x,y)(T_K(x)) \subset T_{F(K)}(y)$$

As a consequence, we deduce that when M is a subset of Y and $y \in M$, then

$$T_{F^{+1}(M)}(x) \subset D^\flat F(x,y)^{+1}(T_M(y)) \tag{5.1}$$

Proof — Take u in $T_K(x)$ and $v \in D^\flat F(x,y)(u)$. Then there exist sequences $h_n > 0$ converging to 0, u_{1n} and u_{2n} converging to u and v_n converging to v such that

$$x + h_n u_{1n} \in K \quad \& \quad y + h_n v_n \in F(x + h_n u_{2n})$$

Since F is Lipschitz around x with a Lipschitz constant l, we deduce that

$$y + h_n v_n \in F(x + h_n u_{1n}) + l h_n \|u_{1n} - u_{2n}\|$$

so that there exists another sequence $v_n^\star$ converging to v such that

$$y + h_n v_n^\star \in F(x + h_n u_{1n}) \subset F(K)$$

This implies that v belongs to the contingent cone to $F(K)$ at y.

Consider now $K := F^{+1}(M)$. Since $F(F^{+1}(M))$ is contained in M, we deduce that

$$D^\flat F(x,y)\left(T_{F^{+1}(M)}(x)\right) \subset T_{F(F^{+1}(M))}(y) \subset T_M(y)$$

from which formula (5.1) ensues. □

Remark — Naturally, we can show in the same way that for a Lipschitz map F the formula

$$DF(x,y)T_K^\flat(x) \subset T_{F(K)}(y)$$

is also true whenever $y \in F(x)$. □

We begin by the following simple result:

Theorem 5.3.2 *Let us consider normed spaces X, Y, Z, a set-valued map $G : X \rightsquigarrow Y$ and a set-valued map $H : Y \rightsquigarrow Z$.*

1. — Let us assume that H is Lipschitz around y, where $y \in G(x)$. Then, for any $z \in H(y)$, we have

$$\begin{cases} D^\flat H(y,z) \circ DG(x,y) \subset D(H \circ G)(x,z) \\ \\ DH(y,z) \circ D^\flat G(x,y) \subset D(H \circ G)(x,z) \end{cases}$$

2. — If $G := g$ is single-valued and Fréchet differentiable at x, we obtain

$$\forall\, z \in H(g(x)),\ D(Hg)(x,z)(u) \subset DH(g(x),z)(g'(x)u)$$

and the equality holds true when H is Lipschitz around $g(x)$.

Proof — We apply Proposition 5.3.1 to equality

$$\text{Graph}(H \circ G) = (\mathbf{1} \times H)\,(\text{Graph}(G))$$

for proving the first statement. The second one follows from

$$\begin{cases} \text{Graph}(H \circ g) = (g \times \mathbf{1})^{-1}\,(\text{Graph}(H)) \\ \\ T_{f^{-1}(K)}(x) \subset f'(x)^{-1}\,(T_K(f(x))) \end{cases} \quad \Box$$

Remark — Let us assume that G is Lipschitz around x. Then, for all $y \in G(x)$ and $z \in (H\Box G)(x)$, we have

$$D(H\Box G)(x,z) \subset DH(y,z)\Box D^\flat G(x,y)$$

because we can apply Proposition 5.3.1 to $M := \mathrm{Graph}(H)$ and because

$$\mathrm{Graph}(H\Box G) = (G \times \mathbf{1})^{+1}(\mathrm{Graph}(H))$$

We state now a more powerful result.

Theorem 5.3.3 *Let X, Z be Banach spaces and Y be a finite dimensional vector-space. Consider set-valued maps $G : X \rightsquigarrow Y$ and $H : Y \rightsquigarrow Z$. Fix $x_0 \in \mathrm{Dom}(G)$, $y_0 \in G(x_0) \cap \mathrm{Dom}(H)$ and $z_0 \in H(y_0)$.*

If G and H are closed and satisfy the transversality condition

$$\mathrm{Im}(CG(x_0,y_0)) - \mathrm{Dom}(CH(y_0,z_0)) = Y$$

then

$$\begin{cases} i) & D^\flat H(y_0,z_0)\circ DG(x_0,y_0) & \subset & D(H\circ G)(x_0,z_0) \\ ii) & D^\flat H(y_0,z_0)\circ D^\flat G(x_0,y_0) & \subset & D^\flat(H\circ G)(x_0,z_0) \\ iii) & CH(y_0,z_0)\circ CG(x_0,y_0) & \subset & C(H\circ G)(x_0,z_0) \end{cases}$$

Proof — If we denote by ω the continuous linear operator from $X \times Y \times Y \times Z$ to Y associating to (x, y_1, y_2, z) the element $y_1 - y_2$ and by $\pi_{X\times Z}$ the canonical projection from $X \times Y \times Y \times Z$ onto $X \times Z$. Observe that

$$\mathrm{Graph}(H\circ G) = \pi_{X\times Z}\Big((\mathrm{Graph}(G)\times \mathrm{Graph}(H))\cap\omega^{-1}(0)\Big)$$

We apply Theorem 4.3.3 with $L = \mathrm{Graph}(G) \times \mathrm{Graph}(H)$, $f = \omega$ and $M = \{0\}$.

Consequently, we deduce that, for instance,

$$\left\{\begin{array}{l}\left(\mathrm{Graph}(D^{\flat}G(x_0,y_0)) \times \mathrm{Graph}(D^{\flat}H(y_0,z_0))\right) \cap \omega^{-1}(0) \\ \\ = \left(T^{\flat}_{\mathrm{Graph}(G)}(x_0,y_0) \times T^{\flat}_{\mathrm{Graph}(H)}(y_0,z_0)\right) \cap \omega^{-1}(0) \\ \\ = T^{\flat}_{\mathrm{Graph}(G)\times\mathrm{Graph}(H)}(x_0,y_0,y_0,z_0) \cap \omega^{-1}(0) \\ \\ = T^{\flat}_{(\mathrm{Graph}(G)\times\mathrm{Graph}(H))\cap\omega^{-1}(0)}(x_0,y_0,y_0,z_0)\end{array}\right. \tag{5.2}$$

By applying the projection $\pi_{X\times Z}$ to both sides of these equalities, we prove that

$$\left\{\begin{array}{l}\mathrm{Graph}\left(D^{\flat}H(y_0,z_0) \circ D^{\flat}G(x_0,y_0)\right) \\ \\ = \pi_{X\times Z}\left(\left(\mathrm{Graph}(D^{\flat}G(x_0,y_0)) \times \mathrm{Graph}(D^{\flat}H(y_0,z_0))\right) \cap \omega^{-1}(0)\right) \\ \\ = \pi_{X\times Z}\left(T^{\flat}_{(\mathrm{Graph}(G)\times\mathrm{Graph}(H))\cap\omega^{-1}(0)}(x_0,y_0,y_0,z_0)\right) \\ \\ \subset T^{\flat}_{\pi_{X\times Z}((\mathrm{Graph}(G)\times\mathrm{Graph}(H))\cap\omega^{-1}(0))}(x_0,z_0) \\ \\ = T^{\flat}_{\mathrm{Graph}(H\circ G)}(x_0,z_0) \; = \; \mathrm{Graph}(D^{\flat}(H\circ G)(x_0,z_0))\end{array}\right.$$

The proof of the other statements is similar and left as an exercise.
□

Remark — We recall that the *norm* of a process (positively homogeneous set-valued map) $A : Y \rightsquigarrow X$ is defined by

$$\|A\| := \sup_{\|y\|\leq 1} \inf_{x\in A(y)} \|x\|$$

If Y is any Banach space, we have to assume that there exist constants

$c > 0$, $\alpha \in [0,1[$ and $\eta > 0$ such that

$$\begin{cases} \forall\,(x,y_1) \in \mathrm{Graph}(G) \cap B((x_0,y_0),\eta), \\ \forall\,(y_2,z) \in \mathrm{Graph}(H) \cap B((y_0,z_0),\eta), \\ B_Y \subset \mathrm{Im}(D^\flat G(x,y_1)) \cap cB_Y - \mathrm{Dom}(DH(y_2,z)) + \alpha B_Y \text{ and} \\ \|D^\flat G(x,y_1)\| \le c \;\&\; \|D(H^{-1})(z,y_2)\| \le c \end{cases}$$

to prove the above formulas.

Indeed, the above assumption implies the second transversality condition of Theorem 4.3.3. For any $v \in Y$, we can find $v_1 \in \mathrm{Im}(D^\flat G(x,y_1))$ and $v_2 \in \mathrm{Dom}(DH(y_2,z))$ such that

$$v = v_1 - v_2 + e, \quad \|v_1\| \le c\|v\|, \quad \|e\| \le \alpha\|v\|$$

Hence $\|v_2\| \le (c+1+\alpha)\|v\|$ and there exist

$$u \in D^\flat G(x,y_1)^{-1}(v_1) \;\;\&\;\; w \in DH(y_2,z)(v_2)$$

such that $\|u\| \le c\|v_1\|$ and $\|w\| \le c\|v_2\|$.

Therefore, $v = \omega(u,v_1,v_2,w) + e$ where (u,v_1,v_2,w) belongs to the contingent cone to the product of the graphs of G and H and $e \in \alpha B_Y$. Consequently, we can derive properties (5.2) and conclude the proof. □

We obtain converse inclusions under more severe assumptions:

Proposition 5.3.4 *Let X, Y, Z be normed spaces. Consider set-valued maps $G : X \rightsquigarrow Y$, $H : Y \rightsquigarrow Z$ and assume that G is pseudo-convex at $(x_0,y_0) \in \mathrm{Graph}(G)$ and H is pseudo-convex at $(y_0,z_0) \in \mathrm{Graph}(H)$. Then*

$$D(H \circ G)(x_0,z_0) \subset \overline{DH(y_0,z_0) \circ DG(x_0,y_0)}$$

Proof — Let ω and $\pi_{X\times Z}$ be the operators defined in the proof of Theorem 5.3.3. When the graph of G is pseudo-convex at (x,y) and the graph of H is pseudo-convex at (y,z), we derive the above property from Proposition 4.2.9 applied in the following situation:

$$\begin{cases} \text{Graph}(D(H \circ G)(x_0, z_0)) = T_{\text{Graph}(H \circ G)}(x_0, z_0) \\ = T_{\pi_{X \times Z}((\text{Graph}(G) \times \text{Graph}(H)) \cap \omega^{-1}(0))}(x_0, z_0) \\ = \text{cl}\left(\pi_{X \times Z}\left(T_{(\text{Graph}(G) \times \text{Graph}(H)) \cap \omega^{-1}(0)}(x_0, y_0, y_0, z_0)\right)\right) \\ \subset \text{cl}\left(\pi_{X \times Z}\left((\text{Graph}(DG(x_0, y_0)) \times \text{Graph}(DH(y_0, z_0))) \cap \omega^{-1}(0)\right)\right) \\ = \text{cl}(\text{Graph}(DH(y_0, z_0) \circ DG(x_0, y_0))) \end{cases}$$

Remark — We can deduce the conclusions of this proposition from any other criterion implying that the contingent cones to the images by $\pi_{X \times Z}$ are the closures of the images of the contingent cones. □

Remark — Let X be a normed space and K, L be subsets of X. Recall that the Bouligand *paratingent cone* $P_K^L(x)$ to K relative to L at $x \in \overline{L} \cap \overline{K}$ is defined by

$$P_K^L(x) := \{v \mid \liminf_{h \to 0+, x' \to_L x} d_K(x' + hv)/h = 0\}$$

and that we set $P_K(x) := P_K^K(x)$.

We can obtain upper estimates of the composition product by using the *paratingent derivatives* defined by saying that *the graph of the paratingent derivative $PF(x, y)$ of F at (x, y) is the paratingent cone to the graph of F at (x, y)*:

$$\text{Graph}(PF(x, y)) := P_{\text{Graph}(F)}(x, y)$$

Hence, for any $u \in X$, the paratingent derivative in the direction u is equal to

$$PF(x, y)(u) = \text{Limsup}_{h \to 0+, u' \to u, (x', y') \to_F (x, y)} \frac{F(x' + hu') - y'}{h}$$

We can also define the *lop-sided paratingent derivatives* $P^x F(x, y)$ and $P^y F(x, y)$ in the following way:

$$\begin{cases} i) \quad P^x F(x, y)(u) = \text{Limsup}_{h \to 0+, u' \to u, y' \to_{F(x)} y} \frac{F(x+hu') - y'}{h} \\ ii) \quad P^y F(x, y)(u) = \text{Limsup}_{h \to 0+, u' \to u, x' \to_{F^{-1}(y)} x} \frac{F(x'+hu') - y}{h} \end{cases}$$

Theorem 5.3.5 *Let X, Y be normed spaces. Assume that $G : X \rightsquigarrow Y$ is Lipschitz around x. If Y is a finite dimensional vector-space and $G(x)$ is bounded, then*

$$D(H \circ G)(x,z) \subset \bigcup_{y \in G(x)} P^z H(y,z) \circ P^x G(x,y)$$

Proof — Let w belong to $D(H \circ G)(x,z)(u)$: there exist sequences $h_n > 0$, u_n and w_n converging to 0, u and w respectively such that

$$\forall\, n, \quad z + h_n w_n \;\in\; H(y_n) \text{ where } y_n \;\in\; G(x + h_n u_n)$$

Since G is Lipschitz around x, there exist $l > 0$ and elements $y_n^0 \in G(x)$ such that, for n large enough,

$$v_n \;:=\; \frac{y_n - y_n^0}{h_n} \text{ satisfies } \|v_n\| \;\leq\; l\|u_n\|$$

Furthermore, $G(x)$ being relatively compact, a subsequence (again denoted) y_n^0 converges to some y. We can also extract a subsequence (again denoted) v_n which converges to some v, since this sequence is bounded and the dimension of Y is finite.

From the relations

$$y_n^0 + h_n v_n \;\in\; G(x + h_n u_n) \;\;\&\;\; z + h_n w_n \;\in\; H(y_n^0 + h_n v_n)$$

we infer that w belongs to $P^z H(y,z)(v)$ and v to $P^x G(x,y)(u)$. □

It is not difficult to check that for all $y \in G(x)$, $u \in \mathrm{Dom}(CG(x,y))$,

$$P(H \square G)(x,z)(u) \;\subset\; \left(PH(y,z) \square CG(x,y)\right)(u)$$

Indeed, let $u \in \mathrm{Dom}(CG(x,y))$ and w belong to $P(H \square G)(x,z)(u)$. Hence there exist a sequence $h_n > 0$ converging to 0 and sequences of elements $(x_n, z_n) \in \mathrm{Graph}(H \square G)$, u_n and w_n converging to (x,z), u and w respectively such that

$$\forall\, n \geq 0, \;\; z_n + h_n w_n \;\in\; \bigcap_{y \,\in\, G(x_n + h_n u_n)} H(y)$$

The set-valued map G being Lipschitz, there exists a sequence of elements $y_n \in G(x_n)$ converging to y. By definition of the square product, we know that $z_n \in H(y_n)$ (because $z_n \in (H \square G)(x_n)$.)

Take now any v in $CG(x,y)(u)$. Since G is Lipschitz around x, there exists a sequence of elements v_n converging to v such that

$$\forall\, n \geq 0, \;\; y_n + h_n v_n \;\in\; G(x_n + h_n u_n)$$

Therefore,

$$\forall\, n \geq 0, \;\; z_n + h_n w_n \;\in\; H(y_n + h_n v_n)$$

so that we infer

$$w \;\in\; PH(y,z)(v)$$

Since this is true for every element v of $CG(x,y)(u)$, we deduce that

$$w \;\in\; \bigcap_{v \in CG(x,y)(u)} PH(y,z)(v) \;=\; (PH(y,z)\square CG(x,y))(u) \qquad \square$$

5.4 Inverse Set-Valued Map Theorem

5.4.1 Stability and Approximation of Inclusions

Let X, Y be Banach spaces and $F : X \leadsto Y$ and $F_n : X \leadsto Y$ be set-valued maps with closed graphs. We are ready to extend Lax's Principle to the approximation of a solution to inclusion

$$\text{find } \; x_0 \;\in\; X \; \text{ such that } \; F(x_0) \;\ni\; y_0$$

by solutions to the approximate inclusions

$$\text{find } \; x_n \;\in\; X \; \text{ such that } \; F_n(x_n) \;\ni\; y_n$$

Definition 5.4.1 *We shall say that the set-valued maps F_n are* consistent *at $(x_0, y_0) \in \text{Graph}(F)$ if there exist sequences of*

$$(x_{0n}, y_{0n}) \;\in\; \text{Graph}(F_n)$$

such that x_{0n} converges to x_0 and y_{0n} converges to y_0.

We shall say that a family of set-valued maps F_n is stable *around (x,y) if there exist constants $c > 0$, $\alpha \in [0,1[$ and $\eta > 0$ such that*

$$\begin{cases} \forall\, (x_n, y_n) \;\in\; \text{Graph}(F_n) \cap B((x_0,y_0),\eta), \; \forall\, v \;\in\; Y, \\ \\ \exists\, u_n \;\in\; X, \; \exists\, w_n \;\in\; Y \;\; \text{with} \;\; v \;\in\; DF_n(x_n,y_n)(u_n) + w_n \\ \\ \text{and} \;\; \|u_n\| \;\leq\; c\|v\| \;\;\&\;\; \|w_n\| \;\leq\; \alpha\|v\| \end{cases}$$

Theorem 5.4.2 *Let us consider Banach spaces X and Y, a sequence of closed set-valued maps $F_n : X \rightsquigarrow Y$, $y_0 \in Y$ and a solution x_0 to the inclusion $F(x_0) \ni y_0$.*

Assume that the set-valued maps F_n are consistent at (x_0, y_0) and that the family of F_n's is stable.

Then there exists a constant l such that, for any sequence of elements $(x_{0n}, y_{0n}) \in \mathrm{Graph}(F_n)$ converging to (x_0, y_0) and any sequence y_n converging to y_0, we have for n large enough

$$d\left(x_{0n}, F_n^{-1}(y_n)\right) \leq l\|y_{0n} - y_n\|$$

As a consequence, we obtain the following error estimate: for any sequence y_n converging to y_0,

$$d\left(x_0, F_n^{-1}(y_n)\right) \leq l\,(d((x_0, y_0), \mathrm{Graph}(F_n)) + \|y_0 - y_n\|)$$

so that x_0 can be approximated by solutions $x_n \in F_n^{-1}(y_n)$.

Proof — We apply Theorem 3.4.5 with X replaced by $X \times Y$, K_n by $\mathrm{Graph}(F_n)$, f_n by the projection Π_Y from $X \times Y$ onto Y. We have to prove that the stability assumption implies transversality assumption (3.9) of Theorem 3.4.5, i.e., that for all $v \in Y$, there exist (u_n, v_n) in the contingent cone $T_{\mathrm{Graph}(F_n)}(x_n, y_n)$ and $w_n \in Y$ satisfying

$$v = v_n + w_n, \quad \max(\|u_n\|, \|v_n\|) \leq c\|v\|, \quad \|w_n\| \leq \alpha\|v\|$$

This information is provided by our stability assumption since the contingent cone to the graph is the graph of the contingent derivative and the norm of $v_n = v - w_n$ is smaller than or equal to $(1+\alpha)\|v\|$. □

When the set-valued maps F_n are all equal to F, we obtain:

Theorem 5.4.3 (Inverse Set-Valued Map Theorem) *Consider a closed set-valued map $F : X \rightsquigarrow Y$, an element (x_0, y_0) of its graph and let us assume that there exist constants $c > 0$, $\alpha \in [0, 1[$ and*

$\eta > 0$ *such that*

$$\begin{cases} \forall\, (x,y) \in \text{Graph}(F) \cap B((x_0,y_0),\eta),\ \forall\, v \in Y, \\ \exists\, u \in X,\ \exists\, w \in Y \text{such that } v \in DF(x,y)(u) + w \\ \text{and } \|u\| \le c\|v\| \ \& \ \|w\| \le \alpha\|v\| \end{cases}$$

Then y_0 belongs to the interior of the image of F and F^{-1} is pseudo-Lipschitz around (y_0, x_0).

We can extend Theorem 3.4.10 to the case of set-valued maps by introducing an adequate definition of uniformly sleek maps, which are the set-valued analogues of the continuously differentiable maps.

Definition 5.4.4 (Strongly Sleek Maps) *We shall say that F is* uniformly sleek *at a point $(x_0, y_0) \in \text{Graph}(F)$ if its graph is uniformly sleek at this point, i.e., if*

$$\sup_{(u,v)\in\text{Graph}(DF(x_0,y_0))\cap(B\times B)} d((u,v), \text{Graph}(DF(x,y)))$$

converges to 0 when (x,y) converges to (x_0,y_0) in $\text{Graph}(F)$.

With this definition, we can state a natural set-valued version of Graves' Theorem

Theorem 5.4.5 *Let us consider Banach spaces X and Y, a closed set-valued map $F : X \rightsquigarrow Y$ and an element (x_0, y_0) of its graph. Let us assume that*

$$DF(x_0,y_0) \text{ is surjective}$$

1. — If F is uniformly sleek at (x_0, y_0), then y_0 belongs to the interior of the image of F and F^{-1} is pseudo-Lipschitz around (y_0, x_0).

2. — If the dimension of Y is finite, it is sufficient to assume that F is sleek at (x_0, y_0) to get the same conclusions.

The proof follows from Theorem 3.4.10. The very same proof yields

Theorem 5.4.6 *Consider Banach spaces X and Y, a closed set-valued map $F : X \rightsquigarrow Y$ and an element (x_0, y_0) of its graph. Let us assume that the dimension of Y is finite and*

$$CF(x_0, y_0) \text{ is surjective}$$

Then y_0 belongs to the interior of the image of F and F^{-1} is pseudo-Lipschitz around (y_0, x_0).

5.4.2 Localization of Inverse Images

In this section we provide some estimates of the inverse image $F^{-1}(y^\star)$.

Theorem 5.4.7 *Let F be a set-valued map from a normed space X to a normed space Y and $(x^\star, y^\star)$ belong to its graph.*

1. — Assume that the dimension of X is finite. Then, for any closed cone P satisfying

$$\ker(DF(x^\star, y^\star)) \cap P = \{0\}$$

there exists $\varepsilon > 0$ such that[2]

$$F^{-1}(y^\star) \cap (x^\star + \varepsilon(P \cap B)) = \{x^\star\}$$

In particular if $\ker(DF(x^\star, y^\star)) = \{0\}$, *then*

$$F^{-1}(y^\star) \cap B(x^\star, \varepsilon) = \{x^\star\}$$

2. — Assume that $F^{-1}(y^\star)$ is convex[3]*. Then*

$$F^{-1}(y^\star) \subset x^\star + \ker(DF(x^\star, y^\star))$$

In particular if $\ker(DF(x^\star, y^\star)) = \{0\}$, *then* $F^{-1}(y^\star) = \{x^\star\}$.

[2] If F^{-1} is pseudo-Lipchitz at (y, x), it is possible to show that conversely, the property

$$F^{-1}(y^\star) \cap (x^\star + \varepsilon(P \cap B)) = \{x^\star\}$$

implies that $\ker(DF(x^\star, y^\star)) \cap \mathrm{Int}(P) = \emptyset$.

[3] or even, pseudo-convex at $x^\star$.

Proof — Assume that the first statement is false. Then there exists a sequence of elements $x_n \in F^{-1}(y^\star)$ converging to $x^\star$ such that $x_n - x^\star \neq 0$ belong to P. Setting $h_n := \|x_n - x^\star\|$ which converges to $0+$ and $u_n := (x_n - x^\star)/h_n$ belonging to the unit sphere and to the cone P, we infer that a subsequence (again denoted) u_n converges to some element $u \in P$ of norm 1 because the dimension of X is finite.

On the other hand, inclusions

$$y^\star + h_n 0 \in F(x_n) = F(x^\star + h_n u_n)$$

imply that $0 \in DF(x^\star, y^\star)(u)$, i.e., that u belongs to the kernel of $DF(x^\star, y^\star)$. This is impossible. By taking $P = X$, we obtain the local uniqueness of a solution $x^\star$ to $y^\star \in F(x)$.

The second statement follows readily from the definition, since we observe that the convexity of the inverse image implies that

$$F^{-1}(y^\star) \subset x^\star + T_{F^{-1}(y^\star)}(x^\star) \subset x^\star + \ker(DF(x^\star, y^\star)) \quad \square$$

5.4.3 The Equilibrium Map

Let us consider two finite dimensional spaces X and Λ and a set-valued map

$$F : \Lambda \times X \rightsquigarrow X$$

We denote by $E : \Lambda \rightsquigarrow X$ the *equilibrium map* defined by

$$\forall \lambda \in \Lambda, \quad E(\lambda) := \{x \in X \mid 0 \in F(\lambda, x)\}$$

which associates with any parameter $\lambda \in \Lambda$ the set of equilibria of $F(\lambda, \cdot)$. We can derive some information on this equilibrium map from the knowledge of its contingent derivative.

Proposition 5.4.8 *Let Λ and X be two finite dimensional vector-spaces and $F : \Lambda \times X \rightsquigarrow X$ be a closed set-valued map. Assume that the circatangent derivatives $CF(\lambda, x, 0)$ are surjective for every (λ, x) belonging to the graph of the equilibrium map E. Then*

1. — The contingent derivative of the equilibrium map is the equilibrium map of the contingent derivative:

$$0 \in DF(\lambda, x, 0)(\mu, v) \iff v \in DE(\lambda, x)(\mu)$$

and

$$0 \in CF(\lambda, x, 0)(\mu, v) \implies v \in CE(\lambda, x)(\mu)$$

2. — *If for any $\mu \in \Lambda$, there exists an equilibrium v of $CF(\lambda, x, 0)(\mu, \cdot)$, then the equilibrium map is pseudo-Lipschitz around (λ, x).*

3. — *Set*

$$Q(\lambda, x) := \{v \in X \mid 0 \in DF(\lambda, x, 0)(0, v)\}$$

Then, for any equilibrium $x \in E(\lambda)$ and any closed cone P satisfying $P \cap Q(\lambda, x) = \{0\}$, there exists $\varepsilon > 0$ such that

$$E(\lambda) \cap (x + \varepsilon(P \cap B)) = \{x\}$$

and in particular, an equilibrium $x \in E(\lambda)$ is locally unique whenever 0 is the only equilibrium of $DF(\lambda, x, 0)(0, \cdot)$.

4. If $E(\lambda)$ is convex, then $E(\lambda) \subset x + Q(\lambda, x)$.

Proof — We observe at once that the graph of the equilibrium map E is the inverse image of the graph of F by the continuous linear operator $\pi \in \mathcal{L}(\Lambda \times X, \Lambda \times X \times X)$ defined by

$$\forall (\lambda, x) \in \Lambda \times X, \quad \pi(\lambda, x) := (\lambda, x, 0)$$

Because the circatangent derivative $CF(\lambda, x, 0)$ is surjective, the transversality condition holds true:

$$\mathrm{Im}(\pi) - C_{\mathrm{Graph}(F)}(\pi(\lambda, x)) = \Lambda \times X - \mathrm{Graph}(CF(\lambda, x, 0)) = \Lambda \times X \times X$$

Therefore Corollary 4.3.4 yields

$$\mathrm{Graph}(DE(\lambda, x)) = \pi^{-1}(\mathrm{Graph}(DF(\lambda, x, 0)))$$

and

$$\mathrm{Graph}(CE(\lambda, x)) \supset \pi^{-1}(\mathrm{Graph}(CF(\lambda, x, 0)))$$

which are the formulas we were looking for.

The second statement follows from Theorem 5.4.6 applied to E^{-1}.

We then apply Theorem 5.4.7 to the inverse of E to derive the localization properties of $E(\lambda)$. □

5.4.4 Local Injectivity

Definition 5.4.9 *Let X, Y be normed spaces and $F : X \rightsquigarrow Y$ be a set-valued map.*

We shall say that it is locally injective *around $x^\star$ if and only if there exists a neighborhood $N(x^\star)$ such that*

$$\forall\, x_1 \neq x_2 \in N(x^\star), \quad F(x_1) \cap F(x_2) = \emptyset$$

It is said to be (globally) injective *if we can take for neighborhood $N(x^\star)$ the whole domain of F.*

Theorem 5.4.10 *Let F be a set-valued map from a finite dimensional vector-space X to a normed space Y and $(x^\star, y^\star)$ belong to its graph. Assume that there exists $\gamma > 0$ such that $F(x^\star + \gamma B)$ is relatively compact and that F has a closed graph.*

If for all $y \in F(x^\star)$ the kernels of the paratingent derivatives $PF(x^\star, y)$ of F at $(x^\star, y)$ are equal to $\{0\}$, then F is locally injective around $x^\star$.

Proof — Assume that F is not locally injective. Then there exist sequences of elements $x_{1n} \neq x_{2n}$, converging to $x^\star$ and y_n satisfying

$$\forall\, n \geq 0, \quad y_n \in F(x_{1n}) \cap F(x_{2n})$$

Let us set $h_n := \|x_{1n} - x_{2n}\|$, which converges to 0, and

$$u_n := (x_{1n} - x_{2n})/h_n$$

The elements u_n do belong to the unit sphere, which is compact. Hence a subsequence (again denoted) u_n does converge to some u different from 0.

Then for all large n

$$y_n \in F(x_{1n}) \cap F(x_{2n}) = F(x_{2n} + h_n u_n) \cap F(x_{2n}) \subset F(x^\star + \gamma B)$$

we deduce that a subsequence (again denoted) y_n converges to some $y \in F(x^\star)$ (because Graph(F) is closed.)

Since the above relation implies

$$\forall\, n \geq 0, \quad y_n + h_n 0 \in F(x_{2n} + h_n u_n)$$

we deduce that

$$0 \in PF(x^\star, y)(u)$$

Hence we have proved the existence of a non zero element of the kernel of $PF(x^\star, y)$, which is a contradiction. □

5.5 Qualitative Solutions

This application deals with problems motivated by economics in the sixties (comparative statics) and, recently, by a domain of Artificial Intelligence known under the name of "qualitative simulation" or "qualitative physics."

It concerns "sign solutions" (or confluence solutions, to adopt the terminology used in Artificial Intelligence) of nonlinear equations and inclusions. We shall provide a criterion involving adequate linearizations of the sign problem, which allows to obtain sign solutions for *both the original problem and its linearization.*

Let us consider a closed set-valued map $F : \mathbf{R}^n \rightsquigarrow \mathbf{R}^m$ and the problem

$$\text{find } x \in \mathbf{R}^n \text{ such that } F(x) \ni y$$

We associate with $\mathbf{R}^n$ the n-dimensional **confluence space** $\mathcal{R}^n$ defined by

$$\mathcal{R}^n := \{ -, 0, + \}^n$$

whose elements are denoted by $a := (a_1, \ldots, a_n)$.

Given a multi-sign $b \in \mathcal{R}^m$, does there exist a "qualitative solution" $a \in \mathcal{R}^n$, defined by:

$$\exists\, x \;\&\; y \in F(x) \text{ such that } \operatorname{sign}(x) = a \;\&\; \operatorname{sign}(y) = b \tag{5.3}$$

We provide a criterion[4] for answering this question.

Before stating our first theorem, it is convenient to introduce the notations

$$\begin{cases} Q_n(a) := \mathbf{R}^n_a := \{v \in \mathbf{R}^n \mid \text{sign of } (v_i) = a_i\} \\ \overline{Q}_n(a) := a\mathbf{R}^n_+ := \{v \in \mathbf{R}^n \mid \text{sign of } (v_i) = a_i \text{ or } 0\} \end{cases}$$

[4] As in numerical analysis, which deals both with approximating infinite-dimensional problems by finite-dimensional ones and with solving them, problems of qualitative analysis arise at two levels: the passage from "quantitative to qualitative" and algorithms for solving the latter. Only the first aspect is investigated here in the framework of solving this inclusion.

The cone $Q_n^\star(a)$ is defined by:

$$v \in Q_n^\star(a) \iff \begin{cases} v_i \geq 0 & \text{if } a_i = + \\ v_i \leq 0 & \text{if } a_i = - \\ v_i \in \mathbf{R} & \text{if } a_i = 0 \end{cases}$$

Observe that $\overline{Q}_n(a)$ is the closure of $Q_n(a)$ and that the positive polar cone of $\overline{Q}_n(a)$ is the cone $Q_n^\star(a)$.

Theorem 5.5.1 *Let $a \in \mathcal{R}^n$, $b \in \mathcal{R}^m$, $F : \mathbf{R}^n \rightsquigarrow \mathbf{R}^m$ be a closed set-valued map. Consider $x_0 \in \mathrm{Dom}(F) \cap \overline{Q}_n(a)$ satisfying $F(x_0) \cap \overline{Q}_m(b) \neq \emptyset$. If the criterion*

$$\begin{cases} (0,0) \text{ is the only solution } (p,q) \text{ to } p \in CF(x_0,y_0)^\star(q) \\ \\ \text{satisfying } p \in Q_n^\star(a) \ \& \ q \in -Q_m^\star(b) \end{cases} \tag{5.4}$$

holds true, then a solves the qualitative inclusion

$$\exists\, x \in Q_n(a) \quad \textit{such that} \quad F(x) \cap Q_m(b) \neq \emptyset \tag{5.5}$$

Proof — Pick $y_0 \in F(x_0) \cap \overline{Q}_m(b)$, then (x_0,y_0) belongs to the intersection of the graph of F and the closed convex cone $\overline{Q}_n(a) \times \overline{Q}_m(b)$, so that $((x_0,y_0),(x_0,y_0))$ is a solution in $\mathrm{Graph}(F) \times (\overline{Q}_n(a) \times \overline{Q}_m(b))$ to the equation

$$(x_1,y_1) - (x_2,y_2) = 0$$

Let $\mathbf{1} \in \mathbf{R}^n$ denote the unit vector $\mathbf{1} := (1,\ldots,1)$ and set

$$a\mathbf{1} := (a_i 1)_{i=1,\ldots n}$$

We shall prove that for some $\varepsilon > 0$ there exists a solution

$$(x_1,y_1) \in \mathrm{Graph}(F), \quad (x_2,y_2) \in \overline{Q}_n(a) \times \overline{Q}_m(b)$$

to the equation

$$(x_1,y_1) - (x_2,y_2) = \varepsilon(a\mathbf{1}, b\mathbf{1})$$

So that $x_1 = x_2 + \varepsilon a\mathbf{1}$ belongs to $Q_n(a)$ and $y_1 = y_2 + \varepsilon b\mathbf{1}$ belongs to $Q_m(b)$ and to $F(x_1)$.

For that purpose, we apply Theorem 5.4.6, which states that a solution to the above equation does exist provided the assumption

$$C_{\mathrm{Graph}(F)}(x_0,y_0) - C_{\overline{Q}_n(a)}(x_0) \times C_{\overline{Q}_m(b)}(y_0) = \mathbf{R}^n \times \mathbf{R}^m$$

is satisfied.

Since x_0 belongs to $\overline{Q}_n(a) := a\mathcal{R}^n_+$, the cone $C_{\overline{Q}_n(a)}(x_0)$ coincides with the tangent cone, which contains $Q_n(a)$ (recall that when Q is a convex cone, $T_Q(x) = \mathrm{cl}(Q + x\mathbf{R}) \supset Q$.) In the same way, $C_{\overline{Q}_m(b)}(y_0) \supset \overline{Q}_m(b)$.

The above assumption would follow from

$$\mathrm{Graph}(CF(x_0, y_0)) - (\overline{Q}_n(a) \times \overline{Q}_m(b)) \;=\; \mathbf{R}^n \times \mathbf{R}^m$$

By polarity, this is equivalent to the condition

$$(\mathrm{Graph}(CF(x_0, y_0)))^- \cap (\overline{Q}_n(a) \times \overline{Q}_m(b))^+ \;=\; \{0\}$$

which is nothing other than condition (5.4.) □

In the single-valued case, we obtain:

Corollary 5.5.2 *Let $a \in \mathcal{R}^n$, $b \in \mathcal{R}^m$, $F : \mathbf{R}^n \rightsquigarrow \mathbf{R}^m$ be a closed set-valued map. Let $x_0 \in \overline{Q}_n(a)$ be such that $f(x_0) \in \overline{Q}_n(b)$ and assume that f is continuously differentiable at x_0. Then the criterion*

$$0 \text{ is the only solution } q \text{ to } q \;\in\; -Q^\star_m(b) \;\;\&\;\; f'(x_0)^\star q \;\in\; Q^\star_n(a)$$

implies that a solves the qualitative equation

$$\exists\, x \text{ such that } \mathrm{sign}(x) \;=\; a \quad \& \quad \mathrm{sign}(f(x)) \;=\; b$$

Remark — When A is a closed convex process, we can apply the above theorem with

$$(x_0, y_0) \;=\; (0,0) \;\in\; \mathrm{Graph}(A) \cap (\overline{Q}_n(a) \times \overline{Q}_m(b))$$

Since $CA(0,0) = A$ (because the tangent cone to the closed convex cone $\mathrm{Graph}(A)$ at the origin is equal to this closed convex cone), we deduce from the above theorem the following consequence:

Corollary 5.5.3 *Let $a \in \mathcal{R}^n$, $b \in \mathcal{R}^m$ and A be a closed convex process*[5] *from $\mathbf{R}^n$ to $\mathbf{R}^m$. If*

$$\begin{cases} (0,0) \text{ is the only solution } (p,q) \text{ to } p \;\in\; A^\star(q) \\[2ex] \text{such that } p \;\in\; Q^\star_n(a) \;\;\&\;\; q \;\in\; -Q^\star_m(b) \end{cases}$$

then a solves the qualitative inclusion

$$\exists\, x \;\in\; Q_n(a) \quad \textit{such that } A(x) \cap Q_m(b) \;\neq\; \emptyset$$

[5]Closed convex processes provide a way to represent uncertainty for linear operators analogous to "sign matrices." They possess better mathematical properties than sign matrices. In particular, since we cannot extend "sign matrix" to nonlinear maps, we can not say that a sign matrix is the derivative of a sign operator. Observe that this corollary provides the same criterion for *both the initial inclusion and its set-valued linearization.*

Therefore, criterion (5.4) implies the existence of a qualitative solution to both problem (5.5) and the linearized problem

$$\exists\, u \in Q_n(a) \text{ such that } CF(x_0, y_0)(u) \cap Q_m(b) \neq \emptyset \quad \square$$

Let us consider the case when $F : \mathbf{R}^n \rightsquigarrow \mathbf{R}^m$ is the set-valued map defined by:

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

where $f : \mathbf{R}^n \mapsto \mathbf{R}^m$ is continuous, L is a closed subset of $\mathbf{R}^n$ and $M : \mathbf{R}^n \rightsquigarrow \mathbf{R}^m$ is a closed set-valued map.

Corollary 5.5.4 *Let $a \in \mathcal{R}^n$, $b \in \mathcal{R}^m$, $M : \mathbf{R}^n \rightsquigarrow \mathbf{R}^m$ be a closed set-valued map and $L \subset \mathbf{R}^n$ be a closed set. Assume that x_0 and y_0 satisfy*

$$x_0 \in \overline{Q}_n(a) \cap L \ \& \ y_0 \in (f(x_0) - M(x_0)) \cap \overline{Q}_m(b)$$

that f is continuously differentiable at x_0, M is Lipschitz at x_0 and that

$$C_L(x_0) - \mathrm{Dom}(CM(x_0, y_0 - f(x_0))) = \mathbf{R}^n$$

The criterion

$$\begin{cases} (0,0) \text{ is the only solution } (p,q) \text{ to} \\ \\ p \in f'(x_0)^\star(q) + CM(x_0, y_0 - f(x_0))^\star(-q) + N_L(x_0) \\ \\ \text{such that } p \in Q_n^\star(a) \quad \& \quad q \in -Q_m^\star(b) \end{cases} \tag{5.6}$$

implies that a solves the qualitative equation

$$\exists\, x \in Q_n(a) \cap L \quad \& \quad y \in Q_m(b) \text{ such that } y \in f(x) - M(x)$$

It follows from Theorem 5.5.1 and Proposition 5.2.3.

The particular case when M is constant yields

Corollary 5.5.5 *Let $a \in \mathcal{R}^n$, $b \in \mathcal{R}^m$ and $L \subset \mathbf{R}^n$, $M \subset \mathbf{R}^m$ be closed sets. Assume that x_0 and y_0 satisfy*

$$x_0 \in \overline{Q}_n(a) \cap L \quad \& \quad y_0 \in (f(x_0) - M) \cap \overline{Q}_m(b)$$

Assume that f is continuously differentiable at x_0. Then the criterion

$$\begin{cases} 0 \text{ is the only solution } q \text{ to} \\ \\ q \in -Q_m^\star(b) \cap N_M(f(x_0) - y_0) \quad \& \quad f'(x_0)^\star q \in Q_n^\star(a) - N_L(x_0) \end{cases}$$

implies that a solves the qualitative equation

$$\exists\, x \in Q_n(a) \cap L \ \& \ y \in Q_m(b) \text{ such that } f(x) \in y + M$$

Another example is provided by qualitative solutions to *variational inequalities.*

Example Variational Inequalities

Let us consider a closed convex subset $K \subset \mathbf{R}^n$ and a $\mathcal{C}^1$-map f from a neighborhood of K to $\mathbf{R}^n$.

An element $y_0 \in \mathbf{R}^n$ being given, we recall that an element $x_0 \in K$ is a solution to the variational inequalities

$$\forall\, x \in K, \quad < f(x_0) - y_0,\ x_0 - x > \leq 0$$

if and only if x_0 is a solution to the inclusion

$$y_0 \in f(x_0) + N_K(x_0)$$

where $N_K(x_0)$ is the normal cone to K at x_0.

Theorem 5.5.6 *Let $a \in \mathcal{R}^n$, $b \in \mathcal{R}^m$ and let $x_0 \in \overline{Q}_n(a) \cap K$, $y_0 \in \overline{Q}_m(b)$ satisfy*

$$\forall\, x \in K, \quad < f(x_0) - y_0, x_0 - x > \leq 0$$

The criterion

$$\left\{ \begin{array}{l} (0,0) \text{ is the only solution } (p,q) \text{ to} \\ p \in f'(x_0)^\star(q) + CN_K(x_0, y_0 - f(x_0))^\star(-q) \\ \text{satisfying } p \in Q_n^\star(a) \ \&\ q \in -Q_m^\star(b) \end{array} \right. \tag{5.7}$$

implies that a solves the qualitative variational inequalities

$$\left\{ \begin{array}{l} \exists\, \overline{x} \in Q_n(a) \cap K \ \&\ \overline{y} \in Q_m(b) \text{ such that} \\ \quad \forall\, x \in K, \ < f(\overline{x}) - \overline{y}, \overline{x} - x > \leq 0 \end{array} \right.$$

Proof — This a consequence of Proposition 5.1.2 with $M(x) := -N_K(x)$ and Theorem 5.5.1. □

Remark — Recall that Proposition 5.2.7 provides a characterization of the circatangent derivative of the normal cone map $K \ni x \rightsquigarrow N_K(x)$ through the circatangent derivative of the best approximation projection π_K to the closed convex subset K:

$$q \in CN_K(x,p)(u) \iff u \in C\pi_K(x+p)(u+q)$$

which can be used to check criterion (5.7.)

5.6 Higher Order Derivatives

Naturally, by taking m^{th}-order contingent sets to a graph of a set-valued map $F : X \rightsquigarrow Y$, we obtain concepts of high-order derivatives.

Definition 5.6.1 *Let X, Y be normed spaces and $F : X \rightsquigarrow Y$ be a set-valued map.*

The m^{th}-order contingent derivative $DF(x,y)$ of F at $(x,y) \in \mathrm{Graph}(F)$ is the set-valued map from X to Y defined by

$$\begin{cases} \mathrm{Graph}(D^{(m)}F(x,y,u_1,v_1,\ldots,u_{m-1},v_{m-1})) \\ \\ := T^{(m)}_{\mathrm{Graph}(F)}(x,y,u_1,v_1,\ldots,u_{m-1},v_{m-1}) \end{cases}$$

We say that it is $m-th$ derivable if its graph is $m-th$ derivable.

This definition implies right away that

$$\begin{cases} D^{(m)}(F^{-1})(y,x,v_1,u_1,\ldots,v_{m-1},u_{m-1}) \\ \\ = (D^{(m)}F(x,y,u_1,v_1,\ldots,u_{m-1},v_{m-1}))^{-1} \end{cases}$$

For simplicity, we shall restrict ourselves to second order contingent derivatives of set-valued maps.

If K is a subset of X and f is a single-valued map which is twice continuously differentiable around a point $x \in K$, then *the second order contingent derivative of the restriction of f to K is given by the formula*:

$$\begin{cases} D^{(2)}(f|_K)(x,u_1)(u_2) = D^{(2)}(f|_K)(x,f(x),u_1,f'(x)u_1)(u_2) \\ \\ = f'(x)u_2 + \frac{1}{2}f''(x)(u_1,u_1) \ \text{ whenever } \ u_2 \in T^{(2)}_K(x,u_1) \end{cases}$$

It is empty when $u_2 \notin T^{(2)}_K(x,u_1)$.

The proof of this fact is similar to the proof of the following proposition.

Proposition 5.6.2 *Let X and Y be normed spaces, f be a twice continuously differentiable single-valued map from an open subset $\Omega \subset X$ to Y, $M : X \rightsquigarrow Y$ be a set-valued map and $L \subset X$. Let $F : X \rightsquigarrow Y$ be the set-valued map defined by:*

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

Then, for each $x \in \Omega \cap \mathrm{Dom}(F)$, $y \in F(x)$ and $v_1 \in DF(x,y)(u_1)$, the second order derivative $D^{(2)}F(x,y,u_1,v_1)(u_2)$ in the direction $u_2 \in T_L^{(2)}(x,u_1)$ is contained in

$$f'(x)u_2 \;+\; \frac{1}{2}f''(x)(u_1,u_1) \;-\; D^{(2)}M(x,f(x)-y,u_1,f'(x)u_1-v_1)(u_2)$$

Furthermore $D^{(2)}F(x,y,u_1,v_1)(u_2)$ is empty whenever $u_2 \notin T_L^{(2)}(x,u_1)$. Equality holds true when we assume that either L or M is twice derivable and M is Lipschitz.

Proof — Let v_2 belong to $D^{(2)}F(x,y,u_1,v_1)(u_2)$. Then there exist $h_n > 0$ converging to 0 and sequences u_{2n} and v_{2n} converging to u_2 and v_2 respectively such that $x + h_n u_1 + h_n^2 u_{2n}$ belongs to L and

$$\forall\, n \ge 0,\;\; y + h_n v_1 + h_n^2 v_{2n} \;\in\; f(x + h_n u_1 + h_n^2 u_{2n}) - M(x + h_n u_1 + h_n^2 u_{2n})$$

This implies that u_2 belongs to $T_L^2(x,u_1)$. Since f is twice continuously differentiable

$$\begin{cases} f(x + h_n u_1 + h_n^2 u_{2n}) \\ = f(x) \;+\; h_n f'(x)u_1 \;+\; h_n^2(f'(x)u_2 \;+\; \frac{1}{2}f''(x)(u_1,u_1)) \;+\; \varepsilon(h_n)) \end{cases}$$

where $\varepsilon(h_n)$ converges to 0 with h_n. Then

$$\begin{cases} f(x) - y + h_n(f'(x)u_1 - v_1) + h_n^2(f'(x)u_2 + \frac{1}{2}f''(x)(u_1,u_1) - v_2 + \varepsilon(h_n)) \\ \in\; M(x + h_n u_1 + h_n^2 u_{2n}) \end{cases}$$

and thus $f'(x)u_2 + \frac{1}{2}f''(x)(u_1,u_1) - v_2$ does belong to

$$D^{(2)}M(x,f(x)-y,u_1,f'(x)u_1-v_1)(u_2)$$

Conversely, assume for instance that M is twice derivable, that u_2 belongs to $T_L^2(x,u_1)$ and that

$$f'(x)u_2 + \frac{1}{2}f''(x)(u_1,u_1) - v_2 \;\in\; D^{(2)}M(x,f(x)-y,u_1,f'(x)u_1-v_1)(u_2)$$

Hence, there exist a sequence $h_n > 0$ converging to 0 and sequences u_{2n} and w_{2n} converging to u_2 and $f'(x)u_2 + \frac{1}{2}f''(x)(u_1,u_1) - v_2$ respectively such that

$$x + h_n u_1 + h_n^2 u_{2n} \;\in\; L$$

and

$$f(x) - y + h_n(f'(x)u_1 - v_1) + h_n^2 w_{2n} \;\in\; M(x + h_n u_1 + h_n^2 u_{2n})$$

(we have used the Lipschitz continuity of M). Then for some $\varepsilon(h_n) \to 0$, the sequence

$$v_{2n} := f'(x)u_2 + \frac{1}{2}f''(x)(u_1, u_1) + \varepsilon(h_n) - w_{2n}$$

converges to v_2 and satisfies

$$y + h_n v_1 + h_n^2 v_{2n} \in f(x + h_n u_1 + h_n^2 u_{2n}) - M(x + h_n u_1 + h_n^2 u_{2n}) \quad \square$$

Chapter 6

Epiderivatives of Extended Functions

Introduction

When we consider real-valued functions, the order relation on $\mathbf{R}$ is involved in optimization (for finding a minimum) or in Lyapunov functions (which decrease along solutions to differential equations or inclusions.) These are reasons why we have to associate with a real-valued function V the set-valued map $\mathbf{V}_{\uparrow}$ defined by

$$\mathbf{V}_{\uparrow}(x) \ := \ V(x) + \mathbf{R}_{+}$$

whose graph is the epigraph of V and consider not only the contingent derivative of V, but also the contingent derivative of $\mathbf{V}_{\uparrow}$. We shall check in Section 1 that its graph is the epigraph of a lower semicontinuous function, which is called the *contingent epiderivative* $D_{\uparrow}V(x)$ of V at x. For instance, the Fermat rule[1] states:

if x minimizes V, then $D_{\uparrow}V(x)(v) \geq 0$ for all $v \in X$,

[1]which should be called *Oresme-Fermat Rule*. Nicholas Oresme (1323-1382), Bishop of Lisieux, 300 years ahead of his time, anticipated the great Fermat himself by observing in his treatises *De Proportionibus Proportionum* and *Algorismus Proportionum* that *near a maximum, the increment of a variable quantity becomes* 0. (He also anticipated the coordinate system in another book *Tractatus de Latitudine Formarum*, Copernicus' views and was only one step to the logarithms.) He did not hesitate to attack repeatedly the extravagant presumptions of astrologers (called at the time mathematicians!..., the mathematicians were then

so that necessary conditions in optimization and optimal control can be derived from this simple rule, provided that a rich and exhaustive calculus of contingent epiderivative is available.

Section 3 is thus devoted to such an epidifferential calculus.

As for derivatives of set-valued maps, an epidifferential calculus involves not only the use of contingent epiderivatives, but of other epiderivatives[2] corresponding to various tangent cones. So, adjacent $D^{\flat}_{\uparrow}V(x)$ and circatangent $C_{\uparrow}V(x)$ epiderivatives are presented in the second section, as well as convex epiderivatives.

We study in Section 4 dual concepts: the *subdifferential* $\partial_0 V(x)$ and the *generalized gradient* $\partial V(x)$ of an extended function defined respectively by

$$\begin{cases} \partial_0 V(x) & := & \{p \in X^\star \mid \forall\, v \in X, \;\; < p, v > \leq \; D_{\uparrow}V(x)(v)\} \\ \partial V(x) & := & \{p \in X^\star \mid \forall\, v \in X, \;\; < p, v > \leq \; C_{\uparrow}V(x)(v)\} \end{cases}$$

They naturally coincide when V is convex or when it is continuously differentiable (in this case, they are made up of the gradient $V'(x)$.)

The subdifferential is always contained in the generalized gradient, and even more, in the case of finite dimensional vector-spaces, the upper limit of the subdifferentials $\partial_0 V(x_n)$ when x_n converges to x is contained in the generalized gradient $\partial V(x)$.

Again in finite dimensional vector-spaces, the subdifferential coincides with the *(local) subdifferential* $\partial_- V(x)$ (used in the concept of viscosity solutions of Hamilton-Jacobi equations), which is the subset of $p \in X^\star$ such that

$$\liminf_{x' \to x} \frac{V(x') - V(x) - < p, x' - x >}{\|x' - x\|} \geq 0$$

This still increases the population of our ménagerie with species doomed to disappear through Darwinian evolution for lack of duality, since their use does not allow recovery of the original information on epiderivatives.

---

called geometers), whose persuasive arguments influenced affairs of the state, and he even criticized the Pope. A clear thinker, he translated Aristotle's books into French and commented on them with extremely original views.

[2]This procedure is also known under the name of the *epigraphical approach.*

The use of generalized gradients for functions is then recommended in a "convex environment, i.e., for convex, and more generally, episleek functions (with convex epigraphs.)

However, there are other reasons than the exploitation of duality for using these dual concepts.

One is their familiarity with more classical concepts. For usual functions on Hilbert spaces, there is a canonical identification between, say, a derivative of a differentiable function and its gradient, and it became traditional to formulate many results in terms of gradients and normal cones. Think of the traditional role of the *Lagrange multipliers* in constrained optimization, for instance, which appeals to generalizations stated in a dual form (Kuhn-Tucker rule, maximum principles in optimal control.)

Another reason lies in the fact that the first attempts to generalize the concept of gradients used limiting procedures. Since it seems easier to take limits of elements (the gradients, for instance) than limits of functionals (the associated directional derivatives, for example), many generalizations of the concept of gradient dealt with set of limits or cluster points of gradients taken in various ways. These approaches are still pursued (in the framework of "proximal analysis", for instance), but we do not dwell on them in this book.

When we deal with convex functions, which are the extended functions with convex epigraphs, all these epiderivatives coincide. Furthermore, lower semicontinuous convex functions enjoy duality properties: We recall in Section 5 what the Fenchel conjugate of a real-valued function is and that this Fenchel conjugacy is a bijective correspondence between lower semicontinuous convex functions on a Banach space and its dual.

It plays a role similar to the one played by the *Legendre transform*, since *the inverse of the subdifferential map of a convex function is the subdifferential map of the conjugate function.*

These facts are at the origin of *duality theory in optimization.* We also recall here the Fenchel Theorem which provides a criterion for the existence of solutions to convex minimization problems, from which one can derive a formula for computing the conjugate of the composition of a function with a continuous linear operator.

We finally point out that the subdifferential map of a convex function is a maximal monotone map. Its Yosida approximation is the gradient of a continuously differentiable convex function, called the Moreau-Yosida approximation.

We conclude this chapter with a presentation of high order epiderivatives, defined from the high-order derivatives of the set-valued map $\mathbf{V}_\uparrow$ associated with an extended function. We shall also estimate the second-order contingent epiderivative of the Moreau-Yosida approximation of a convex function.

6.1 Contingent Epiderivatives

6.1.1 Extended Functions and their Epigraphs

Let us consider an extended function $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ whose domain

$$\mathrm{Dom}(V) := \{x \in X \mid V(x) \neq \pm\infty\}$$

is not empty. (Such a function is said to be *proper* in convex and non smooth analysis. We shall rather say that it is *nontrivial* in order to avoid confusion with proper maps.)

Any function V defined on a subset $K \subset X$ can be regarded as the extended function V_K equal to V on K and to $+\infty$ outside of K, whose domain is K.

An extended function is characterized by its *epigraph*

$$\mathcal{E}p(V) := \{(x,\lambda) \in X \times \mathbf{R} \mid V(x) \leq \lambda\}$$

An extended function V is convex (resp. positively homogeneous) if and only if its epigraph is convex (resp. a cone.)

We also observe that any positively homogeneous extended function is non trivial whenever $V(0) \neq -\infty$. In this case, $V(0) = 0$.

The *hypograph* of a function $W : X \to \mathbf{R} \cup \{\pm\infty\}$ is defined in a symmetric way:

$$\mathcal{H}p(W) := \{(x,\lambda) \in X \times \mathbf{R} \mid W(x) \geq \lambda\} = -\mathcal{E}p(-W)$$

Such a function is concave if and only if $-W$ is convex, i.e., if and only if its hypograph is convex.

The main examples of extended functions are the *indicators* ψ_K *of subsets* K defined by

$$\psi_K(x) := \begin{cases} 0 & \text{if } x \in K \\ +\infty & \text{if not} \end{cases}$$

The indicator ψ_K is lower semicontinuous if and only if K is closed and ψ_K is convex if and only if K is convex. One can regard the sum $V + \psi_K$ as the restriction of V to K.

We recall the convention $\inf(\emptyset) := +\infty$.

Lemma 6.1.1 *Consider a function* $\delta : X \mapsto \mathbf{R} \cup \{\pm\infty\}$. *Its epigraph is closed if and only if*

$$\forall\, v \in X, \;\; \delta(v) = \liminf_{v' \to v} \delta(v')$$

Assume that the epigraph of δ *is a closed cone. Then the following conditions are equivalent:*

$$\begin{cases} i) & \forall\, v \in X, \;\; \delta(v) > -\infty \\ ii) & \delta(0) = 0 \\ iii) & (0,-1) \notin \mathcal{E}p(\delta) \end{cases}$$

Proof — Assume that the epigraph of δ is closed and pick $v \in X$. There exists a sequence of elements v_n converging to v such that

$$\lim_{n \to \infty} \delta(v_n) = \liminf_{v' \to v} \delta(v')$$

Hence, for any $\lambda > \liminf_{v' \to v} \delta(v')$, there exist N such that, for all $n \geq N$, $\delta(v_n) \leq \lambda$, i.e., such that $(v_n, \lambda) \in \mathcal{E}p(\delta)$. By taking the limit, we infer that $\delta(v) \leq \lambda$, and thus, that $\delta(v) \leq \liminf_{v' \to v} \delta(v')$. The converse statement is obvious.

Suppose next that the epigraph of δ is a cone. Then it contains $(0,0)$ and $\delta(0) \leq 0$. The statements *ii*) and *iii*) are clearly equivalent.

If *i*) holds true and $\delta(0) < 0$, then

$$(0,-1) = \frac{1}{-\delta(0)}(0, \delta(0))$$

belongs to the epigraph of δ, as well as all $(0,-\lambda)$, and (by letting $\lambda \to +\infty$) we deduce that $\delta(0) = -\infty$, so that *i*) implies *ii*).

To end the proof, assume that $\delta(0) = 0$ and that for some v, $\delta(v) = -\infty$. Then, for any $\varepsilon > 0$, the pair $(v, -1/\varepsilon)$ belongs to the epigraph of δ, as well as the pairs $(\varepsilon v, -1)$. By letting ε converge to 0, we infer that $(0, -1)$ belongs also to the epigraph, since it is closed. Hence $\delta(0) < 0$, a contradiction. □

6.1.2 Contingent Epiderivatives

We can also regard an extended function V as the set-valued map $\mathbf{V} : X \rightsquigarrow \mathbf{R}$ defined by

$$\mathbf{V}(x) := \begin{cases} V(x) & \text{if } x \in \mathrm{Dom}(V) \\ \emptyset & \text{if } x \notin \mathrm{Dom}(V) \end{cases}$$

so that we can define in the usual way the contingent derivatives of $\mathbf{V}$ at $x \in \mathrm{Dom}(\mathbf{V})$:

$$\begin{cases} D\mathbf{V}(x)(u) := D\mathbf{V}(x, V(x))(u) \\ \\ = \{v \mid \liminf_{h \to 0+, u' \to u} |V(x + hu') - V(x) - hv|/h = 0\} \end{cases}$$

However, minimization problems and Lyapunov functions involve obviously the order relation on $\mathbf{R}$. Hence, when dealing with such problems, it is convenient to consider two new set-valued maps $\mathbf{V}_\uparrow$ and $\mathbf{V}_\downarrow$ defined in the following way:

$$\begin{cases} i) \quad \mathbf{V}_\uparrow(x) := \begin{cases} V(x) + \mathbf{R}_+ & \text{if } x \in \mathrm{Dom}(V) \\ \emptyset & \text{if } V(x) = +\infty \\ \mathbf{R} & \text{if } V(x) = -\infty \end{cases} \\ ii) \quad \mathbf{V}_\downarrow(x) := \begin{cases} V(x) - \mathbf{R}_+ & \text{if } x \in \mathrm{Dom}(V) \\ \emptyset & \text{if } V(x) = -\infty \\ \mathbf{R} & \text{if } V(x) = +\infty \end{cases} \end{cases}$$

We see at once that

$$\mathrm{Graph}(\mathbf{V}_\uparrow) = \mathcal{E}p(V) \quad \& \quad \mathrm{Graph}(\mathbf{V}_\downarrow) = \mathcal{H}p(V)$$

Therefore, we are led naturally to associate with these two set-valued maps $\mathbf{V}_\uparrow$ and $\mathbf{V}_\downarrow$ their contingent derivatives. We observe that the values of the contingent derivative of $\mathbf{V}_\uparrow$ are obviously half lines in the sense that

$$\begin{cases} \forall \lambda \geq V(x), \forall u \in \mathrm{Dom}(D\mathbf{V}_\uparrow(x,\lambda)), \\ D\mathbf{V}_\uparrow(x,\lambda)(u) = D\mathbf{V}_\uparrow(x,\lambda)(u) + \mathbf{R}_+ \end{cases}$$

Definition 6.1.2 *Let $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ be a nontrivial extended function and x belong to its domain. We shall say that the function $D_\uparrow V(x)$ from X to $\mathbf{R} \cup \{\pm\infty\}$ defined by*

$$\forall u \in X, \ D_\uparrow V(x)(u) := \inf\{v \mid v \in D\mathbf{V}_\uparrow(x, V(x))(u)\}$$

is the contingent epiderivative *of V at x in the direction u.*

The function V is said to be contingently epidifferentiable *at x if its contingent epiderivative never takes the value $-\infty$.*

It is said to be episleek *(at x) if its epigraph is sleek (at $(x, V(x))$).*

Actually, the contingent epiderivative can be characterized as a limit of differential quotients:

Proposition 6.1.3 *Let $V : X \mapsto \mathbf{R}\cup\{\pm\infty\}$ be a nontrivial extended function and x belong to its domain. Then*

$$D_\uparrow V(x)(u) = \liminf_{h \to 0+, u' \to u} \frac{V(x+hu') - V(x)}{h}$$

The function V is contingently epidifferentiable at x if and only if $D_\uparrow V(x)(0) = 0$.

Proof — Assume first that $D_\uparrow V(x)(u) < +\infty$. By the very definition of the contingent epiderivative, for any $\lambda > D_\uparrow V(x)(u)$, there exists some $v \in D\mathbf{V}_\uparrow(x, V(x))(u)$ smaller than λ. Therefore, we can associate with any $\varepsilon > 0$ elements $h \in]0, \varepsilon[$ and $(u', v') \in B((u, v), \varepsilon)$ satisfying $V(x) + hv' \geq V(x + hu')$. So that

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

Hence the liminf of the differential quotients $(V(x + hu') - V(x))/h$ is smaller or equal to $D_\uparrow V(x)(u)$.

Conversely, let $h_n > 0$ converging to 0, u_n converging to u and $\lambda_n \geq V(x + h_n u_n)$ be such that the differential quotients

$$v_n := (\lambda_n - V(x))/h_n$$

converge to

$$v := \liminf_{h \to 0+,\, u' \to u} \frac{V(x + hu') - V(x)}{h}$$

Then the pairs $(x + h_n u_n, V(x) + h_n v_n)$ belong to the graph of $\mathbf{V}_\uparrow$, so that $v \in D\mathbf{V}_\uparrow(x)(u)$. Hence

$$\liminf_{h \to 0+, u' \to u} \frac{V(x + hu') - V(x)}{h} \geq D_\uparrow V(x)(u)$$

Consider now the case when $D_\uparrow V(x)(u) = +\infty$, i.e., the case when the subset $D\mathbf{V}_\uparrow(x, V(x))(u) = \emptyset$. Then for any $v \in \mathbf{R}$, for any $h_n > 0$ converging to 0 and any sequence u_n converging to u, we have

$$V(x + h_n u_n) \geq V(x) + h_n v$$

for n large enough, so that the liminf of the differential quotients is equal to $+\infty$. $\square$

We define in a symmetric way the *contingent hypoderivative* $D_\downarrow V(x)$ from X to $\mathbf{R} \cup \{\pm\infty\}$ of $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ at a point x of its domain by

$$D_\downarrow V(x)(u) := -D_\uparrow(-V)(x)(u) = \limsup_{h \to 0+, u' \to u} \frac{V(x + hu') - V(x)}{h}$$

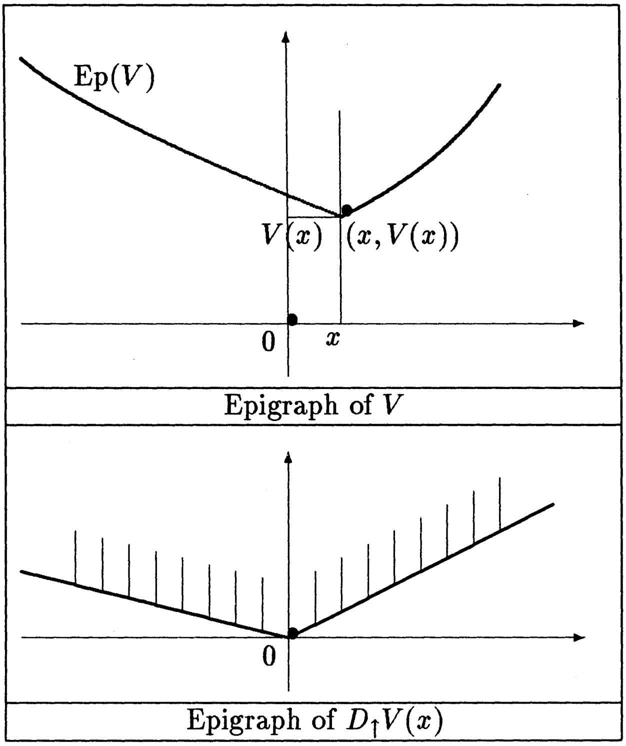
We could also have defined the contingent epiderivative of a function by taking the contingent cone to its epigraph:

Proposition 6.1.4 *Let $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ be a nontrivial extended function and x belong to its domain.*

Then the contingent cone to the epigraph of V at $(x, V(x))$ is the epigraph of the contingent epiderivative of V at x:

$$\mathcal{E}p(D_\uparrow V(x)) = T_{\mathcal{E}p(V)}(x, V(x))$$

Figure 6.1: Epigraph of the Contingent Derivative



Epigraph of V

Epigraph of $D_{\uparrow}V(x)$

Consequently, the epigraph of the contingent epiderivative at x is a closed cone. The contingent epiderivative $D_\uparrow V(x)$ is then lower semicontinuous and positively homogeneous whenever V is contingently epidifferentiable at x.

Furthermore,

$$\forall\, w \ \geq\ V(x),\ \ T_{\mathcal{E}p(V)}(x,w)\ \subset\ T_{\mathrm{Dom}(V)}(x)\times\mathbf{R}$$

and

$$\forall\, w\ >\ V(x),\ \ \mathrm{Dom}(D_\uparrow V(x))\times\mathbf{R}\ \subset\ T_{\mathcal{E}p(V)}(x,w)$$

If the restriction of V to its domain is upper semicontinuous, then, for all $w > V(x)$,

$$T_{\mathcal{E}p(V)}(x,w)\ =\ T_{\mathrm{Dom}(V)}(x)\times\mathbf{R}$$

Proof

1. — The first claim follows from the very definition of the contingent epiderivative. Fix $w \geq V(x)$. Let us assume that (u,v) belongs to $T_{\mathcal{E}p(V)}(x,w)$. We infer that there exist sequences u_n, v_n and $h_n > 0$ converging to u, v and 0 such that

$$w + h_n v_n\ \geq\ V(x + h_n u_n)$$

We thus deduce that u belongs to the contingent cone to the domain of V at x, and thus, that $T_{\mathcal{E}p(V)}(x,w) \subset T_{\mathrm{Dom}(V)}(x)\times\mathbf{R}$.

2. — If u belongs to the domain of the contingent epiderivative of V at x, if $w > V(x)$ and if v is any real number, we check that (u,v) belongs to $T_{\mathcal{E}p(V)}(x,w)$.

Indeed, there exist sequences of elements $h_n > 0$, u_n and v_n converging to 0, u and $D_\uparrow V(x)(u)$ respectively such that

$$(x + h_n u_n, V(x) + h_n v_n)\ \in\ \mathcal{E}p(V)$$

But we can write

$$(x+h_nu_n, w+h_nv) = (x+h_nu_n, V(x)+h_nv_n)+(0, w-V(x)+h_n(v-v_n))$$

Since $w - V(x) + h_n(v - v_n)$ is strictly positive when h_n is small enough, we infer that $(x + h_nu_n, w + h_nv)$ belongs to the epigraph of V, i.e., that (u,v) belongs to the cone $T_{\mathcal{E}p(V)}(x,w)$.

3. — Let w be strictly larger than $V(x)$ and u belong to $T_{\mathrm{Dom}(V)}(x)$. Then there exist sequences u_n and $h_n > 0$ converging to u and 0 such that $V(x + h_n u_n) < +\infty$ for all n.

When V is upper semicontinuous on its domain, for all

$$0 < \varepsilon < (w - V(x))/2$$

there exists $\eta > 0$ such that, for all $h_n \|u_n\| < \eta$, we obtain

$$V(x + h_n u_n) \leq V(x) + \varepsilon < w - \varepsilon$$

Let v be given arbitrarily in $\mathbf{R}$. Then, for any $h_n > 0$ when $v \geq 0$ or for any $h_n \in]0, \frac{\varepsilon}{-v}[$ when $v < 0$, inequality $w - \varepsilon \leq w + h_n v$ implies that $V(x + h_n u_n) \leq w + h_n v$, i.e., that the pair (u, v) belongs to $T_{\mathcal{E}p(V)}(x, w)$. $\square$

Proposition 6.1.5 *Let $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ be an extended function and x belong to its domain. Take any*

$$u \in \mathrm{Dom}(D_\uparrow V(x)) \cap \mathrm{Dom}(D_\downarrow V(x))$$

Then

$$\{D_\uparrow V(x)(u), D_\downarrow V(x)(u)\} \subset D\mathbf{V}(x)(u) \subset [D_\uparrow V(x)(u), D_\downarrow V(x)(u)]$$

Proof — Since $\mathrm{Graph}(V) = \mathcal{E}p(V) \cap \mathcal{H}p(V)$, we deduce that the inclusions

$$T_{\mathrm{Graph}(V)}(x, V(x)) \subset T_{\mathcal{E}p(V)}(x, V(x)) \cap T_{\mathcal{H}p(V)}(x, V(x))$$

can be translated into

$$\mathrm{Graph}(D\mathbf{V}(x)) \subset \mathcal{E}p(D_\uparrow V(x)) \cap \mathcal{H}p(D_\downarrow V(x))$$

from which the inclusion

$$D\mathbf{V}(x)(u) \subset [D_\uparrow V(x)(u), D_\downarrow V(x)(u)]$$

follows. Since the contingent epiderivative of V at x in the direction u is equal to

$$D_\uparrow V(x)(u) := \liminf_{h \to 0+, u' \to u} \frac{V(x + hu') - V(x)}{h}$$

we see that $D_{\uparrow}V(x)(u)$ is the limit of a subsequence of differential quotients $\left(\frac{V(x+hu')-V(x)}{h}\right)$ when $h \to 0+$ and thus, that $D_{\uparrow}V(x)(u)$ belongs to $D\mathbf{V}(x)(u)$. The same is true for the contingent hypoderivative. □

Theorem 6.1.6 *Let $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ be an extended function which is continuous on a neighborhood of $x \in \mathrm{Dom}(V)$. Then for every*

$$u \in \mathrm{Dom}(D_{\uparrow}V(x)) \cap \mathrm{Dom}(D_{\downarrow}V(x))$$

we have

$$D\mathbf{V}(x)(u) = [D_{\uparrow}V(x)(u), D_{\downarrow}V(x)(u)]$$

Proof — By Proposition 6.1.6, it remains to show that when V is continuous on a neighborhood of x, the inclusion

$$T_{\mathrm{Ep}(V)}(x, V(x)) \cap T_{\mathrm{Hp}(V)}(x, V(x)) \subset T_{\mathrm{Graph}(V)}(x, V(x))$$

holds true. Pick any (u, v) in the intersection: There exist sequences $h_n^{\uparrow} > 0$ and $h_n^{\downarrow} > 0$ converging to $0+$, and sequences of pairs $(u_n^{\uparrow}, v_n^{\uparrow})$ and $(u_n^{\downarrow}, v_n^{\downarrow})$ converging to (u, v) such that

$$\frac{V(x + h_n^{\uparrow}u_n^{\uparrow}) - V(x)}{h_n^{\uparrow}} \leq v_n^{\uparrow} \quad \& \quad \frac{V(x + h_n^{\downarrow}u_n^{\downarrow}) - V(x)}{h_n^{\downarrow}} \leq v_n^{\downarrow}$$

We claim that it is enough to show that for every n large enough, there exists $\lambda_n \in [0, 1]$ such that, setting

$$h_n := (1 - \lambda_n)h_n^{\uparrow} + \lambda_n h_n^{\downarrow}$$

and

$$u_n := \frac{(1 - \lambda_n)h_n^{\uparrow}u_n^{\uparrow} + \lambda_n h_n^{\downarrow}u_n^{\downarrow}}{h_n} \quad \& \quad v_n := \frac{(1 - \lambda_n)h_n^{\uparrow}v_n^{\uparrow} + \lambda_n h_n^{\downarrow}v_n^{\downarrow}}{h_n}$$

we have the following equality

$$\frac{V(x + h_n u_n) - V(x)}{h_n} = v_n$$

Indeed $h_n > 0$ converges to $0+$ and (u_n, v_n) converges to (u, v). So, (u, v) belongs to the contingent cone to the graph of V.

To find such λ_n, consider for all large n the continuous functions $\varphi_n : [0,1] \to \mathbf{R}$ defined by

$$\begin{cases} \varphi_n(\lambda) := \\ V\left(x + (1-\lambda)h_n^{\uparrow}u_n^{\uparrow} + \lambda h_n^{\downarrow}u_n^{\downarrow}\right) - \left(V(x) + (1-\lambda)h_n^{\uparrow}v_n^{\uparrow} + \lambda h_n^{\downarrow}v_n^{\downarrow}\right) \end{cases}$$

and observe that

$$\varphi_n(0) \leq 0 \quad \& \quad \varphi_n(1) \geq 0$$

Since φ_n is continuous, there exists $\lambda_n \in [0,1]$ such that $\varphi_n(\lambda_n) = 0$. Then λ_n satisfies the required property. □

The contingent epiderivative coincides with the directional derivative $< V'(x), u >$ when V is Fréchet differentiable.

If V is Fréchet differentiable at a point $x \in K$, then the *contingent epiderivative of the restriction is the restriction of the derivative to the contingent cone*:

$$D_{\uparrow}(V|_K)(x)(u) \; := \; \begin{cases} < V'(x), u > & \text{if } \; u \in T_K(x) \\ +\infty & \text{if not} \end{cases}$$

The formulas become much more simple when V is Lipschitz: the contingent epiderivative coincides with the *lower Dini derivative* :

Proposition 6.1.7 *Let us assume that $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ is Lipschitz at a point x of its domain. Then*

$$D_{\uparrow}V(x)(u) \; = \; \liminf_{h \to 0+} \frac{V(x+hu) - V(x)}{h} \quad \text{(the lower Dini derivative)}$$

and satisfies for some $l > 0$:

$$\forall\, u \in X, \; |D_{\uparrow}V(x)(u)| \; \leq \; l\|u\|$$

Remark — There are other intimate connections between contingent cones and contingent epiderivatives.

Let ψ_K be the *indicator* of a subset K. Then it is easy to check that

$$D_{\uparrow}(\psi_K)(x) \; = \; \psi_{T_K(x)}$$

Therefore we can either derive properties of the epiderivatives from properties of the tangent cones to epigraphs, as we did, or take the opposite approach by using the above formula. □

There is also an obvious link between the contingent cone and the contingent epiderivative of the distance function to K since we can write for every $x \in K$:

$$T_K(x) \;=\; \{v \;\in\; X \;\mid\; D_\uparrow d_K(x)(v) \;=\; 0\}$$

We also mention the following estimate of the contingent epiderivative of the distance function to K:

Proposition 6.1.8 *Let K be a closed subset of a normed vector-space and $\Pi_K(y)$ be the set of projections of y onto K, i.e., the subset of $z \in K$ such that $\|y - z\| = d_K(y)$. Then when $\Pi_K(y) \neq \emptyset$, we have the following inequalities:*

$$D_\uparrow d_K(y)(v) \;\leq\; d\left(v, T_K(\Pi_K(y))\right)$$

Proof — We begin by proving this inequality when y belongs to K. Indeed, for all $w \in X$, inequality

$$d_K(y + hv) \;\leq\; d_K(y + hw) + h\|v - w\|$$

implies that

$$D_\uparrow d_K(y)(v) \;=\; \liminf_{h \to 0+} d_K(y + hv)/h \;\leq\; \|v - w\| \text{ when } w \;\in\; T_K(y)$$

Assume next that $y \notin K$. Choose $z \in \Pi_K(y)$. Then

$$d_K(y + hv) - d_K(y) \;\leq\; \|y - z\| + d_K(z + hv) - d_K(y) \;=\; d_K(z + hv)$$

From the first part of the proof, we deduce that $D_\uparrow d_K(z)(v) \leq d(v, T_K(z))$, and consequently, that $D_\uparrow d_K(y)(v) \leq d(v, T_K(z).)$ $\square$

6.1.3 Fermat and Ekeland Rules

Since we can define the contingent epiderivative of any extended function $V : X \mapsto \mathbf{R} \cup \{+\infty\}$, we can extend the "Fermat rule" to any minimization problem.

Theorem 6.1.9 (Fermat Rule) *Let $V : X \mapsto \mathbf{R} \cup \{+\infty\}$ be a nontrivial extended function on a normed space X and $x \in \mathrm{Dom}(V)$ a local minimizer of V on X.*

Then x is a solution to the variational inequalities:

$$\forall\, u \;\in\; X, \;\; 0 \;\leq\; D_\uparrow V(x)(u)$$

Proof — The proof is naturally obvious: We write that for all $u' \in X$ and all $h > 0$,

$$0 \leq (V(x+hu') - V(x))/h$$

and we take the lim inf when h converges to 0 and u' to u. □

The converse statement is not true without further assumptions such as convexity or, more generally, pseudo-convexity, or second order conditions:

Definition 6.1.10 *An extended function $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ is called* pseudo-convex *at $x \in \mathrm{Dom}(V)$ if its epigraph is pseudo-convex at $(x, V(x))$, i.e., if*

$$\forall\, y \in X, \;\; D_\uparrow V(x)(y-x) \leq V(y) - V(x)$$

We infer easily from this definition the converse of the Fermat Rule:

Proposition 6.1.11 *Let $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ and $x \in \mathrm{Dom}(V)$ satisfy $0 \leq D_\uparrow V(x)(u)$ for all $u \in X$. If V is pseudo-convex at x, then x achieves the minimum of V.*

What is not obvious is the use of this Fermat rule for more and more general problems, when the function V is built from other simpler functions and involves constraints.

The search for necessary conditions for a minimum requires quite a rich calculus of contingent epiderivatives which provides estimates of $D_\uparrow V(x)(u)$. In particular, when constraints (of the type $x \in K$) are involved, the fact that the epiderivative of the restriction to K is the restriction of the epiderivative to the contingent cone $T_K(x)$, allows one to write necessary conditions using also contingent cones to constraint sets (or in the dual form, using gradients and polars of the contingent cones.)

In the same way, it is easy to derive an epidifferential version of Ekeland's Variational Principle:

Theorem 6.1.12 *Let X be a Banach space, $V : X \mapsto \mathbf{R}_+ \cup \{+\infty\}$ be a nontrivial lower semicontinuous bounded from below function* *and* $x_0 \in \mathrm{Dom}(V)$ *be a given point of its domain. Then, for any* $\varepsilon > 0$*, there exists a solution* $x_\varepsilon \in \mathrm{Dom}(V)$ *to:*

$$\begin{cases} i) \quad V(x_\varepsilon) + \varepsilon\|x_\varepsilon - x_0\| \ \leq \ V(x_0) \\ \\ ii) \quad \forall\, u \in X, \ 0 \ \leq \ D_\uparrow V(x_\varepsilon)(u) + \varepsilon\|u\| \end{cases} \tag{6.1}$$

6.1.4 Elementary Properties

We present below some formulas concerning epiderivatives which allow exploitation of the Fermat and Ekeland rules.

1. — **Sum and composition**

Proposition 6.1.13 *Let us consider two Banach spaces* X, Y, *a single-valued map* $f : X \mapsto Y$ *and two extended functions* V *and* W *from* X *and* Y *to* $\mathbf{R} \cup \{+\infty\}$ *respectively.*

Let x_0 *belong to the domain of the function* $V + W \circ f$ *and assume that* f *is Fréchet differentiable at* x_0*. Then*

$$D_\uparrow V(x_0)(u) + D_\uparrow W(f(x_0))(f'(x_0)u) \ \leq \ D_\uparrow(V + W \circ f)(x_0)(u)$$

In particular, if K *is a subset of* X *and if* $x_0 \in K \cap \mathrm{Dom}(V)$*, then*

$$\forall\, u \in T_K(x_0), \ D_\uparrow V(x_0)(u) \ \leq \ D_\uparrow(V|_K)(x_0)(u)$$

2. — **Supremum of Functions**

Let us consider now a family of functions

$$V_i : X \ \mapsto \ \mathbf{R} \cup \{\pm\infty\}, \ (i \in I)$$

and let us associate with it the function U defined by

$$U(x) \ := \ \sup_{i \in I} V_i(x)$$

We set $I(x) := \{i \in I \mid V_i(x) = U(x)\}$. The following estimate is obvious whenever $I(x)$ is not empty and $x \in \mathrm{Dom}(U)$:

$$\forall\, u \in X, \ \sup_{i \in I(x)} D_\uparrow V_i(x)(u) \ \leq \ D_\uparrow U(x)(u) \tag{6.2}$$

because, for all $i \in I(x)$,

$$\frac{V_i(x+hu) - V_i(x)}{h} \leq \frac{U(x+hu) - U(x)}{h}$$

3. — Marginal Function

Let us consider normed spaces X and Y and an extended function $U : X \times Y \mapsto \mathbf{R} \cup \{\pm\infty\}$.

We associate with it the marginal function $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ defined by

$$V(x) := \inf_{y \in Y} U(x,y)$$

Let π denote the projection from $X \times Y \times \mathbf{R}$ to $X \times \mathbf{R}$. We observe that:

$$\pi \mathcal{E}p(U) \subset \mathcal{E}p(V) \subset \overline{\pi \mathcal{E}p(U)}$$

The first inclusion is obvious. The very definition of the infimum implies that for every $\varepsilon > 0$ and every $(x, \lambda) \in \mathcal{E}p(V)$, there exists $y_\varepsilon \in Y$ such that $(x, y_\varepsilon, \lambda + \varepsilon)$ belongs to $\mathcal{E}p(U)$. Taking projections and passing to the limit, the second inclusion ensues.

Let $x_0 \in X$ be given. Suppose that there exists $y_0 \in Y$ which achieves the minimum of $U(x_0, \cdot)$ on Y:

$$V(x_0) = U(x_0, y_0) \neq \pm\infty$$

By taking contingent cones to the above inclusions, we obtain the inequality

$$\forall\, u \in X, \;\; D_\uparrow V(x_0)(u) \leq \liminf_{u' \to u} \left(\inf_{v \in Y} D_\uparrow U(x_0, y_0)(u', v) \right)$$

because

$$\begin{cases} \overline{\pi \mathcal{E}p(D_\uparrow U(x_0, y_0))} = \overline{\pi T_{\mathcal{E}p(U)}(x_0, y_0, U(x_0, y_0))} \\ \\ \subset T_{\pi(\mathcal{E}p(U))}(x_0, V(x_0)) = T_{\mathcal{E}p(V)}(x_0, V(x_0)) = \mathcal{E}p(D_\uparrow V(x_0)) \end{cases}$$

can be easily translated into the desired inequality. □

In order to obtain the equality sign in some of the above formulas, we need further assumptions, such as the functions are episleek or pseudo-convex, which actually require the introduction of adjacent and circatangent epiderivatives. These are the topics of the next section.

6.2 Other Epiderivatives

We recall the definition of "lim sup inf" (or Γ-convergence) of functions of two variables introduced in Definition 5.2.4:

$$\liminf\sup_{x'\to x}\inf_{y'\to y}\phi(x',y') := \sup_{\varepsilon>0}\inf_{\eta>0}\sup_{x'\in B(x,\eta)}\inf_{y'\in B(y,\varepsilon)}\phi(x',y')$$

In the same way, we set

$$\liminf_{x'\to x}\sup_{y'\to y}\phi(x',y') := \inf_{\varepsilon>0}\sup_{\eta>0}\inf_{x'\in B(x,\eta)}\sup_{y'\in B(y,\varepsilon)}\phi(x',y')$$

6.2.1 Adjacent and Circatangent Epiderivatives

We can associate with the adjacent and Clarke tangent cones the adjacent and circatangent derivatives of V at $x \in \mathrm{Dom}(V)$ of an extended function $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ regarded as the set-valued map $\mathbf{V} : X \rightsquigarrow \mathbf{R}$ taking empty values outside of the domain of V:

$$\begin{cases} D^\flat\mathbf{V}(x)(u) := D^\flat\mathbf{V}(x,V(x))(u) = \\ \{v \mid \limsup_{h\to 0+}\inf_{u'\to u} |V(x+hu') - V(x) - hv|/h = 0\} \end{cases}$$

and

$$\begin{cases} C\mathbf{V}(x)(u) := C\mathbf{V}(x,V(x))(u) = \\ \{v \mid \limsup_{h\to 0+,x'\to x}\inf_{u'\to u} |V(x'+hu') - V(x) - hv|/h = 0\} \end{cases}$$

We are also naturally led to associate with the two set-valued maps $\mathbf{V}_\uparrow$ and $\mathbf{V}_\downarrow$, defined in the preceding section, their adjacent and circatangent derivatives at points $(x, V(x))$.

Definition 6.2.1 (Epiderivatives) *Let $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ be a nontrivial extended function and x belong to its domain. We shall say that the functions $D^\flat_\uparrow V(x)$ and $C_\uparrow V(x)$ from X to $\mathbf{R} \cup \{\pm\infty\}$ defined respectively by*

$$\begin{cases} i) \quad \forall\, u \in X, \ D^\flat_\uparrow V(x)(u) := \inf\left\{v \mid v \in D^\flat\mathbf{V}_\uparrow(x,V(x))(u)\right\} \\ ii) \quad \forall\, u \in X, \ C_\uparrow V(x)(u) := \inf\{v \mid v \in C\mathbf{V}_\uparrow(x,V(x))(u)\} \end{cases}$$

are the adjacent and circatangent epiderivatives *of V at x in the direction u.*

These epiderivatives can also be characterized as limits of differential quotients.

Proposition 6.2.2 *Let $V : X \mapsto \mathbf{R}\cup\{\pm\infty\}$ be a nontrivial extended function and x belong to its domain. Then*

$$D_\uparrow^\flat V(x)(u) \;=\; \lim\,\sup_{h\to 0+}\,\inf_{u'\to u}\frac{V(x+hu')-V(x)}{h}$$

and

$$C_\uparrow V(x)(u) = \lim\,\sup_{h\to 0+,x'\to x,V(x')\leq\lambda'\to V(x)}\inf_{u'\to u}\frac{V(x'+hu')-\lambda'}{h}$$

We define in a symmetric way the *adjacent and circatangent hypoderivatives* $D_\downarrow^\flat V(x)$ and $C_\downarrow V(x)$ from X to $\mathbf{R}\cup\{\pm\infty\}$

$$D_\downarrow^\flat V(x)(u) = -D_\uparrow^\flat(-V)(x)(u) := \lim\,\inf_{h\to 0+}\,\sup_{u'\to u}\frac{V(x+hu')-V(x)}{h}$$

and

$$\begin{cases} C_\downarrow V(x)(u) \;=\; -C_\uparrow(-V)(x)(u) \\ := \;\lim\,\inf_{h\to 0+,x'\to x,V(x')\geq\lambda'\to V(x)}\,\sup_{u'\to u}(V(x'+hu')-\lambda')/h \end{cases}$$

Naturally, when V is contingently epidifferentiable, the adjacent and circatangent epiderivatives are lower semicontinuous and positively homogeneous, and the circatangent epiderivative is moreover convex. They coincide with the directional derivatives $< V'(x), u >$ when V is respectively Fréchet and continuously differentiable.

Remark — It is also possible to derive another interpretation of the epiderivatives in terms of epilimits of the following differential quotients

$$\nabla_h V(x) : u \;\mapsto\; \nabla_h V(x)(u) \;:=\; \frac{V(x+hu)-V(x)}{h}$$

as we shall see in Chapter 7. □

Let $K \subset X$, $x \in K$ and V be Fréchet differentiable at x. Then the *adjacent epiderivative of the restriction* $V|_K$ *is the restriction of the derivative to the adjacent cone.*

When V is continuously differentiable at a point $x \in K$, the similar statement holds true for the circatangent derivative:

$$\begin{cases} i) & D_\uparrow^\flat (V|_K)(x)(u) & := \begin{cases} < V'(x), u > & \text{if } u \in T_K^\flat(x) \\ +\infty & \text{if not} \end{cases} \\ ii) & C_\uparrow (V|_K)(x)(u) & := \begin{cases} < V'(x), u > & \text{if } u \in C_K(x) \\ +\infty & \text{if not} \end{cases} \end{cases}$$

The formulas become much more simple when V is Lipschitz.

Proposition 6.2.3 *Let us assume that* $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ *is Lipschitz around a point* x *of its domain. Then*

$$D_\uparrow^\flat V(x)(u) \;=\; \limsup_{h \to 0+} \frac{V(x+hu) - V(x)}{h}$$

is the upper Dini derivative and[3]

$$C_\uparrow V(x)(u) \;=\; \limsup_{h \to 0+, x' \to x} \frac{V(x'+hu) - V(x')}{h}$$

Furthermore, for some neighborhood $\mathcal{U}$ *of* x,

$$\begin{cases} i) & \text{the map } (y,u) \in \mathcal{U} \times X \;\mapsto\; C_\uparrow V(y)(u) \\ & \textit{is upper semicontinuous} \\ ii) & \text{the map } u \mapsto C_\uparrow V(x)(u) \;\textit{ is Lipschitz on } X \\ iii) & \forall\, u \in X, \;\; C_\uparrow(-V)(x)(u) \;=\; C_\uparrow V(x)(-u) \end{cases}$$

We also observe the following relations between tangent cones to epigraphs and epigraphs of epiderivatives:

Proposition 6.2.4 *Let* $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ *be a nontrivial extended function and* x *belong to its domain.*

[3]Used also with the notation $V^o(x,u)$ by Clarke and $V^\uparrow(x,u)$ by Rockafellar.

Then the tangent cones to the epigraph of V at $(x, V(x))$ are the epigraphs of the corresponding epiderivatives of V at x:

$$\mathcal{E}p(D_{\uparrow}^{\flat}V(x)) = T_{\mathcal{E}p(V)}^{\flat}(x, V(x)) \ \ \& \ \ \mathcal{E}p(C_{\uparrow}V(x)) = C_{\mathcal{E}p(V)}(x, V(x))$$

and, when $x \in K \subset X$,

$$D_{\uparrow}^{\flat}(\psi_K)(x) \ = \ \psi_{T_K^{\flat}(x)} \ \ \& \ \ C_{\uparrow}(\psi_K)(x) \ = \ \psi_{C_K(x)}$$

The relations between tangent cones and epiderivatives of the distance functions are given in:

Proposition 6.2.5 *Let K be a subset of a normed space X, $y \in X$ and $\Pi_K(y)$ be the set of projections of y onto K. Then if $\Pi_K(y) \neq \emptyset$, we have*

$$\left\{ \begin{array}{ll} i) & D_{\uparrow}^{\flat}d_K(y)(v) \ \leq \ d\left(v, T_K^{\flat}(\Pi_K(y))\right) \\ \\ ii) & C_{\uparrow}d_K(y)(v) \ \leq \ d\left(v, C_K(\Pi_K(y))\right) \end{array} \right. \tag{6.3}$$

Proof — The proof of the first statement is analogous to the proof of Proposition 6.1.8. For proving the second inequality, we take $z \in \Pi_K(y)$ and $w \in C_K(z)$. We observe that when $y \notin K$,

$$\forall \, z \in \Pi_K(y), \ \ \forall \, x \in K, \ \ \|z - x\| \ \leq \ 2\|y - x\|$$

Hence

$$\left\{ \begin{array}{l} \sup_{h \leq \alpha, \ \|y-x\| \leq \beta} (d_K(y + hv) - d_K(y))/h \\ \\ \leq \ \sup_{h \leq \alpha, \ \|z-x\| \leq 2\beta} d_K(z + hw)/h + \|v - w\| \end{array} \right.$$

so that for every $w \in C_K(z)$, $C_{\uparrow}d_K(y)(v) \leq \|v - w\|$. Since $z \in \Pi_K(y)$ and $w \in C_K(z)$ were chosen arbitrarily, our proof ensues. □

The adjacent derivatives enjoy

Proposition 6.2.6 *Let $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ be an extended function continuous on a neighborhood of $x \in \mathrm{Int}(\mathrm{Dom}(V))$. The values of the adjacent derivative at x are convex.*

Furthermore, for any

$$u \ \in \ \mathrm{Dom}(D_{\uparrow}^{\flat}V(x)) \cap \mathrm{Dom}(D_{\downarrow}^{\flat}V(x))$$

we have

$$D^{\flat}\mathbf{V}(x)(u) \ = \ [D_{\uparrow}^{\flat}V(x)(u), D_{\downarrow}^{\flat}V(x)(u)]$$

Proof — As for the contingent case, we can check that for any $u \in \mathrm{Dom}(D_\uparrow^\flat V(x)) \cap \mathrm{Dom}(D_\downarrow^\flat V(x))$, we have

$$\{D_\uparrow^\flat V(x)(u), D_\downarrow^\flat V(x)(u)\} \subset D^\flat \mathbf{V}(x)(u) \subset [D_\uparrow^\flat V(x)(u), D_\downarrow^\flat V(x)(u)]$$

Let us take now two elements v_1 and v_2 in $D^\flat \mathbf{V}(x)(u)$. Pick any $\lambda \in [0,1]$. We shall show that $(1-\lambda)v_1 + \lambda v_2$ belongs to $D^\flat \mathbf{V}(x)(u)$ and thus, that $D^\flat \mathbf{V}(x)(u)$ is an interval.

By definition of the adjacent derivative, we know that there exist sequences u_{1h} and u_{2h} converging to u when $h \to 0+$ such that

$$\frac{V(x+hu_{1h}) - V(x)}{h} \to v_1 \quad \& \quad \frac{V(x+hu_{2h}) - V(x)}{h} \to v_2$$

Since V is continuous on a neighborhood of x, for h small enough, it maps the interval $x + h[u_{1h}, u_{2h}]$ to a connected subset, i.e., an interval which contains $V(x+hu_{1h})$ and $V(x+hu_{2h})$. Then there exists $w_h \in [u_{1h}, u_{2h}]$ such that

$$(1-\lambda)V(x+hu_{1h}) + \lambda V(x+u_{2h}) = V(x+hw_h)$$

Hence w_h converges to u and

$$\lim_{h\to 0+} \frac{V(x+hw_h) - V(x)}{h} = (1-\lambda)v_1 + \lambda v_2$$

This ends the proof. □

6.2.2 Other Convex Epiderivatives

Convexity of the circatangent epiderivative counts among its attractive features. But, when the function is not episleek, there are other positively homogeneous convex lower semicontinuous functions lying between the contingent epiderivative and the circatangent epiderivative, among which one can choose whenever convexity is a prerequisite to apply duality arguments.

Definition 6.2.7 *Let $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ be a nontrivial extended function and x belong to its domain. We shall say that a lower semicontinuous convex positively homogeneous function*

$$\delta V(x) : X \mapsto \mathbf{R} \cup \{\pm\infty\}$$

satisfying

$$D_\uparrow V(x)(u) \;\leq\; \delta V(x)(u) \;\leq\; C_\uparrow V(x)(u) \tag{6.4}$$

is a convex epiderivative. *It is said to be minimal if there is no other strictly smaller convex epiderivative.*

All these convex epiderivatives coincide when V is episleek.

Convex epiderivatives exist whenever V is contingently epidifferentiable at x and do not exist if the circatangent epiderivative at x takes the value $-\infty$ (or, equivalently, $C_\uparrow V(x)(0) < 0$.)

The epigraphs of the convex epiderivatives are the closed convex cones lying between the Clarke tangent cone and the contingent cone to the epigraph of V at x.

Among the possible choices of such convex epiderivatives, we can single out the convex epiderivative $D_\uparrow^\infty V(x)$ whose epigraph is the convex kernel of the epigraph of the contingent epiderivative. It satisfies

$$\begin{cases} \forall\, u \in X, \;\; D_\uparrow V(x)(u) \;\leq\; D_\uparrow^\infty V(x)(u) \\ \\ = \sup_{v \in \mathrm{Dom}(D_\uparrow V(x))} (D_\uparrow V(x)(u+v) - D_\uparrow V(x)(v)) \;\leq C_\uparrow V(x)(u) \end{cases}$$

Denote by $\mathcal{D}_V(x)$ the family of minimal convex epiderivatives $\delta V(x)$ of V at x. Since the convex kernel of the contingent cone is the intersection of its maximal closed convex subcones (by Proposition 4.5.2), we deduce that $D_\uparrow^\infty V(x)$ *is the supremum of the minimal convex epiderivatives*:

$$\forall\, u \in X, \;\; D_\uparrow^\infty V(x)(u) \;=\; \sup_{\delta V(x) \in \mathcal{D}_V(x)} \delta V(x)(u)$$

Also, the convex kernel of a cone P being equal to the Clarke tangent cone to P at the origin (Proposition 4.5.3), we obtain the formula

$$\forall\, u \in X, \;\; D_\uparrow^\infty V(x)(u) \;=\; C_\uparrow(D_\uparrow V(x))(0)(u)$$

Remark — Paratingent epiderivatives

We can also define the *paratingent epiderivative* $P_\uparrow V(x)$ of an extended function $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ at $x \in \mathrm{Dom}(V)$ by saying that the *epigraph of the paratingent epiderivative is the paratingent cone* $P_{\mathrm{Ep}(V)}^{\mathrm{Graph}(V)}(x, V(x))$.

One can then check that

$$P_\uparrow V(x)(v) \; = \; \liminf_{h\to 0+, y\to x, V(y)\to V(x), u'\to u} \frac{V(y+hu')-V(y)}{h}$$

We deduce from Choquet's Theorem 4.5.7 the following result due to Shi Shuzhong:

Theorem 6.2.8 *Assume that $V : X \mapsto \mathbf{R}$ is continuous. Then there exists a residual R on which the contingent and paratingent epiderivatives do coincide.*

Proof — Since Graph(V) is closed, Choquet's Theorem implies that there exists a residual $\mathcal{R}$ of the graph of V on which the contingent and paratingent cones coincide. Let R denote the projection $\Pi\mathcal{R}$ of the residual $\mathcal{R} \subset \mathrm{Graph}(V) \subset X \times \mathbf{R}$ to X. The continuity of V implies that the restriction of the projector Π to the graph of V is bicontinuous, so that R is a residual of the domain of V. $\square$

6.3 Epidifferential Calculus

Is the epiderivative of the sum of two functions equal to the epiderivative of the sum? We answer here this question and study as well the epiderivative of the composition product.

Theorem 6.3.1 *Let us consider two finite dimensional vector-spaces X and Y, a continuous single-valued map $f : X \mapsto Y$ and two extended lower semicontinuous functions V and W from X and Y to $\mathbf{R} \cup \{+\infty\}$ respectively. Let x_0 belong to the domain of the function*

$$U \; := \; V + W \circ f$$

We assume that f is continuously differentiable around x_0, that V and W are contingently epidifferentiable at x_0 and $f(x_0)$ respectively and that the following transversality condition:

$$\mathrm{Dom}(C_\uparrow W(f(x_0))) - f'(x_0)\,(\mathrm{Dom}(C_\uparrow V(x_0))) \; = \; Y$$

holds true. Then the epiderivatives of U satisfy the estimates:

$$\begin{cases} i) & D_\uparrow U(x_0)(u) \; \leq \; D_\uparrow^\flat V(x_0)(u) \; + \; D_\uparrow W(f(x_0))(f'(x_0)u) \\ ii) & D_\uparrow^\flat U(x_0)(u) \; = \; D_\uparrow^\flat V(x_0)(u) \; + \; D_\uparrow^\flat W(f(x_0))(f'(x_0)u) \\ iii) & C_\uparrow U(x_0)(u) \; \leq \; C_\uparrow V(x_0)(u) \; + \; C_\uparrow W(f(x_0))(f'(x_0)u) \end{cases}$$

In particular, if K is a closed subset of X and if $x_0 \in K \cap \mathrm{Dom}(V)$ satisfies

$$\mathrm{Dom}(C_\uparrow V(x_0)) - C_K(x_0) \;=\; X$$

then, if V is contingently epidifferentiable at x_0,

$$\begin{cases} i) & \forall\, u \in T_K^\flat(x_0), \;\; D_\uparrow(V|_K)(x_0)(u) \;\leq\; D_\uparrow V(x_0)(u) \\ ii) & \forall\, u \in T_K^\flat(x_0), \;\; D_\uparrow^\flat(V|_K)(x_0)(u) \;=\; D_\uparrow^\flat V(x_0)(u) \\ iii) & \forall\, u \in C_K(x_0), \;\; C_\uparrow(V|_K)(x_0)(u) \;\leq\; C_\uparrow V(x_0)(u) \end{cases}$$

Proof — We shall prove the formula only in the case of circatangent epiderivatives.

If we set

$$\begin{cases} K \;:=\; \mathcal{E}p(V) \times \mathcal{E}p(W) \times \mathbf{R} \;\subset\; X \times \mathbf{R} \times Y \times \mathbf{R} \times \mathbf{R} \\ G(x,a,y,b,c) \;:=\; (f(x) - y, a + b - c) \\ H(x,a,y,b,c) \;:=\; (x,c) \end{cases}$$

we can write

$$\mathcal{E}p(U) \;=\; H(K \cap G^{-1}(0,0))$$

We shall use Theorem 4.3.3 to estimate tangent cones to $K \cap G^{-1}(0,0)$. We first observe that the assumptions of our theorem imply the corresponding transversality conditions of Theorem 4.3.3.

Let us set

$$z_0 \;=\; (x_0, V(x_0), f(x_0), W(f(x_0)), U(x_0))$$

Then we deduce that

$$C_K(z_0) \cap G'(z_0)^{-1}(0,0) \;\subset\; C_{K \cap G^{-1}(0)}(z_0)$$

It remains to show that this inclusion implies the desired inequality. Fix $u \in \mathrm{Dom}(C_\uparrow V(x_0))$ such that $f'(x_0)u \in \mathrm{Dom}(C_\uparrow W(f(x_0)))$ and define

$$\lambda \;=\; C_\uparrow V(x_0)(u), \;\; \mu \;=\; C_\uparrow W(f(x_0))(f'(x_0)u)$$

Hence

$$(u, \lambda, f'(x_0)u, \mu, \lambda + \mu) \in C_K(z_0) \cap G'(z_0)^{-1}(0,0)$$

and thus, it belongs to the Clarke tangent cone to $K \cap G^{-1}(0,0)$ at z_0.

This means that for any sequence $h_n > 0$ converging to 0 and any sequence

$$z_n := (x_n, a_n, y_n, b_n, c_n) \in K \cap G^{-1}(0,0)$$

converging to z_0, there exist elements $w_n := (u_n, \lambda_n, v_n, \mu_n, \nu_n)$ converging to $(u, \lambda, f'(x_0)u, \mu, \lambda + \mu)$ satisfying

$$\forall\, n \geq 0, \quad z_n + h_n w_n \in K \cap G^{-1}(0,0)$$

Since z_n and $z_n + h_n w_n$ belong to $G^{-1}(0,0)$, we infer that

$$f(x_n) = y_n, \; f(x_n + h_n u_n) = y_n + h_n v_n, \; c_n = a_n + b_n, \; \nu_n = \lambda_n + \mu_n$$

and using that $z_n + h_n w_n$ belongs to K, we deduce that

$$\frac{V(x_n + h_n u_n) - a_n}{h_n} \leq \lambda_n \;\&\; \frac{W(y_n + h_n v_n) - b_n}{h_n} \leq \mu_n$$

Consequently,

$$\frac{U(x_n + h_n u_n) - c_n}{h_n} \leq \nu_n$$

Since (u_n, ν_n) converges to $(u, \lambda + \mu)$, we finally obtain

$$C_\uparrow U(x_0)(u) \leq \lambda + \mu = C_\uparrow V(x_0)(u) + C_\uparrow W(f(x_0))(f'(x_0)u) \quad \square$$

Remark — If X, Y are Banach spaces, the conclusions remain true when we replace the transversality assumption by the following *stability assumption*: there exist constants $c > 0$, $\alpha \in [0, 1[$ and $\eta > 0$ such that, for all n,

$$\left\{\begin{array}{ll} i) & \forall\, x \in \mathrm{Dom}(V) \cap B(x_0, \eta), \; \forall\, y \in \mathrm{Dom}(W) \cap B(f(x_0), \eta) \\ & B_Y \subset f'(x_0)\left(\mathrm{Dom}(D_\uparrow^\flat V(x)) \cap cB_X\right) - \mathrm{Dom}(D_\uparrow W(y)) + \alpha B_Y \\ ii) & \sup_{u \in \mathrm{Dom}(D_\uparrow^\flat V(x))} |D_\uparrow^\flat V(x)(u)| / \|u\| \leq c \\ iii) & \sup_{v \in \mathrm{Dom}(D_\uparrow W(y))} |D_\uparrow W(y)(v)| / \|v\| \leq c \end{array}\right. \tag{6.5}$$

For that purpose, we have to check that the second transversality assumption of Theorem 4.3.3 is satisfied, i.e., that there exists a constant $c' > 0$ such that, for all n, for all

$$(x,a,y,b,c) \in K \text{ close to } z_0 = (x_0, V(x_0), f(x_0), W(f(x_0)), U(x_0))$$

for all $(z,\lambda) \in Y \times \mathbf{R}$, there exist $(u,\mu,v,\nu,\delta) \in T_K(x,a,y,b,c)$ and e such that

$$\begin{cases} i) \quad z = f'(x)u - v + e \quad \& \quad \lambda = \mu + \nu - \delta \\ ii) \quad \|e\| \le \alpha(\|z\| + |\lambda|) \quad \& \quad \|u\| + \|v\| + |\mu| + |\nu| + |\delta| \le c'(\|z\| + |\lambda|) \end{cases}$$

Assumptions (6.5) imply right away that there exist $u \in \mathrm{Dom}(D_\uparrow^\flat V(x))$, $v \in \mathrm{Dom}(D_\uparrow W(y))$ and e satisfying

$$\begin{cases} i) \quad z = f'(x)u - v + e, \\ ii) \quad \|e\| \le \alpha\|z\| \quad \& \quad \|u\| \le c\|z\| \quad \& \quad \|v\| \le (1 + \alpha + c\|f'(x)\|)\|z\| \end{cases}$$

Let us now take

$$\mu := c\|u\|, \quad \nu := c\|v\|, \quad \delta := \mu + \nu - \lambda$$

We deduce from (6.5)*iii*) that (u,μ) belongs to $\mathcal{E}p(D_\uparrow^\flat V(x))$, that (v,ν) belongs to $\mathcal{E}p(D_\uparrow W(y))$ and that

$$D_\uparrow^\flat V(x)(u) + D_\uparrow W(y)(v) \le c(\|u\| + \|v\|) = \mu + \nu = \lambda + \delta$$

Consequently

$$|\delta| \le (|\lambda| + c(\|u\| + \|v\|)) \le c'(\|z\| + |\lambda|) \quad \square$$

We now provide formulas for computing the epiderivative of the supremum of a finite number of extended functions.

Let us consider a finite family of functions

$$V_i : X \mapsto \mathbf{R} \cup \{\pm\infty\}, \quad (i \in I)$$

with which we associate the function U defined by

$$U(x) := \max_{i \in I} V_i(x)$$

We set

$$I(x) := \{i \in I \mid V_i(x) = U(x)\}$$

Proposition 6.3.2 *Consider a finite dimensional vector-space X and n extended functions $V_i : X \mapsto \mathbf{R} \cup \{+\infty\}$. If the following transversality assumption at $x_0 \in \mathrm{Dom}(U)$ holds true*

$$\forall\, u_i \in X, \ \bigcap_{i=1}^{n} (\mathrm{Dom}(C_\uparrow V_i(x_0)) - u_i) \neq \emptyset$$

then

$$\forall\, u \in \bigcap_{i=1}^{n} \mathrm{Dom}(D_\uparrow^\flat V_i(x_0)), \ \ D^\flat U(x_0)(u) = \max_{i \in I(x_0)} D_\uparrow^\flat V_i(x_0)(u)$$

and

$$\forall\, u \in \bigcap_{i=1}^{n} \mathrm{Dom}(C_\uparrow V_i(x_0)), \ \ CU(x_0)(u) \leq \max_{i \in I(x_0)} C_\uparrow V_i(x_0)(u)$$

Proof — Since the dimension of X is finite and since the epigraph of U is the intersection of the epigraphs of the n functions V_i, we can use Corollary 4.3.6, stating that if for all pairs (u_i, λ_i),

$$\bigcap_{i=1}^{n} \left(C_{\mathcal{E}p(V_i)}(x_0, U(x_0)) - (u_i, \lambda_i)\right) \neq \emptyset$$

then

$$T^\flat_{\mathcal{E}p(U)}(x_0, U(x_0)) = \bigcap_{i=1}^{n} T^\flat_{\mathcal{E}p(V_i)}(x_0, U(x_0))$$

The former property follows immediately from the assumption

$$\bigcap_{i=1}^{n} (\mathrm{Dom}(C_\uparrow V_i(x_0)) - u_i) \neq \emptyset$$

The left-hand side of the latter formula is the epigraph of the adjacent epiderivative of U at x_0. Since

$$T^\flat_{\mathcal{E}p(V_i)}(x_0, U(x_0)) \supset T^\flat_{\mathrm{Dom}(V_i)}(x_0) \times \mathbf{R}$$

whenever $V_i(x_0) < U(x_0)$, i.e., whenever $i \notin I(x_0)$, we deduce that

$$\forall\, u \in \bigcap_{i=1}^{n} \mathrm{Dom}(D_\uparrow^\flat V_i(x_0)), \ \ D^\flat U(x_0)(u) \leq \max_{i \in I(x_0)} D_\uparrow^\flat V_i(x_0)(u)$$

The proof of the second statement is analogous. □

We now complete formulas for the epiderivatives of marginal functions: Consider two normed vector spaces X and Y and an extended function

$$U : X \times Y \mapsto \mathbf{R} \cup \{\pm\infty\}$$

with which we associate the marginal function $V : X \mapsto \mathbf{R} \cup \{+\infty\}$ defined by

$$V(x) := \inf_{y \in Y} U(x, y)$$

Proposition 6.3.3 *Let us consider two normed vector spaces X and Y, an extended function $U : X \times Y \mapsto \mathbf{R} \cup \{\pm\infty\}$ and its marginal function V. Let $x_0 \in \mathrm{Dom}(V)$ and suppose that there exists $y_0 \in Y$ which achieves the minimum of $U(x_0, \cdot)$ on Y:*

$$V(x_0) = U(x_0, y_0)$$

If U is pseudo-convex at (x_0, y_0), then

$$\forall\, u \in X, \;\; D_\uparrow V(x_0)(u) = \liminf_{u' \to u} \left(\inf_{v \in Y} D_\uparrow U(x_0, y_0)(u', v) \right)$$

Proof — Let π denote the projection from $X \times Y \times \mathbf{R}$ to $X \times \mathbf{R}$. We recall that:

$$\pi \mathcal{E}p(U) \subset \mathcal{E}p(V) \subset \overline{\pi \mathcal{E}p(U)}$$

We deduce these statements from any criterion implying that the contingent cones to the images are the closures of the images of the contingent cones. This is the case of the epigraph of a pseudo-convex function, which is pseudo-convex. Then, equalities

$$\begin{cases} \overline{\pi \mathcal{E}p(D_\uparrow U(x_0, y_0))} = \overline{\pi T_{\mathcal{E}p(U)}(x_0, y_0, U(x_0, y_0))} \\ = T_{\pi(\mathcal{E}p(U))}(x_0, V(x_0)) = T_{\mathcal{E}p(V)}(x_0, V(x_0)) = \mathcal{E}p(D_\uparrow V(x_0)) \end{cases}$$

can be easily translated into the desired result. □

6.4 Generalized Gradient

6.4.1 Subdifferentials and Generalized Gradients

We devote this section to dual concepts of epiderivatives of extended functions.

When a function V is differentiable at x, its gradient $V'(x)$, being a continuous linear functional, is therefore an element $V'(x) \in X^\star$ of the dual of X:

$$\forall\, v \in X, \;\; < V'(x), v > = \; D_\uparrow V(x)(v)$$

When V is no longer differentiable, but contingently epidifferentiable, we can still introduce *subgradients* of V at x, which are those continuous linear functionals $p \in X^\star$ satisfying

$$\forall\, v \in X, \;\; < p, v > \leq \; D_\uparrow V(x)(v)$$

which constitute the (possibly empty) closed convex subset

$$\partial_0 V(x) \; := \; \{p \in X^\star \mid \forall\, v \in X, \;\; < p, v > \; \leq \; D_\uparrow V(x)(v)\}$$

Definition 6.4.1 (Subdifferential) *Let $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ be a nontrivial extended function and x belong to its domain. Assume that V is contingently epidifferentiable at x. The subset*

$$\partial_0 V(x) \; \subset \; X^\star$$

is called the subdifferential *of V at x (or* contingent generalized gradient*) and its elements are called the* subgradients.

We observe that when the contingent epiderivative $D_\uparrow V(x)(\cdot)$ is convex, it is the support function of the subdifferential $\partial_0 V(x)$.

When it is not convex, it is convenient to regard any convex epiderivative $\delta V(x)(\cdot)$ (see Definition 6.2.7) as the support function

$$\forall\, v \in X, \;\; \delta V(x)(v) \; = \; \sigma(\partial_\delta V(x), v)$$

of the closed convex subset $\partial_\delta V(x)$ defined by

$$\partial_\delta V(x) \; := \; \{p \in X^\star \mid \; \forall\, v \in X, \;\; < p, v > \leq \; \delta V(x)(v)\} \tag{6.6}$$

Definition 6.4.2 *Let $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ be a nontrivial extended function and x belong to its domain. Assume that V is contingently epidifferentiable at x.*

The (Clarke) generalized gradient *of V at x is defined by*

$$\partial V(x) \; := \; \{p \in X^\star \mid \forall \, u \in X, \;\; < p, u > \leq \; C_\uparrow V(x)(u)\}$$

More generally, we say that the closed convex subset $\partial_\delta V(x)$ associated to a convex epiderivative $\delta V(x)$ through formula (6.6) is the δ-generalized gradient of V at x*.*

Naturally, we observe that

$$\partial_0 V(x) \; \subset \; \partial_\delta V(x) \; \subset \; \partial V(x) \tag{6.7}$$

Another way to state that $p \in \partial_0 V(x)$ is a subgradient is to say that the pair $(p, -1)$ belongs to the polar cone to the contingent cone to the epigraph of V at $(x, V(x))$:

$$(p, -1) \; \in \; \left(T_{\mathcal{E}p(V)}(x, V(x))\right)^-$$

In the same way, to state that $p \in \partial V(x)$ amounts to saying that the pair $(p, -1)$ belongs to the (Clarke) normal cone $N_{\mathcal{E}p(V)}(x, V(x))$ to the epigraph of V at $(x, V(x))$.

Remark — Recall that $\mathcal{D}_V(x)$ denotes the family of minimal convex epiderivatives $\delta V(x)$ of V at x. The δ-generalized gradients $\partial_\delta V(x)$ associated with minimal convex epiderivatives $\delta V(x)$ are the minimal δ-generalized gradients containing the subdifferential $\partial_0 V(x)$.

If we denote by $\partial_\infty V(x)$ the generalized gradient associated with the convex kernel of the contingent epiderivative, which is the supremum of the minimal convex epiderivatives, we obtain the formula

$$\partial_0 V(x) \; \subset \; \partial_\infty V(x) \; = \; \overline{co} \left(\bigcup_{\delta V(x) \in \mathcal{D}_V(x)} \partial_\delta V(x) \right) \; \subset \; \partial V(x)$$

Naturally, when V is Fréchet differentiable at x, then

$$D_\uparrow V(x)(v) \; = < V'(x), v > \leq \; \delta V(x)(v)$$

so that *the subdifferential $\partial_0 V(x)$ is reduced to the gradient $V'(x)$.*

When V is continuously differentiable at x, the circatangent epiderivative coincides with the directional derivative $< V'(x), \cdot >$, so that the δ-generalized gradients are reduced to the only gradient:

$$\partial_\delta V(x) \;=\; \{V'(x)\} \text{ when } V \text{ is continuously differentiable at } x$$

If V is continuously differentiable around a point $x \in K$, then the *generalized gradient of the restriction is the sum of the gradient and the normal cone*:

$$\partial(V|_K)(x) \;=\; V'(x) + N_K(x)$$

We also note that the generalized gradient of the indicator of a subset is the normal cone:

$$\partial\psi_K(x) \;=\; N_K(x)$$

The Fermat and Ekeland rules can be formulated in a δ-generalized gradient way:

Theorem 6.4.3 (Fermat Rule) *Let $V : X \mapsto \mathbf{R}\cup\{+\infty\}$ be a nontrivial extended function defined on a normed space and $x \in \mathrm{Dom}(V)$ be a local minimizer of V on X. Then it is a solution to the inclusions:*

$$0 \;\in\; \partial_0 V(x) \;\subset\; \partial_\delta V(x) \;\subset\; \partial V(x)$$

for any convex epiderivative $\delta V(x)$ of V at x.

Theorem 6.4.4 (Ekeland Rule) *Consider a Banach space X, a nontrivial lower semicontinuous bounded from below function*

$$V : X \mapsto \mathbf{R}_+ \cup \{+\infty\}$$

and let $x_0 \in \mathrm{Dom}(V)$ be a given point of its domain. Assume that V is contingently epidifferentiable and let $\delta V(x)$ denote any convex epiderivative of V at x.

Then, for any $\varepsilon > 0$, there exists a solution $x_\varepsilon \in \mathrm{Dom}(V)$ to:

$$\left\{\begin{array}{ll} i) & V(x_\varepsilon) + \varepsilon\|x_\varepsilon - x_0\| \;\leq\; V(x_0) \\ \\ ii) & 0 \;\in\; \partial_\delta V(x_\varepsilon) + \varepsilon B_\star \;\subset\; \partial V(x_\varepsilon) + \varepsilon B_\star \end{array}\right. \qquad (6.8)$$

Proof — Indeed, Theorem 6.1.12 implies the existence of $x_\varepsilon \in X$ satisfying (6.8) *i*) and

$$\forall\, v \in X,\ \ 0 \ \le\ D_\uparrow V(x_\varepsilon)(v) + \varepsilon\|v\|$$

Therefore, for any convex epiderivative $\delta V(x_\varepsilon)$, we deduce that

$$\begin{cases} 0 \ \le\ \delta V(x_\varepsilon)(v) + \varepsilon\|v\| \\ =\ \sigma(\partial_\delta V(x_\varepsilon), v) + \sigma(\varepsilon B_\star, v) \ =\ \sigma(\partial_\delta V(x_\varepsilon) + \varepsilon B_\star, v) \ \square \end{cases}$$

6.4.2 Limits of Subdifferentials and Gradients

Not only the subdifferential $\partial_0 V(x)$ is contained in the Clarke generalized gradient $\partial V(x)$, but, in a finite dimensional vector-space, the upper limit of subdifferentials $\partial_0 V(x_n)$ when $x_n \to x$ is also a subset of this generalized gradient:

Theorem 6.4.5 *Let X be a finite dimensional vector-space and*

$$V : X \mapsto \mathbf{R} \cup \{+\infty\}$$

an extended lower semicontinuous function contingently epidifferentiable in a neighborhood of $x \in \mathrm{Dom}(V)$. Then

$$\mathrm{Limsup}_{(x',V(x'))\to(x,V(x))}\ \partial_0 V(x') \ \subset\ \partial V(x)$$

Proof — Indeed, let x_n be any subsequence converging to x such that $V(x_n)$ converges to $V(x)$ and p be the limit of a subsequence of subgradients $p_n \in \partial_0 V(x_n)$. Since

$$(p_n, -1) \ \in\ \left(T_{\mathcal{E}p(V)}(x_n, V(x_n))\right)^-$$

we deduce from Theorem 4.4.3 that $(p, -1)$ belongs to the (Clarke) normal cone to the epigraph of V at $(x, V(x))$. This means that p belongs to the generalized gradient $\partial V(x)$. $\square$

In Hilbert spaces and some Banach spaces, one can prove that the limits of the gradients at points where V is differentiable belong to the generalized gradient:

Theorem 6.4.6 *Consider a uniformly smooth Banach space X such that the norm of $X^\star$ is Fréchet differentiable off the origin and an extended lower semicontinuous function $V : X \mapsto \mathbf{R} \cup \{+\infty\}$ contingently epidifferentiable on a neighborhood of $x \in \mathrm{Dom}(V)$.*

Let a sequence of elements $(x_n)_{n\geq 1}$ converging to x be such that V is Fréchet differentiable at x_n and $V(x_n)$ converges to $V(x)$. Then

$$\sigma - \mathrm{Limsup}_{n\to\infty}\{V'(x_n)\} \subset \partial V(x)$$

Proof — We claim that the pair $(V'(x_n), -1)$ belongs to the polar cone to the *weak contingent cone* $T^{\sigma}_{\mathcal{E}p(V)}(x_n, V(x_n))$, defined in Section 4.1.5. by

$$\begin{cases} (u,v) \in T^{\sigma}_{\mathcal{E}p(V)}(x_n, V(x_n)) \iff (u,v) \text{ is a weak cluster point of} \\ (u_h, v_h) \in (\mathcal{E}p(V) - (x_n, V(x_n)))/h \text{ when } h \to 0+ \end{cases}$$

Indeed, fix $(u,v) \in T^{\sigma}_{\mathcal{E}p(V)}(x_n, V(x_n))$. Since V is Fréchet differentiable at x_n,

$$V(x_n + hu_h) = V(x_n) + h < V'(x_n), u_h > + h\varepsilon(h) \leq V(x_n) + hv_h$$

where $\varepsilon(h)$ converges to zero with h. Since continuous linear functionals are weakly continuous, we deduce that

$$< V'(x_n), u > -v = \langle (V'(x_n), -1), (u,v) \rangle \leq 0$$

Now, assume that p is the weak-$\star$ limit of a subsequence of the gradients $V'(x_n)$. Then the pair $(p, -1)$ belongs to the sequentially weak upper limit of the polar cones to the *weak contingent cones*:

$$(p, -1) \in \sigma - \mathrm{Limsup}_{n\to\infty} \left(T^{\sigma}_{\mathcal{E}p(V)}(x_n, V(x_n))\right)^-$$

By Theorem 4.4.4, it is contained in the normal cone $N_{\mathcal{E}p(V)}(x, V(x))$. This means that p belongs to the generalized gradient $\partial V(x)$. □

6.4.3 Local Subdifferentials and Superdifferentials

In finite dimensional vector-spaces, the subdifferential coincides with the local subdifferential $\partial_- V(x)$ defined in the following way:

Definition 6.4.7 *Let X be a normed space, $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ and $x_0 \in \mathrm{Dom}(V)$. We shall say that the subset $\partial_- V(x_0)$ of elements $p \in X^\star$ satisfying*

$$\liminf_{x \to x_0} \frac{V(x) - V(x_0) - < p, x - x_0 >}{\|x - x_0\|} \geq 0 \tag{6.9}$$

is the local subdifferential[4] of V at x_0.

In a symmetric way, the local superdifferential $\partial_+ V(x_0)$ *of V at x_0 is defined by*

$$\partial_+ V(x_0) := -\partial_-(-V)(x_0)$$

It is the subset of elements $p \in X^\star$ satisfying

$$\limsup_{x \to x_0} \frac{V(x) - V(x_0) - < p, x - x_0 >}{\|x - x_0\|} \leq 0$$

Clearly super- and subdifferential are closed (possibly empty) convex sets. When V is Fréchet differentiable, they coincide with the gradient of V.

Proposition 6.4.8 *Let X be a normed space, $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ be a nontrivial extended function and $x_0 \in \mathrm{Dom}(V)$. Then the local subdifferential $\partial_- V(x)$ is contained in the subdifferential $\partial_0 V(x)$.*

These two subsets coincide when the dimension of X is finite.

Proof — Let us take $p \in \partial_- V(x_0)$ and $v \in X$. There exist sequences of elements $h_n > 0$ and $v_n \in X$ converging to 0 and v respectively such that $(V(x_0 + h_n v_n) - V(x_0))/h_n$ converges to $D_\uparrow V(x_0)(v)$. Therefore

$$\begin{cases} D_\uparrow V(x_0)(v) - < p, v > \\ = \lim_{n \to \infty} (V(x_0 + h_n v_n) - V(x_0) - h_n < p, v_n >)/h_n \\ \geq \|v\| \liminf_{x \to x_0} (V(x) - V(x_0) - < p, x - x_0 >)/\|x - x_0\| \geq 0 \end{cases}$$

[4]This concept, used by Crandall & P.-L. Lions for defining *viscosity solutions* to Hamilton-Jacobi equations, has been called "subdifferential." We add here the adjective "local" to avoid confusion with the concept of subdifferential (or contingent generalized gradient), although these two concepts are equivalent in finite dimensional vector-spaces.

so that $< p, v > \leq D_\uparrow V(x_0)(v)$ for every $v \in X$.

Conversely, let p satisfy $< p, v > \leq D_\uparrow V(x_0)(v)$ for every $v \in X$ and assume that X is finite dimensional. There exists a sequence of elements x_n converging to x_0 such that

$$\left\{\begin{array}{l} \liminf_{x \to x_0} (V(x) - V(x_0) - < p, x - x_0 >)/\|x - x_0\| \\ \\ = \lim_{n \to \infty} (V(x_n) - V(x_0) - < p, x_n - x_0 >)/\|x_n - x_0\| \end{array}\right.$$

Taking a subsequence if needed, we may assume that the sequence of elements $v_n := (x_n - x_0)/|x_n - x_0\|$ of the unit sphere (which is compact) converges to an element v. Then for $h_n := \|x_n - x_0\|$, we have

$$\left\{\begin{array}{l} \lim_{n \to \infty} (V(x_n) - V(x_0) - < p, x_n - x_0 >)/\|x_n - x_0\| \\ \\ = \lim_{n \to \infty} (V(x_0 + h_n v_n) - V(x_0) - h_n < p, v_n >)/h_n \\ \\ \geq D_\uparrow V(x_0)(v) - < p, v > \geq 0 \end{array}\right.$$

so that p satisfies property (6.9.) □

6.4.4 Remarks

The table of formulas on support functions allows one to translate the properties of the circatangent epiderivatives into corresponding properties of the generalized gradient and vice-versa.

For instance, Proposition 6.2.3 can be restated in the following form.

Proposition 6.4.9 *When $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ is locally Lipschitz on the interior of its domain, then the generalized gradient satisfies:*

$$\left\{\begin{array}{l} \bullet \text{ the function } (x, u) \in X \times \text{Int}(\text{Dom}(V)) \mapsto \sigma(\partial V(x), u) \text{ is} \\ \quad \text{upper semicontinuous, so that } \partial V(\cdot) \text{ is upper hemicontinuous} \\ \\ \bullet \ \forall\, x \in \text{Int}(\text{Dom}(V)),\ \partial V(x) \text{ is nonempty bounded closed convex} \\ \\ \bullet \ \forall\, x \in \text{Int}(\text{Dom}(V)),\ \partial V(x) = -\partial(-V(x)) \end{array}\right.$$

Remark — When a function V is continuously differentiable at x, its gradient $V'(x)$ being a continuous linear functional, it is both an element $V'(x) \in X^\star$ of the dual and the image $V'(x)^\star(+1)$ of $+1$ by the transpose $V'(x)^\star$ of $V'(x)$, a linear operator from $\mathbf{R}$ to $X^\star$.

When V is no longer continuously differentiable, the generalized gradient remains intimately connected with the transpose of the circatangent derivative $C\mathbf{V}_\uparrow(x, V(x))$, (which, shall we recall, is the codifferential $C\mathbf{V}_\uparrow(x, V(x))^\star$) of the set-valued map $\mathbf{V}_\uparrow$ at $(x, V(x))$. Namely:

Proposition 6.4.10 *Let $V : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ and $x \in \mathrm{Dom}(V)$. The generalized gradient is the value at* 1 *of the codifferential of $\mathbf{V}_\uparrow$ at $(x, V(x))$.*

Proof — Indeed, the codifferential is a closed convex process from $\mathbf{R}$ to $X^\star$ which, being positively homogeneous, needs to be defined only at the points -1, 0 and $+1$.

We obtain from the definitions

$$\left\{\begin{array}{lll} i) & CV_\uparrow(x, V(x))^\star(-1) & = & \emptyset \\ ii) & CV_\uparrow(x, V(x))^\star(0) & = & (\mathrm{Dom}(C_\uparrow V)(x))^- =: \partial^\infty V(x) \\ iii) & CV_\uparrow(x, V(x))^\star(+1) & = & \partial V(x) \ \square \end{array}\right. \tag{6.10}$$

6.5 Convex Functions

Convex functions enjoy further properties. We already mentioned that an extended function is convex (respectively lower semicontinuous) if and only if its epigraph is convex (respectively closed.) A convex function defined on a finite dimensional vector-space is locally Lipschitz on the interior of its domain and a convex lower semicontinuous function on a Banach space is also locally Lipschitz on the interior of its domain.

Proposition 6.5.1 *When the function $V : X \mapsto \mathbf{R} \cup \{+\infty\}$ is convex, the contingent, adjacent and circatangent epiderivatives coincide and are equal to*

$$D_\uparrow V(x)(u) \;=\; \liminf_{u' \to u} \left(\inf_{h>0} \frac{V(x+hu') - V(x)}{h} \right)$$

Furthermore, when x belongs to the interior of the domain of V,

$$\forall\, u \in X, \;\; D_\uparrow V(x)(u) \;=\; \inf_{h>0} \frac{V(x+hu) - V(x)}{h} \;\; \text{is finite}$$

When V is convex, the generalized gradient coincides with the subdifferential introduced by Moreau and Rockafellar for convex functions:

$$\partial V(x) = \{p \in X^\star \mid \forall y \in X, < p, y - x > \leq V(y) - V(x)\}$$

There is more to that: lower semicontinuous convex functions enjoy duality properties. In the same way that we associated with cones their polar cones, with closed convex processes their transposes, we can, following Fenchel, associate with lower semicontinuous convex functions *conjugate functions* for the same reasons, and with the same success.

Definition 6.5.2 *Let $V : X \to \mathbf{R} \cup \{+\infty\}$ be any nontrivial extended function defined on a Banach space X. We associate with it its conjugate function $V^\star : X^\star \to \mathbf{R} \cup \{+\infty\}$ defined on the dual of X by*

$$\forall p \in X^\star, \ V^\star(p) := \sup_{x \in X}(< p, x > -V(x))$$

Its biconjugate $V^{\star\star} : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ is defined by

$$V^{\star\star}(x) := \sup_{p \in X^\star} (< p, x > -V^\star(p))$$

For instance, the conjugate function of the indicator ψ_K of a subset K is the support function σ_K.

We deduce from the definition the following convenient inequality

$$\forall x \in X, \ p \in X^\star, \ < p, x > \leq V(x) + V^\star(p)$$

known as *Fenchel's Inequality.* The epigraphs of the conjugate and biconjugate functions being closed convex subsets, the conjugate function is lower semicontinuous and convex and so is its biconjugate when it never takes the value $-\infty$. We observe that

$$\forall x \in X, \ V^{\star\star}(x) \leq V(x)$$

If equality holds, then V is convex and lower semicontinuous. The converse statement, a consequence of the Hahn-Banach Separation Theorem, is the first basic theorem of convex analysis:

Theorem 6.5.3 *A nontrivial extended function $V : X \to \mathbf{R} \cup \{+\infty\}$ is convex and lower semicontinuous if and only if it coincides with its biconjugate. In this case, the conjugate function $V^\star$ is non trivial.*

So, the Fenchel correspondence associating with any function V its conjugate $V^\star$ is a one to one correspondence between the sets of nontrivial lower semicontinuous convex functions defined on X and its dual $X^\star$. This fact is at the root of duality theory in convex optimization.

We deduce at once the following characterization of the subdifferential:

Proposition 6.5.4 *Let $V : X \to \mathbf{R} \cup \{+\infty\}$ be a nontrivial extended convex function defined on a Banach space X. Then*

$$p \in \partial V(x) \iff <p, x> = V(x) + V^\star(p)$$

If moreover the function V is lower semicontinuous, then the inverse of the subdifferential $\partial V(\cdot)$ is the subdifferential $\partial V^\star(\cdot)$ of the conjugate function*:*

$$p \in \partial V(x) \iff x \in \partial V^\star(p)$$

Since $-V^\star(0) = \inf_{x \in X} V(x)$, the Fermat rule becomes:

Theorem 6.5.5 *Let $V : X \mapsto \mathbf{R} \cup \{+\infty\}$ be a nontrivial convex extended function defined on a normed space X. Then $\overline{x} \in X$ achieves the minimum of V on X if and only if*

$$0 \in \partial V(\overline{x}) \text{ or, equivalently, } \forall\, u \in X,\ 0 \le D_\uparrow V(\overline{x})(u)$$

If moreover the function V is lower semicontinuous, then $\partial V^\star(0)$ is the set of minimizers of V.

The second fundamental theorem of convex analysis, also a consequence of the Separation Theorem, provides a criterion (weak constraint qualification) for computing the conjugate function of the sum and the composition product of convex functions:

Theorem 6.5.6 (Fenchel) *Let X and Y be reflexive Banach spaces, $A \in \mathcal{L}(X, Y)$ a continuous linear operator and $V : X \to \mathbf{R} \cup \{+\infty\}$*

and $W : Y \to \mathbf{R} \cup \{+\infty\}$ *be convex lower semicontinuous functions. Let* $U := V + W \circ A$.

We posit the weak constraint qualification assumption[5]*:*

$$0 \in \operatorname{Int}(A(\operatorname{Dom}(V)) - \operatorname{Dom}(W))$$

Then

$$\operatorname{Dom}(U^\star) = \operatorname{Dom}(V^\star) + A^\star \operatorname{Dom}(W^\star)$$

and, for every $p \in \operatorname{Dom}(U^\star)$*, there exists a solution* $\overline{q} \in Y^\star$ *to the minimization problem*

$$U^\star(p) = (V^\star(p - A^\star\overline{q}) + W^\star(\overline{q}) = \inf_{q \in Y^\star} (V^\star(p - A^\star q) + W^\star(q))$$

When $p = 0$, we deduce from the fact that

$$U^\star(0) = - \inf_{x \in X} U(x)$$

and from the above theorem the *duality relation*: for any $x \in X$, $q \in Y^\star$,

$$\inf_{x \in X} (V(x) + W(Ax)) + \inf_{q \in Y^\star} (V^\star(-A^\star q) + W^\star(q)) = 0$$

We observe that the subdifferential map $\partial V(\cdot)$ of a nontrivial lower semicontinuous convex function is obviously *monotone*. When X is a Hilbert space, we can prove that $\partial V(\cdot)$ *is maximal monotone*. This is a consequence of Minty's Theorem 3.5.8 and the following Moreau Theorem:

Theorem 6.5.7 (Moreau) *Let* X *be a Hilbert space identified with its dual and*

$$V : X \mapsto \mathbf{R} \cup \{+\infty\}$$

be a nontrivial lower semicontinuous convex function. For every $\lambda > 0$*, define the* Moreau-Yosida approximation V_λ *by*

$$\forall\, x \in X, \;\; V_\lambda(x) := \inf_{y \in Y} \left(V(y) + \frac{1}{2\lambda}\|x - y\|^2 \right) \leq V(x)$$

[5] We observe that U is non trivial if and only if $0 \in A(\operatorname{Dom}(V)) - \operatorname{Dom}(W)$. In this case, inequality $U^\star(p) \leq \inf_{q \in Y^\star}(V^\star(p - A^\star q) + W^\star(q))$ is always true.

This minimization problem has a unique solution $J_\lambda x$:

$$V_\lambda(x) \; = \; V(J_\lambda x) + \frac{1}{2\lambda}\|x - J_\lambda x\|^2$$

which is the resolvent of the maximal monotone map $\partial V(\cdot)$.

Furthermore, V_λ is convex, continuously differentiable, and converges pointwise to V on $\mathrm{Dom}(V)$ when $\lambda \to 0+$.

The Yosida approximation defined by $A_\lambda x := (x - J_\lambda x)/\lambda$, is both equal to the gradient of V_λ at x and to a subgradient of V at $J_\lambda x$:

$$A_\lambda x \; = \; V_\lambda'(x) \quad \& \quad A_\lambda x \; \in \; \partial V(J_\lambda x)$$

Finally, this implies that the domain of $\partial V(\cdot)$ is dense in $\mathrm{Dom}(V)$.

6.6 Higher Order Epiderivatives

Let us consider a normed space X and a nontrivial extended function

$$V : X \;\mapsto\; \mathbf{R} \cup \{+\infty\}$$

We can define m^{th}-order epiderivatives by taking m^{th}-order derivatives of the set-valued map $\mathbf{V}_\uparrow$.

We observe that v_m belongs to

$$D^{(m)}\mathbf{V}_\uparrow(x, y, u_1, v_1, \ldots, u_{m-1}, v_{m-1})(u_m)$$

if and only if

$$\begin{cases} v_m \; \geq \; \liminf_{h \to 0+, u_m' \to u_m} \frac{1}{h^m} \\ \left(V(x + hu_1 + \ldots + h^{m-1}u_{m-1} + h^m u_m') - y - hv_1 - \ldots - h^{m-1}v_{m-1}\right) \end{cases}$$

Denote by $D_\uparrow^{(j)}V(x, u_1, \ldots, u_j)$ the infimum of the subset

$$D^{(j)}\mathbf{V}_\uparrow\left(x, V(x), u_1, D_\uparrow V(x)(u_1), \ldots, D_\uparrow^{(j-1)}V(x, u_1, \ldots, u_{j-1})(u_j)\right)(u_j)$$

which is equal to

$$\begin{cases} D_\uparrow^{(j)}V(x, u_1, \ldots, u_j) \; := \; \liminf_{h \to 0+, \, u_j' \to u_j} \frac{1}{h^j} \\ \left(V(x + hu_1 + \ldots + h^{j-1}u_{j-1} + h^j u_j') - \sum_{l=0}^{j-1} h^l D_\uparrow^{(l)}V(x, u_1, \ldots, u_l)\right) \end{cases}$$

By taking successively $y = V(x)$, $v_1 := D_\uparrow V(x, u_1)$, etc., we can check recursively that for any j,

$$\left\{ \begin{array}{l} D^{(j)}\mathbf{V}_\uparrow \left(x, V(x), u_1, D_\uparrow V(x)(u_1), \dots, D_\uparrow^{(j-1)}V(x, u_1, \dots, u_{j-1})\right)(u_j) \\ \\ = \left[D_\uparrow^{(j)}V(x, u_1, \dots, u_j), +\infty\right[\end{array} \right.$$

Definition 6.6.1 *Let us consider a normed space X and a nontrivial extended function $V : X \mapsto \mathbf{R} \cup \{+\infty\}$. Let $x \in \mathrm{Dom}(V)$ and*

$$u_1 \in \mathrm{Dom}(D_\uparrow V(x)), \dots, u_{m-1} \in \mathrm{Dom}\left(D_\uparrow^{(m-1)}V(x, u_1, \dots, u_{m-2}, \cdot)\right)$$

be given. Then

$$\left\{ \begin{array}{l} D_\uparrow^{(m)}V(x, u_1, \dots, u_m) := \liminf_{h \to 0+, u'_m \to u_m} \frac{1}{h^m} \\ \\ \left(V(x + hu_1 + \dots h^{m-1}u_{m-1} + h^m u'_m) - \sum_{l=0}^{m-1} h^l D_\uparrow^{(l)}V(x, u_1, \dots, u_l)\right) \end{array} \right.$$

is called the m^{th}-order contingent epiderivative *of V at x in the directions u_1, u_2, ..., u_m.*

We define in the same way m^{th}-order adjacent and circatangent epiderivatives

$$D_\uparrow^{\flat(m)}V(x, u_1, \dots, u_m) \quad \& \quad C_\uparrow^{(m)}V(x, u_1, \dots, u_m)$$

using the corresponding $m-th$ order derivatives of set-valued maps.

For $m = 2$, we obtain

$$\left\{ \begin{array}{l} D_\uparrow^2 V(x, u_1, u_2) \\ := \liminf_{h \to 0+, u'_2 \to u_2} \left(V(x + hu_1 + h^2 u'_2) - V(x) - hD_\uparrow V(x)(u_1)\right)/h^2 \end{array} \right.$$

We observe right away that

$$D_\uparrow^{(m)}\psi_K(x, u_1, \dots, u_{m-1}, \cdot) = \psi_{T_K^{(m)}(x, u_1, \dots, u_{m-1})}(\cdot)$$

and that, for every $x \in K$

$$u_m \in T_K^{(m)}(x, u_1, \dots, u_{m-1}) \iff D_\uparrow^{(m)}d_K(x, u_1, \dots, u_{m-1}, u_m) = 0$$

The m^{th}-order contingent epiderivative enjoys the same kind of calculus as those of first-order.

Theorem 6.6.2 *Let us consider two finite dimensional vector-spaces X and Y, a continuous single-valued map $f : X \mapsto Y$, and two nontrivial extended functions V and W from X and Y to $\mathbf{R} \cup \{+\infty\}$ respectively.*

Let x_0 belong to the domain of the function $U := V + W \circ f$. We assume that f is twice continuously differentiable around x_0, that V and W are contingently epidifferentiable at x_0 and $f(x_0)$ respectively and that the following transversality condition:

$$\mathrm{Dom}\,(C_\uparrow W(f(x_0))) - f'(x_0)\,(\mathrm{Dom}\,(C_\uparrow V(x_0))) \;=\; Y$$

holds true.

Then the second-order epiderivatives of U satisfy the estimates:

$$\begin{cases} D_\uparrow^{(2)} U(x_0,u_1,u_2) \;\leq\; D_\uparrow^{\flat(2)} V(x_0,u_1,u_2) \\ \\ +\, D_\uparrow^{(2)} W(f(x_0), f'(x_0)u_1, f'(x_0)u_2 \;+\; \frac{1}{2} f''(x_0)(u_1,u_1)) \end{cases}$$

and

$$\begin{cases} D_\uparrow^{\flat(2)} U(x_0,u_1,u_2) \;=\; D_\uparrow^{\flat(2)} V(x_0,u_1,u_2) \\ \\ +\, D_\uparrow^{\flat(2)} W(f(x_0), f'(x_0)u_1, f'(x_0)u_2 \;+\; \frac{1}{2} f''(x_0)(u_1,u_1)) \end{cases}$$

In particular, if K is a closed subset of X and if $x_0 \in K \cap \mathrm{Dom}(V)$ satisfies

$$\mathrm{Dom}(C_\uparrow V(x_0)) - C_K(x_0) \;=\; X$$

then, if V is contingently epidifferentiable at x_0,

$$\begin{cases} i) \quad \forall\, u \in T_K^{\flat(2)}(x_0,u_1), \;\; D_\uparrow^{(2)}(V|_K)(x_0,u_1,u_2) \;=\; D_\uparrow^{(2)} V(x_0,u_1,u_2) \\ \\ ii) \quad \forall\, u \in T_K^{\flat(2)}(x_0,u_1), \;\; D_\uparrow^{\flat(2)}(V|_K)(x_0,u_1,u_2) \;=\; D_\uparrow^{\flat(2)} V(x_0,u_1,u_2) \end{cases}$$

Proof — The proof is exactly the same as the first-order Theorem 6.3.1, but we use Theorem 4.7.4 instead of Theorem 4.3.3. □

We can make more precise the Fermat rule by using m^{th}-order contingent epiderivatives.

Indeed, we saw that if $V : X \to \mathbf{R} \cup \{+\infty\}$ achieves its minimum at some $x_0 \in \mathrm{Dom}(V)$, then for every $u_1 \in X$, we have $D_\uparrow V(x_0)(u_1) \geq 0$. We can now distinguish the directions u_1 at which $D_\uparrow V(x_0)(u_1) = 0$. Then we see that for such directions, $D_\uparrow^{(2)} V(x_0)(u_1, u_2) \geq 0$, and so on.

Proposition 6.6.3 *Let us consider a normed space X and a nontrivial extended function $V : X \mapsto \mathbf{R} \cup \{+\infty\}$. Assume that x_0 achieves the minimum of V. Then*

$$\begin{cases} \forall\, u_1, \ldots, u_{m-1} \text{ such that } \forall\, j = 1, \ldots, m-1,\ D_\uparrow^{(j)} V(x_0, u_1, \ldots, u_j) = 0, \\ \text{we have } \forall\, u_m \in X,\ \ 0 \ \leq\ D_\uparrow^{(m)} V(x_0, u_1, \ldots, u_{m-1}, u_m) \end{cases}$$

We can prove that second order conditions are sufficient in finite dimensional vector-spaces:

Theorem 6.6.4 *Consider a normed space X and a nontrivial extended function $V : X \mapsto \mathbf{R} \cup \{+\infty\}$.*

Assume that $x_0 \in \mathrm{Dom}(V)$ satisfies

$$\begin{cases} i) & \forall\, u \in X,\ \ 0 \ \leq\ D_\uparrow V(x_0)(u) \\ ii) & \forall\, u \neq 0 \text{ such that } D_\uparrow V(x_0)(u) = 0,\ D_\uparrow^{(2)} V(x_0, u, u) > 0 \end{cases} \tag{6.11}$$

Then x_0 is a weak local minimizer of V *in the sense that for every u in the unit sphere S, there exists $\varepsilon_u > 0$ such that, for any $h \in [0, \varepsilon_u]$ and any $e \in \varepsilon_u B$,*

$$V(x_0) \ \leq\ V(x_0 + hu + h^2 u + h^2 e)$$

Proof — Assume the contrary: there exist $u \in S$ and sequences of $h_n > 0$ and $e_n \in X$ converging to 0 such that

$$V(x_0 + h_n u + h_n^2 u + h_n^2 e_n) - V(x_0) \ <\ 0$$

By letting $n \to \infty$, we deduce that $D_\uparrow V(x_0)(u) \leq 0$. Together with assumption (6.11) i), this implies that $D_\uparrow V(x_0)(u) = 0$. We then infer from the definition of the second-order contingent epiderivative that

$$D_\uparrow^{(2)} V(x_0, u, u) \ \leq\ 0$$

which is a contradiction of assumption (6.11) ii). □

Remark — A local minimizer is always a weak local one.

Stronger conditions on the second order epiderivative imply stronger conditions. □

6.6.1 Second Order Epiderivatives of Moreau-Yosida Approximations

Let X be a Hilbert space identified with its dual and

$$V : X \mapsto \mathbf{R} \cup \{+\infty\}$$

be a nontrivial lower semicontinuous convex function. We recall that it is the pointwise limit when $\lambda \to 0+$ of $\mathcal{C}^1$ convex functions V_λ, the *Moreau-Yosida approximations*, defined by

$$\forall\, x \in X,\;\; V_\lambda(x) \;:=\; \inf_{y\in Y}\left(V(y) + \frac{1}{2\lambda}\|x-y\|^2\right)$$

This minimization problem has a unique solution $J_\lambda x$:

$$V_\lambda(x) \;=\; V(J_\lambda x) + \frac{1}{2\lambda}\|x - J_\lambda x\|^2$$

Furthermore, $A_\lambda x := (x - J_\lambda x)/\lambda$ is both equal to the gradient of V_λ at x and belongs to a subgradient of V at $J_\lambda x$:

$$A_\lambda x \;=\; V_\lambda'(x) \quad \& \quad A_\lambda x \;\in\; \partial V(J_\lambda x)$$

Finally, we associate with any $p \in \partial V(x)$ the normal cone $N_{\partial V(x)}(p)$ to the subdifferential of V at x at some $p \in \partial V(x)$, equal to

$$\{v \;\in\; X \mid \;\; < p, v > = \; D_\uparrow V(x)(v)\}$$

For instance, when $V = \psi_K$ is the indicator of a closed convex subset K, then

$$V_\lambda(x) \;:=\; d_K^2(x)/2\lambda \;\;\&\;\; J_\lambda x \;=\; \pi_K(x)$$

is the projection of best approximation of x to K and

$$N_{N_K(\pi_K x)}(x - \pi_K x) \;:=\; \{v \;\in\; T_K(\pi_K x) \mid \;\; < x - \pi_K(x), v > = \; 0\}$$

Theorem 6.6.5 *Let X be a Hilbert space identified with its dual and $V : X \mapsto \mathbf{R} \cup \{+\infty\}$ be a nontrivial lower semicontinuous convex function. Then, for any $x \in X$, $\lambda > 0$, for every u_1, u_2, $v_2 \in X$ and*

$$v_1 \;\in\; N_{\partial V}(J_\lambda(x))(A_\lambda(x))$$

we have

$$D_\uparrow^2 V_\lambda(x, u_1, u_2) \;\le\; D_\uparrow^2 V(J_\lambda x, v_1, v_2) + < A_\lambda x, u_2 - v_2 > + \frac{1}{2\lambda}\|u_1 - v_1\|^2$$

In particular, by taking $V = \psi_K$ we obtain

Corollary 6.6.6 *Let K be a closed convex subset of a Hilbert space X. Then, for any u_1, u_2, $v_1 \in T_K(\pi_K(x))$ satisfying*

$$< x - \pi_K(x), v_1 > = 0$$

and $v_2 \in T_K^{(2)}(x, v_1)$, inequality

$$D_\uparrow^2 d_K^2(x, u_1, u_2) \leq 2 < x - \pi_K(x), u_2 - v_2 > + \|u_1 - v_1\|^2$$

holds true.

Proof — We observe that

$$\begin{cases} V_\lambda(x + hu_1 + h^2u_2) \\ \leq V(J_\lambda x + hv_1 + h^2v_2) + \|x - J_\lambda x + h(u_1 - v_1) + h^2(u_2 - v_2)\|^2/2\lambda \\ = V(J_\lambda x + hv_1 + h^2v_2) + \|x - J_\lambda x\|^2/2\lambda + h < (x - J_\lambda x)/\lambda, u_1 - v_1 > \\ +h^2 \left(< (x - J_\lambda x), u_2 - v_2 > /\lambda + \|u_1 - v_1\|^2/2\lambda + \varepsilon(h)\right) \end{cases}$$

where $\varepsilon(h)$ converges to 0 with h.

We now take into account that

$$A_\lambda = (1 - J_\lambda)/\lambda \quad \& \quad D_\uparrow V_\lambda(x)(u_1) = < A_\lambda x, u_1 >$$

and that

$$D_\uparrow V(J_\lambda x)(v_1) =< A_\lambda x, v_1 > \quad \text{whenever} \quad v_1 \in N_{\partial V(J_\lambda(x))}(A_\lambda(x))$$

Therefore, we obtain

$$\begin{cases} \left(V_\lambda(x + hu_1 + h^2u_2) - V_\lambda(x) - hD_\uparrow V_\lambda(x)(u_1)\right)/h^2 \\ \leq \left(V(J_\lambda x + hv_1 + h^2v_2) - V(J_\lambda x) - hD_\uparrow V(J_\lambda x)(v_1)\right)/h^2 \\ + < A_\lambda x, u_2 - v_2 > +\|u_1 - v_1\|^2/2\lambda + \varepsilon(h) \end{cases}$$

We then take the liminf when $h \to 0+$ in the above inequalities to obtain the formula. □

Chapter 7

Graphical & Epigraphical Convergence

Introduction

Here are some basic notions on *graphical limits* of single-valued and/or set-valued maps and *epigraphical limits* (*or epilimits*) of extended functions.

It is one of the basic themes of this book to regard maps, and more generally, set-valued maps, not only as maps from one space to another, but as graphs, in order to characterize them in an intrinsic and symmetric way.

This point of view, which goes back to the prehistory of analysis with Fermat and Descartes dealing with curves rather than functions, has been let aside for a long time. In particular, as far as limits of functions and maps are concerned, generations of mathematicians have been accustomed to deal with many concepts of convergence of functions, from pointwise to uniform, but all based on the fact that a map is a map, and not a graph.

When dealing with limits of maps, either single-valued or set-valued, it is quite advantageous to replace pointwise convergence by *graphical convergence*:

Instead of studying (more or less uniform) limits of images of maps, we consider the limits of their graphs[1]. Then the graphs of

[1]This point of view, regarding maps as graphs, has already been used in

the *upper and lower graphical limits* of a sequence of set-valued maps F_n are the upper and lower limits of the graphs of the F_n's.

One of the main reasons for doing so is to treat on the same footing a map and its inverse. This is quite important in approximation theory, where the problem is to derive pointwise convergence of the inverses from the pointwise convergence of the maps.

Indeed, we already proved an extension of Lax's principle to the general case of inclusions (Theorem 5.4.2) stating that stability, convergence of the data and *consistency* imply the convergence of the solutions:

Consistency of set-valued maps F_n at (x_0, y_0) means simply that (x_0, y_0) belongs to their graphical lower limit.

Hence, a first problem is to investigate the connections between graphical and pointwise convergence.

We then relate in Section 3 the concepts of graphical convergence of set-valued maps to the concepts of epigraphical limits of functions. These concepts have recently been successful in overcoming the failure of pointwise convergence in many problems of calculus of variations, optimization, stochastic programming[2], etc.

The use of this concept is mandatory whenever the order relation of the real line comes into play, as in optimization or Lyapunov stability for instance. In such cases, a function is replaced by the set-valued map obtained by adding to it the positive cone $\mathbf{R}_+$ (for minimization), whose graph is thus the epigraph of the function. Therefore, the convergence of the graphs of such set-valued maps is the convergence of the epigraphs of the associated functions.

We present just a selection of issues on epiconvergence, among which are some formulas dealing with epilimits of sums and composition products of functions (Section 4) and conjugates of epilimits (Section 5.)

We answer in Section 6 the problem of the graphical convergence of gradients of differentiable functions and generalized gradients of lower semicontinuous functions. In the last section, we study the "asymptotic derivatives" of a sequence of functions.

Chapter 5 to build a differential calculus of set-valued maps based on "graphical derivatives."

[2]see for instance the book [30] and its bibliography.

7.1 Graphical Limits

7.1.1 Definitions

We shall use the concepts of upper and lower limits of sets to define *graphical convergence* of set-valued maps.

Definition 7.1.1 (Graphical Convergence) *Let us consider metric spaces X, Y and a sequence of set-valued maps $F_n : X \rightsquigarrow Y$. The set-valued maps* $\mathrm{Lim}^{\sharp}{}_{n\to\infty}F_n$ *and* $\mathrm{Lim}^{\flat}{}_{n\to\infty}F_n$ *from X to Y defined by*

$$\begin{cases} i) & \mathrm{Graph}(\mathrm{Lim}^{\sharp}{}_{n\to\infty}F_n) := \mathrm{Limsup}_{n\to\infty}\mathrm{Graph}(F_n) \\ ii) & \mathrm{Graph}(\mathrm{Lim}^{\flat}{}_{n\to\infty}F_n) := \mathrm{Liminf}_{n\to\infty}\mathrm{Graph}(F_n) \end{cases}$$

are called the (graphical) upper and lower limits *of the set-valued maps F_n respectively.*

If the graphical lower and upper limits coincide, it is called the graphical limit.

To say that (x, y) belongs to the graphical lower limit of set-valued maps F_n amounts to saying that they are *consistent* at (x, y) (see Definition 5.4.1.)

We provide a more explicit characterization of these graphical upper and lower limits, which follows immediately from the definitions:

Proposition 7.1.2 *Consider metric spaces X, Y and a sequence of set-valued maps $F_n : X \rightsquigarrow Y$. Then*

$$y \in \left(\mathrm{Lim}^{\sharp}{}_{n\to\infty}F_n\right)(x)$$

if and only if y is the limit of a subsequence of elements $y_{n'} \in F(x_{n'})$ where $x_{n'}$ converges to x. Furthermore

$$y \in (\mathrm{Lim}^{\flat}{}_{n\to\infty}F_n)(x)$$

if and only if y is the limit of a sequence of elements $y_n \in F(x_n)$, where x_n converges to x.

The pointwise limit f of single-valued maps f_n (when it exists) is a selection of the graphical upper limit of the f_n. *The latter is equal to f when f_n remains in an equicontinuous subset*: Indeed, in this case, any limit of elements $(x_n, f_n(x_n))$ being of the form $(x, f(x))$ belongs to the graph of f.

We point out two useful formulas:

Proposition 7.1.3 *Let us consider a sequence of set-valued maps $F_n : X \rightsquigarrow Y$. Then*

$$\begin{cases} (\mathrm{Lim}^\sharp{}_{n\to\infty} F_n)(x) & \supset & \bigcap_{\varepsilon>0} \mathrm{Limsup}_{n\to\infty} F_n(B(x,\varepsilon)) \\ (\mathrm{Lim}^\flat{}_{n\to\infty} F_n)(x) & \subset & \bigcap_{\varepsilon>0} \mathrm{Liminf}_{n\to\infty} F_n(B(x,\varepsilon)) \end{cases}$$

These formulas can be regarded as relating graphical convergence with some kind of "almost pointwise convergence." But can we compare the graphical convergence of F_n and the "pointwise convergence" of F_n, i.e., the upper and lower limits of the subsets $F_n(x)$? The following statement provides the easy answers.

Proposition 7.1.4 *Let us consider metric spaces X, Y and a sequence of set-valued maps $F_n : X \rightsquigarrow Y$. Then the relations*

$$\begin{cases} \left(\mathrm{Lim}^\sharp{}_{n\to\infty} F_n\right)(x) = \mathrm{Limsup}_{n\to\infty, x_n\to x} F_n(x_n) \\ = \bigcap_{N>0,\eta>0} cl\left(\bigcup_{n\geq N, x_n\in B(x,\eta)} F_n(x_n)\right) \end{cases}$$

and, setting $K_n := \mathrm{Dom}(F_n)$,

$$\begin{cases} \left(\mathrm{Lim}^\flat{}_{n\to\infty} F_n\right)(x) \supset \mathrm{Liminf}_{n\to\infty, x_n\to_{K_n} x} F_n(x_n) \\ = \bigcap_{\varepsilon>0} \bigcup_{N>0,\eta>0} \left(\bigcap_{n\geq N, x_n\in B_{K_n}(x,\eta)} B(F_n(x_n),\varepsilon)\right) \end{cases}$$

hold true.

The missing equality holds true under more assumptions:

Theorem 7.1.5 *Let X and Y be Banach spaces. Consider a sequence of closed set-valued maps $F_n : X \rightsquigarrow Y$ and its lower graphical limit $\mathrm{Lim}^\flat{}_{n\to\infty} F_n$. Let*

$$y_0 \in (\mathrm{Lim}^\flat{}_{n\to\infty} F_n)(x_0)$$

and assume that there exist constants $c > 0$ and $\eta > 0$ such that

$$\begin{cases} \forall\ (x_n, y_n) \ \in \ \mathrm{Graph}(F_n) \cap B((x_0, y_0), \eta), \\ \|DF_n(x_n, y_n)\| \ := \ \sup_{u \in X} \inf_{v \in DF_n(x_n, y_n)(u)} \|v\|/\|u\| \ \leq \ c \end{cases}$$

Then, for any sequence x_n converging to x_0, we have

$$y_0 \ \in \ \mathrm{Liminf}_{n\to\infty} F_n(x_n)$$

Consequently, whenever this assumption is satisfied for every $y_0 \in (\mathrm{Lim}^{\flat}{}_{n\to\infty} F_n)(x_0)$, then

$$(\mathrm{Lim}^{\flat}{}_{n\to\infty} F_n)(x_0) \ = \ \mathrm{Liminf}_{n\to\infty, x_n \to x_0} F_n(x_n)$$

Proof — This is just Theorem 5.4.2 of Chapter 5 applied to the set-valued maps F_n^{-1} with $\alpha = 0$. □

7.1.2 Graphical Convergence of Closed Convex Processes

Let us consider a sequence of closed convex processes F_n from a Banach space X to a Banach space Y and their transposes $F_n^{\star} : Y^{\star} \leadsto X^{\star}$. Graphical lower and upper convergences are exchanged by transposition:

Proposition 7.1.6 *Let X, Y be Banach spaces and $F_n : X \leadsto Y$ be closed convex processes. We supply $X^{\star}$ and $Y^{\star}$ with the weak-$\star$ topologies. Then the graphical lower limit $\mathrm{Lim}^{\flat}{}_{n\to\infty} F_n$ of the sequence $(F_n)_n$ is the transpose of the sequentially weak graphical upper limit of the transposes $F_n^{\star}$:*

$$\mathrm{Lim}^{\flat}{}_{n\to\infty} F_n \ = \ \left(\sigma - \mathrm{Lim}^{\sharp}_{\uparrow}{}_{n\to\infty} F_n^{\star}\right)^{\star}$$

Proof — It follows from Proposition 1.1.8 applied to the closed convex cones $K_n := \mathrm{Graph}(F_n)$:

The lower limit of the K_n's is equal to the negative polar cone of the sequentially weak upper limit of the cones K_n^{-}. This can be translated by saying that the graph of the graphical lower limit is the negative polar of the graph of the sequentially weak graphical upper limit of the transposes $F_n^{\star}$. □

7.1.3 Monotone and Maximal Monotone Maps

Let X be a Hilbert space, the dual of which is identified with X and $F : X \leadsto X$ be a *monotone map*.

If the graphical lower limit $\mathrm{Lim}^\flat{}_{n\to\infty} F_n$ of a sequence of monotone maps F_n is maximal monotone, then it is actually the graphical limit of this sequence:

Proposition 7.1.7 (Convergence of Monotone Maps) *Let X be a Hilbert space and let its dual $X^\star$ be supplied with the weak-$\star$ topology. Consider monotone set-valued maps $F_n : X \leadsto X^\star$ and a maximal monotone map $F : X \leadsto X^\star$.*

If F is contained in the graphical lower limit $\mathrm{Lim}^\flat{}_{n\to\infty} F_n$ of the F_n's, then F is actually the graphical limit of the F_n's.

Proof — We have to prove that the graphical upper limit $\mathrm{Lim}^\sharp{}_{n\to\infty} F_n$ of the set-valued maps F_n is contained in F.

Let p belong to $(\mathrm{Lim}^\sharp{}_{n\to\infty} F_n)(x)$. Hence the pair (x,p) is a cluster point of a sequence of elements (x_n, p_n) of the graph of F_n.

Take now any pair (y,q) in the graph of F. Since

$$\mathrm{Graph}(F) \subset \mathrm{Graph}\left(\mathrm{Lim}^\flat{}_{n\to\infty} F_n\right)$$

we know that there exists a sequence of elements (y_n, q_n) of the graph of F_n converging to (y,q). The monotonicity of the set-valued maps F_n implies the inequalities

$$< p_n - q_n \, , \, x_n - y_n > \geq 0$$

Going to the limit, we deduce that

$$\forall \, (y,q) \in \mathrm{Graph}(F), \; < p - q \, , \, x - y > \geq 0$$

Therefore p belongs to $F(x)$ because of the maximality of the graph of F among monotone graphs. $\square$

7.2 Convergence Theorems

We present in this section two versions of a Convergence Theorem which plays a key role in the study of differential inclusions (see Chapter 10 below.)

Let $a(\cdot)$ be a measurable strictly positive function from an open subset $\Omega \subset \mathbf{R}^n$ to $\mathbf{R}_+$. We denote by $L^1(\Omega; Y, a)$ the space of classes of measurable functions integrable for the measure $a(\omega)d\omega$.

Theorem 7.2.1 (Convergence Theorem I) *Let X be a topological vector space, Y be a finite dimensional vector space and F_n be a sequence of nontrivial set-valued maps from X to Y. We assume that for any $x \in \mathrm{Dom}(F)$, there exists a neighborhood $\mathcal{V}$ of x such that $\bigcup_{n\geq 1} F_n(\mathcal{V})$ is bounded.*

Let us consider measurable *functions x_m and y_m from Ω to X and Y respectively, satisfying:*

for almost all $\omega \in \Omega$ and for all neighborhood $\mathcal{U}$ of 0 in the product space $X \times Y$, there exists $M := M(\omega, \mathcal{U})$ such that

$$\forall\, m \;>\; M,\; (x_m(\omega), y_m(\omega)) \;\in\; \mathrm{Graph}(F_m) + \mathcal{U} \tag{7.1}$$

If we assume that

$$\begin{cases} i) & x_m(\cdot) \text{ converges almost everywhere to a function } x(\cdot) \\ ii) & y_m(\cdot) \;\in\; L^1(\Omega; Y, a) \text{ and converges } \textit{weakly} \text{ in } L^1(\Omega; Y, a) \\ & \textit{to a function } \; y(\cdot) \in L^1(\Omega; Y, a) \end{cases}$$

then for almost all $\omega \in \Omega,\; y(\omega) \in \overline{co}F^\sharp(x(\omega))$.

Proof — Let us recall that in a Banach space ($L^1(\Omega; Y, a)$ in our case), the closure of a convex subset (for the normed topology) coincides with its weak closure for the weakened topology[3]

$$\sigma\left(L^1(\Omega; Y, a), L^\infty(\Omega; Y^\star, a^{-1})\right)$$

We apply this result: for every m, the function $y(\cdot)$ belongs to the weak closure of the convex hull $\mathrm{co}(\{y_p(\cdot)\}_{p\geq m}.)$ It coincides with the

[3]By definition of the weakened topology $\sigma(X, X^\star)$, the continuous linear functionals and the weakly continuous linear functionals coincide. Therefore, the closed half-spaces and weakly closed half-spaces are the same. The Hahn-Banach Separation Theorem, which holds true in Hausdorff locally convex topological vector spaces, states that closed convex subsets are the intersections of the closed half-spaces containing them. Since the weakened topology is locally convex, we then deduce that closed convex subsets and weakly closed convex subsets do coincide. This result is known as *Mazur's Theorem.*

When X is not reflexive, closed convex subsets of the dual supplied with the weak topology $\sigma(X^\star, X)$ may be different from their closure for the weak topology $\sigma(X^\star, X^{\star\star})$.

(strong) closure of $\mathrm{co}(\{y_p(\cdot)\}_{p\geq m})$. Hence we can choose functions

$$v_m(\cdot) \ := \ \sum_{p=m}^{\infty} a_m^p y_p(\cdot) \ \in \ \mathrm{co}(\{y_p(\cdot)\}_{p\geq m})$$

(where the coefficients a_m^p are positive or equal to 0 but for a finite number of them, and where $\sum_{p=m}^{\infty} a_m^p = 1$) *which converge strongly* to $y(\cdot)$ in $L^1(\Omega; Y, a)$. This implies that the sequence $a(\cdot)v_m(\cdot)$ converges strongly to the function $a(\cdot)y(\cdot)$ in $L^1(\Omega; Y)$, since the operator of multiplication by $a(\cdot)$ is continuous from $L^1(\Omega; Y, a)$ to $L^1(\Omega; Y)$.

Thus, there exists another subsequence (again denoted by) $v_m(\cdot)$ such that[4]

$$\text{for almost all } \omega \in \Omega, \ \ a(\omega)v_m(\omega) \ \ \text{converges to } a(\omega)y(\omega)$$

Since the function $a(\cdot)$ is strictly positive, we deduce that

$$\text{for almost all } \omega \in \Omega, \ \ v_m(\omega) \ \ \text{converges to } y(\omega)$$

Let $\omega \in \Omega$ such that $x_m(\omega)$ converges to $x(\omega)$ in X and $v_m(\omega)$ converges to $y(\omega)$ in Y.

By assumption (7.1), we know that for any neighborhood

$$\mathcal{U} \ := \ \mathcal{V} \times \mathcal{W}$$

---

[4] *Strong convergence of a sequence in Lebesgue spaces L^p $(1 \leq p < +\infty)$ implies that some subsequence converges almost everywhere.* Let us consider indeed a sequence of functions f_n converging *strongly* to a function f in L^p. We can associate with it a subsequence f_{n_k} satisfying

$$\|f_{n_k} - f\|_{L^p} \ \leq \ 2^{-k}; \ \ \cdots \ < \ n_k \ < \ n_{k+1} \ < \ \cdots$$

Therefore, the series of integrals

$$\sum_{k=1}^{\infty} \int \|f_{n_k}(\omega) - f(\omega)\|_Y^p d\omega \ < \ +\infty$$

is convergent. The Monotone Convergence Theorem implies that the series

$$\sum_{k=1}^{\infty} \|f_{n_k}(\omega) - f(\omega)\|_Y^p$$

converges almost everywhere. For every ω where this series converges, we infer that the general term converges to 0.

of $(0,0) \in X \times Y$, there exists N such that for all $q \geq N$,

$$x_q(\omega) \in x(\omega) + \mathcal{V}/2 \quad \& \quad (x_q(\omega), y_q(\omega)) \in \mathrm{Graph}(F_q) + \mathcal{U}/2$$

Hence

$$y_q(\omega) \in F_q(x(\omega) + \mathcal{V}) + \mathcal{W}$$

Consequently,

$$\forall\, m \geq N, \; v_m(\omega) \in \overline{co}\left(\bigcup_{q \geq N} (F_q(x(\omega) + \mathcal{V}) + \mathcal{W})\right)$$

Then its limit $y(\omega)$ belongs to these subsets for all N and all neighborhoods of the origin $\mathcal{V}$, $\mathcal{W}$ in X and Y respectively, so that

$$y(\omega) \in \bigcap_{N>0,\, \mathcal{V},\, \mathcal{W}} \overline{co}\left(\bigcup_{q \geq N} (F_q(x(\omega) + \mathcal{V}) + \mathcal{W})\right)$$

By assumption, there exist neighborhoods of the origin $\mathcal{V}$ and $\mathcal{W}$ such that the subsets $F_q(x(\omega) + \mathcal{V}) + \mathcal{W}$ are contained in a bounded subset. Hence

$$y(\omega) \in \bigcap_{N>0} \overline{co}\left(\bigcup_{q \geq N} (F_q(x(\omega) + \mathcal{V}/N) + \mathcal{W}/N)\right)$$

Since the dimension of Y is finite, Lemma 1.1.9 implies that $y(\omega)$ belongs to the closed convex hull of the upper limit of the subsets

$$F_q(x(\omega) + \mathcal{V}/N) + \mathcal{W}/N$$

i.e., the closed convex hull of $F^\sharp(x(\omega))$. $\square$

The "continuous version" of the convergence theorem can be stated for any Banach space in the following way:

Theorem 7.2.2 (Convergence Theorem II) *Let X be a topological vector space, Y a Banach space and F be a nontrivial set-valued map from X to Y. We assume that F* is upper semicontinuous on its domain.

Let us consider measurable *functions* $x_m(\cdot)$ *and* $y_m(\cdot)$ *from* Ω *to* X *and* Y *respectively, satisfying:*

for almost all $\omega \in \Omega$ *and for all neighborhoods* $\mathcal{U}$ *of* 0 *in the product space* $X \times Y$*, there exists* $M := M(\omega, \mathcal{U})$ *such that*

$$\forall\, m > M, \; (x_m(\omega), y_m(\omega)) \in \mathrm{Graph}(F) + \mathcal{U}$$

If we assume that

$$\begin{cases} i) & x_m(\cdot) \text{ converges almost everywhere to a function } x(\cdot) \\ ii) & y_m(\cdot) \in L^1(\Omega; Y, a) \text{ and converges } \textit{weakly} \text{ in } L^1(\Omega; Y, a) \\ & \textit{to a function } y \in L^1(\Omega; Y, a) \end{cases}$$

then

for almost all $\omega \in \Omega$ such that $x(\omega) \in \mathrm{Dom}(F)$, $y(\omega) \in \overline{co}F(x(\omega)).)$

Proof — The proof is analogous to that of the previous theorem with $F_n := F$. With the same arguments, we arrive at the conclusion that for almost all $\omega \in \Omega$ such that $x(\omega) \in \mathrm{Dom}(F)$, all neighborhoods $\mathcal{V}$ and $\mathcal{W}$ of the origin in X and Y respectively,

$$y(\omega) \in \overline{co}(F(x(\omega) + \mathcal{V}) + \mathcal{W})$$

Since F is upper semicontinuous at $x(\omega)$, for every $\varepsilon > 0$, there exists a neighborhood $\mathcal{V}$ such that $F(x(\omega) + \mathcal{V}) \subset F(x(\omega)) + \varepsilon B$. Taking $\mathcal{W} = \varepsilon B$, we obtain

$$y(\omega) \in \overline{co}(F(x(\omega)) + 2\varepsilon B$$

Hence, $\varepsilon > 0$ being arbitrary, $y(\omega)$ belongs to $\overline{co}F(x(\omega))$. □

7.3 Epilimits

7.3.1 Definitions and Elementary Properties

Let us consider a metric space X and a sequence of extended functions

$$V_n : X \mapsto \mathbf{R} \cup \{\pm\infty\}$$

whose domains

$$\mathrm{Dom}(V_n) := \{ x \in X \mid V_n(x) \neq \pm\infty \}$$

are not empty.

Taking into account the order relation of $\mathbf{R}$, as we did in Chapter 6, when we introduced epiderivatives, we associate with each extended function V_n two new set-valued maps $\mathbf{V_{n\uparrow}}$ and $\mathbf{V_{n\downarrow}}$ defined in the following way:

$$\left\{\begin{array}{lll} i) & \mathbf{V_{n\uparrow}} := & \left\{\begin{array}{ll} V_n(x)+\mathbf{R}_+ & \text{if } x \in \mathrm{Dom}(V_n) \\ \emptyset & \text{if } V_n(x) = +\infty \\ \mathbf{R} & \text{if } V_n(x) = -\infty \end{array}\right. \\ ii) & \mathbf{V_{n\downarrow}} := & \left\{\begin{array}{ll} V_n(x)-\mathbf{R}_+ & \text{if } x \in \mathrm{Dom}(V_n) \\ \emptyset & \text{if } V_n(x) = -\infty \\ \mathbf{R} & \text{if } V_n(x) = +\infty \end{array}\right. \end{array}\right.$$

We see at once that

$$\mathrm{Graph}(\mathbf{V_{n\uparrow}}) = \mathcal{E}p(V_n) \ \& \ \mathrm{Graph}(\mathbf{V_{n\downarrow}}) = \mathcal{H}p(V_n)$$

Using the concept of graphical upper and lower limits with these associated set-valued maps, we come up with the concepts of epi and hypo convergence, which are thus obtained by taking upper and lower limits of their epigraphs and hypographs.

Definition 7.3.1 (Epilimits) *Let us consider a metric space X and a sequence of extended functions $V_n : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ whose domains are not empty. We shall say that*

1. — the function $\lim^{\flat}_{\uparrow n\to\infty} V_n$ whose epigraph is the lower limit of the epigraphs of the functions V_n

$$\mathcal{E}p(\lim{}^{\flat}_{\uparrow n\to\infty} V_n) := \mathrm{Liminf}_{n\to\infty}\mathcal{E}p(V_n)$$

is the upper epilimit *of the functions V_n*

2. — the function $\lim^{\sharp}_{\uparrow n\to\infty} V_n$ whose epigraph is the upper limit of the epigraphs of the functions V_n

$$\mathcal{E}p(\lim{}^{\sharp}_{\uparrow n\to\infty} V_n) := \mathrm{Limsup}_{n\to\infty}\mathcal{E}p(V_n)$$

is the lower epilimit *of the functions* V_n

3. — the function $\lim^{\flat}_{\downarrow n\to\infty} V_n$ *whose hypograph is the lower limit of the hypographs of the functions* V_n

$$\mathcal{H}p(\lim^{\flat}_{\downarrow n\to\infty} V_n) \; := \; \mathrm{Liminf}_{n\to\infty}\mathcal{H}p(V_n)$$

is the lower hypo-limit *of the functions* V_n

4. — the function $\lim^{\sharp}_{\downarrow n\to\infty} V_n$ *whose hypograph is the upper limit of the hypographs of the functions* V_n

$$\mathcal{H}p(\lim^{\sharp}_{\downarrow n\to\infty} V_n) \; := \; \mathrm{Limsup}_{n\to\infty}\mathcal{H}p(V_n)$$

is the upper hypo-limit *of the functions* V_n.

If the upper and lower epilimits coincide at some point x_0, *then we shall say that the common value*

$$(\lim_{\uparrow n\to\infty} V_n)(x_0) \; := \; \left(\lim^{\sharp}_{\uparrow n\to\infty} V_n\right)(x_0) \; = \; \left(\lim^{\flat}_{\uparrow n\to\infty} V_n\right)(x_0)$$

is the epilimit *of the sequence of functions* V_n *at* x_0, *and we define the* hypolimit $\lim_{\downarrow n\to\infty} V_n$ *at* x_0 *in the same way.*

The terminology concerning the epilimits seems at odds with the choice of the limits; the *upper* epilimit is associated with the *lower* limit[5]. However, they are consistent in the case of hypo-limits. This is due to the analytical definitions of these epilimits, involving the concepts of Γ-convergence and *lim sup inf* introduced in Definition 5.2.4:

$$\limsup_{x'\to x} \inf_{y'\to y} \phi(x',y') \; := \; \sup_{\varepsilon>0} \inf_{\eta>0} \sup_{x'\in B(x,\eta)} \inf_{y'\in B(y,\varepsilon)} \phi(x',y')$$

The concept of *lim inf sup* is defined in a symmetric way, and the adaptation to sequences (or filters) is straightforward: For instance

$$\limsup_{n\to\infty} \inf_{y'\to y} \phi_n(y') \; := \; \sup_{\varepsilon>0} \inf_{N>0} \sup_{n\geq N} \inf_{y'\in B(y,\varepsilon)} \phi_n(y')$$

[5]This would never have happened had *"Convex Analysis"* been *"Concave Analysis"*, and then, had maximization problems taken precedence over minimization ones. Which is not the case, despite economists who prefer to maximize profits, utilities and all these nice things.

Proposition 7.3.2 *Let us consider a metric space X and a sequence of extended functions $V_n : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ whose domains are not empty. We obtain the following formulas:*

$$\left\{\begin{array}{llll} i) & \left(\mathrm{lim}^{\flat}_{\uparrow n\to\infty} V_n\right)(x_0) & = & \limsup_{n\to\infty} \inf_{x\to x_0} V_n(x) \\ ii) & \left(\mathrm{lim}^{\sharp}_{\uparrow n\to\infty} V_n\right)(x_0) & = & \liminf_{n\to\infty, x\to x_0} V_n(x) \\ iii) & \left(\mathrm{lim}^{\flat}_{\downarrow n\to\infty} V_n\right)(x_0) & = & \liminf_{n\to\infty} \sup_{x\to x_0} V_n(x) \\ iv) & \left(\mathrm{lim}^{\sharp}_{\downarrow n\to\infty} V_n\right)(x_0) & = & \limsup_{n\to\infty, x\to x_0} V_n(x) \end{array}\right.$$

Proof — We shall check these formulas for epigraphical convergence only.

1. — We set $V_{\uparrow}^{\flat} := \mathrm{lim}^{\flat}_{\uparrow n\to\infty} V_n$ and show first that

$$\limsup_{n\to\infty} \inf_{x\to x_0} V_n(x) \leq V_{\uparrow}^{\flat}(x_0)$$

It is enough to consider the case when $V_{\uparrow}^{\flat}(x_0) < +\infty$. To compute the value of $V_{\uparrow}^{\flat}$ at x_0, we use the fact that for every $\lambda \geq V_{\uparrow}^{\flat}(x_0)$, there exist sequences of elements x_n converging to x_0 and λ_n to λ such that $\lambda_n \geq V_n(x_n)$.

Therefore, for all $\varepsilon > 0$ and $\eta > 0$, there exists N such that, for all $n \geq N$, we have

$$\inf_{\|x-x_0\|\leq\eta} V_n(x) \leq V_n(x_n) \leq \lambda_n \leq \lambda + \varepsilon$$

and thus

$$\sup_{n\geq N} \inf_{\|x-x_0\|\leq\eta} V_n(x) \leq \lambda + \varepsilon$$

from which we deduce that

$$\forall\, \lambda \geq V_{\uparrow}^{\flat}(x_0), \; \limsup_{n\to\infty} \inf_{x\to x_0} V_n(x) \leq \lambda$$

and thus, that

$$\limsup_{n\to\infty} \inf_{x\to x_0} V_n(x) \leq V_{\uparrow}^{\flat}(x_0)$$

2. — Conversely, to prove the other inequality, we have to show that for any

$$\lambda_0 \ \geq \ \limsup_{n \to \infty} \inf_{x \to x_0} V_n(x) \ \neq \ +\infty$$

the pair (x_0, λ_0) belongs to the epigraph of $V_\uparrow^\flat$, i.e., to the lower limit of the epigraphs of the functions V_n. But, by the very definition of the infimum, we deduce that for all $\varepsilon > 0$ and $\eta > 0$, there exist N such that, for all $n \geq N$, we can find an element $x_n \in B(x_0, \eta)$ satisfying

$$V_n(x_n) \ \leq \inf_{\|x - x_0\| \leq \eta} V_n(x) + \varepsilon \ \leq \ \sup_{n \geq N} \ \inf_{\|x - x_0\| \ \leq \ \eta} V_n(x) + \varepsilon \ \leq \ \lambda_0 + \varepsilon$$

By taking $\varepsilon = \eta = 1/n$ and setting $\lambda_n := \lambda_0 + 1/n$, we have proved that x_n converges to x_0, λ_n to λ_0 and that $V_n(x_n) \leq \lambda_n$ for all n.

3. — We set $V_\uparrow^\sharp := \lim_\uparrow^\sharp{}_{n \to \infty} V_n$ and prove next that

$$\liminf_{n \to \infty, x \to x_0} V_n(x) \ \leq \ V_\uparrow^\sharp(x_0)$$

It is obvious when $V_\uparrow^\sharp(x_0) = +\infty$. Otherwise, let us estimate any $\lambda \geq V_\uparrow^\sharp(x_0)$. We know that for any $\varepsilon > 0$, $\eta > 0$ and $N > 0$, there exist $n' \geq N$, $(x_{n'}, \lambda_{n'})$ in the epigraph of $V_{n'}$ satisfying

$$\|x_{n'} - x_0\| \ \leq \ \eta \ \& \ V_{n'}(x_{n'}) \ \leq \ \lambda_{n'} \ \leq \ \lambda + \varepsilon$$

We deduce that

$$\inf_{n \geq N, \|x - x_0\| \leq \eta} V_n(x) \ \leq \ \lambda + \varepsilon$$

Hence

$$\liminf_{n \to \infty, x \to x_0} V_n(x) \ \leq \ \lambda$$

By letting $\lambda \to V_\uparrow^\sharp(x_0)$ we obtain

$$\liminf_{n \to \infty, x \to x_0} V_n(x) \ \leq \ V_\uparrow^\sharp(x_0)$$

4. — To show the opposite inequality

$$V_{\uparrow}^{\sharp}(x_0) \leq \liminf_{n\to\infty, x\to x_0} V_n(x)$$

we begin by noting that it is always true when

$$\liminf_{n\to\infty, x\to x_0} V_n(x) = +\infty$$

We conclude by observing that the pair (x_0, λ_1) where

$$\lambda_1 \geq \liminf_{n\to\infty, x\to x_0} V_n(x)$$

belongs to the epigraph of $V_{\uparrow}^{\sharp}$, because, by the very definition of the *lim inf*, we can construct a subsequence of elements (again denoted) (x_n, λ_n) of the epigraphs of V_n converging to (x_0, λ_1). □

It may be useful to store the following inequalities:

$$\begin{cases} i) & \lim^{\sharp}_{\uparrow n\to\infty} V_n \leq \lim^{\flat}_{\uparrow n\to\infty} V_n \leq \lim^{\sharp}_{\downarrow n\to\infty} V_n \\ ii) & \lim^{\sharp}_{\uparrow n\to\infty} V_n \leq \lim^{\flat}_{\downarrow n\to\infty} V_n \leq \lim^{\sharp}_{\downarrow n\to\infty} V_n \end{cases} \tag{7.2}$$

Remark — We have defined the concepts of epilimits from the limits of the epigraphs. Conversely, we can recover the limits of subsets from the epilimits of their indicators:

Proposition 7.3.3 *Let us consider a metric space X and a sequence of subsets $K_n \subset X$ and their indicators ψ_{K_n}. Let $K^\sharp$ and $K^\flat$ denote the upper and lower limits of the K_n's. Then*

$$\begin{cases} i) & \psi_{K^\flat} = \lim^{\flat}_{\uparrow n\to\infty} \psi_{K_n} \text{ is the upper epilimit of } \psi_{K_n} \\ ii) & \psi_{K^\sharp} = \lim^{\sharp}_{\uparrow n\to\infty} \psi_{K_n} \text{ is the lower epilimit of } \psi_{K_n} \end{cases}$$

The proof is left as an exercise. □

From the Inverse Stability Theorem 7.1.5, we can derive criteria which imply some equalities in formulas (7.2.)

Proposition 7.3.4 *Let us consider a metric space X and a sequence of extended lower semicontinuous functions $V_n : X \mapsto \mathbf{R}\cup\{+\infty\}$ whose domains are not empty.*

Let $x_0 \in X$ and assume that there exist constants $c > 0$, $N > 0$ and $\eta > 0$ such that, for all $n \geq N$, $x \in B(x_0, \eta)$, the domains of the contingent epiderivatives $D_\uparrow V_n(x)$ are equal to the whole space and

$$\sup_{\|u\|=1} |D_\uparrow V_n(x)(u)| \leq c \tag{7.3}$$

Then

$$\left(\mathrm{lim}^\flat_{\uparrow n\to\infty} V_n\right)(x_0) = \left(\mathrm{lim}^\sharp_{\downarrow n\to\infty} V_n\right)(x_0)$$

or, equivalently,

$$\limsup_{n\to\infty, x\to x_0} V_n(x) = \limsup_{n\to\infty} \inf_{x\to x_0} V_n(x)$$

Proof — Set $V^\flat_\uparrow := \mathrm{lim}^\flat_{\uparrow n\to\infty} V_n$. We apply Theorem 7.1.5 to the set-valued maps F_n defined by

$$F_n(x) := \mathbf{V}_{n\uparrow}(x)$$

It is easy to check that assumption (7.3) implies the stability assumption of Theorem 7.1.5. This is straightforward when $\lambda_n = V(x_n)$, since in this case

$$\|DF_n(x, \lambda_n)\| := \sup_{\|u\|=1} \inf_{\mu\in DF_n(x,\lambda_n)(u)} |\mu| = \sup_{\|u\|=1} |D_\uparrow V_n(x)(u)| \leq c$$

When $\lambda_n > V(x_n)$, Proposition 6.1.4 implies that

$$\mathrm{Dom}(D_\uparrow V_n(x)) \times \mathbf{R} \subset \mathrm{Graph}(DF_n(x, \lambda_n))$$

so that

$$\|DF_n(x, \lambda_n)\| := \sup_{\|u\|=1} \inf_{\mu\in DF_n(x,\lambda_n)(u)} |\mu| = 0$$

Hence, there exists a constant c such that, for all $n \geq N$, $x \in \mathrm{Dom}(V_n)$ close to x_0 and $\lambda_n \geq V_n(x)$ close to $V^\flat_\uparrow(x_0)$, we have

$$\|DF_n(x, \lambda_n)\| \leq c$$

Now, Theorem 7.1.5 implies that $V^\flat_\uparrow(x_0)$, which by definition belongs to the graphical lower limit of the F_n's, also belongs to the lower limit of the $\mathbf{V}_{n\uparrow}(x)$ when $n \to \infty$ and $x \to x_0$. Therefore:

$$\limsup_{n\to\infty, x\to x_0} V_n(x) \leq V^\flat_\uparrow(x_0)$$

The left hand side of this inequality is equal to $V^\sharp_\downarrow(x_0)$. This, together with inequalities (7.2), implies that $V^\sharp_\downarrow(x_0) = V^\flat_\uparrow(x_0)$. □

7.3.2 Convergence of Infima and Minimizers

We expect that the infimum of an epilimit is closely related to the limit of the infima.

To prove this fact, we recall that an extended function V is said to be *inf-compact or lower semicompact* if for any $\lambda \in \mathbf{R}$, the level set

$$\{x \in X \mid V(x) \leq \lambda\}$$

is relatively compact. A sequence of functions V_n is called *sequentially inf-compact* if for any $\lambda \in \mathbf{R}$, every subsequence of elements

$$\{x_{n'} \in X \mid V_{n'}(x_{n'}) \leq \lambda\}$$

remains in a relatively compact set.

Proposition 7.3.5 (Limits of Infima) *Let us consider a metric space X and a sequence of extended functions $V_n : X \mapsto \mathbf{R} \cup \{+\infty\}$ whose domains are not empty. Then*

$$\limsup_{n\to\infty}(\inf V_n) \leq \inf\left(\lim{}^{\flat}_{\uparrow n\to\infty} V_n\right)$$

If we assume that the functions V_n are sequentially inf-compact, *then*

$$\inf\left(\lim{}^{\sharp}_{\uparrow n\to\infty} V_n\right) \leq \liminf_{n\to\infty}(\inf V_n)$$

Consequently, if V is the epilimit of a sequence of sequentially inf-compact functions V_n, then

$$\begin{cases} i) & \inf V = \lim_{n\to\infty}(\inf V_n) \\ ii) & \operatorname{Limsup}_{n\to\infty}\{y \mid V_n(y) = \inf V_n\} \subset \{x_0 \mid V(x_0) = \inf V\} \end{cases}$$

Proof

1. — Set

$$V_{\uparrow}^{\flat} := \lim{}^{\flat}_{\uparrow n\to\infty} V_n$$

The first inequality holds true whenever $V_{\uparrow}^{\flat} \equiv +\infty$. Otherwise, let $\lambda > \inf V_{\uparrow}^{\flat}$ be fixed and x_0 be chosen such that $V_{\uparrow}^{\flat}(x_0) < \lambda$. We know that for all $\eta > 0$, there exists $N > 0$ such that

$$\sup_{n\geq N}\inf_{x\in B(x_0,\eta)} V_n(x) \leq \lambda$$

Therefore,

$$\sup_{n\geq N}(\inf V_n) \;\leq\; \sup_{n\geq N}\inf_{x\in B(x_0,\eta)} V_n(x) \;\leq\; \lambda$$

Hence it is enough to let N go to ∞ and λ to $V_\uparrow^\flat(x_0)$.

2. — Set

$$V_\uparrow^\sharp \;:=\; \lim{}_\uparrow^\sharp{}_{n\to\infty} V_n$$

Taking a subsequence and keeping the same notations, we may assume that

$$\liminf_{n\to\infty}(\inf V_n) \;=\; \lim_{n\to\infty}(\inf V_n)$$

Consider first the case when

$$\liminf_{n\to\infty}(\inf V_n) \;>\; -\infty$$

Let us choose for any n an element x_n such that

$$V_n(x_n) \;\leq\; \inf V_n + 1/n$$

They remain in a relatively compact subset since the functions V_n are sequentially inf-compact. So a subsequence of x_n does converge to some x_0. Therefore,

$$\begin{cases} \inf V_\uparrow^\sharp \;\leq\; V_\uparrow^\sharp(x_0) \;=\; \liminf_{n\to\infty, x\to x_0} V_n(x) \\ \\ \leq\; \liminf_{n\to\infty} V_n(x_n) \;=\; \liminf_{n\to\infty}(\inf V_n) \end{cases}$$

When

$$\liminf_{n\to\infty}(\inf V_n) \;=\; -\infty$$

we can choose $x_n \in X$ in such a way that $V_n(x_n) \to -\infty$. Sequential inf-compactness thus implies that a subsequence of x_n converges to some x_0. Then, for all $\mu > 0$, the pair $(x_0, -\mu)$ belongs to the upper limit of the epigraphs of the V_n's, i.e.,

$$(\lim{}_\uparrow^\sharp{}_{n\to\infty} V_n)(x_0) \;=\; -\infty$$

3. — Equality *i)* is a consequence of the first two inequalities. Observe that the last statement holds true if the left-hand side of inclusion *ii)* is empty. Otherwise, consider an element

$$x_0 \;\in\; \mathrm{Limsup}_{n\to\infty}\{y \mid V_n(y) \;=\; \inf V_n\}$$

which is the limit of a subsequence of elements $x_{n'}$ such that $V_{n'}(x_{n'}) = \inf V_{n'}$. Then *i)* implies that

$$V_{\uparrow}^{\sharp}(x_0) \;\leq\; \liminf_{n'\to\infty} V_{n'}(x_{n'}) \;=\; \inf V$$

so that $V_{\uparrow}^{\sharp}(x_0) \leq \inf V$. Since $V(x_0) = V_{\uparrow}^{\sharp}(x_0)$ by assumption, we have proved that x_0 minimizes V. $\square$

Unfortunately, there are counter-examples for the property that the set of minimizers of the upper epilimit is the lower limit of the sets of minimizers of the functions V_n. However, Theorem 7.1.5 provides some results about the lower limits of level sets:

Proposition 7.3.6 *Let us consider a metric space X and a sequence of extended lower semicontinuous functions $V_n : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ whose domains are not empty.*

Let x_0 belong to the domain of their upper epilimit and assume that there exist $N > 0$ and constants $c > 0$, $\eta > 0$ such that, for all $n \geq N$, $x \in B(x_0, \eta) \cap \mathrm{Dom}(V_n)$,

$$\left\{ \begin{array}{lllll} i) & \exists\, u_n^- \in cB_X & \textit{such that} & D_{\uparrow}V_n(x)(u_n^-) & = \; -1 \\ & & & & \\ ii) & \exists\, u_n^+ \in cB_X & \textit{such that} & D_{\uparrow}V_n(x)(u_n^+) & = \; +1 \end{array} \right. \qquad (7.4)$$

Then there exists a constant l such that, for any sequence of elements x_{0n} converging to x_0 such that $V_n(x_{0n})$ converges to $\left(\mathrm{lim}^{\flat}_{\uparrow n\to\infty} V_n\right)(x_0)$, and for any λ_n converging to $V_{\uparrow}^{\flat}(x_0)$, there exist $\widehat{x}_n \in X$ satisfying

$$V_n(\widehat{x}_n) \;\leq\; \lambda_n, \;\; \|\widehat{x}_n - x_{0n}\| \;\leq\; l|\lambda_n - V_n(x_{0n})|$$

Proof — Set

$$V_{\uparrow}^{\flat} \;:=\; \mathrm{lim}^{\flat}_{\uparrow n\to\infty} V_n$$

We apply Theorem 5.4.2 to the set-valued maps $F_n(x) := \mathbf{V}_{n\uparrow}(x)$, which are obviously consistent at $(x_0, V_{\uparrow}^{\flat}(x_0))$ by definition of $V_{\uparrow}^{\flat}$. Set $\lambda_0 = V_{\uparrow}^{\flat}(x_0)$.

Assumption (7.4) implies that the set-valued maps F_n are stable: There exist constants $\eta > 0$ and $c > 0$ such that, for any

$$(x_n, \mu_n) \;\in\; \mathrm{Graph}(F_n) \cap B((x_0, \lambda_0), \eta)$$

for any $v \in \mathbf{R}$, there exists $u_n \in X$ such that

$$v \in DF_n(x_n, \mu_n)(u_n) \ \& \ \|u_n\| \le c|v|$$

Indeed, if $\mu_n > V_n(x_n)$, we can take $u_n = 0$. If $\mu_n = V_n(x_n)$, we take $v = vu_n^+$ if $v \ge 0$ and $v = -vu_n^-$ if $v < 0$.

Therefore, Theorem 5.4.2 implies that for any λ_n approximating λ_0, we can find $\widehat{x}_n \in F_n^{-1}(\lambda_n)$ satisfying the required error estimates. □

Remark — Observe that stability assumptions (7.4) imply that for all $n \ge N$ and all $x_n \in B(x_0, \eta)$,

$$\inf_{x \in X} V_n(x) < V_n(x_n)$$

since the Fermat rule is violated. This excludes the case when the level sets are set of minimizers. □

7.3.3 Variational Systems

Naturally, we can introduce the same definitions for "continuous" parameters $u \in U$, where U is another metric space.

In such a framework, we consider a family of extended functions

$$\forall u \in U, \ V(u, \cdot) : X \mapsto \mathbf{R} \cup \{\pm\infty\}$$

depending upon the parameter u, with which we associate the set-valued maps

$$\left\{ \begin{array}{lll} i) & \mathbf{V}(u)_\uparrow(x) := & \left\{ \begin{array}{ll} V(u,x) + \mathbf{R}_+ & \text{if } (u,x) \in \mathrm{Dom}(V) \\ \emptyset & \text{if } V(u,x) = +\infty \\ \mathbf{R} & \text{if } V(u,x) = -\infty \end{array} \right. \\ ii) & \mathbf{V}(u)_\downarrow(x) := & \left\{ \begin{array}{ll} V(u,x) - \mathbf{R}_+ & \text{if } (u,x) \in \mathrm{Dom}(V) \\ \emptyset & \text{if } V(u,x) = -\infty \\ \mathbf{R} & \text{if } V(u,x) = +\infty \end{array} \right. \end{array} \right.$$

There is a subtle, but important, difference between the extended function

$$V : (u, x) \in U \times X \mapsto \mathbf{R} \cup \{\pm\infty\}$$

for which the variables u and x are on the same footing, and the set-valued maps

$$\mathbf{V}(\cdot)_\uparrow : u \in U \rightsquigarrow \mathbf{V}(u)_\uparrow$$

for which the variables play a different role: u, the role of a parameter and x the role of the variable.

This type of situation occurs for instance in optimization where the order relation on $\mathbf{R}$ involves the variable x, when we minimize the function depending on the parameter u with respect to x.

To emphasize this difference, we borrow from Rockafellar and Wets the following terminology: a set-valued map $\mathbf{V}(\cdot)_\uparrow$ is called a *variational system.*

Definition 7.3.7 (Variational system) *Let us consider metric spaces X, U and a variational system $\mathbf{V}(\cdot)_\uparrow$ defined on $U \times X$. We shall say that $V(\cdot,\cdot)$ (or $\mathbf{V}(\cdot)_\uparrow$, to be precise), is*

1. — upper epicontinuous *at u if the set-valued map $u' \rightsquigarrow \mathbf{V}(u')_\uparrow$ is lower semicontinuous at u*

2. — lower epicontinuous *at u if $\mathbf{V}(u)_\uparrow = \limsup_{u' \to u} \mathbf{V}(u')_\uparrow$*

3. — epicontinuous *at u if it is both lower and upper epicontinuous at u.*

The definition of upper and lower hypocontinuity are naturally symmetric.

As in the discrete case, these concepts can be characterized in an analytic way:

Proposition 7.3.8 *Let us consider metric spaces X and U and a variational system $\mathbf{V}(\cdot)_\uparrow$ defined on $U \times X$. Then V is*

1. — *upper epicontinuous at u if and only if for any $x \in \mathrm{Dom}(V(u,\cdot))$*

$$V(u,x) \;:=\; \limsup_{u' \to u} \inf_{x' \to x} V(u',x')$$

2. — *lower epicontinuous at u if and only if for any $x \in \mathrm{Dom}(V(u,\cdot))$*

$$V(u,x) \;=\; \liminf_{u' \to u, x' \to x} V(u',x')$$

3. — *epicontinuous at u if and only if for any $x \in \mathrm{Dom}(V(u,\cdot))$*

$$V(u,x) \;=\; \limsup_{u' \to u} \inf_{x' \to x} V(u',x') \;=\; \liminf_{u' \to u, x' \to x} V(u,x)$$

Example — Definition 6.1.2 of epiderivatives provides another interpretation of the epiderivatives in terms of epilimits of the difference quotients

$$\nabla_h V(x) \;:=\; u \;\mapsto\; \frac{V(x+hu) - V(x)}{h}$$

1. — the contingent epiderivative is the lower epilimit of the difference quotients $\nabla_h V(x)$ when $h \to 0+$

2. — the adjacent epiderivative is the upper epilimit of the difference quotients $\nabla_h V(x)$ when $h \to 0+$

3. — the circatangent epiderivative is the upper epilimit of the difference quotients

$$u \mapsto (V(x' + hu) - \lambda')/h$$

when $h \to 0+$ and $(x', \lambda') \in \mathcal{E}p(V)$ converges to $(x, V(x))$. $\square$

7.4 Epilimits of Sums and Composition Products

This section is devoted to the upper and lower epilimits of the functions

$$U_n := V_n + W_n \circ A$$

in terms of the upper and lower epilimits of the functions V_n and W_n.

Theorem 7.4.1 *Let us consider Banach spaces X, Y, a continuous linear operator $A \in \mathcal{L}(X, Y)$, two sequences of extended functions V_n and W_n from X and Y to $\mathbf{R} \cup \{+\infty\}$ respectively and define*

$$U_n := V_n + W_n \circ A$$

1. — Then for every $x_0 \in X$, the inequality

$$(\lim{}^{\sharp}_{\uparrow n\to\infty} V_n)(x_0) + (\lim{}^{\sharp}_{\uparrow n\to\infty} W_n)(Ax_0) \leq (\lim{}^{\sharp}_{\uparrow n\to\infty} U_n)(x_0) \quad (7.5)$$

holds true whenever the left hand side is well defined.

2. — If in addition V_n and W_n are lower semicontinuous and the following stability assumption *holds true:*

there exist constants $c > 0$, $\alpha \in [0, 1[$ and $\eta > 0$ such that, for all n,

$$\left\{\begin{array}{ll} i) & \forall\, x \in \mathrm{Dom}(V_n) \cap B(x_0, \eta), \ \forall\, y \in \mathrm{Dom}(W_n) \cap B(Ax_0, \eta) \\ & B_Y \subset A\left(\mathrm{Dom}(D_\uparrow^\flat V_n(x)) \cap cB_X\right) - \mathrm{Dom}(D_\uparrow W_n(y)) + \alpha B_Y \\ ii) & \sup_{u \in \mathrm{Dom}(D_\uparrow^\flat V_n(x))} |D_\uparrow^\flat V_n(x)(u)|/\|u\| \leq c \\ iii) & \sup_{v \in \mathrm{Dom}(D_\uparrow W_n(y))} |D_\uparrow W_n(y)(v)|/\|v\| \leq c \end{array}\right. \quad (7.6)$$

then the upper epilimit $\lim^{\flat}_{\uparrow n\to\infty} U_n$ *satisfies:*

$$(\lim^{\flat}_{\uparrow n\to\infty} U_n)(x_0) \;\leq\; (\lim^{\flat}_{\uparrow n\to\infty} V_n)(x_0) + (\lim^{\flat}_{\uparrow n\to\infty} W_n)(Ax_0)$$

whenever the right hand side is well defined.

Consequently, if the sequences of functions V_n *and* W_n *have epilimits* $V_\uparrow$ *and* $W_\uparrow$ *at* x_0 *and* Ax_0 *respectively and if* $V_\uparrow(x_0) + W_\uparrow(Ax_0)$ *is well defined, then the functions* U_n *have an epilimit* $U_\uparrow$ *at* x_0 *and*

$$U_\uparrow(x_0) \;=\; V_\uparrow(x_0) + W_\uparrow(Ax_0)$$

Proof — For simplicity, we adopt the notations

$$V^{\sharp}_{\uparrow} \;:=\; \lim^{\sharp}_{\uparrow n\to\infty} V_n \quad \& \quad V^{\flat}_{\uparrow} \;:=\; \lim^{\flat}_{\uparrow n\to\infty} V_n$$

for the functions V_n as well as for the functions U_n and W_n.

Since inequality (7.5) holds true when one of the values $V^{\sharp}_{\uparrow}(x_0)$ and $W^{\sharp}_{\uparrow}(Ax_0)$ is equal to $-\infty$, or $U^{\sharp}_{\uparrow}(x_0)$ is equal to $+\infty$, we have to check the formula when the two first values are larger than $-\infty$ and the last one is smaller than $+\infty$. Then, by definition of the *lim inf*, there is a finite number ρ, an integer N and a positive number η such that

$$\forall\, n \;\geq\; N,\;\; \forall\, x \;\in\; B(x_0,\eta),\; V_n(x) \;\geq\; \rho,\;\; W_n(Ax) \;\geq\; \rho$$

By definition of the lower epilimit, the pair $\left(x_0, U^{\sharp}_{\uparrow}(x_0)\right)$ is the limit of a subsequence (again denoted by) (x_n, c_n) satisfying

$$V_n(x_n) + W_n(Ax_n) \;\leq\; c_n$$

The two above inequalities imply that the sequences of real numbers $a_n := V_n(x_n)$ and $b_n := c_n - a_n$ are bounded. Hence, subsequences (again denoted) a_n and b_n do converge to a and b satisfying

$$U^{\sharp}_{\uparrow}(x_0) \;=\; a+b,\;\; a \;\geq\; V^{\sharp}_{\uparrow}(x_0),\;\; b \;\geq\; W^{\sharp}_{\uparrow}(Ax_0)$$

To prove the second statement, we begin by observing that if we set

$$\begin{cases} i) & K_n \;:=\; \mathcal{E}p(V_n)\times\mathcal{E}p(W_n)\times\mathbf{R} \;\subset\; X\times\mathbf{R}\times Y\times\mathbf{R}\times\mathbf{R} \\[2ex] ii) & G(x,a,y,b,c) \;:=\; (Ax-y, a+b-c) \\[2ex] iii) & H(x,a,y,b,c) \;:=\; (x,c) \end{cases}$$

then we can write

$$\mathcal{E}p(U_n) \;=\; H(K_n\cap G^{-1}(0,0))$$

Therefore, it is sufficient to show that the lower limit of the subsets $K_n \cap G^{-1}(0,0)$ contains the intersection of the lower limit of the subsets K_n with $G^{-1}(0,0)$. For that purpose, we shall use Theorem 3.4.5.

Let us consider any sequence of elements

$$(x_{0n}, a_{0n}) \in \mathcal{E}p(V_n), \ (y_{0n}, b_{0n}) \in \mathcal{E}p(W_n)$$

converging respectively to $(x_0, V_\uparrow^\flat(x_0))$ and $(Ax_0, W_\uparrow^\flat(Ax_0))$ and let us set $c_{0n} := a_{0n} + b_{0n}$.

We observe that the elements $(x_{0n}, a_{0n}, y_{0n}, b_{0n}, c_{0n})$ belong to K_n and that

$$G(x_{0n}, a_{0n}, y_{0n}, b_{0n}, c_{0n}) = (Ax_{0n} - y_{0n}, 0)$$

converges to $(0,0)$.

We begin by checking that the assumptions (7.6) of our Theorem imply the stability assumption (3.9) of Theorem 3.4.5, i.e., that there exists a constant $c > 0$ such that, for all n,

$$\forall\, (x,a,y,b,c) \in K_n \ \text{ close to } \ \left(x_0, V_\uparrow^\flat(x_0), y_0, W_\uparrow^\flat(y_0), 0\right)$$

for all $(z,\lambda) \in X \times \mathbf{R}$, there exist $(u,\mu,v,\nu,\delta) \in T_{K_n}(x,a,y,b,c)$ and $e \in X$ satisfying

$$\begin{cases} i) & z = Au - v + e \ \ \& \ \ \lambda = \mu + \nu - \delta \\ ii) & \|e\| \leq \alpha(\|z\| + |\lambda|) \\ iii) & \|u\| + \|v\| + |\mu| + |\nu| + |\delta| \leq c(\|z\| + |\lambda|) \end{cases}$$

Assumptions (7.6) $i)$ and $ii)$ imply right away that there exist

$$u \in \mathrm{Dom}(D_\uparrow^\flat V_n(x)), \ \ v \in \mathrm{Dom}(D_\uparrow W_n(y)) \ \ \& \ \ e \in X$$

satisfying

$$z = Au - v + e, \ \ \|e\| \leq \alpha(\|z\| + |\lambda|), \ \ \|u\| \leq c\|z\|$$

Then

$$\|v\| \leq \|z\| + \|e\| + \|A\|\|u\| \leq \gamma(\|z\| + \lambda)$$

for some positive γ independent of (z,λ).

Let us take now

$$\mu := c\|u\|, \ \ \nu := c\|v\|, \ \ \delta := c(\|u\| + \|v\|) - \lambda$$

We deduce from (7.6) $iii)$ that

$$(u,\mu) \in \mathcal{E}p(D_\uparrow V_n)(x), \ \ (v,\nu) \in \mathcal{E}p(D_\uparrow W_n)(y)$$

and that

$$D_\uparrow V_n(x)(u) + D_\uparrow W_n(y)(v) \;\leq\; c(\|u\| + \|v\|) \;=\; \lambda + \delta$$

Consequently

$$|\delta| \;\leq\; |\lambda| + c(\|u\| + \|v\|) \;\leq\; c'(\|z\| + |\lambda|)$$

The conclusion of Theorem 3.4.5 being available, we then know that there exist elements $(\widehat{x}_n, \widehat{a}_n, \widehat{y}_n, \widehat{b}_n, \widehat{c}_n) \in K_n$ satisfying

$$G(\widehat{x}_n, \widehat{a}_n, \widehat{y}_n, \widehat{b}_n, \widehat{c}_n) \;=\; 0$$

and

$$\|\widehat{x}_n - x_{0n}\| + \|\widehat{y}_n - y_{0n}\| + |\widehat{a}_n - a_{0n}| + |\widehat{b}_n - b_{0n}| + |\widehat{c}_n - c_{0n}| \;\leq\; l\|Ax_{0n} - y_{0n}\|$$

Hence $\lim_{n\to\infty} |\widehat{a}_n - a_{0n}| = \lim_{n\to\infty} |\widehat{b}_n - b_{0n}| = 0$, so that $\widehat{a}_n$ and $\widehat{b}_n$ converge to $V_\uparrow^\flat(x_0)$ and $W_\uparrow^\flat(x_0)$ respectively. Furthermore, inequalities $U_n(\widehat{x}_n) \leq \widehat{a}_n + \widehat{b}_n$ imply that $U_\uparrow^\flat(x_0) \leq V_\uparrow^\flat(x_0) + W_\uparrow^\flat(Ax_0)$. □

7.5 Conjugate Functions of Epilimits

Let us consider a Banach space X, a sequence of nontrivial convex lower semicontinuous functions

$$V_n : X \;\mapsto\; \mathbf{R} \cup \{+\infty\}$$

and their conjugate functions

$$V_n^\star : X^\star \;\mapsto\; \mathbf{R} \cup \{+\infty\}$$

The purpose of this section is to characterize the upper epilimit of the functions V_n.

We can still define conjugate of an extended function

$$U : X \;\mapsto\; \mathbf{R} \cup \{\pm\infty\}$$

by the formula

$$U^\star(p) \;=\; \sup_{x\in X}(< p, x > - U(x))$$

We observe that $U^\star \equiv +\infty$ when $U(x) = -\infty$ for some $x \in X$ and that if $U \equiv +\infty$, its conjugate $U^\star \equiv -\infty$.

When $X^\star$ is supplied with the weak-$\star$ topology, we naturally adapt the definition of lower epilimit of a sequence of extended functions $W_n : X^\star \mapsto \mathbf{R} \cup \{\pm\infty\}$ to the concept of sequentially weak lower epilimit: *The epigraph of* $\sigma - \lim^{\sharp}_{\uparrow n\to\infty} W_n$ *is by definition the sequentially weak upper limit of the epigraphs of the functions* W_n.

Theorem 7.5.1 (Mosco) *Consider a Banach space X and a sequence of nontrivial convex lower semicontinuous functions*

$$V_n : X \mapsto \mathbf{R} \cup \{+\infty\}$$

Assume that the upper epilimit of the functions V_n is not trivial. Then it is equal to the conjugate of the sequentially weak lower epilimit of the sequence of conjugate functions $V_n^\star$:

$$\lim^{\flat}_{\uparrow n\to\infty} V_n \;=\; \left(\sigma - \lim^{\sharp}_{\uparrow n\to\infty} V_n^\star\right)^\star$$

Proof — We set

$$V^\flat_\uparrow \;:=\; \lim^{\flat}_{\uparrow n\to\infty} V_n, \quad W \;:=\; \sigma - \lim^{\sharp}_{\uparrow n\to\infty} V_n^\star$$

Then for any $p \in X^\star$, there exists a subsequence of elements $(p_{n'}, \lambda_{n'}) \in \mathcal{E}p(V_{n'}^\star)$ converging weakly to $(p, W(p))$.

We observe first that inequality

$$W^\star(x) \;\leq\; V^\flat_\uparrow(x)$$

always holds true: Indeed, it holds true when $V^\flat_\uparrow(x) = +\infty$ or $W \equiv +\infty$, because, in this case, $W^\star \equiv -\infty$. Assume now that $V^\flat_\uparrow(x) < +\infty$ and pick p such that $W(p) < +\infty$. Then there exist sequences of elements x_n and $\mu_n \geq V_n(x_n)$ converging (strongly) to x and $V^\flat_\uparrow(x)$ respectively and subsequences $p_{n'}$ and $\lambda_{n'} \geq V_{n'}^\star(p_{n'})$ converging weakly to p and $W(p)$ respectively. Therefore, by taking the limit when $n' \to \infty$, Fenchel inequalities

$$< p_{n'}, x_{n'} > \leq \; V_{n'}(x_{n'}) + V_{n'}^\star(p_{n'}) \;\leq\; \mu_{n'} + \lambda_{n'}$$

imply

$$< p, x > -W(p) \;\leq\; V^\flat_\uparrow(x)$$

from which our claim follows.

If for some $q \in X^\star$, $W(q) = -\infty$, then for all $x \in X$, $W^\star(x) = +\infty$. Consequently, in this case, inequality

$$V_\uparrow^\flat(x) \leq W^\star(x) \tag{7.7}$$

is satisfied. It still holds true when $V_\uparrow^\flat(x) = -\infty$.

It remains to prove that it is verified when W never takes the value $-\infty$ and $V_\uparrow^\flat(x) > -\infty$. To this end, we shall check that $\lambda \leq W^\star(x)$ for all $\lambda < V_\uparrow^\flat(x)$. Fix such (x, λ). The latter statement means that the pair (x, λ) does not belong to the epigraph of $V_\uparrow^\flat$, which is, by definition, the lower limit of the epigraphs of functions V_n. Therefore, there exist $\varepsilon > 0$ and a subsequence of functions (again denoted by) V_n such that

$$((x + \varepsilon B) \times (\lambda + \varepsilon[-1, +1])) \cap \mathcal{E}p(V_n) = \emptyset$$

We infer from the Hahn-Banach Separation Theorem that there exist elements $(p_n, -\alpha_n) \in X^\star \times \mathbf{R}$ satisfying $\|p_n\| + |\alpha_n| = 1$ and

$$\forall\ (x_n, \mu_n) \in \mathcal{E}p(V_n),\ \langle p_n, x_n\rangle - \alpha_n\mu_n \leq \langle p_n, x\rangle - \alpha_n\lambda - \varepsilon \tag{7.8}$$

This inequality implies that $\alpha_n \geq 0$ because, if $\alpha_n < 0$, it is enough to take $x_n \in \mathrm{Dom}(V_n)$ and let

$$\mu_n^k := V_n(x_n) + k \to +\infty \text{ with } k$$

to derive the contradiction $\infty \leq cst$.

We proceed now in three steps.

1. — Assume that $\liminf_{n\to\infty} \alpha_n$ is strictly positive. Then there exist $\delta > 0$ and a subsequence (again denoted by) α_n satisfying $\alpha_n \geq \delta$. Therefore, we can divide inequality (7.8) by $\alpha_n > 0$. Setting $q_n := p_n/\alpha_n$ and taking the supremum over the $x_n \in \mathrm{Dom}(V_n)$, we deduce that,

$$V_n^\star(q_n) \leq < q_n, x > -\lambda - \varepsilon/\alpha_n \leq < q_n, x > -\lambda$$

On the other hand, $\|q_n\| \leq 1/\delta$, so that there exists a subsequence (again denoted) q_n converging weakly to some cluster point q. Since $< q_n, x > -\lambda$ converges to $< q, x > -\lambda$, we infer that

$$W(q) \leq \langle q, x\rangle - \lambda$$

and thus, that $\lambda \leq W^\star(x)$.

2. — This happens at least when x belongs to the domain of $V_\uparrow^\flat$, because, in this case, we claim that $\alpha_n > \varepsilon/4\rho$ for n large enough (where $\rho := V_\uparrow^\flat(x) - \lambda > 0$.) Indeed, there exist sequences of elements x_{0n} and $\mu_n \geq V(x_{0n})$ converging strongly to x and $V_\uparrow^\flat(x)$ respectively. Therefore, for n large enough, $\|x - x_{0n}\| < \varepsilon/2$ and $\mu_n \leq V_\uparrow^\flat(x) + \rho$. By taking such x_{0n} and μ_n in inequality (7.8), we obtain

$$\begin{cases} 0 \leq \langle p_n, x - x_{0n}\rangle + \alpha_n(\mu_n - \lambda) - \varepsilon \\ \\ \leq \|x - x_{0n}\| + \alpha_n(V_\uparrow^\flat(x) - \lambda + \rho) - \varepsilon \leq 2\alpha_n\rho - \varepsilon/2 \end{cases}$$

Therefore, inequality (7.7) holds true whenever $V_\uparrow^\flat(x) < +\infty$. Furthermore, we deduce from the first step that *the domain of W is not empty.*

3. — Consider now the case when $V_\uparrow^\flat(x) = +\infty$.

Either $\liminf_{n\to\infty} \alpha_n > 0$ and, by the first step, $\lambda \leq W^\star(x)$, or $\liminf_{n\to\infty} \alpha_n = 0$, and a subsequence (again denoted by) α_n converges to 0.

In this case, we take an element q in the domain of W, which is not empty by the second step. Then there exist subsequences (again denoted) q_n and $\rho_n \geq V_n^\star(q_n)$ which converge weakly to some q and $W(q)$ respectively. Next, we add Fenchel's inequality

$$\langle q_n, x_n\rangle \leq V_n(x_n) + V_n^\star(q_n)$$

to inequalities (7.8) multiplied by $\mu > 0$ and we obtain

$$\begin{cases} \forall\, x_n \in \mathrm{Dom}(V_n),\ < q_n + \mu p_n, x_n > \\ \\ \leq (1 + \mu\alpha_n)V_n(x_n) + \rho_n + < \mu p_n, x > -\mu\alpha_n\lambda - \mu\varepsilon \end{cases}$$

Dividing by $1 + \mu\alpha_n$ and taking the supremum over the domain of V_n, yields

$$V_n^\star\left(\frac{q_n + \mu p_n}{1 + \mu\alpha_n}\right) \leq \frac{\rho_n + < \mu p_n, x > -\mu\alpha_n\lambda - \mu\varepsilon}{1 + \mu\alpha_n}$$

On the other hand, knowing that $\|p_n\| \leq 1$, there exists still a subsequence (again denoted by) p_n converging weakly to some cluster point p. The right hand side of the above inequality converges to $W(q) + < \mu p, x > -\mu\varepsilon$. Hence, by going to the limit, we infer that

$$W(q + \mu p) \leq W(q) + < \mu p, x > -\mu\varepsilon$$

Therefore, taking into account that

$$< q + \mu p, x > - W(q + \mu p) \leq W^\star(x)$$

we obtain

$$\mu\varepsilon - W(q) + < q, x > \leq W^\star(x)$$

Since $\varepsilon > 0$, we deduce that $W^\star(x) \geq \lambda$ by taking μ large enough. Therefore, in both cases, $\lambda \leq W^\star(x)$, so that, by letting $\lambda \rightarrow +\infty$, we infer that $W^\star(x) = +\infty$. $\square$

By taking the conjugates, we obtain the formula

$$(\mathrm{lim}^\flat_{\uparrow n\rightarrow\infty} V_n)^\star = \left(\sigma - \mathrm{lim}^\sharp_{\uparrow n\rightarrow\infty} V_n^\star\right)^{\star\star}$$

Applying it to the indicators $V_n := \psi_{K_n}$ of closed convex subsets K_n, the conjugate of which are the support functions of these subsets, we obtain a formula for the support function of the lower limit of a sequence of subsets K_n:

•

Corollary 7.5.2 *Let X be a Banach space and $K_n \subset X$ be a sequence of nonempty closed convex subsets. Then the support function of their lower limit $K^\flat$ is the biconjugate of the sequentially weak lower limit of the support functions σ_{K_n} of the subsets K_n:*

$$\sigma(K^\flat, \cdot) = (\sigma - \mathrm{lim}^\sharp_{\uparrow n\rightarrow\infty} \sigma_{K_n}(\cdot))^{\star\star}$$

When the space X is reflexive, we can apply Theorem 7.5.1 to the conjugate functions $V_n^\star$ to obtain the following consequence:

Theorem 7.5.3 *Consider a reflexive Banach space X and a sequence of nontrivial convex lower semicontinuous functions $V_n : X \mapsto \mathbf{R} \cup \{+\infty\}$.*

Assume that the upper epilimit of the conjugates $V_n^\star$ is not trivial. Then the conjugate of the sequentially weak lower epilimit of the sequence of functions V_n is equal to the upper epilimit of the conjugates $V_n^\star$:

$$\left(\sigma - \lim{}^{\sharp}_{\uparrow n\to\infty} V_n\right)^\star \;=\; \lim{}^{\flat}_{\uparrow n\to\infty} V_n^\star$$

Consequently, we deduce the following:

Corollary 7.5.4 *Consider a reflexive Banach space X and a sequence of nontrivial lower semicontinuous convex functions*

$$V_n : X \;\mapsto\; \mathbf{R} \cup \{+\infty\}$$

Let us assume that its upper epilimit and the sequentially weak lower epilimit do coincide with a nontrivial function V.

Then for any $x \in \mathrm{Dom}(V)$ and $p \in \mathrm{Dom}(V^\star)$, there exist sequences $x_n \in X$ and $p_n \in X^\star$ converging to x and p respectively and satisfying

$$\limsup_{n\to\infty} V_n(x_n) \;\le\; V(x), \quad \limsup_{n\to\infty} V_n^\star(p_n) \;\le\; V^\star(p)$$

Applying Theorem 7.5.3 to the indicators $V_n := \psi_{K_n}$ of subsets K_n, we obtain a formula for the support function of the sequentially weak upper limit of subsets:

Corollary 7.5.5 *Let K_n be a sequence of nonempty closed convex subsets of a reflexive Banach space X. Then the support function of the sequentially weak upper limit of the subsets K_n is equal to the upper epilimit of the support functions of the subsets K_n.*

7.6 Graphical Convergence of Gradients

We study in this section the convergence properties of gradients and generalized gradients of a sequence of functions V_n.

7.6.1 Convergence of Gradients of Smooth Functions

We begin with

Theorem 7.6.1 *Let Ω be an open subset of a Banach space X. Consider a sequence of Fréchet differentiable functions $V_n : \Omega \mapsto \mathbf{R}$ converging to a function V uniformly on bounded subsets of Ω. Assume also that V is bounded from below on bounded subsets of Ω.*

Then the set-valued map $\partial_- V : \Omega \rightsquigarrow X^\star$ is contained in the graphical lower limit of the gradients V_n' of the functions V_n:

for any $x \in \Omega$ and any $p \in \partial_- V(x)$, there exists a sequence x_n converging to x such that $V_n'(x_n)$ converges to p.

If V is bounded from above on bounded subsets of Ω, the set-valued map $\partial_+ V$ is also contained in the graphical lower limit of the gradients V_n' of the functions V_n.

The same holds true if the functions V_n are locally Lipschitz and Gâteaux differentiable on Ω.

Proof — Let us fix $\varepsilon > 0$ and $x \in \Omega$. To say that $p \in \partial_- V(x)$ amounts to saying that there exists $\mu \in]0, \varepsilon[$ such that

$$\forall\, y \in B(x, \mu), \;\; V(x) - V(y) - < p, x - y > \leq \frac{\varepsilon}{2}\|x - y\|$$

We may assume that $B(x, \mu) \subset \Omega$. Then, the functions V_n converge uniformly to V on the ball $B(x, \mu)$: For any $\delta > 0$, there exists N_δ such that, for all $n \geq N_\delta$ and for all $y \in B(x, \mu)$, $V_n(y) - V(y) \leq \delta/2$. We take $\delta = \mu\varepsilon/4$ and fix $n \geq N_\delta$.

We now apply Ekeland's Theorem to the continuous function

$$y \;\mapsto\; V_n(y) - < p, y >$$

on the ball $B(x, \mu)$: There exists $x_n \in B(x, \mu)$ satisfying

$$\begin{cases} i) \quad V_n(x_n) - < p, x_n > + \varepsilon\|x - x_n\| \;\leq\; V_n(x) - < p, x > \\[2ex] ii) \quad \forall\, y \in B(x, \mu), \\ \qquad V_n(x_n) - < p, x_n > \leq\; V_n(y) - < p, y > + \varepsilon\|x_n - y\| \end{cases}$$

We infer from the first inequality that

$$\varepsilon\|x - x_n\| \;\leq\; \delta + V(x) - V(x_n) - < p, x - x_n > \leq\; \delta + \frac{\varepsilon}{2}\|x - x_n\|$$

from which we deduce that

$$\|x - x_n\| \leq 2\delta/\varepsilon = \mu/2$$

Since x_n belongs to the interior of $B(x, \mu)$ and minimizes on this ball the function

$$V_n(y) - < p, y > + \varepsilon\|x_n - y\|$$

the Fermat Rule implies that

$$\forall u \in X, \ 0 \leq \langle V_n'(x_n) - p, u \rangle + \varepsilon\|u\|$$

which means that $\|V_n'(x_n) - p\| \leq \varepsilon$. Hence the pair (x, p) is the limit of a sequence of elements $(x_n, V_n'(x_n))$ of the graph of V_n'. □

This statement is actually a consequence of a more precise

Theorem 7.6.2 *Let us consider a Banach space X and a sequence of extended lower semicontinuous functions $V_n : X \mapsto \mathbf{R} \cup \{+\infty\}$ whose domains contain a ball $B(x, \alpha)$. Assume that the functions V_n converge uniformly to an extended function V on this ball. Assume also that V is bounded from below on $B(x, \alpha)$.*

Choose any kind of δ-generalized gradients $\partial_\delta V_n$. Then for any $p \in \partial_- V(x)$, there exists a sequence x_n converging to x and a sequence of elements $p_n \in \partial_\delta V_n(x_n)$ converging to p.

Proof — We begin the proof as in Theorem 7.6.1, but this time the Fermat Rule implies that

$$\forall u \in X, \ 0 \leq D_\uparrow V_n(x_n)(u) - < p, u > + \varepsilon\|u\|$$

Consequently, the function V_n is contingently epidifferentiable at x_n. Furthermore, we deduce that for any convex epiderivative $\partial_\delta V_n(x_n)$ and for all

$$\forall u \in X, \ 0 \leq \delta V_n(x_n)(u) - < p, u > + \varepsilon\|u\|$$

This can be written as

$$0 \in \partial_\delta V_n(x_n) - p + \varepsilon B_\star$$

where $B_\star$ denotes the unit ball of $X^\star$. This means that there exists $p_n \in \partial_\delta V_n(x_n)$ such that $\|p - p_n\| \leq \varepsilon$. Hence the pair (x, p) is the limit of a sequence of elements (x_n, p_n) of the graph of $\partial_\delta V_n$. □

7.6.2 Convergence of Subdifferentials of Convex Functions

Theorem 7.6.3 *Consider a reflexive Banach space X and a sequence of nontrivial lower semicontinuous convex functions $V_n : X \mapsto \mathbf{R} \cup \{+\infty\}$. Let us assume that its upper epilimit and the sequentially weak lower epilimit coincide with a nontrivial function V.*

Then the subdifferential map ∂V is contained in the graphical lower limit of the subdifferential maps ∂V_n:

$$\partial V \ \subset \ \mathrm{Lim}^{\flat}{}_{n\to\infty} \partial V_n$$

More precisely, let $p \in \partial V(x)$. Consider sequences x_{0n} and p_{0n} converging to x and p respectively satisfying

$$\limsup_{n\to\infty} V_n(x_{0n}) \ \le \ V(x), \quad \limsup_{n\to\infty} V_n^\star(p_{0n}) \ \le \ V^\star(p)$$

(which exist by Corollary 7.5.4.) Then the error estimate

$$d((x_{0n}, p_{0n}), \mathrm{Graph}(\partial V_n)) \ \le \ \sqrt{V_n(x_{0n}) + V_n^\star(p_{0n}) - < p_{0n}, x_{0n} >} \tag{7.9}$$

holds true.

Proof — We introduce the "duality lack"

$$\delta_n \ := \ V_n(x_{0n}) + V_n^\star(p_{0n}) - < p_{0n}, x_{0n} > \ge \ 0$$

which converges to 0, since by assumption

$$\begin{cases} 0 \ \le \ \limsup_{n\to\infty} \delta_n \\ \le \ \limsup_{n\to\infty} V_n(x_{0n}) + \limsup_{n\to\infty} V_n^\star(p_{0n}) - < p, x > \\ \le \ V(x) + V^\star(p) - < p, x > = \ 0 \end{cases}$$

If $\delta_n = 0$, this means that p_{0n} belongs to $\partial V_n(x_{0n})$. If not, we apply Ekeland's Theorem to the lower semicontinuous function

$$y \ \mapsto \ V_n(y) - < p_{0n}, y >$$

Since $\limsup_{n\to\infty} V_n^\star(p_{0n}) < +\infty$, there exist constants c and N such that

$$\forall\, n \;\geq\; N, \;\; V_n^\star(p_{0n}) \;\leq\; c$$

This implies that the functions $y \mapsto V_n(y) - < p_{0n}, y >$ are bounded from below by $-c$.

Hence, we can apply Ekeland's Variational Principle in the generalized gradient form (Theorem 6.4.4): there exists a solution x_n satisfying

$$\left\{\begin{array}{ll} i) & V_n(x_n) - < p_{0n}, x_n > + \sqrt{\delta_n}\|x_n - x_{0n}\| \\ & \leq V_n(x_{0n}) - < p_{0n}, x_{0n} > \\ & \\ ii) & 0 \;\in\; \partial V_n(x_n) - p_{0n} + \sqrt{\delta_n} B_\star \end{array}\right. \tag{7.10}$$

Inequality (7.10) $ii)$ tells us that there exists p_n in $\partial V_n(x_n)$ satisfying

$$\|p_{0n} - p_n\| \;\leq\; \sqrt{\delta_n}$$

On the other hand, inequality (7.10) $i)$ yields

$$\|x_{0n} - x_n\| \;\leq\; \frac{1}{\sqrt{\delta_n}}\,(V_n(x_{0n}) - V_n(x_n) + < p_{0n}, x_n > - < p_{0n}, x_{0n} >)$$

Taking into account that

$$< p_{0n}, x_n > \leq\; V_n(x_n) + V_n^\star(p_{0n})$$

we obtain

$$\|x_{0n} - x_n\| \;\leq\; \frac{1}{\sqrt{\delta_n}}\,(V_n(x_{0n}) + V_n^\star(p_{0n}) - < p_{0n}, x_{0n} >) \;=\; \sqrt{\delta_n}$$

Hence inequality (7.9) is proved.

Since (x, p) is the limit of (x_{0n}, p_{0n}) by definition, we deduce that it is also the limit of the sequence of elements (x_n, p_n) of the graph of ∂V_n. □

Since ∂V is a maximal monotone set-valued map when X is a Hilbert space, Proposition 7.1.7 implies:

Theorem 7.6.4 (Attouch) *Let X be a Hilbert space. We supply X with the norm topology and $X^\star$ with the weak-$\star$ topology. Consider a sequence of nontrivial lower semicontinuous convex functions $V_n : X \mapsto \mathbf{R} \cup \{+\infty\}$.*

Let us assume that its upper epilimit and the sequentially weak lower epilimit do coincide with a nontrivial function denoted by V. Then

$$\partial V \;=\; \mathrm{Lim}^\flat{}_{n\to\infty}\partial V_n \;=\; \mathrm{Lim}^\sharp{}_{n\to\infty}\partial V_n$$

In particular, we deduce that under the assumptions of Theorem 7.6.4, the graphical convergence of the subdifferentials ∂V_n to ∂V implies that for every $x \in X$,

$$\partial V(x) \;=\; \sigma - \mathrm{Limsup}_{n\to\infty, x_n \to x}\partial V_n(x_n)$$

When the functions V_n are the indicators of nonempty closed convex subsets K_n, we obtain the following theorem on the convergence of the sequence of normal cones to these subsets:

Corollary 7.6.5 *Let X be a Hilbert space and $K_n \subset X$ be a sequence of nonempty closed convex subsets. We supply X with the norm topology and $X^\star$ with the weak-$\star$ topology.*

Assume that the lower limit K of the sequence $(K_n)_n$ coincides with its sequentially weak upper limit. Then

$$N_K(x) \;=\; \sigma - \mathrm{Limsup}_{n\to\infty, x_n \to x}N_{K_n}(x_n)$$

In order to state that some $p \in \partial V(x)$ belongs to the lower limit of $\partial V_n(x_n)$ for any sequence $x_n \in \mathrm{Dom}(\partial V_n)$ converging to x, we need some stability assumptions on the contingent second derivatives of the functions V_n.

Naturally, the *contingent Hessian* $\partial^2 V(x,p) := D(\partial V)(x,p)$ of V at some point (x,p) in the graph of ∂V is defined as the contingent derivative of the set-valued map ∂V at (x,p).

Proposition 7.6.6 *Let X be a Hilbert space. Consider a sequence of nontrivial lower semicontinuous convex functions $V_n : X \mapsto \mathbf{R} \cup \{+\infty\}$.*

Let us assume that their upper epilimit and their sequentially weak lower epilimit coincide with a nontrivial function denoted by V.

Let p belong to $\partial V(x)$. We posit the following stability assumption: There exist constants $c > 0$ and $\eta > 0$ such that

$$\begin{cases} \forall\ (x_n, p_n) \in \text{Graph}(\partial V_n) \cap B((x,p),\eta),\ \text{Dom}(\partial^2 V_n(x_n,p_n)) = X \\ \text{and} \\ \|\partial^2 V_n(x_n,p_n)\| \ := \ \sup_{u\in X} \inf_{\pi\in\partial^2 V_n(x_n,p_n)(u)} \|\pi\|/\|u\| \ \leq \ c \end{cases}$$

Then $p \in \text{Liminf}_{n\to\infty, x_n\to x} \partial V_n(x_n)$.

Proof — We apply Theorem 5.4.2 to the set-valued maps $F_n := \partial V_n$. $\square$

7.7 Asymptotic Epiderivatives

We are now able to define asymptotic epiderivatives of a sequence of functions V_n, by taking the asymptotic tangent cones to their epigraphs. We denote by

$$V_\uparrow^\flat \ := \ \lim{}^\flat_{\uparrow n\to\infty} V_n$$

their upper limit.

Definition 7.7.1 (Asymptotic Epiderivatives) *Let us consider a sequence of extended functions $V_n : X \mapsto \mathbf{R} \cup \{+\infty\}$ whose domains are not empty and an element x_0 in the domain of the upper epilimit of the functions V_n (i.e., $V_\uparrow^\flat(x_0) \neq \pm\infty$.) We shall say that the function $C_\uparrow^\flat\{V_n\}(x_0)$ defined by*

$$C_\uparrow^\flat\{V_n\}(x_0)(u) := \limsup_{n\to\infty, x\to x_0, V_n(x)\leq\lambda\to V_\uparrow^\flat(x_0), h\to 0+} (V_n(x+hu') - \lambda)/h$$

is the asymptotic circatangent epiderivative *of the sequence of functions V_n at x_0 in the direction u.*

We see at once that the epigraph of $C_\uparrow^\flat\{V_n\}(x_0)$ is the asymptotic circatangent cone to the epigraphs of the functions V_n at $(x_0, V_\uparrow^\flat(x_0))$, or, equivalently, that $C_\uparrow^\flat\{V_n\}(x_0)$ is the upper epilimit of the difference quotients

$$u' \ \mapsto \ (V_n(x+hu') - \lambda)/h$$

when

$$n \to \infty \ \& \ (x,\lambda,h) \ \in \ \mathcal{E}p(V_n) \times \mathbf{R}_+ \ \text{ converges to } \ (x_0, V_\uparrow^\flat(x_0), 0)$$

We deduce that the asymptotic circatangent epiderivative is a positively homogeneous, lower semicontinuous and convex function from X to $\mathbf{R} \cup \{\pm\infty\}$.

We shall estimate the asymptotic circatangent epiderivative of a family of functions $U_n := V_n + W_n \circ A$.

Theorem 7.7.2 *Consider two Banach spaces X and Y, a continuous linear operator $A \in \mathcal{L}(X,Y)$ and two sequences of extended lower semicontinuous functions V_n and W_n from X and Y to $\mathbf{R} \cup \{+\infty\}$ respectively. Assume that x_0 belongs to the domain of $V_\uparrow^\flat$ and Ax_0 to the domain of $W_\uparrow^\flat$.*

We posit the following stability assumption*: there exist constants $c > 0$, $\alpha \in [0,1[$ and $\eta > 0$ such that, for all n, (7.6) holds true.*

Then, the asymptotic circatangent epiderivative of the sequence of functions $U_n := V_n + W_n \circ A$ satisfies the estimate:

$$C_\uparrow^\flat\{U_n\}(x_0)(u) \;\leq\; C_\uparrow^\flat\{V_n\}(x_0)(u) + C_\uparrow^\flat\{W_n\}(Ax_0)(Au) \tag{7.11}$$

whenever the right hand side is well defined.

Remark — Since the assumptions of this theorem are the same as those of Theorem 7.4.1, we know that $U_\uparrow^\flat(x_0) \leq V_\uparrow^\flat(x_0) + W_\uparrow^\flat(Ax_0)$ whenever the right hand side is well defined. □

Proof — We apply Theorem 3.4.5, since we have seen in the proof of Theorem 7.4.1 that if we set

$$\begin{cases} i) & K_n \;:=\; \mathcal{E}p(V_n) \times \mathcal{E}p(W_n) \times \mathbf{R} \;\subset\; X \times \mathbf{R} \times Y \times \mathbf{R} \times \mathbf{R} \\ ii) & G(x,a,y,b,c) \;:=\; (Ax - y, a + b - c) \\ iii) & H(x,a,y,b,c) \;:=\; (x,c) \end{cases}$$

we can write

$$\mathcal{E}p(U_n) \;=\; H(K_n \cap G^{-1}(0,0))$$

The stability assumption being the same as the one of Theorem 7.4.1, we already know that they imply the stability assumptions of Theorem 3.4.5.

Hence, we deduce exactly as in the proof of Theorem 7.4.1 that

$$\begin{cases} C_{K^\flat}^\flat\left(x_0, V_\uparrow^\flat(x_0), Ax_0, W_\uparrow^\flat(Ax_0), U_\uparrow^\flat(x_0)\right) \cap G^{-1}(0,0) \\ \subset\; C_{K^\flat \cap G^{-1}(0,0)}^\flat\left(x_0, V_\uparrow^\flat(x_0), Ax_0, W_\uparrow^\flat(Ax_0), U_\uparrow^\flat(x_0)\right) \end{cases}$$

It remains to show that this inclusion implies inequality (7.11.)

Inequality (7.11) being satisfied when one of the members of the right hand side is $+\infty$, we assume that both of them are different from $+\infty$. Let us consider

$$\lambda \geq C_{\uparrow}^{\flat}\{V_n\}(x_0)(u), \quad \mu \geq C_{\uparrow}^{\flat}\{W_n\}(Ax_0)(Au), \quad \nu := \lambda + \mu$$

Hence the element $(u, \lambda, Au, \mu, \nu)$ belongs to

$$C_{K^\flat}^{\flat}\left(x_0, V_{\uparrow}^{\flat}(x_0), Ax_0, W_{\uparrow}^{\flat}(Ax_0), U_{\uparrow}^{\flat}(x_0)\right) \cap G^{-1}(0,0)$$

It then belongs to the asymptotic circatangent cone to $K_n \cap G^{-1}(0,0)$. Then, for all sequence $h_n > 0$ converging to zero, there exist elements (u_n, λ_n, μ_n) converging to (u, λ, μ) such that, for all $n \geq N$,

$$(x_n + h_n u_n, a_n + h_n \lambda_n, Ax_n + h_n Au_n, b_n + h_n \mu_n, a_n + b_n + h_n \nu_n)$$

belongs to $K_n \cap G^{-1}(0,0)$.

Therefore, the pairs $(x_n + h_n u_n, a_n + b_n + h_n(\lambda_n + \mu_n))$ belong to the epigraph of U_n. Since $(u_n, \lambda_n + \mu_n)$ converges to (u, ν), we deduce that

$$C_{\uparrow}^{\flat}\{U_n\}(x_0)(u) \leq \lambda + \mu$$

By letting λ and μ converge to $C_{\uparrow}^{\flat}\{V_n\}(x_0)(u)$ and $C_{\uparrow}^{\flat}\{W_n\}(Ax_0)(Au)$ respectively, we infer that

$$C_{\uparrow}^{\flat}\{U_n\}(x_0)(u) \leq C_{\uparrow}^{\flat}\{V_n\}(x_0)(u) + C_{\uparrow}^{\flat}\{W_n\}(Ax_0)(Au) \quad \square$$

Chapter 8

Measurability and Integration of Set-Valued Maps

Introduction

We meet measurable maps whenever we deal with models of systems having measurable data. Integrals of set-valued maps are involved in many convexification (also called relaxation) problems, since roughly speaking, the integral of a measurable set-valued map is always convex.

Another important instance where measurable set-valued maps do arise is related to linearization of differential inclusions along a solution. Consider indeed a Lipschitz set-valued map $F : \mathbf{R}^n \rightsquigarrow \mathbf{R}^n$ and the differential inclusion

$$x' \in F(x)$$

Let $\overline{x} : [0, T] \mapsto \mathbf{R}^n$ be a solution to this differential inclusion, i.e., an absolutely continuous map satisfying $\overline{x}'(t) \in F(\overline{x}(t))$ almost everywhere in $[0, T]$. Linearization of this inclusion along $\overline{x}(\cdot)$ leads to the new differential inclusion:

$$w'(t) \in DF(\overline{x}(t), \overline{x}'(t))(w(t)) \text{ almost everywhere in } [0, T]$$

where DF denotes the contingent derivative of F. Setting

$$G(t,w) := DF(\overline{x}(t),\overline{x}'(t))(w)$$

we obtain a set-valued map G merely measurable in time and Lipschitz in w (see Chapter 10.)

We investigate in this chapter some properties of measurable set-valued maps with closed images, which we define in Section 1. Under adequate assumptions, *a set-valued map is measurable if and only if its graph is measurable.*

Measurable maps with nonempty closed images have measurable selections. This result was born as a lemma by von Neumann in 1949 and has found different extensions and applications. It was given in a stronger form by Kuratowski and Ryll-Nardzewski. Castaing has proved that there exists a countable dense family of measurable selections, a useful theorem since countability plays a major role in many problems of measurability.

These two statements constitute the core of what we call the Characterization Theorem for measurable maps, that we prove in Section 3.

But it is already used in Section 2 which is devoted to the calculus of measurable maps. It is shown there that the main operations, such as union, intersection, direct and inverse image, lower and upper limits etc., do preserve measurability. The convex hull of a measurable map is measurable, and measurable selections from the convex hull have a Carathéodory representation. This is particularly useful in control theory, when we have to deal with relaxed controls.

In Section 4 we investigate selections from limits of measurable maps and in Section 5 tangent cones in Lebesgue spaces.

Section 6 is devoted to the concept of integral of set-valued maps introduced by Aumann, who proved that the integral is closed and convex when the set-valued map is measurable, integrably bounded and takes closed values in a finite dimensional space. (He was motivated at the time by problems of cooperative game theory and mathematical economics where convexity properties are missing when the set of players is finite. He then introduced the concept of continuum of players and replaced the sum by the integral.)

Convexity of the integral is an immediate consequence of Lyapunov's Convexity Theorem on convexity of the range of a nonatomic vector measure. Similar results concerning convexity of the integral are still valid in the infinite dimensional case (*the closure of the integral is convex.*) They are proved in Section 7.

The integral of a time-dependent set-valued map is also equal to the integral of finite concatenations of extremal selections. This result is sometimes referred as the *bang-bang principle.* The bang-bang principle we state here is due to Olech. It says in essence, that to integrate a map $F : [a,b] \rightsquigarrow \mathbf{R}^n$, we can restrict the attention to piecewise extremal trajectories with the number of switchings not greater than $n+1$. This is the topic of Section 8, which does not use the Lyapunov Theorem.

The last section deals with the extension to upper semicontinuous maps of the Poincaré Recurrence Theorem: *Let X be a compact metric space, $\mathcal{P}(X)$ denote the set of probability measures on X, $F : X \rightsquigarrow X$ be a closed set-valued map and $\mu \in \mathcal{P}(X)$ an invariant measure of F. For any Borelian subset B, let*

$$B_\infty := \bigcap_{N\geq 0} \bigcup_{n\geq N} F^{-n}(B)$$

be the subset of points x such that for all N, there exists $n \geq N$ such that $F^{-n}(x) \cap B \neq \emptyset$. Then the measure of $B \cap B_\infty$ is equal to the measure of B.

The statement of this theorem is clear as soon as we have defined the notion of an invariant measure of a set-valued map F:

Let $\mathcal{B}$ denote the σ-algebra of Borel subsets of X. We recall that if $F \equiv f$ is single-valued, an invariant probability measure μ is defined by:

$$\forall\, A \in \mathcal{B}, \quad \mu(A) = \mu(f^{-1}(A))$$

When F is set-valued, we cannot extend this definition as it is because $A \mapsto \mu(F^{-1}(A))$ is no longer a measure. However, we shall introduce the following definition: *We say that μ is an* invariant measure *of a closed set-valued map $F : X \rightsquigarrow X$ if and only if*

$$\forall\, A \in \mathcal{B}, \ \ \mu(A) \leq \mu(F^{-1}(A)) \tag{8.1}$$

Indeed we see that, for single-valued maps f, this definition coincides with the classical one when we apply it to both A and its complement. We shall prove the following result:

Let X be a compact metric space and $F : X \rightsquigarrow X$ be a closed set-valued map with nonempty values. Then there exists an invariant probability measure.

8.1 Measurable Set-Valued Maps

We define in this section measurable set-valued maps taking their values in a *complete separable metric space*, (a *Polish space*[1].) This contains separable Banach spaces and, in particular, Lebesgue spaces L^p and Sobolev spaces $W^{m,p}$ with $1 \le p < \infty$.

Consider a set Ω and a family $\mathcal{A}$ of subsets A of Ω. Recall that $\mathcal{A}$ is called a *$\sigma-$algebra* if it verifies the following properties:

$$\begin{cases} i) & \emptyset \in \mathcal{A} \\ ii) & A \in \mathcal{A} \Longrightarrow \Omega \setminus A \in \mathcal{A} \\ iii) & A_n \in \mathcal{A}, \ n = 1, 2, \ldots \Longrightarrow \bigcup_{n \ge 1} A_n \in \mathcal{A} \end{cases}$$

It is clear that $i)$ and $ii)$ imply that $\Omega \in \mathcal{A}$ and from $ii)$, $iii)$ follows that the intersection of a countable family of elements of $\mathcal{A}$ is still an element of $\mathcal{A}$.

We call the pair $(\Omega, \mathcal{A})$ a *measurable space* and the elements of $\mathcal{A}$ *measurable sets*.

If Ω is a topological space, then the smallest $\sigma-$algebra containing all open sets is called the *Borel $\sigma-$algebra*. We denote it by $\mathcal{B}(\Omega)$ or simply by $\mathcal{B}$, when it is clear from the context.

A map $\mu : \mathcal{A} \mapsto \mathbf{R} \cup \{+\infty\}$ is called a *positive measure* if for any sequence of disjoint sets $A_n \in \mathcal{A}$

$$\mu(\cup_{n \ge 1} A_n) = \sum_{n \ge 1} \mu(A_n)$$

A positive measure is *σ-finite* if Ω is the union of a (countable) sequence of measurable sets of finite measure. The $\sigma-$algebra $\mathcal{A}$ is

[1]Polish space is the name given by Bourbaki to a topological space which is metrizable in such a way that it becomes complete and separable. This space is particularly often discussed in mathematics. Opial wrote: "...this touching tribute to the Polish School of Mathematics testifies most expressively to its merits and to its status in mathematics."

complete (or, to be more precise, *$\mu-$complete*) if for every set $A \in \mathcal{A}$ satisfying $\mu(A) = 0$ any subset $A_1 \subset A$ is an element of $\mathcal{A}$.

For instance for every open (or closed) subset $\Omega \subset \mathbf{R}^n$, the set of all Lebesgue measurable subsets of Ω is complete (with respect to the Lebesgue measure.) The Lebesgue measure is also $\sigma-$finite.

In summary, we shall say that the triple $(\Omega, \mathcal{A}, \mu)$ is a *complete $\sigma-$finite measure space* if μ is a positive $\sigma-$finite measure such that $\mathcal{A}$ is $\mu-$complete.

Consider a complete separable metric space X. A single-valued $f : \Omega \mapsto X$ is said to be *measurable* ($\mathcal{A}$-measurable, to be more precise) if for every open subset $\mathcal{O} \subset X$, $f^{-1}(\mathcal{O}) \in \mathcal{A}$ (or equivalently for every closed subset $C \subset X$, $f^{-1}(C) \in \mathcal{A}$.) We say that a map is *simple* if it takes only a finite number of values. A simple map is measurable if and only if for every $x \in X$, $f^{-1}(x) \in \mathcal{A}$.

The following equivalences provide in many instances convenient tools to check measurability of a map:

$$\left\{\begin{array}{ll} i) & f \text{ is } \mathcal{A}-\text{measurable} \\ ii) & f \text{ is the pointwise limit of simple measurable maps} \\ iii) & f \text{ is the uniform limit of measurable maps} \\ & \text{assuming a countable number of values} \end{array}\right.$$

One immediately deduces that the *pointwise limit of measurable maps is measurable*.

Measurable set-valued maps with closed images are defined in a similar way:

Definition 8.1.1 *Consider a measurable space $(\Omega, \mathcal{A})$, a complete separable metric space X and a set-valued map $F : \Omega \rightsquigarrow X$ with closed images.*

The map F is called measurable *if the inverse image of each open set is a measurable set: for every open subset $\mathcal{O} \subset X$, we have*

$$F^{-1}(\mathcal{O}) \; := \; \{\omega \in \Omega \mid F(\omega) \cap \mathcal{O} \neq \emptyset\} \in \mathcal{A}$$

Remark — The notion of measurability that we use here is sometimes called in the literature *weak measurability*: a set-valued map is weakly measurable if inverse images of open sets are measurable, in comparison to *strong measurability*: inverse images of closed sets are measurable. However in the framework we shall deal with (complete σ-finite measure spaces), these two notions do coincide. That is why we simply use the word "measurable." □

The domain of a measurable map is measurable as well as its complement: $\{\omega \in \Omega \mid F(\omega) = \emptyset\}$.

A measurable set-valued map has a measurable selection:

Definition 8.1.2 *Let $(\Omega, \mathcal{A})$ be a measurable space and X a complete separable metric space. Consider a set-valued map $F : \Omega \rightsquigarrow X$. A measurable map $f : \Omega \mapsto X$ satisfying*

$$\forall\, \omega \in \Omega, \;\; f(\omega) \;\in\; F(\omega)$$

is called a measurable selection *of F.*

Theorem 8.1.3 (Measurable Selection) *Let X be a complete separable metric space, $(\Omega, \mathcal{A})$ a measurable space, F a measurable set-valued map from Ω to closed nonempty subsets of X. Then there exists a measurable selection of F.*

Proof — Fix a countable dense subset $\{x_n\}_{n\geq 1}$ of X. We construct a sequence of measurable maps $f_k : \Omega \mapsto X$, $k \geq 0$ taking values in $\{x_n\}_{n\geq 1}$ and converging uniformly to a selection f of F, so that f is measurable. We proceed by induction to construct such a sequence.

Denote by d the distance in X. For every $\omega \in \Omega$, let $n \geq 1$ be the smallest integer such that $F(\omega) \cap \overset{o}{B}(x_n, 1) \neq \emptyset$. We set $f_0(\omega) = x_n$. Then $f_0(\cdot)$ is measurable. Furthermore

$$\forall\, \omega \in \Omega, \;\; d(f_0(\omega),\; F(\omega)) \;<\; 1$$

Assume that we already constructed measurable maps

$$f_k : \Omega \;\mapsto\; \{x_n\}_{n\geq 1}, \quad k \;=\; 0, ..., m$$

satisfying

$$\forall\, 0 \;\leq\; k \;\leq\; m, \;\; \forall\, \omega \;\in\; \Omega, \;\; d(f_k(\omega), F(\omega)) \;<\; \frac{1}{2^k} \tag{8.2}$$

and

$$\forall\, 0 \;\leq\; k \;<\; m-1, \;\; d(f_k(\omega), f_{k+1}(\omega)) \;<\; \frac{1}{2^{k-1}} \tag{8.3}$$

For every n, define

$$S_n = \{ \omega \in \Omega \mid f_m(\omega) = x_n \}$$

The sets S_n are mutually disjoint and $\bigcup_{n \geq 1} S_n = \Omega$. Furthermore (8.2) implies that

$$\forall \omega \in S_n, \quad F(\omega) \cap \overset{o}{B}(x_n, 2^{-m}) \neq \emptyset$$

Fix $\omega \in \Omega$ and let n be such that $\omega \in S_n$. Consider the smallest integer r such that

$$F(\omega) \cap \overset{o}{B}(x_n, 2^{-m}) \cap \overset{o}{B}(x_r, 2^{-(m+1)}) \neq \emptyset$$

and set $f_{m+1}(\omega) = x_r$. Then

$$d(f_m(\omega), f_{m+1}(\omega)) \leq 2^{-m} + 2^{-(m+1)} < 2^{-m+1}$$

Moreover

$$d(f_{m+1}(\omega), F(\omega)) < 2^{-(m+1)}$$

Clearly, this defines a measurable map

$$f_{m+1} : \Omega \mapsto \{x_n\}_{n \geq 1}$$

and (8.2), (8.3) hold true with m replaced by $m+1$.

From inequality (8.3) follows that for all $\omega \in \Omega$, $(f_k(\omega))_{k \geq 1}$ is a Cauchy sequence in a complete metric space X. Let $f(\omega)$ denote the limit of $f_k(\omega)$. From (8.3) we deduce that f_k converges uniformly to f and from (8.2) that

$$d(f(\omega), F(\omega)) = 0$$

Thus f is measurable and for every $\omega \in \Omega$, $f(\omega) \in F(\omega)$. □

The next theorem provides several useful characterizations of measurability. For that purpose, let $\mathcal{A} \otimes \mathcal{B}$ denote the $\sigma-$algebra generated by products $A \times B$, where $A \in \mathcal{A}$ and $B \in \mathcal{B}$ ($\mathcal{B}$ is the Borel $\sigma-$algebra of a metric space X.)

Theorem 8.1.4 (Characterization Theorem) *Let $(\Omega, \mathcal{A}, \mu)$ be a complete σ−finite measure space, X a complete separable metric space and $F : \Omega \rightsquigarrow X$ a set-valued map with non empty closed images. Then the following properties are equivalent:*

$$\left\{\begin{array}{ll} i) & F \text{ is measurable} \\ ii) & \text{The graph of } F \text{ belongs to } \mathcal{A}\otimes\mathcal{B} \\ iii) & F^{-1}(C) \in \mathcal{A} \text{ for every closed set } C \subset X \\ iv) & F^{-1}(B) \in \mathcal{A} \text{ for every Borel set } B \subset X \\ v) & \text{For all } x \in X \text{ the map } d(x, F(\cdot)) \text{ is measurable} \\ vi) & \text{There exists a sequence of measurable selections } (f_n)_{n\geq 1} \\ & \text{of } F \text{ such that } \forall\, \omega \in \Omega,\ \ F(\omega) = \overline{\bigcup_{n\geq 1} f_n(\omega)} \end{array}\right.$$

A countable family of selections satisfying the latter property is said to be *dense*.

The equivalence of statements $i)$ and $ii)$ is fundamental in the framework of our *graphical approach*, since it characterizes measurable set-valued maps through the measurability of their graphs. Here too, this happens to be quite important.

The equivalence of claims $i)$ and $vi)$, due to Castaing, is also very useful, since countability is basic in many measurability questions.

We shall prove this theorem in the third section. But first we use it to provide some properties of measurable set-valued maps.

8.2 Calculus of Measurable Maps

We develop in this section the calculus of measurable set-valued maps in *complete σ-finite measure spaces*. However the reader should be aware that some of the results below can be extended to more general cases.

First, we note that semicontinuous maps are measurable:

Proposition 8.2.1 (Continuity and Measurability) *Consider a metric space Ω and a complete σ-finite measure space $(\Omega, \mathcal{A}, \mu)$. We assume that $\mathcal{A}$ contains all open subsets of Ω. Let X be a complete separable metric space and $F : \Omega \leadsto X$ a set-valued map with closed nonempty images. If F is upper semi-continuous (or lower semi-continuous), then F is measurable.*

Proof — If F is upper semicontinuous, then for every closed set $C \subset X$, $F^{-1}(C)$ is closed. We deduce from Characterization Theorem 8.1.4 that F is measurable. If F is lower semicontinuous, then for every open set $\mathcal{O} \subset X$, $F^{-1}(\mathcal{O})$ is open. $\square$

Theorem 8.2.2 (Convex Hull) *Let $(\Omega, \mathcal{A}, \mu)$ be a complete σ-finite measure space, X a separable Banach space and $F : \Omega \leadsto X$ a set-valued map with closed images. Then the map*

$$\Omega \ni \omega \leadsto \overline{co} F(\omega)$$

is measurable.

To prove this theorem we need the following well known result:

We recall that a map φ from $\Omega \times X$ to a metric space Y is called *Carathéodory*, if for every $x \in X$, $\varphi(\cdot, x)$ is measurable and for every $\omega \in \Omega$, $\varphi(\omega, \cdot)$ is continuous.

Lemma 8.2.3 *Consider two complete separable metric spaces X, Y, a measurable space $(\Omega, \mathcal{A})$ and a Carathéodory map $\varphi : \Omega \times X \mapsto Y$. Then for every measurable $f : \Omega \mapsto X$, the map $\omega \mapsto \varphi(\omega, f(\omega))$ is measurable.*

We provide its proof for convenience of the reader.

Proof — Since $f(\cdot)$ is measurable, there exists a sequence of simple measurable maps $f_n : \Omega \mapsto X$ converging pointwise to f. Then $\omega \mapsto \varphi(\omega, f_n(\omega))$ is measurable. Since φ is continuous with respect to the second variable,

$$\forall\, \omega \in \Omega, \ \lim_{n \to \infty} \varphi(\omega, f_n(\omega)) = \varphi(\omega, f(\omega))$$

Hence $\varphi(\omega, f(\omega))$ is the pointwise limit of measurable maps and thus, is measurable. $\square$

Proof of Theorem 8.2.2 — It is not restrictive to assume that F has nonempty images. Characterization Theorem 8.1.4 implies that there exists a dense sequence of measurable selections $(f_n)_{n\geq 1}$ of F. Let $\mathbf{Q}_+$ denote all nonnegative rational numbers. Consider the set

$$\Lambda := \left\{(\lambda_1, ..., \lambda_n) \mid \lambda_i \in \mathbf{Q}_+, \sum_{i=1}^{n} \lambda_i = 1, n \geq 1\right\}$$

and the countable family of maps $\sum_{i=1}^{n} \lambda_i f_i$, where $(\lambda_1, ..., \lambda_n) \in \Lambda$. By Lemma 8.2.3 these maps are measurable selections of $\overline{co}F$. They are also dense in $\overline{co}F$. Characterization Theorem 8.1.4 ends the proof. □

Theorem 8.2.4 (Union and Intersection) *Let $(\Omega, \mathcal{A}, \mu)$ be a complete σ-finite measure space, X a complete separable metric space and $F_n : \Omega \rightsquigarrow X$ set-valued maps with closed images. Then the maps*

$$\Omega \ni \omega \rightsquigarrow G(\omega) := \overline{\bigcup_{n\geq 1} F_n(\omega)} \;\; \& \;\; \Omega \ni \omega \rightsquigarrow H(\omega) := \bigcap_{n\geq 1} F_n(\omega)$$

are measurable.

Proof — Fix an open set $\mathcal{O} \subset X$. Then

$$G^{-1}(\mathcal{O}) = \{\omega \in \Omega \mid \bigcup_{n\geq 1} (F_n(\omega) \cap \mathcal{O}) \neq \emptyset\} = \bigcup_{n\geq 1} F_n^{-1}(\mathcal{O})$$

To prove the second statement observe that H has closed images and

$$\mathrm{Graph}(H) = \bigcap_{n\geq 1} \mathrm{Graph}\,(F_n)$$

By Characterization Theorem 8.1.4, Graph (F_n) belongs to $\mathcal{A} \otimes \mathcal{B}$ and so does the graph of H, which is measurable, thanks to the same theorem. □

Theorem 8.2.5 (Lower and Upper Limits) *Let $(\Omega, \mathcal{A}, \mu)$ be a complete σ-finite measure space, X a complete separable metric space and $F_n : \Omega \rightsquigarrow X$, $n \geq 1$ set-valued maps with closed images. Then the maps*

$$\Omega \ni \omega \rightsquigarrow \mathrm{Liminf}_{n\to\infty} F_n(\omega) \;\; \& \;\; \Omega \ni \omega \rightsquigarrow \mathrm{Limsup}_{n\to\infty} F_n(\omega)$$

are measurable.

Consequently, if for every $\omega \in \Omega$, the images $F_n(\omega)$ converge to a subset $F(\omega)$, the set-valued map F (called the pointwise limit*) is measurable.*

We need the following

Lemma 8.2.6 *Consider a measurable space $(\Omega, \mathcal{A})$, complete separable metric spaces X, Y and let $g : \Omega \times X \mapsto Y$ be a Carathéodory map. Then g is $\mathcal{A} \otimes \mathcal{B}$-measurable.*

Proof — Indeed it is enough to show that there exists a sequence of maps $g_n : \Omega \times X \mapsto Y$ measurable on $\mathcal{A} \otimes \mathcal{B}$, converging pointwise to g. Consider a dense sequence $x_k \in X$, $k \geq 1$. Fix $(\omega, x) \in \Omega \times X$. For every $n \geq 1$, let k be the smallest integer satisfying

$$x \in \overset{o}{B}\left(x_k, \frac{1}{n}\right)$$

and set $g_n(\omega, x) = g(\omega, x_k)$. Then g_n converge pointwise to g. Furthermore

$$\forall\, y \in Y_k := \overset{o}{B}\left(x_k, \frac{1}{n}\right) \setminus \bigcup_{m<k} \overset{o}{B}\left(x_m, \frac{1}{n}\right), \quad g_n(\omega, y) = g(\omega, x_k)$$

But $\bigcup_{n \geq 1} Y_k = X$ and therefore for each n, g_n is $\mathcal{A} \otimes \mathcal{B}$-measurable. $\square$

Proof of Theorem 8.2.5 — By Lemma 8.2.6 and Theorem 8.1.4 the map $(\omega, x) \mapsto d(F_n(\omega), x)$ is $\mathcal{A} \otimes \mathcal{B}(X)$-measurable. Thus for every $\varepsilon > 0$,

$$\mathrm{Graph}(B(F_n(\cdot), \varepsilon)) := \{(\omega, x) \in \Omega \times X \mid d(F_n(\omega), x) \leq \varepsilon\} \subset \mathcal{A} \otimes \mathcal{B}(X)$$

This and Characterization Theorem 8.1.4 yield that the map

$$\omega \rightsquigarrow B(F_n(\omega), \varepsilon) \text{ is measurable}$$

Then the result follows from Theorem 8.2.4 and equations

$$\left\{ \begin{array}{lll} i) \ \mathrm{Liminf}_{n \to \infty} F_n(\omega) & = & \bigcap_{k \geq 1} \overline{\bigcup_{N \geq 1} \bigcap_{n \geq N} B(F_n(\omega), \frac{1}{k})} \\ ii) \ \mathrm{Limsup}_{n \to \infty} F_n(\omega) & = & \bigcap_{N \geq 1} \overline{\bigcup_{n \geq N} F_n(\omega)} \ \square \end{array} \right.$$

Definition 8.2.7 *Consider metric spaces X, Y and a measurable space $(\Omega, \mathcal{A})$.*

A set-valued map $G : \Omega \times X \rightsquigarrow Y$ with closed values is called a Carathéodory map *if for every $x \in X$ the map $\omega \rightsquigarrow G(\omega, x)$ is measurable and for every $\omega \in \Omega$ the map $x \rightsquigarrow G(\omega, x)$ is continuous.*

Theorem 8.2.8 (Direct Image) *Let $(\Omega, \mathcal{A}, \mu)$ be a complete σ-finite measure space, X a complete separable metric space and $F : \Omega \rightsquigarrow X$ a measurable set-valued map with closed images.*

Consider a Carathéodory set-valued map G from $\Omega \times X$ to a complete separable metric space Y. Then, the map

$$\Omega \ni \omega \rightsquigarrow \overline{G(\omega, F(\omega))}$$

is measurable.

In particular for every measurable single-valued map $z : \Omega \mapsto X$, the set-valued map $\omega \rightsquigarrow G(\omega, z(\omega))$ is measurable and for every Carathéodory single-valued map φ from $\Omega \times X$ to Y, the set-valued map

$$\Omega \ni \omega \rightsquigarrow \overline{\varphi(\omega, F(\omega))}$$

is measurable.

Proof — It is not restrictive to assume that F has nonempty images. From Characterization Theorem 8.1.4, there exists a dense sequence $(f_n)_{n\geq 1}$ of measurable selections of F.

We claim that the map $\omega \rightsquigarrow G(\omega, f_n(\omega))$ is measurable. Indeed consider a sequence of measurable simple maps f_{nk} from Ω to X, converging pointwise to f_n when $k \to \infty$. Then, since f_{nk} are simple, for every k the set-valued map $\omega \rightsquigarrow G(\omega, f_{nk}(\omega))$ is measurable. On the other hand, since $G(\omega, \cdot)$ is continuous,

$$\forall\, \omega \in \Omega, \quad \mathrm{Lim}_{k\to\infty} G(\omega, f_{nk}(\omega)) = G(\omega, f_n(\omega))$$

and from Theorem 8.2.5, we deduce that this limit is again a measurable map.

On the other hand, since G is continuous with respect to the second variable and $(f_n)_{n\geq 1}$ is dense, for every $\omega \in \Omega$

$$\overline{G(\omega, F(\omega))} = \overline{\bigcup_{n\geq 1} G(\omega, f_n(\omega))}$$

Theorem 8.2.4 ends the proof. $\square$

Theorem 8.2.9 (Inverse Image) *Consider a complete σ-finite measure space $(\Omega, \mathcal{A}, \mu)$, complete separable metric spaces X, Y, measurable set-valued maps $F : \Omega \rightsquigarrow X$, $G : \Omega \rightsquigarrow Y$ with closed images. Let $g : \Omega \times X \mapsto Y$ be a Carathéodory map. Then the set-valued map H defined by*

$$H(\omega) := \{x \in F(\omega) \mid g(\omega, x) \in G(\omega)\}$$

is measurable. Consequently, if

$$\forall\, \omega \in \Omega, \quad g(\omega, F(\omega)) \cap G(\omega) \neq \emptyset$$

then there exists a measurable selection f of F such that for every $\omega \in \Omega$, $g(\omega, f(\omega))$ belongs to $G(\omega)$.

Proof — Observe that the map H has closed images. Define the single-valued map $\varphi : \Omega \times X \mapsto \Omega \times Y$ by $\varphi(\omega, x) = (\omega, g(\omega, x))$. Then, by Lemma 8.2.6,

$$\forall\, B \in \mathcal{B}(Y), \quad \{(\omega, x) \mid g(\omega, x) \in B\} \in \mathcal{A} \otimes \mathcal{B}(X)$$

Observe that

$$\text{Graph}(H) = \text{Graph}(F) \cap \varphi^{-1}(\text{Graph(G)})$$

By Characterization Theorem 8.1.4,

$$\text{Graph}(F) \in \mathcal{A} \otimes \mathcal{B}(X) \quad \& \quad \text{Graph}(G) \in \mathcal{A} \otimes \mathcal{B}(Y)$$

On the other hand for every $A \times B \in \mathcal{A} \otimes \mathcal{B}(Y)$,

$$\varphi^{-1}(A \times B) = \{(\omega, x) \mid g(\omega, x) \in B\} \cap (A \times X) \in \mathcal{A} \otimes \mathcal{B}(X)$$

Hence for every $C \subset \mathcal{A} \otimes \mathcal{B}(Y)$, $\varphi^{-1}(C) \in \mathcal{A} \otimes \mathcal{B}(X)$. Thus Graph$(H)$ belongs to $\mathcal{A} \otimes \mathcal{B}(X)$ and from Characterization Theorem 8.1.4, we deduce that H is measurable. Finally the Measurable Selection Theorem implies that it has a measurable selection, whenever H has nonempty images. $\square$

As a consequence, we obtain the very useful

Theorem 8.2.10 (Filippov) *Consider a complete σ-finite measure space $(\Omega, \mathcal{A}, \mu)$, complete separable metric spaces X, Y and a measurable set-valued map $F : \Omega \rightsquigarrow X$ with closed nonempty images. Let $g : \Omega \times X \mapsto Y$ be a Carathéodory map. Then for every measurable map $h : \Omega \mapsto Y$ satisfying*

$$h(\omega) \in g(\omega, F(\omega)) \text{ for almost all } \omega \in \Omega$$

there exists a measurable selection $f(\omega) \in F(\omega)$ such that

$$h(\omega) = g(\omega, f(\omega)) \text{ for almost all } \omega \in \Omega$$

Marginal functions and maps are measurable under the following conditions:

Theorem 8.2.11 (Marginal Map) *Consider a complete σ-finite measure space $(\Omega, \mathcal{A}, \mu)$, complete separable metric spaces X, Y, a measurable set-valued map $F : \Omega \rightsquigarrow X$ with closed nonempty images and a Carathéodory function $f : \Omega \times X \mapsto \mathbf{R}$. Then the marginal function $v : \Omega \mapsto \mathbf{R} \cup \{-\infty\}$ defined by*

$$\forall\, \omega \in \Omega, \quad v(\omega) := \inf_{x \in F(\omega)} f(\omega, x)$$

is measurable. Furthermore, the marginal map R defined by

$$\forall\, \omega \in \Omega, \quad R(\omega) := \{x \in F(\omega) \mid f(\omega, x) = \inf_{y \in F(\omega)} f(\omega, y)\}$$

is also measurable.

We recall the following:

Lemma 8.2.12 *Consider a measurable space $(\Omega, \mathcal{A})$. Then for every sequence of measurable real-valued functions $f_n : \Omega \mapsto \mathbf{R}$, the extended function*

$$\omega \mapsto f(\omega) := \inf_{n \geq 1} f_n(\omega) \in \mathbf{R} \cup \{-\infty\}$$

has a measurable domain of definition and is measurable on it. A similar statement holds true for $\sup_{n \geq 1} f_n$.

Proof of Theorem 8.2.11 — Let us consider a dense family of measurable selections $(g_n)_{n\geq 1}$ of the map F. Since the maps $f(\omega, \cdot)$ are continuous, we infer that

$$\forall\, \omega \in \Omega, \quad v(\omega) \;=\; \inf_{n\geq 1} f(\omega, g_n(\omega))$$

Since the maps $\omega \mapsto f(\omega, g_n(\omega))$ are measurable by Lemma 8.2.3, we deduce from Lemma 8.2.12 that v is also measurable.

Theorem 8.2.9 on inverses of measurable maps imply that the marginal map R is measurable, since

$$R(\omega) \;=\; \{x \in F(\omega) \mid f(\omega, x) = v(\omega)\}$$

and because v is measurable. □

Corollary 8.2.13 *Let us consider a complete σ-finite measure space $(\Omega, \mathcal{A}, \mu)$, a complete separable metric space X, a measurable set-valued map $F : \Omega \rightsquigarrow X$ with closed images and measurable single-valued maps $f : \Omega \mapsto X$, $\rho : \Omega \mapsto \mathbf{R}_+$. Then the following maps are measurable:*

1. — *the map $\omega \rightsquigarrow B(f(\omega), \rho(\omega))$*
2. — *the distance function $\Omega \ni \omega \mapsto d(f(\omega), F(\omega))$*
3. — *the projection map $\Omega \ni \omega \rightsquigarrow \Pi_{F(\omega)}(f(\omega))$ defined by*

$$\Pi_{F(\omega)}(f(\omega)) \;:=\; \{x \in F(\omega) \mid\; d(x, f(\omega)) \;=\; d(f(\omega), F(\omega))\}$$

Consequently if for every $\omega \in \Omega$, $\Pi_{F(\omega)}(f(\omega)) \neq \emptyset$, then there exists a measurable selection $g(\omega) \in F(\omega)$ such that

$$d(f(\omega), g(\omega)) \;=\; d(f(\omega), F(\omega))$$

Proof — Consider the function $g(\omega, x) = d(x, f(\omega))$. By Lemma 8.2.3 it is measurable in ω. Since it is also continuous in x we may use Theorem 8.2.9 with $F \equiv X$, $G(\omega) = [0, \rho(\omega)]$. To prove the second and third statements, we apply Theorem 8.2.11 to the function $(\omega, x) \mapsto d(x, f(\omega))$ □

Theorem 8.2.14 (Support Function) *Let us consider a complete σ-finite measure space $(\Omega, \mathcal{A}, \mu)$ and let X be a separable Banach* *space, $F : \Omega \rightsquigarrow X$ be a measurable set-valued map with nonempty closed images. Then it has measurable support functions: for every $p \in X^\star$,*

$$\text{the function} \quad \omega \mapsto \sigma(F(\omega), p) := \sup_{f \in F(\omega)} < p, f > \quad \text{is measurable}$$

The converse statement holds true if the dual of X is separable and images of F are convex and bounded.

Proof — Theorem 8.2.11 implies that the support functions $\sigma(F(\cdot), p)$ are measurable.

Assume next that for every $p \in X^\star$, $\sigma(p, F(\cdot))$ is measurable, that $X^\star$ is separable and F has closed convex and bounded images. To prove the last statement is enough to verify that for all x, the map $\omega \mapsto d(x, F(\omega))$ is measurable.

We first observe that since $F(\omega)$ is bounded, its support function $\sigma(F(\omega), \cdot)$ is continuous. Let $p_n \in X^\star$, $n \geq 1$ be a dense set of points of the unit sphere of $X^\star$. By our assumption, for every $n \geq 1$, $\sigma(F(\cdot), p_n)$ is measurable. Fix $x \in X$. Then, using that $F(\omega)$ is closed and convex, we get

$$\left\{ \begin{array}{l} d(x, F(\omega)) \;=\; d(0, F(\omega) - x) \;=\; -\inf_{\|p\|_\star \leq 1} \sigma(F(\omega) - x, p) \\ = \sup_{\|p\|_\star \leq 1}(\langle p, x\rangle - \sigma(F(\omega), p)) = \sup_{n \geq 1}(\langle p_n, x\rangle - \sigma(F(\omega), p_n)) \end{array} \right.$$

Hence $d(x, F(\cdot))$ is the supremum of measurable functions. Consequently it is measurable. □

The following result is very useful to study relaxation problems in control theory.

Theorem 8.2.15 (Carathéodory Representation) *Consider a complete σ-finite measure space $(\Omega, \mathcal{A}, \mu)$, a measurable set-valued map $G : \Omega \rightsquigarrow \mathbf{R}^n$ with nonempty closed images and a measurable selection $f(\omega) \in \mathrm{co}(G(\omega))$. Then there exist measurable functions $\lambda_k : \Omega \mapsto \mathbf{R}_+$ and measurable selections $f_k(\omega) \in G(\omega)$, $k = 0, ...n$ such that*

$$f(\omega) \;=\; \sum_{k=0}^{n} \lambda_k(\omega) f_k(\omega) \quad \& \quad \sum_{k=0}^{n} \lambda_k(\omega) \;=\; 1$$

Proof — Let S^{n+1} denote the set of all points $(\lambda_0, ..., \lambda_n) \in \mathbf{R}_+^{n+1}$ such that $\sum_{k=0}^n \lambda_k = 1$. Consider the continuous map

$$h : \mathbf{R}_+^{n+1} \times (\mathbf{R}^n)^{n+1} \mapsto \mathbf{R}^n$$

defined by

$$h(\lambda_0, ..., \lambda_n, x_0, ..., x_n) = \sum_{k=0}^n \lambda_k x_k$$

and the measurable set-valued map

$$F(\omega) = S^{n+1} \times (G(\omega))^{n+1}$$

By the Carathéodory theorem, $f(\omega) \in h(F(\omega))$. We conclude the proof by applying Theorem 8.2.9 to this set-valued map F with $g(\omega, x) = h(x)$ and the map G equal to f. □

8.3 Proof of the Characterization Theorem

Characterization Theorem 8.1.4 is the consequence of both a Theorem due to Castaing and a classical result on the projection of certain classes of Borel subsets.

We show first that a map is measurable if and only if it has a countable dense subset of measurable selections.

Theorem 8.3.1 (Castaing) *Let $(\Omega, \mathcal{A})$ be a measurable space, X a complete separable metric space and $F : \Omega \rightsquigarrow X$ be a set-valued map with non empty closed images. Then the statements $i)$, $v)$, $vi)$ of Characterization Theorem 8.1.4 are equivalent.*

Proof

1. — We begin by proving that $i) \Longrightarrow vi)$. Consider a dense countable subset $\{x_n\}_{n\geq 1}$ of X. For all $n, k \geq 1$ define the set-valued map $G_{nk} : \Omega \rightsquigarrow X$ by

$$G_{nk}(\omega) = \begin{cases} F(\omega) \cap \overset{o}{B}(x_n, \frac{1}{k}) & \text{if } F(\omega) \cap \overset{o}{B}(x_n, \frac{1}{k}) \neq \emptyset \\ F(\omega) & \text{otherwise} \end{cases}$$

and set $F_{nk}(\omega) = \overline{G_{nk}(\omega)}$. Then F_{nk} has nonempty closed images and is measurable. This follows from the very definition of G_{nk} and equality

$$\text{for every open } \mathcal{O} \subset X, \quad F_{nk}^{-1}(\mathcal{O}) = G_{nk}^{-1}(\mathcal{O})$$

Hence, by Theorem 8.1.3, F_{nk} has a measurable selection f_{nk}. It remains to verify that for every $\omega \in \Omega$, the values $f_{nk}(\omega)$ are dense in $F(\omega)$.

Fix $x \in F(\omega)$ and $\varepsilon > 0$. Let $k \geq 1$ be such that $1/k \leq \varepsilon/2$ and n such that $d(x_n, x) < 1/k$. Hence

$$F(\omega) \cap B(x_n, 1/k) \neq \emptyset$$

and $f_{nk}(\omega) \in B(x_n, 1/k)$. This yields

$$d(f_{nk}(\omega), x) \leq d(f_{nk}(\omega), x_n) + d(x_n, x) < \varepsilon$$

Since $\varepsilon > 0$ is arbitrary, the proof of this implication ensues.

2. — We prove now that $vi) \Longrightarrow v)$.

Let $f_n(\cdot)$ be as in $vi)$. Fix $x \in X$ and let d denote the metric of X. Then for every n, the function $\omega \mapsto d(x, f_n(\omega))$ is measurable (thanks to Lemma 8.2.3.) Consequently, the function

$$\omega \mapsto d(x, F(\omega)) = \inf_{n \geq 1} d(x, f_n(\omega))$$

is also measurable and $v)$ follows.

3. — Finally assume that $v)$ holds true. Then for every $x \in X$ and $r > 0$ the set

$$\{\omega \in \Omega \mid d(x, F(\omega)) < r\} = F^{-1}\left(\overset{o}{B}(x, r)\right) \in \mathcal{A}$$

Fix an open set $\mathcal{O} \subset X$. Since X is separable, $\mathcal{O}$ is a countable union of balls $\overset{o}{B}(x_n, r_n)$. Hence

$$F^{-1}(\mathcal{O}) = \cup_{n \geq 1} F^{-1}\left(\overset{o}{B}(x_n, r_n)\right) \in \mathcal{A}$$

The proof is complete. $\square$

To prove the rest of Characterization Theorem 8.1.4. we need a very useful projection property enjoyed by complete measure spaces:

Theorem 8.3.2 (Measurable Projection) *Let $(\Omega, \mathcal{A}, \mu)$ be a complete σ-finite measure space, X a complete separable metric space and $G \in \mathcal{A} \otimes \mathcal{B}(X)$. Then its projection is measurable:*

$$\pi_\Omega(G) := \{\omega \in \Omega \mid \exists\, x \in X,\ (\omega, x) \in G\} \in \mathcal{A}$$

The proof can be found in [91].[2]

Proof of Characterization Theorem 8.1.4 — We already know $i) \iff v) \iff vi)$. It is clear that $iv) \Longrightarrow iii)$.

To show that $iii) \Longrightarrow i)$, fix an open subset $\mathcal{O} \subset X$ and define the closed sets

$$C_n = \{x \in X \mid d(x, X \backslash \mathcal{O}) \geq 1/n\}$$

where d states for the distance. Then $\mathcal{O} = \bigcup_{n \geq 1} C_n$. Consequently $F(\omega) \cap \mathcal{O} \neq \emptyset$ if and only if for some $n \geq 1$, $F(\omega) \cap C_n \neq \emptyset$. This yields

$$F^{-1}(\mathcal{O}) = \bigcup_{n \geq 1} F^{-1}(C_n) \in \mathcal{A}$$

and ends the proof of $i)$.

We show next that $v)$ implies $ii)$. Observe that

$$\text{Graph}\,(F) = \{\, (\omega, x) \in \Omega \times X \mid d(x, F(\omega)) = 0 \,\}$$

By Lemma 8.2.6 the function $(\omega, x) \mapsto d(x, F(\omega))$ is $\mathcal{A} \otimes \mathcal{B}$ measurable and therefore

$$\{(\omega, x) \in \Omega \times X \mid d(x, F(\omega)) = 0\} \in \mathcal{A} \otimes \mathcal{B}$$

Thus Graph(F) belongs to $\mathcal{A} \otimes \mathcal{B}$.

[2] It is quite tempting to think that the projection (and therefore any continuous image) of a Borel set should be still a Borel set. This wishful thinking was made once by Lebesgue. Suslin, who had discovered that error, could not believe that the great Lebesgue would commit such a mistake. The discovery of Borel sets whose continuous image is not Borel started the intensive study of *analytic sets*, i.e., continuous images of Borel sets, which have many interesting properties.

As "tartes Tatin", "Rocquefort" and "Sauterne wine", this example shows how frequently mistakes and errors can trigger unexpected advances in both everyday life and sciences.

We end the proof by assuming that *ii*) holds true and observing that for every Borel set $B \in \mathcal{B}$ we have

$$F^{-1}(B) = \pi_{\Omega}(\mathrm{Graph}(F) \cap (\Omega \times B))$$

But $\mathrm{Graph}(F) \cap (\Omega \times B)$ is an element of $\mathcal{A} \otimes \mathcal{B}$. Therefore, using Theorem 8.3.2, we deduce *iv*).

8.4 Limits of Measurable Maps and Selections

Let $(\Omega, \mathcal{S}, \mu)$ be a complete σ–finite measure space and X be a separable Banach space.

For a measurable map $f : \Omega \mapsto X$ such that $\|f(\cdot)\|$ belongs to $L^1(X; \mathbf{R}, \mu)$, denote by $\int_{\Omega} f d\mu$ or $\int_{\Omega} f(\omega)\mu(d\omega)$ the *integral* of f.

Recall that two measurable single-valued maps are equal almost everywhere if the set where they are different is of zero measure.

We denote by $L^p(\Omega; X, \mu)$, $1 \leq p < \infty$ the Banach space of (classes) of measurable maps $f : \Omega \mapsto X$ such that $\int_{\Omega} \|f\|^p \, d\mu < \infty$.

Consider a sequence of measurable set-valued maps

$$K_n : \Omega \ni \omega \rightsquigarrow K_n(\omega) \subset X$$

We associate with it the subsets $\mathcal{K}_n \subset L^p(\Omega; X, \mu)$ $(1 \leq p < \infty)$ of selections, defined by

$$\mathcal{K}_n := \{x(\cdot) \in L^p(\Omega; X, \mu) \mid \text{for almost all } \omega \in \Omega, \ x(\omega) \in K_n(\omega)\}$$

The purpose of the next theorem is to compare the limits of the sets $\mathcal{K}_n$ and the sets of selections $x(\cdot)$ of the limits of the sets $K_n(\omega)$.

Theorem 8.4.1 *Let us assume that the set-valued maps K_n are measurable, have closed images and that $\omega \mapsto \sup_{n \geq 1} d(0, K_n(\omega))$ belongs to $L^p(\Omega; X, \mu)$, where $1 \leq p < \infty$. Then*

$$\begin{cases} \{x(\cdot) \in L^p(\Omega; X, \mu) \mid \text{for almost all } \omega, \ x(\omega) \in \mathrm{Liminf}_{n \to \infty} K_n(\omega)\} \\ \subset \mathrm{Liminf}_{n \to \infty} \mathcal{K}_n \subset \mathrm{Limsup}_{n \to \infty} \mathcal{K}_n \subset \\ \{x(\cdot) \in L^p(\Omega; X, \mu) \mid \text{for almost all } \omega, \ x(\omega) \in \mathrm{Limsup}_{n \to \infty} K_n(\omega)\} \end{cases}$$

Consequently, if the subsets $K_n(\omega)$ have a limit $K(\omega)$ for almost all $\omega \in \Omega$, then the subsets $\mathcal{K}_n$ converge to the subset

$$\mathcal{K} := \{x(\cdot) \in L^p(\Omega; X, \mu) \mid \text{ for almost all } \omega,\ x(\omega) \in K(\omega)\}$$

Furthermore, if the dimension of X is finite and if $L^p(\Omega; X, \mu)$ is supplied with the weak topology, then

$$\begin{cases} \sigma - \mathrm{Limsup}_{n\to\infty}\mathcal{K}_n \subset \\ \{x(\cdot) \in L^p(\Omega; X, \mu) \mid \text{for a.a. } \omega,\ \ x(\omega) \in \overline{co}\,(\mathrm{Limsup}_{n\to\infty}K_n(\omega))\} \end{cases}$$

Proof — Let $x(\cdot)$ belong to the first subset. Then the functions $a_n(\cdot)$ defined by

$$a_n(\omega) := d(x(\omega), K_n(\omega))$$

are measurable and converge to 0 almost everywhere. Since

$$\text{for almost all } \omega \in \Omega,\ \ a_n(\omega) \le \|x(\omega)\| + \sup_{n\ge 1} d(0, K_n(\omega))$$

and since the right-hand side of this inequality belongs to $L^p(\Omega; \mathbf{R}, \mu)$, we deduce from Lebesgue's Theorem that the functions $a_n(\cdot)$ do converge to 0 in $L^p(\Omega; \mathbf{R}, \mu)$. Let $k \in L^p(\Omega; \mathbf{R}, \mu)$ be a function with strictly positive values. Corollary 8.2.13 allows us to choose a measurable selection $z_n(\cdot)$ of $K_n(\cdot)$ such that

$$\|x(\omega) - z_n(\omega)\| \le a_n(\omega) + k(\omega)/n$$

It belongs to $L^p(\Omega; \mathbf{R}, \mu)$ since

$$\text{for almost all } \omega \in \Omega,\ \ \|z_n(\omega)\| \le \|x(\omega)\| + a_n(\omega) + k(\omega)/n$$

Therefore $z_n(\cdot)$ belongs to $\mathcal{K}_n$ and converges to $x(\cdot)$ in $L^p(\Omega; X, \mu)$, i.e., $x(\cdot)$ does belong to the lower limit of the subsets $\mathcal{K}_n$.

Let us choose some $x(\cdot)$ in the upper limit of the subsets $\mathcal{K}_n$. Then there exists a subsequence of elements $z_{n'}(\cdot)$ of $\mathcal{K}_{n'}$ converging to $x(\cdot)$ in $L^p(\Omega; X, \mu)$. It has a subsequence converging almost everywhere to $x(\cdot)$ and consequently, for almost all ω, $x(\omega)$ belongs to the upper limit of the subsets $K_n(\omega)$.

Finally, assume that $x(\cdot)$ belongs to the sequentially weak upper limit of the subsets $\mathcal{K}_n$. It is a weak limit in $L^p(\Omega; X, \mu)$ of a subsequence of functions $z_n(\cdot) \in \mathcal{K}_n$ and also, thanks to Mazur's Theorem, the strong limit of convex combinations $v_n(\cdot)$ of elements of the sequence $z_n(\cdot)$. Then a subsequence (again denoted by) v_n converges almost everywhere to $x(\cdot)$. We conclude as in the proof of Convergence Theorem 7.2.1 that

$$\text{for almost all } \omega \in \Omega, \;\; x(\omega) \in \overline{co}\,(\mathrm{Limsup}_{n\to\infty} K_n(\omega)) \quad \square$$

8.5 Tangent Cones in Lebesgue Spaces

Let $(\Omega, \mathcal{S}, \mu)$ be a complete σ–finite measure space and X be a separable Banach space. Let us consider a measurable set-valued map

$$K : \Omega \ni \omega \rightsquigarrow K(\omega) \subset X$$

We associate with it the subset $\mathcal{K} \subset L^p(\Omega; X, \mu)$ of selections defined by

$$\mathcal{K} := \{\, x(\cdot) \in L^p(\Omega; X, \mu) \mid \text{for almost all } \omega \in \Omega, \;\; x(\omega) \in K(\omega) \,\}$$

We shall characterize the contingent and adjacent tangent cones to $\mathcal{K}$ in terms of the tangent cones to the subsets $K(\omega)$.

Theorem 8.5.1 *Let us assume that the set-valued map K is measurable and has closed images. Then for every $x \in \mathcal{K}$, the set-valued maps:*

$$\Omega \ni \omega \rightsquigarrow T_{K(\omega)}(x(\omega)) \quad \& \quad \Omega \ni \omega \rightsquigarrow T^\flat_{K(\omega)}(x(\omega))$$

are measurable. Furthermore

$$\begin{cases} \Big\{v(\cdot) \in L^p(\Omega; X, \mu) \mid \text{ for almost all } \omega, \;\; v(\omega) \in T^\flat_{K(\omega)}(x(\omega))\Big\} \\ \subset T^\flat_{\mathcal{K}}(x(\cdot)) \subset T_{\mathcal{K}}(x(\cdot)) \\ \subset \Big\{v(\cdot) \in L^p(\Omega; X, \mu) \mid \text{ for almost all } \omega, \;\; v(\omega) \in T_{K(\omega)}(x(\omega))\Big\} \end{cases}$$

This theorem and its numerous applications motivate the introduction of adjacent tangent cones and derivable sets.

Proof — Fix $x \in \mathcal{K}$ and $\omega \in \Omega$.

Let $\mathbf{Q}_+$ denote the set of all strictly positive rationals. It is countable and from the very definition of the tangent cones, we can write

$$T_{K(\omega)}(x(\omega)) = \bigcap_{\alpha \in \mathbf{Q}_+} cl \left(\bigcup_{h \in]0,\alpha] \cap \mathbf{Q}_+} \frac{K(\omega) - x(\omega)}{h} \right)$$

and

$$T^\flat_{K(\omega)}(x(\omega)) = \bigcap_{n>0} cl \left(\bigcup_{\alpha \in \mathbf{Q}_+} \bigcap_{h \in]0,\alpha] \cap \mathbf{Q}_+} \left(\frac{K(\omega) - x(\omega)}{h} + \frac{1}{n} B \right) \right)$$

We deduce the first assertion from Theorem 8.2.4.

To prove the second statement, consider $v(\cdot)$ in the first subset. We have to prove that when $h > 0$ goes to 0, there exist maps $v_h(\cdot) \in L^p(\Omega; X, \mu)$ converging to $v(\cdot)$ such that

$$\text{for almost all } \omega \in \Omega,\ \ x(\omega) + hv_h(\omega) \in K(\omega)$$

Let us set:

$$a_h(\omega) := d\left(v(\omega), \frac{K(\omega) - x(\omega)}{h}\right)$$

Functions a_h are measurable and converge to 0 almost everywhere because for almost all $\omega \in \Omega$, $v(\omega)$ belongs to the adjacent cone to $K(\omega)$ at $x(\omega)$.

Since for almost all ω, $a_h(\omega) \leq \|v(\omega)\|$ and $v \in L^p(\Omega; \mathbf{R}, \mu)$, by the Lebesgue dominated convergence theorem, the functions $a_h(\cdot)$ do converge to 0 in $L^p(\Omega; \mathbf{R}, \mu)$. Fix $k \in L^p(\Omega; \mathbf{R}, \mu)$ with strictly positive values. Corollary 8.2.13 yields that for some $z_h \in \mathcal{K}$

$$\forall\, \omega \in \Omega,\ d\left(v(\omega), \frac{z_h(\omega) - x(\omega)}{h}\right) \leq a_h(\omega) + hk(\omega)$$

We define now the maps $v_h(\cdot)$ by

$$v_h(\omega) := (z_h(\omega) - x(\omega))/h$$

They are measurable, satisfy

$$\|v_h(\omega) - v(\omega)\| \leq a_h(\omega) + hk(\omega)$$

and thus, converge to $v(\cdot)$ in $L^p(\Omega; X, \mu)$ since $a_h(\cdot)$ converges to 0 in $L^p(\Omega; \mathbf{R}, \mu)$. We infer that $v(\cdot)$ belongs to $T^\flat_{\mathcal{K}}(x(\cdot))$ because

$$\text{for almost all } \omega \in \Omega, \quad x(\omega) + hv_h(\omega) \in K(\omega)$$

Let us choose next some $v(\cdot)$ in the contingent cone to the subset $\mathcal{K}$. Then there exist sequences $h_n > 0$ and $v_n(\cdot)$ converging respectively to 0 and to $x(\cdot)$ in $L^p(\Omega; X)$ and satisfying

$$\text{for almost all } \omega \in \Omega, \quad x(\omega) + h_n v_n(\omega) \in K(\omega)$$

Then a subsequence (again denoted) $v_n(\cdot)$ converges almost everywhere to $v(\cdot)$ and consequently, for almost all ω, $v(\omega)$ belongs to the contingent cone to the subset $K(\omega)$ at $x(\omega)$. □

As an obvious consequence, we obtain

Corollary 8.5.2 *Let us assume that the set-valued map K is measurable and has closed images. Let $x(\cdot) \in \mathcal{K}$. If for almost every $\omega \in \Omega$, the subsets $K(\omega)$ are derivable at $x(\omega)$, so is $\mathcal{K}$ at x and*

$$T_{\mathcal{K}}(x) = \left\{ v(\cdot) \in L^p(\Omega; X, \mu) \mid \text{for almost all } \omega, \ v(\omega) \in T_{K(\omega)}(x(\omega)) \right\}$$

8.6 Integral of Set-Valued Maps

Let $(\Omega, \mathcal{A}, \mu)$ be a complete $\sigma-$finite measure space and X a separable Banach space supplied with the norm $\|\cdot\|$.

In this section we investigate some properties of the integral of a set-valued map F from Ω into closed nonempty subsets of X.

We denote by $\mathcal{F}$ *the set of all integrable selections of* F:

$$\mathcal{F} = \{ f \in L^1(\Omega; X, \mu) \mid f(\omega) \in F(\omega) \text{ almost everywhere in } \Omega \}$$

A set-valued map $F : \Omega \rightsquigarrow X$ is called *integrably bounded* if there exists a nonnegative function $k \in L^1(\Omega; \mathbf{R}, \mu)$ such that

$$F(\omega) \subset k(\omega)B \text{ almost everywhere in } \Omega$$

In this case every measurable selection of F is an element of $\mathcal{F}$ thanks to Lebesgue's Theorem.

Aumann did suggest definition of the integral of a set-valued map in the following way:

Definition 8.6.1 *The integral of F on Ω is the set of integrals of integrable selections of F:*

$$\int_\Omega F d\mu := \left\{ \int_\Omega f d\mu \mid f \in \mathcal{F} \right\}$$

The integral is convex whenever F has convex images and we shall prove later that its closure is still convex even when the images of F are no longer convex.

It is also clear that

$$\forall \lambda \in \mathbf{R}, \quad \int_\Omega \lambda F d\mu = \lambda \int_\Omega F d\mu$$

Proposition 8.6.2 *Let us consider measurable, integrably bounded set-valued maps F, $F_i : \Omega \rightsquigarrow X$, $i = 1, 2$ with nonempty closed images and set $G(\omega) := \overline{F_1(\omega) + F_2(\omega)}$. Then*

1. — $\overline{\int_\Omega G d\mu} = \overline{\int_\Omega F_1 \mu + \int_\Omega F_2 d\mu}$

2. — $\overline{\int_\Omega \overline{co} F d\mu} = \overline{co} \int_\Omega F d\mu$

3. — $\forall p \in X^\star, \ \sigma\left(\int_\Omega F d\mu, p\right) = \int_\Omega \sigma(F(\omega), p)\mu(d\omega)$

4. — *If for some $x \in \int_\Omega F d\mu$ and $p \in X^\star$*

$$< p, x > = \sigma\left(\int_\Omega F d\mu, p\right)$$

then for every $\overline{f} \in \mathcal{F}$ satisfying $x = \int_\Omega \overline{f} d\mu$, we obtain

$$\text{for almost all } \omega \in \Omega, \quad < p, \overline{f}(\omega) > = \sigma(F(\omega), p)$$

Proof — Theorems 8.2.8 and 8.2.2 imply that G and $\overline{co}F$ are measurable. To prove the first equality it is enough to show that

$$\int_\Omega G d\mu \subset \overline{\int_\Omega F_1 d\mu + \int_\Omega F_2 d\mu}$$

since the other inclusion is obvious.

Consider a measurable selection g of G and let $(f_{in})_{n \geq 1}$ be a dense sequence of measurable selections of F_i (which exist by Characterization Theorem 8.1.4.) Then $(f_{1n} + f_{2m})_{n,m \geq 1}$ is a dense sequence of measurable selections of G. Set

$$G_{nm}(\omega) = \{f_{1i}(\omega) + f_{2j}(\omega) \mid 1 \leq i \leq n, \ 1 \leq j \leq m\}$$

Then

$$\forall\, \omega \in \Omega, \quad \lim_{n\to\infty, m\to\infty} d(g(\omega), G_{nm}(\omega)) \; = \; 0$$

Thanks to Corollary 8.2.13, there exists a measurable selection g_{nm} of G_{nm} satisfying

$$\|g(\omega) \; - \; g_{nm}(\omega)\| \; = \; d(g(\omega), G_{nm}(\omega))$$

Thus $g_{nm}(\omega) \in F_1(\omega) + F_2(\omega)$. Applying Theorem 8.2.9, we obtain measurable selections f^i_{nm} of F_i such that $f^1_{nm} + f^2_{nm} = g_{nm}$. This yields

$$d\left(\int_\Omega g_{nm} d\mu, \int_\Omega F_1 d\mu \; + \; \int_\Omega F_2 d\mu\right) \; = \; 0$$

By taking the limit when $n, m \to \infty$, the first statement ensues. It is clear that

$$\overline{co} \int_\Omega F d\mu \; \subset \; \overline{\int_\Omega \overline{co} F d\mu}$$

To prove the converse inclusion, fix $k \in L^1(\Omega; \mathbf{R}, \mu)$ with strictly positive values and a measurable selection g of $\overline{co}F$. Recall that Theorem 8.2.14 states that the support function

$$\omega \; \mapsto \; \sigma(\overline{co}F(\omega), p) \; = \; \sigma(F(\omega), p)$$

is measurable and observe that for every $p \in X^\star$

$$\left\langle p, \int_\Omega g d\mu \right\rangle \; \leq \; \int_\Omega \sigma(\overline{co}F(\omega), p)\mu(d\omega) \; = \; \int_\Omega \sigma(F(\omega), p)\mu(d\omega)$$

Fix $\varepsilon > 0$ and set $\varphi(x) =< p, x >$. Theorem 8.2.9 on the measurability of inverse images applied with

$$G(\omega) = [\sigma(F(\omega), p) - \varepsilon k(\omega), \sigma(F(\omega), p)]$$

yields that there exists a measurable selection $f \in \mathcal{F}$ such that

$$< p, f(\omega) > \geq \; \sigma(F(\omega), p) - \varepsilon k(\omega)$$

Consequently

$$\int_\Omega \sigma(F(\omega), p)\mu(d\omega) \leq \int_\Omega (< p, f > + \varepsilon k) d\mu \leq \sigma\left(\int_\Omega F d\mu, p\right) + \varepsilon \, \|k\|_{L^1}$$

Since $\varepsilon > 0$ and $p \in X^\star$ are arbitrary, using the separation theorem, we end the proof of the second and third statements.

Fix x, $\overline{f}$ as in the claim 4. From the third statement

$$\int_\Omega \sigma(F(\omega), p)\mu(d\omega) = \sigma(\int_\Omega F d\mu, p) = \int_\Omega < p, \overline{f} > d\mu$$

Consequently we deduce that

$$\int_\Omega \left(\sigma(F(\omega), p) - < p, \overline{f}(\omega) >\right) \mu(d\omega) = 0$$

But $\sigma(F(\omega), p) \geq < p, \overline{f}(\omega) >$. This achieves the proof. □

The famous Lyapunov's theorem on convexity of the range of a vector measure implies that the integral of a set-valued map is convex when X is a finite dimensional vector-space.

Recall that a set $A \in \mathcal{A}$ is called an *atom* (for the measure μ) if $\mu(A) > 0$ and for every measurable subset $A_1 \subset A$, $\mu(A_1)$ is equal to either 0 or $\mu(A)$. A measure is *nonatomic* if $\mathcal{A}$ does not contain atoms. Dirac and discrete measures are atomic. *The Lebesgue measure is nonatomic.*

A point z of a convex set K is called *extremal* if there is no x, $y \in K$ and $0 < \lambda < 1$ such that $z = \lambda x + (1 - \lambda)y$. We denote by $ext(K)$ the set of all extremal points of K.

When $X = \mathbf{R}^n$, then the integral of any set-valued map is convex (even when the values of F are not convex) and is closed when the map is integrably bounded:

Theorem 8.6.3 (Convexity of the Integral) *Let $F : \Omega \rightsquigarrow \mathbf{R}^n$ be a measurable set-valued map with nonempty closed images. If μ is nonatomic, then*

1. — The integral $\int_\Omega F d\mu$ is convex and extremal points of $\overline{co}\,(\int_\Omega F d\mu)$ are contained in $\int_\Omega F d\mu$.

2. — If a sequence $x_k \in \int_\Omega F d\mu$, $k = 1, 2, \ldots$ converges to an extremal point x of $\overline{co}\,(\int_\Omega F d\mu)$, then every sequence $f_k \in \mathcal{F}$ such that

$$\int_\Omega f_k d\mu = x_k$$

converges in $L^1(\Omega; \mathbf{R}^n, \mu)$ to some $f \in \mathcal{F}$ satisfying $\int_\Omega f d\mu = x$.

In particular for all $x \in ext\,(\overline{co} \int_\Omega F d\mu)$ there exists a unique $f \in \mathcal{F}$ with $x = \int_\Omega f d\mu$.

If in addition F is integrably bounded, then the integral of F is also compact.

When the dimension of the Banach space X is no longer finite, the integral may no longer be convex, but its closure is convex:

Theorem 8.6.4 (Convexity of the Closure of Integral) *Let X be a separable Banach space and $F : \Omega \rightsquigarrow X$ be a measurable set-valued map with nonempty closed images. If μ is nonatomic, then the closure of its integral is convex:*

1. — $\overline{\int_\Omega F d\mu} = \overline{co}\,(\int_\Omega F d\mu)$.

2. — If $x \in ext \int_\Omega F d\mu$, then $f \in \mathcal{F}$ satisfying $\int_\Omega f d\mu = x$ is unique.

3. — If F is integrably bounded, then

$$\int_\Omega \overline{co} F d\mu = \overline{\int_\Omega F d\mu}$$

Furthermore when X is reflexive and F has convex images and is integrably bounded, then the integral $\int_\Omega F d\mu$ is closed.

Remark — It is not difficult to check that the third statement remains true if we assume that $\mathcal{F} \neq \emptyset$ and $\mu(\Omega) < \infty$ instead of the integrable boundedness of F. □

It is convenient to introduce the following

Definition 8.6.5 *Consider a map $F : \Omega \rightsquigarrow X$. We say that $f \in \mathcal{F}$ is an extremal selection of F if $\int_\Omega f d\mu$ is an extremal point of $\overline{co}\,(\int_\Omega F d\mu)$. We denote by $\mathcal{F}_e$ the set of extremal selections:*

$$\mathcal{F}_e := \left\{ f \in \mathcal{F} \mid \int_\Omega f d\mu \in ext\left(\overline{co} \int_\Omega F d\mu\right)\right\}$$

We deduce from Theorem 8.6.3 and the Carathéodory Theorem:

Corollary 8.6.6 *Let $F : \Omega \rightsquigarrow \mathbf{R}^n$ be a measurable integrably bounded set-valued map with nonempty closed images. If μ is nonatomic, then for every*

$$x \in \int_\Omega F d\mu$$

there exist $n+1$ extremal selections $f_k \in \mathcal{F}_e$ and $n+1$ measurable sets $A_k \in \mathcal{A}$, $k = 0, ..., n$, such that

$$x = \int_\Omega \left(\sum_{k=0}^{n} \chi_{A_k} f_k \right) d\mu$$

where χ_{A_k} is the characteristic function of A_k.

Proof — Fix $x \in \int_\Omega F d\mu$ and let $\lambda_k \geq 0$, $x_k \in ext\,(\overline{co} \int_\Omega F d\mu)$, $k = 0, ..., n$ be such that

$$\sum_{k=0}^{n} \lambda_k x_k = x, \ \sum_{k=0}^{n} \lambda_k = 1$$

which exist by the Carathéodory theorem. Theorem 8.6.3 implies the existence of $f_k \in \mathcal{F}_e$ such that $\int_\Omega f_k d\mu = x_k$ for every $k = 0, ..., n$.

Then the set-valued map $G(\omega) := \{f_k(\omega)\}_{k=0,1,...,n}$ has closed nonempty images and is measurable. Since each x_k belongs to its integral, which is convex by Theorem 8.6.3, so does x. Hence there exists an integrable selection $f := \sum_{k=0}^{n} \chi_{A_k} f_k$ of G such that $x = \int_\Omega f d\mu$. $\square$

Theorems 8.4.1 and 8.6.3 imply the set-valued version of the Lebesgue Dominated Convergence Theorem:

Theorem 8.6.7 *Let $(\Omega, \mathcal{A}, \mu)$ be a complete σ–finite measure space and X a separable Banach space. Consider measurable set-valued maps*

$$K_n : \Omega \ni \omega \rightsquigarrow K_n(\omega) \subset X$$

with closed images.

1. — If the function $\omega \mapsto \sup_{n \geq 1} d(0, K_n(\omega))$ is integrable, then

$$\int_\Omega (\mathrm{Liminf}_{n\to\infty} K_n)\, d\mu \subset \mathrm{Liminf}_{n\to\infty} \left(\int_\Omega K_n d\mu \right)$$

2. — If the dimension of X is finite and $\omega \mapsto \sup_{n\geq 1}\|K_n(\omega)\|$ is integrable, then

$$\mathrm{Limsup}_{n\to\infty}\left(\int_\Omega K_n d\mu\right) \subset \int_\Omega (\mathrm{Limsup}_{n\to\infty} K_n)\, d\mu$$

Therefore, under the last asumptions, if the subsets $K_n(\omega)$ have a limit when $n \to \infty$ for almost every $\omega \in \Omega$, then

$$\mathrm{Lim}_{n\to\infty}\left(\int_\Omega K_n d\mu\right) = \int_\Omega (\mathrm{Lim}_{n\to\infty} K_n)\, d\mu$$

Proof — We begin by proving the first inclusion. An element v of the integral of the lower limit can be written $v = \int_\Omega f d\mu$ where f is the integrable selection of the lower limit of the maps $K_n(\cdot)$. By Theorem 8.4.1, it is the limit in $L^1(\Omega; X, \mu)$ of maps $f_n \in \mathcal{K}_n$, which are integrable selections of the set-valued maps $K_n(\cdot)$. Hence v is the limit of the integrals

$$v_n = \int_\Omega f_n d\mu \in \int_\Omega K_n d\mu$$

To prove the second inclusion, assume that

$$v := \lim_{j\to\infty} \int_\Omega f_{n_j}(\omega)\mu(d\omega) \in \mathrm{Limsup}_{n\to\infty} \int_\Omega K_n(\omega)\mu(d\omega)$$

where $f_{n_j} \in \mathcal{K}_{n_j}$. Since $\sup_{n\geq 0}\|K_n(\cdot)\|$ is integrable, the maps $f_{n_j}(\cdot)$ are integrally bounded. By the Dunford-Pettis Theorem, a subsequence (again denoted by) $f_{n_j}(\cdot)$ converges weakly to a map $f(\cdot)$, which belongs to the sequentially weak upper limit of the subsets $\mathcal{K}_n$.

The second statement of Theorem 8.4.1 implies that

$$\text{for almost all } \omega \in \Omega,\ \ f(\omega) \in \overline{co}\,(\mathrm{Limsup}_{n\to\infty} K_n(\omega))$$

Therefore, by Theorem 8.6.3, we infer that

$$v \in \int_\Omega \overline{co}\,(\mathrm{Limsup}_{n\to\infty} K_n(\omega))\, \mu(d\omega) = \int_\Omega \mathrm{Limsup}_{n\to\infty} K_n(\omega)\mu(d\omega)$$

This completes the proof of the second inclusion. □

8.7 Proofs of the Convexity of the Integral

We start with the more simple finite dimensional case.

8.7.1 Finite dimensional case

We assume here that $X = \mathbf{R}^n$ and F, μ verify all the assumptions of Theorem 8.6.3.

There exist many proofs of the first statement of Theorem 8.6.3. We have chosen among them the one based on a lemma which gives an estimate of the norm $\|f - g\|_{L^1}$ for $f, g \in \mathcal{F}$ in terms of the integrals of f and g. This lemma is intrinsically interesting since it has found several applications in control theory, for instance in the investigation of the Hölder behavior of controls[3].

We associate with any $p \in \mathbf{R}^n$ and $f, g \in \mathcal{F}$

$$\delta_p(f,g) := \sup\left\{\left\|\int_\Omega f d\mu - y\right\| \mid y \in \int_\Omega F d\mu \ \&\ \left\langle p, \int_\Omega g d\mu\right\rangle \leq \langle p, y\rangle\right\}$$

Lemma 8.7.1 *Let f and g belong to $\mathcal{F}$. Then for all $p \in \mathbf{R}^n$*

$$\left\langle p, \int_\Omega g d\mu\right\rangle \leq \left\langle p, \int_\Omega f d\mu\right\rangle \implies \|f - g\|_{L^1} \leq 4n\delta_p(f,g) \quad (8.4)$$

Furthermore if

$$\left\langle p, \int_\Omega f d\mu\right\rangle = \sigma\left(p, \int_\Omega F d\mu\right)$$

then

$$\forall\, g \in \mathcal{F}, \quad \|f - g\|_{L^1} \leq (2n-1)\delta_p(f,g)$$

Proof — We first observe that

$$\|f - g\|_{L^1} \leq 2n \sup_{A \in \mathcal{A}} \left\|\int_A (f - g) d\mu\right\| \quad (8.5)$$

Indeed, if f_i denote the components of f, let us define

$$A_i := \{\omega \mid f_i(\omega) - g_i(\omega) \geq 0\}$$

[3] See for instance [171].

We observe that

$$\|f-g\|_{L^1} \leq \sum_{i=1}^{n} \int_{\Omega} |f_i-g_i| d\mu = \sum_{i=1}^{n} \int_{A_i} (f_i-g_i) d\mu + \sum_{i=1}^{n} \int_{\Omega \backslash A_i} (g_i-f_i) d\mu$$

Then inequalities

$$\int_{A_i} (f_i - g_i) d\mu \leq \left\| \int_{A_i} (f-g) d\mu \right\| \leq \sup_{A \in \mathcal{A}} \left\| \int_{A} (f-g) d\mu \right\|$$

and

$$\int_{\Omega \backslash A_i} (g_i - f_i) d\mu \leq \left\| \int_{\Omega \backslash A_i} (g-f) d\mu \right\| \leq \sup_{A \in \mathcal{A}} \left\| \int_{A} (f-g) d\mu \right\|$$

imply estimate (8.5.)

It remains to show that for any $A \in \mathcal{A}$ and any $p \in X^\star$ satisfying $\langle p, \int_\Omega (f-g) d\mu \rangle \geq 0$, we have

$$\left\| \int_{A} (f-g) d\mu \right\| \leq 2\delta_p(f,g) \tag{8.6}$$

in order to prove (8.4.) We note that we can write

$$\int_A (f-g) d\mu = \int_\Omega f d\mu - y_A \text{ where } y_A := \int_A g d\mu + \int_{\Omega \backslash A} f d\mu \in \int_\Omega F d\mu$$

Therefore

$$\int_\Omega f d\mu - y_A + \int_\Omega f d\mu - y_{\Omega \backslash A} = \int_\Omega f d\mu - \int_\Omega g d\mu \tag{8.7}$$

Consequently assumption $\langle p, \int_\Omega (f-g) d\mu \rangle \geq 0$ and equation

$$\left\langle p, \int_\Omega f d\mu - y_A \right\rangle + \left\langle p, \int_\Omega f d\mu - y_{\Omega \backslash A} \right\rangle = \left\langle p, \int_\Omega (f-g) d\mu \right\rangle \tag{8.8}$$

imply that either $< p, \int_\Omega f d\mu - y_A >$ or $< p, \int_\Omega f d\mu - y_{\Omega \backslash A} >$ is not larger than $< p, \int_\Omega (f-g) d\mu >$. Say for instance that

$$\left\langle p, \int_\Omega f d\mu - y_A \right\rangle \leq \left\langle p, \int_\Omega f d\mu - \int_\Omega g d\mu \right\rangle$$

Then

$$\left\langle p, \int_\Omega g d\mu \right\rangle \leq \langle p, y_A \rangle \tag{8.9}$$

and therefore

$$\left\| \int_A (f-g) d\mu \right\| = \left\| \int_\Omega f d\mu - y_A \right\| \leq \delta_p(f,g) \tag{8.10}$$

Otherwise

$$\left\langle p, \int_\Omega f d\mu - y_{\Omega \backslash A} \right\rangle \leq \left\langle p, \int_\Omega f d\mu - \int_\Omega g d\mu \right\rangle$$

and interchanging the roles of A and $\Omega \backslash A$, we obtain

$$\left\| \int_{\Omega \backslash A} (f-g) d\mu \right\| = \left\| \int_\Omega f d\mu - y_{\Omega \backslash A} \right\| \leq \delta_p(f,g)$$

On the other hand we always have

$$\left\| \int_\Omega (f-g) d\mu \right\| \leq \delta_p(f,g)$$

Equation (8.7) implies that

$$\begin{cases} \|\int_A (f-g) d\mu\| = \|\int_\Omega f d\mu - y_A\| \\ \leq \|\int_\Omega f d\mu - \int_\Omega g d\mu\| + \left\|\int_\Omega f d\mu - y_{\Omega \backslash A}\right\| \leq 2\delta_p(f,g) \end{cases}$$

In summary, we proved inequality (8.6) and thus, the first statement of the lemma.

To prove the second one, let p be such that

$$\left\langle p, \int_\Omega f d\mu \right\rangle = \sigma\left(p, \int_\Omega F d\mu\right)$$

Changing coordinates, we may assume without any loss of generality that $p = (0, ..., 0, 1)$. Then Proposition 8.6.2 implies that

$$f_n(\omega) - g_n(\omega) \geq 0$$

almost everywhere in Ω. Hence in this case, proceeding as in the proof of inequality (8.5) we get

$$\left\{\begin{array}{lll} \|f-g\|_{L^1} & \leq & 2(n-1)\sup_{A\in\mathcal{A}}\|\int_A(f-g)d\mu\| + \|\int_\Omega(f-g)d\mu\| \\ & \leq & (2n-1)\sup_{A\in\mathcal{A}}\|\int_A(f-g)d\mu\| \end{array}\right. \tag{8.11}$$

Fix a measurable $A \subset \Omega$ and define y_A, $y_{\Omega\backslash A}$ as before. Both terms in the left hand side of (8.8) are nonnegative. Thus they are not greater than $\langle p, \int_\Omega(f-g)d\mu\rangle$. Hence (8.9) holds true and (8.10) follows. Consequently

$$\sup_{A\in\mathcal{A}}\left\|\int_\Omega (f-g)d\mu\right\| \leq \delta_p(f,g)$$

and from (8.11), the second statement ensues. □

Convexity Theorem 8.6.3 is the consequence of both the following Theorem 8.7.2 and the Lyapunov Convexity Theorem 8.7.3 we recall below.

We begin by proving

Theorem 8.7.2 *Any extremal point of $\overline{co}\,(\int_\Omega F d\mu)$ belongs to $\int_\Omega F d\mu$*
Furthermore if $f_k \in \mathcal{F}$, $k = 1, 2, ...$ are such that $\int_\Omega f_k d\mu$ converge to an extremal point of $\overline{co}\,(\int_\Omega F d\mu)$, then f_k converges to a selection $f \in \mathcal{F}$ in $L^1(\Omega; \mathbf{R}^n, \mu)$.

Proof — Set $K := \overline{co}\,(\int_\Omega F d\mu)$ and let $e \in K$ be an extremal point of K. Observe that to prove this theorem it is enough to show that every sequence $f_k \in \mathcal{F}$ such that

$$\lim_{k\to\infty}\int_\Omega f_k d\mu = e$$

converges in $L^1(\Omega; \mathbf{R}^n, \mu)$ to a selection $f \in \mathcal{F}$. We use Lemma 8.7.1 which allows derivation that it is a Cauchy sequence from the fact that the sequence of their integrals is a Cauchy sequence.

Step 1. We claim that *for every $\varepsilon > 0$ there is $\delta > 0$ such that*

$$\forall\, f_1, f_2 \in \mathcal{F} \text{ with } \left\|\int_\Omega f_i d\mu - e\right\| \leq \delta,\ \ i = 1, 2$$

we have $\|f_1 - f_2\|_{L^1} \leq \varepsilon$.

Indeed it is enough to consider the case $K \neq \{e\}$. Fix $\varepsilon > 0$ and pick $0 < \eta \leq \varepsilon/8n$ such that

$$Q \; := \; K \cap (e + \eta \Sigma^{n-1}) \; \neq \; \emptyset$$

where Σ^{n-1} denotes the unit sphere of $\mathbf{R}^n$. Since Q is compact, so is $coQ \subset B(e, \eta) \cap K$. The point e being extremal, it does not belong to coQ. Thus we can separate e from coQ: there exists $p \in \mathbf{R}^n$ with $\|p\| = 1$ such that

$$\langle p, e \rangle \;\; > \; \sigma_Q(p) := \max\{ < p, x > \; | \; x \in coQ \; \}$$

Define the open halfspace

$$H \; := \; \{ \, x \; \in \; X \, | \; < p, x > > \sigma_Q(p) \, \}$$

Then

$$H \cap (K \backslash B(e, \eta)) \; = \; \emptyset \tag{8.12}$$

Indeed, if $x \in H \cap (K \backslash B(e, \eta))$, then

$$y \; := \; e + \frac{\eta}{\|x - e\|}(x - e) \; = \; \left(1 - \frac{\eta}{\|x - e\|}\right) e + \frac{\eta}{\|x - e\|} x$$

belongs to Q, so that $< p, y > \leq \sigma_Q(p)$. But

$$< p, y > > \; \left(1 - \frac{\eta}{\|x - e\|}\right) \sigma_Q(p) \; + \; \frac{\eta}{\|x - e\|} \sigma_Q(p) \; = \; \sigma_Q(p)$$

so that we get a contradiction.

Therefore $H \cap K \subset B(e, \eta)$. Since e belongs to the interior of H, there exists $\delta > 0$ such that $B(e, \delta) \subset H$. Then

$$K \cap B(e, \delta) \; \subset \; H \cap K \; \subset \; B(e, \eta)$$

Consider $f_1, f_2 \in \mathcal{F}$ such that $|\int_\Omega f_i d\mu - e| < \delta$ for $i = 1, 2$. Then both $\int_\Omega f_i d\mu$ belong to $K \cap H$ so that they satisfy

$$\left\langle p, \int_\Omega f_i d\mu \right\rangle \; > \; \sigma_Q(p)$$

Changing notations if needed, we may assume that

$$\left\langle p, \int_\Omega f_2 d\mu \right\rangle \leq \left\langle p, \int_\Omega f_1 d\mu \right\rangle$$

Therefore

$$\begin{cases} \{ y \in \int_\Omega F d\mu \mid \langle p, \int_\Omega f_2 d\mu \rangle \leq < p, y > \} \\ \subset \{ y \in \int_\Omega F d\mu \mid \sigma_Q(p) < < p, y > \} \subset K \cap H \end{cases}$$

By applying Lemma 8.7.1 with $f = f_1$ and $g = f_2$ and using the inclusion $K \cap H \subset B(e, \eta)$, we deduce that

$$\begin{cases} \|f_1 - f_2\|_{L^1} \leq \\ 4n \sup \{ \|\int_\Omega f_1 d\mu - y\| \mid y \in \int_\Omega F d\mu, < p, \int_\Omega f_2 d\mu > \leq < p, y > \} \\ \leq 4n \sup \{ \|\int_\Omega f_1 d\mu - y\| \mid y \in H \cap K \} \leq 4n(2\eta) \leq \varepsilon \end{cases}$$

and our claim follows.

Step 2. We next claim that the extremal point e of K belongs to the closure of the integral of F:

$$\forall \eta > 0, \quad \left(\int_\Omega F d\mu \right) \cap B(e, \eta) \neq \emptyset$$

Indeed suppose for a moment that it does not hold true for some $\eta > 0$. Taking $\eta > 0$ small enough, we may assume that the set Q defined in Step 1 is nonempty. Let H be as in Step 1. If $B(e, \eta) \cap \int_\Omega F d\mu = \emptyset$ we deduce from (8.12) that

$$H \cap \int_\Omega F d\mu = \emptyset$$

and consequently $H \cap K = \emptyset$, because K is the convex hull of $\int_\Omega F d\mu$. This contradicts the fact that $e \in K \cap H$.

Step 3. Finally, we show that the extremal point e of K belongs to the integral of F: Indeed, by Step 2, it belongs to the closure of $\int_\Omega F d\mu$. Consider a sequence of elements $x_k := \int_\Omega f_k d\mu$ of $\int_\Omega F d\mu$ converging to e.

By Step 1, $(f_k)_{k\geq 1}$ is a Cauchy sequence in $L^1(\Omega; \mathbf{R}^n, \mu)$. Thus it converges to some f and has a subsequence $(f_{k_j})_{j\geq 1}$ converging almost everywhere to f. The values of F being closed, we finally obtain that for almost every $\omega \in \Omega$, $f(\omega) \in F(\omega)$. Therefore f is an integrable selection of F satisfying $\int_\Omega f d\mu = e$, so that e belongs to the integral of F. □

We recall now Lyapunov's Convexity Theorem on the range of a vector valued measure which we need to prove the convexity of the integral in Theorem 8.6.3.

Theorem 8.7.3 (Lyapunov's Convexity Theorem) *We assume that μ is nonatomic and let $f \in L^1(\Omega; \mathbf{R}^n, \mu)$. Then the set*

$$\nu(\mathcal{A}) := \left\{ \int_A f d\mu \right\}_{A\in\mathcal{A}}$$

is convex and compact.

The proof can be found in [45, Appendix C] and [8].

Proof of Theorem 8.6.3 — Fix $f_i \in \mathcal{F}$, $i = 1, 2$ and $\lambda \in [0, 1]$. We have to prove the existence of $f \in \mathcal{F}$ satisfying

$$\lambda \int_\Omega f_1 d\mu + (1-\lambda) \int_\Omega f_2 d\mu = \int_\Omega f d\mu$$

Define the vector measure $\nu : \mathcal{A} \mapsto \mathbf{R}^n \times \mathbf{R}^n$ by

$$\nu(A) = \left(\int_A f_1 d\mu, \int_A f_2 d\mu \right) \tag{8.13}$$

Then ν is finite and nonatomic. Lyapunov's Theorem 8.7.3 implies that the range of ν is convex. Since

$$\nu(\emptyset) = \{0\} \ \& \ \nu(\Omega) = \left(\int_\Omega f_1 d\mu, \int_\Omega f_2 d\mu \right)$$

there exists $A \in \mathcal{A}$ satisfying $\nu(A) = \lambda\nu(\Omega)$, i.e.,

$$\lambda \int_\Omega f_1 d\mu = \int_A f_1 d\mu, \ \lambda \int_\Omega f_2 d\mu = \int_A f_2 d\mu$$

Therefore the map $f = \chi_A f_1 + \chi_{\Omega \backslash A} f_2$ is an integrable selection of F we were looking for.

From Theorem 8.7.2, we know that the integral $\int_\Omega F d\mu$ contains all extremal points of the convex set $\overline{co} \int_\Omega F d\mu$. This proves the second part of the first statement. Theorem 8.7.2 implies also the second assertion.

Finally, if F is integrably bounded, then the set $\overline{co} \int_\Omega F d\mu$ is compact. Since $\int_\Omega F d\mu$ is convex and contains all the extremal points of its closed convex hull, we deduce from the Carathéodory theorem that

$$\int_\Omega F d\mu \;=\; co\left(\int_\Omega F d\mu\right) \;=\; \overline{co}\left(\int_\Omega F d\mu\right)$$

The proof is completed. □

8.7.2 Infinite Dimensional Case

In this section we impose on F the assumptions of the infinite dimensional Convexity Theorem 8.6.4.

To prove it, we need the following extension of Lyapunov's theorem to the infinite dimensional case.

Theorem 8.7.4 *Let Y be a separable Banach space. Consider an integrable map $f \in L^1(\Omega; Y, \mu)$ and define the map (vector measure) $\nu : \mathcal{A} \mapsto Y$ by*

$$\nu(A) \;=\; \int_A f d\mu$$

If μ is nonatomic, then $\overline{\nu(\mathcal{A})}$ is convex and compact.

Proof — Let $f_n \in L^1(\Omega; Y, \mu)$, $n = 1, 2, ...$ be simple maps converging to f in $L^1(\Omega; Y, \mu)$. Define continuous linear operators $T, T_n : L^\infty(\Omega; \mathbf{R}, \mu) \mapsto Y$, $n = 1, 2, ...$ by

$$\forall\, g \in L^\infty(\Omega; \mathbf{R}, \mu),\;\; T_n(g) \;=\; \int_\Omega g f_n d\mu, \;\; T(g) \;=\; \int_\Omega g f d\mu$$

which satisfy

$$\lim_{n \to \infty} \|T_n - T\| \;\leq\; \lim_{n \to \infty} \|f_n - f\|_{L^1} \;=\; 0$$

Since f_n is simple, the range of T_n is a finite dimensional space. Thus T_n is compact[4] and the limit T of the T_n's is compact as well.

To prove that $\overline{\nu(\mathcal{A})}$ is convex it is enough to check that for any $x, y \in \nu(\mathcal{A})$ and $\lambda \in [0,1]$, $\lambda x + (1-\lambda)y \in \overline{\nu(\mathcal{A})}$. Fix such x, y, λ and let $C, D \in \mathcal{A}$ be such that

$$x = \int_C f d\mu, \quad y = \int_D f d\mu$$

By the Lyapunov Convexity Theorem, for every n

$$\{ T_n(\chi_A) \mid A \in \mathcal{A}\} \text{ is convex}$$

Hence there exist $A_n \in \mathcal{A}$, $n = 1, \ldots$ such that

$$\forall n \geq 1, \quad \lambda T_n(\chi_C) + (1-\lambda)T_n(\chi_D) = T_n(\chi_{A_n})$$

Hence

$$\begin{cases} \|\nu(A_n) - (\lambda x + (1-\lambda)y)\| \leq \|T(\chi_{A_n}) - T_n(\chi_{A_n})\| \\ + \|\lambda T_n(\chi_C) + (1-\lambda)T_n(\chi_D) - (\lambda T(\chi_C) + (1-\lambda)T(\chi_D))\| \\ \leq \|T_n - T\| + \lambda\|T_n - T\| + (1-\lambda)\|T_n - T\| = 2\|T_n - T\| \end{cases}$$

converges to 0 since $\|T_n - T\|$ converges to 0. □

The above infinite version of Lyapunov's Convexity Theorem implies Theorem 8.6.4:

Proof of Theorem 8.6.4 — The proof is similar to the one of Theorem 8.6.3. We have to show that for all $f_i \in \mathcal{F}$, $i = 1, 2$, $\varepsilon > 0$ and $\lambda \in [0,1]$, there exists $f \in \mathcal{F}$ satisfying

$$\left\|\lambda \int_\Omega f_1 d\mu + (1-\lambda)\int_\Omega f_2 d\mu - \int_\Omega f d\mu\right\| \leq \varepsilon$$

Define the vector measure $\nu : \mathcal{A} \mapsto X \times X$ by (8.13.) By Theorem 8.7.4 the closure of the range of ν is convex. Since

$$\nu(\emptyset) = \{0\} \;\; \& \;\; \nu(\Omega) = \left(\int_\Omega f_1 d\mu, \int_\Omega f_2 d\mu\right)$$

[4] A linear operator is compact if it sends the unit ball into a relatively compact set. The set of compact operators is closed in the uniform operator topology ([148, Lemma VI.5.3]) or [44, Chapter 10].

we have for some $A \in \mathcal{A}$

$$\left\|\lambda \int_\Omega f_1 d\mu - \int_A f_1 d\mu\right\| + \left\|\lambda \int_\Omega f_2 d\mu - \int_A f_2 d\mu\right\| \leq \varepsilon$$

Setting $f = \chi_A f_1 + \chi_{\Omega\setminus A} f_2$, we deduce that $\overline{\int_\Omega F d\mu}$ is closed and convex.

To prove the second assertion, fix an extremal point x of $\int_\Omega F d\mu$ and let $f, g \in \mathcal{F}$ be such that $x = \int_\Omega f d\mu = \int_\Omega g d\mu$. The space X being separable, there exists a sequence of $p_n \in X^\star$, $n \geq 1$, which separates points of X. Define

$$A_n = \{\omega \in \Omega \mid < p_n, f(\omega) > > < p_n, g(\omega) >\}$$

and set

$$f_1 = \chi_{A_n} f + \chi_{\Omega\setminus A_n} g \quad \& \quad f_2 = \chi_{A_n} g + \chi_{\Omega\setminus A_n} f$$

Then $f_1, f_2 \in \mathcal{F}$. Observe that if $\mu(A_n) > 0$, then

$$\left\langle p_n, \int_\Omega f_1 d\mu \right\rangle > \left\langle p_n, \int_\Omega f_2 d\mu \right\rangle$$

and therefore $\int_\Omega f_1 d\mu \neq \int_\Omega f_2 d\mu$. On the other hand

$$\frac{1}{2}\int_\Omega f_1 d\mu + \frac{1}{2}\int_\Omega f_2 d\mu = x$$

Since x is extremal, we deduce that for every n, $\mu(A_n) = 0$. Similarly the set

$$A'_n = \{\omega \in \Omega \mid < p_n, g(\omega) > > < p_n, f(\omega) >\}$$

has zero measure. Since $(p_n)_{n\geq 1}$ separates points of X,

$$\mu(\{\, \omega \in \Omega \mid f(\omega) \neq g(\omega) \,\}) \leq \sum_{n\geq 1} (\mu(A_n) + \mu(A'_n)) = 0$$

The third statement follows from the first one and Proposition 8.6.2. To prove the last claim, assume that X is reflexive and that F is integrably bounded and has nonempty closed convex images. Then the set $\mathcal{F}$ is weakly compact in $\sigma(L^1(\Omega; X, \mu), L^\infty(\Omega; X, \mu))$ and the proof follows. □

8.8 The Bang-Bang Principle

When in addition Ω is an interval $[a, b] \subset \mathbf{R}$ and μ is the Lebesgue measure, then we have a Carathéodory type result for measurable selections known under the name of the Bang-Bang Principle:

Recall that for a closed convex set $K \subset \mathbf{R}^n$, a convex subset $Q \subset K$ is called an *extremal face* of K if for every $x, y \in K$ and $0 < \lambda < 1$

$$\lambda x + (1 - \lambda)y \in Q \Longrightarrow x,\ y \in Q$$

Every extremal point is an extremal face. The *dimension of the extremal face* Q is equal to the dimension of the affine space spanned by Q. The dimension of an extremal point is zero.

Definition 8.8.1 *Consider a map $F : [a, b] \rightsquigarrow \mathbf{R}^n$. A selection $f \in \mathcal{F}$ is called piecewise extremal if there is a partition*

$$a = t_0 < t_1 < \ldots < t_k = b$$

such that $\int_{t_i}^{t_{i+1}} f d\mu$ is an extremal point of $\overline{co}\left(\int_{t_i}^{t_{i+1}} F d\mu\right)$.

In the next theorem we assume that $\Omega = [a, b]$, $\mathcal{A}$ is the $\sigma-$algebra of Lebesgue measurable subsets of $[a, b]$ and μ is the Lebesgue measure.

Theorem 8.8.2 (Bang-Bang Principle) *Consider a measurable, integrably bounded set-valued map $F : [a, b] \rightsquigarrow \mathbf{R}^n$ with nonempty closed images. Then*

1. — $\int_a^b F d\mu$ is convex and compact.

2. — If x belongs to an extremal face of $\int_a^b F d\mu$ of dimension k, then there exists a piecewise extremal selection $f \in \mathcal{F}$ such that $x = \int_a^b f d\mu$ and the number of subintervals on which f is extremal can be made not greater than $k + 1$.

The above result in particular implies that for every $x \in \int_a^b F d\mu$, there exist $f \in \mathcal{F}$ and a partition of $[a, b]$ in $n + 1$ subintervals

$$a = t_0 < \ldots < t_n < t_{n+1} = b$$

such that for every $0 \leq i \leq n$, the restriction $f|_{[t_i,t_{i+1}]}$ is extremal for the restriction of F to $[t_i, t_{i+1}]$ and $x = \int_a^b f d\mu$.

Although the first assertion of the Bang-Bang Principle follows from Theorem 8.6.3, we obtain it directly as a by-product of the proof without any help of the Lyapunov Theorem. In this way the proof given below is self-contained.

Proof of Theorem 8.8.2 — Since μ is the Lebesgue measure we shall adopt the usual notation dt for $d\mu$ and $\mu(d\omega)$.

Observe that convexity and closedness of the integral $\int_a^b Fdt$ follows from the fact that every extremal face of $\overline{co}\int_a^b Fdt$ is a subset of $\int_a^b Fdt$. Thus our first statement would result from the second one.

To prove it, we proceed by induction with respect to the dimension m of the face Q.

We already know, thanks to Theorem 8.7.2, that for every measurable, integrably bounded set-valued map $G : [a,c] \rightsquigarrow \mathbf{R}^p$ with nonempty closed images, every extremal point of $\overline{co}\,(\int_a^c Gdt)$ belongs to $\int_a^c Gdt$ (where c is any number larger than a and $p \geq 1$ is any integer.)

Thus every extremal face of $\overline{co}\,(\int_a^c Gdt)$ of dimension $m = 0$ is a subset of $\int_a^c Gdt$.

Assume next that the above is valid for any such map G and every extremal face of dimension smaller than k. Fix a measurable, integrably bounded map $G : [a,c] \rightsquigarrow \mathbf{R}^p$ with nonempty closed images such that a nonempty set Q is a k-dimensional extremal face of $\overline{co}\,(\int_a^c Gdt)$. Since Q is extremal, we obtain

$$Q \cap ext\left(\overline{co}\int_a^c Gd\mu\right) = ext\, Q \tag{8.14}$$

Denote by $\mathcal{G}$ the set of all integrable selections of G. Pick an extremal point e of Q and let $g_e \in \mathcal{G}$ be an (extremal) selection of G such that $\int_a^c g_e(t)dt = e$. (It exists by (8.14) and Theorem 8.7.2.) Replacing G by $G - g_e$, we may assume that $e = 0$ and $g_e = 0$. Let P be the vector space spanned by Q.

We claim that

$$\forall\, g \in \mathcal{G} \text{ such that } \int_a^c g(t)dt \in P,\ g(t) \in P \text{ a.e. in } [a,c] \tag{8.15}$$

Indeed it is enough to show that for every measurable set A, we have $\int_A g(t)dt \in P$ (P is the intersection of a finite number of hyperplanes.) Fix a measurable A and define $g_1 = \chi_A g$, $g_2 = (\chi_{[a,c]\setminus A})g$. Then $g_1, g_2 \in \mathcal{G}$ and

$$\frac{1}{2}\int_a^c g_1(t)dt + \frac{1}{2}\int_a^c g_2(t)dt = \frac{1}{2}\int_a^c g(t)dt + \frac{1}{2}0 \in Q$$

Hence, because Q is an extremal face,

$$\int_a^c g_1(t)dt = \int_A g(t)dt \in Q \subset P$$

This ends the proof of our claim.

Consider the set-valued map $G_1(t) = G(t) \cap P$ with nonempty closed images. For every $\tau \in [a, c]$, set

$$\mathcal{R}(\tau) = \int_0^\tau G_1 dt$$

Then

$$\forall\, \tau < \tau',\ \mathcal{R}(\tau) \subset \mathcal{R}(\tau') \subset P$$

and the map $\tau \rightsquigarrow \overline{co}\mathcal{R}(\tau)$ is continuous, because G_1 is integrably bounded. Fix $x_0 \in Q$ and let $\bar{t} \in [a, c]$ be such that $x_0 \in \overline{co}\mathcal{R}(\bar{t})$ and for every $t < \bar{t}$, $x_0 \notin \overline{co}\mathcal{R}(t)$.

Then x_0 is a boundary point of $\overline{co}\mathcal{R}(\bar{t})$. Consequently it belongs to an extremal face Q' of $\overline{co}\left(\int_a^{\bar{t}} G_1 dt\right)$ of dimension smaller than k.

Since $\overline{co}\left(\int_a^{\bar{t}} G dt\right) \subset \overline{co}\left(\int_a^c G dt\right)$ and $Q' \subset Q$, (8.15) imply that Q' is also an extremal face of $\overline{co}\left(\int_a^{\bar{t}} G dt\right)$. Thus there exist $a = t_0 < t_1 < \ldots < t_k = \bar{t}$ and a selection g of G such that $\int_a^{\bar{t}} g(t)dt = x_0$ and for all $0 \le i \le k-1$, $\int_{t_i}^{t_{i+1}} g(t)dt$ is an extremal point of $\overline{co}\int_{t_i}^{t_{i+1}} G dt$. Set

$$g_0(t) = \begin{cases} g(t) & \text{if } t \in [a, \bar{t}] \\ 0 & \text{otherwise} \end{cases}$$

Then $g_0 \in \mathcal{G}$ and $\int_a^c g_0(t)dt = x_0$. Furthermore, since

$$\overline{co}\int_a^c G dt = \overline{co}\int_a^{\bar{t}} G dt + \overline{co}\int_{\bar{t}}^c G dt$$

$\int_t^c g_0(t)dt$ is an extremal point of $\overline{co} \int_t^c Gdt$. □

Remark — Note that if x_0 is not an extremal point of Q, then we have a choice of e in the proof above and, therefore, such x_0 can be reached by, at least, two different piecewise extremal elements. □

8.9 Invariant Measures & Poincaré's Recurrence Theorem

8.9.1 Linear Extension of Set-Valued Maps

Let X and Y be two compact metric spaces and $\mathcal{B}(X)$ and $\mathcal{B}(Y)$ their Borel σ-algebras. We recall that the dual $\mathcal{C}^\star(X)$ of the space of continuous functions is isomorphic to the space of Radon measures on X and that a continuous single-valued $f : X \mapsto Y$ can be extended to a continuous linear operator $\mathcal{F}$ from $\mathcal{C}^\star(X)$ on X to $\mathcal{C}^\star(Y)$ by the formula

$$\forall\, \mu \in \mathcal{C}^\star(X),\;\; \forall\, B \in \mathcal{B}(Y),\;\; \mathcal{F}(\mu)(B) \;:=\; \mu(f^{-1}(B))$$

This fact can be extended to set-valued maps $F : X \rightsquigarrow Y$. We denote by $\mathcal{P}(X) \subset \mathcal{C}^\star(X)$ the (weakly compact convex set) of probability measures on X.

Definition 8.9.1 *Let $F : X \rightsquigarrow Y$ be a set-valued map. Denote by $\mathcal{F}$, the* linear extension of F *— the set-valued map from $\mathcal{P}(X)$ to $\mathcal{P}(Y)$ defined in the following way:*

$\nu \in \mathcal{P}(Y)$ belongs to $\mathcal{F}(\mu)$ if and only if

$$\forall\, B \in \mathcal{B}(Y),\;\; \nu(B) \;\le\; \mu(F^{-1}(B))$$

We extend it as a set-valued map from $\mathcal{C}^\star(X)$ to $\mathcal{C}^\star(Y)$ by setting

$$\mathcal{F}(\mu) \;:=\; \begin{cases} \emptyset & \text{if } \mu \text{ is nonpositive,} \\ \{0\} & \text{if } \mu \;=\; 0 \\ \mu(X)\mathcal{F}(\mu/\mu(X)) & \text{if } \mu \text{ is positive} \end{cases}$$

Proposition 8.9.2 *Consider compact metric spaces X, Y and a closed set-valued map with nonempty values $F : X \rightsquigarrow Y$. Then $\mathcal{F}$ is a closed convex process with nonempty values.*

Furthermore, for any $\mu \in \mathcal{P}(X)$, ν belongs to $\mathcal{F}(\mu)$ if and only if

$$\text{for every open subset } \mathcal{O} \subset Y, \quad \nu(\mathcal{O}) \leq \mu(F^{-1}(\mathcal{O}))$$

The closed convex process $\mathcal{F}$ is a (set-valued linear) extension of F in the sense that for Dirac measures, $\delta_y \in \mathcal{F}(\delta_x)$ if and only if $y \in F(x)$.

If $G : Y \rightsquigarrow Z$ is a closed set-valued map with nonempty values from Y to a compact metric space Z, then the extension $\mathcal{H}$ of the product $H := G \circ F$ contains the product of the extensions: $\mathcal{G} \circ \mathcal{F} \subset \mathcal{H}$.

Proof — Let us consider a measure $\mu \in \mathcal{P}(X)$. The image $\mathcal{F}(\mu)$ is not empty, Indeed, by the Measurable Selection Theorem 8.1.3, F, being upper semicontinuous with closed images, is measurable, so that there exists at least one measurable selection f of F. Define ν_f by the formula $\nu_f(A) := \mu(f^{-1}(A))$, which is a probability measure. Since $f^{-1}(A) \subset F^{-1}(A)$, we infer that $\mu(f^{-1}(A)) \leq \mu(F^{-1}(A))$ so that ν_f belongs to $\mathcal{F}(\mu)$.

We prove now that $\mathcal{F}(\mu)$ can be defined as the set of measures $\nu \in \mathcal{P}(X)$ satisfying

$$\text{for every open subset } \mathcal{O} \subset Y, \quad \nu(\mathcal{O}) \leq \mu(F^{-1}(\mathcal{O}))$$

We first extend this formula to compact subsets $K \subset Y$. Since the graph of F, and thus, of F^{-1}, is closed and X is compact, F^{-1} is also upper semicontinuous. We then know that for any neighborhood $\mathcal{O}_n \supset F^{-1}(K)$, there exists an open neighborhood $\mathcal{M}_n \supset K$ satisfying $F^{-1}(\mathcal{M}_n) \subset \mathcal{O}_n$. Let us choose open subsets $\mathcal{O}_n$ such that

$$\mu(\mathcal{O}_n) \searrow \mu(F^{-1}(K))$$

Hence inequalities

$$\nu(K) \leq \nu(\mathcal{M}_n) \leq \mu(F^{-1}(\mathcal{M}_n)) \leq \mu(\mathcal{O}_n)$$

imply by going to the limit that $\nu(K) \leq \mu(F^{-1}(K))$.

Take now any measurable subset $B \in \mathcal{B}(Y)$. There exists a sequence of compact subsets $K_n \subset B$ such that

$$\nu(K_n) \nearrow \nu(B)$$

Then inequalities

$$\nu(K_n) \leq \mu(F^{-1}(K_n)) \leq \mu(F^{-1}(B))$$

imply that $\nu(B) \leq \mu(F^{-1}(B))$.

Assume that $y \in F(x)$. Then $\delta_y \in \mathcal{F}(\delta_x)$ since, for any open subset $\mathcal{O} \subset Y$, $\delta_y(\mathcal{O}) \leq \delta_x(F^{-1}(\mathcal{O}))$. This is obvious when $y \notin \mathcal{O}$. If not, the left-hand side is equal to 1, and so is the right-hand side, because

$$x \in F^{-1}(y) \subset F^{-1}(\mathcal{O})$$

Conversely, if $y \notin F(x)$, there exists an open subset $\mathcal{O} \ni y$ such that $F(x) \cap \mathcal{O} = \emptyset$, i.e., $x \notin F^{-1}(\mathcal{O})$. Then $\delta_y(\mathcal{O}) = 1$ and $\delta_x\left((F^{-1}(\mathcal{O}))\right) = 0$, so that $\delta_y \notin \mathcal{F}(\delta_x)$.

Formula $\mathcal{G} \circ \mathcal{F} \subset \mathcal{H}$ is obvious as well as convexity of the graph of $\mathcal{F}$.

It remains to prove that Graph($\mathcal{F}$) is closed when the spaces of Radon measures are supplied with the weak-$\star$ topology.

For that purpose, consider a sequence of measures $(\mu_n, \nu_n) \in$ Graph($\mathcal{F}$) converging to (μ, ν) in the weak-$\star$ topologies of the duals $\mathcal{C}^\star(X)$ and $\mathcal{C}^\star(Y)$ respectively.

It is sufficient to prove that the graph of the restriction of $\mathcal{F}$ to $\mathcal{P}(X)$ is weakly closed. Indeed, when the measures μ_n and $\nu_n \in \mathcal{F}(\mu_n)$ are positive and converge weakly to μ and ν, subsequences of the probability measures $\overline{\mu}_n := \mu_n/\mu_n(X)$ and

$$\overline{\nu}_n := \nu_n/\mu_n(X) \in \mathcal{F}(\overline{\mu}_n)$$

converge weakly to probability measures $\overline{\mu}$ and $\overline{\nu} \in \mathcal{F}(\overline{\mu})$ respectively, because the sets of probability measures are weakly compact and the graph of the restriction of $\mathcal{F}$ to $\mathcal{P}(X)$ is assumed to be weakly closed. Since the measures $\mu_n(X)$ converge to $\mu(X)$, we deduce that $\mu = \mu(X)\overline{\mu}$ and $\nu = \mu(X)\overline{\nu}$. Then $\nu = 0$ when $\mu = 0$ and otherwise,

$$\nu = \mu(X)\overline{\nu} \in \mu(X)\mathcal{F}(\mu/\mu(X))$$

Hence we consider a sequence $(\mu_n, \nu_n) \in \mathcal{P}(X) \times \mathcal{P}(Y)$ of the graph of $\mathcal{F}$ converging to (μ, ν). In order to prove that $\nu \in \mathcal{F}(\mu)$, it is enough to check that for any open subset $\mathcal{O} \subset X$, inequality

$$\nu(\mathcal{O}) \leq \mu(F^{-1}(\mathcal{O}))$$

holds true thanks to the first part of the proposition.

Fix an open subset $\mathcal{O} \subset X$. Since X is compact, $\mathcal{O}$ is the union of an increasing sequence of open subsets $\mathcal{O}_p \subset \mathcal{O}, p \geq 1$ such that for every $p \geq 1$, $\overline{\mathcal{O}_p} \subset \mathcal{O}$. Let us fix p and observe that $F^{-1}(\overline{\mathcal{O}_p})$ is compact, as the image of a compact by the upper semicontinuous set-valued map F^{-1}.

Inequalities

$$\forall\, n \geq 1, \;\; \nu_n(\mathcal{O}_p) \;\leq\; \mu_n\left(F^{-1}(\mathcal{O}_p)\right) \;\leq\; \mu_n\left(F^{-1}(\overline{\mathcal{O}_p})\right)$$

imply, thanks to a version of Alexandroff's Theorem recalled just after the end of the proof, that

$$\left\{ \begin{array}{l} \nu(\mathcal{O}_p) \;\leq\; \liminf_{n\to\infty} \nu_n(\mathcal{O}_p) \\ \leq\; \limsup_{n\to\infty} \mu_n\left(F^{-1}(\overline{\mathcal{O}_p})\right) \;\leq\; \mu\left(F^{-1}(\overline{\mathcal{O}_p})\right) \;\leq\; \mu\left(F^{-1}(\mathcal{O})\right) \end{array} \right.$$

It remains to observe that $\nu(\mathcal{O}) = \sup_{p\geq 1} \nu(\mathcal{O}_p)$ to conclude. $\square$

We prove now the version of Alexandroff's Theorem (see [148, Theorem 4.9.15]) we needed:

Theorem 8.9.3 (Alexandroff) *Let X be a compact metric space. Consider a sequence of Radon probability measures $\mu_n \in \mathcal{P}(X)$ converging weakly to μ. Then, for any open subset $\mathcal{O} \subset X$,*

$$\mu(\mathcal{O}) \;\leq\; \liminf_{n\to\infty} \mu_n(\mathcal{O})$$

and for any closed subset $K \subset X$,

$$\limsup_{n\to\infty} \mu_n(K) \;\leq\; \mu(K)$$

Proof — Let $\mathcal{O} \subset X$ be an open subset. Since a Radon measure μ is regular, we can associate with any $\varepsilon > 0$ a compact subset $K \subset \mathcal{O}$ such that $\mu(\mathcal{O}\backslash K) \leq \varepsilon$.

Let $f \in \mathcal{C}(X)$ be a continuous function equal to 1 on the closed subset K and to 0 on the complement of $\mathcal{O}$. We observe that $\chi_K \leq f \leq \chi_{\mathcal{O}}$ and thus, that

$$\mu(\mathcal{O}) \;\leq\; \mu(K) + \varepsilon \;\leq\; \int f d\mu + \varepsilon$$

Therefore, inequalities $\int f d\mu_n \leq \mu_n(\mathcal{O})$ imply that

$$\mu(\mathcal{O}) \leq \int f d\mu + \varepsilon \leq \liminf_{n\to\infty} \mu_n(\mathcal{O}) + \varepsilon$$

We conclude by letting ε converge to 0.

The proof of the second inequality for compact subsets K is analogous, since for any $\varepsilon > 0$, there exists an open $\mathcal{O} \supset K$ such that $\mu(\mathcal{O}\backslash K) \leq \varepsilon$. Define the continuous function f as above. Inequalities $\mu_n(K) \leq \int f d\mu_n$ imply that

$$\limsup_{n\to\infty} \mu_n(K) \leq \int f d\mu \leq \mu(\mathcal{O}) \leq \mu(K) + \varepsilon \quad \square$$

Remark — We deduce that for any open subset $\mathcal{O}$ such that $\mu(\mathcal{O}) = \mu(\overline{\mathcal{O}})$, weak convergence of probability measures μ_n to μ imply that $\mu_n(\mathcal{O})$ converges to $\mu(\mathcal{O})$. The converse is also true. (See the proof of [148, Theorem 4.9.15]) for instance.) □

8.9.2 Invariant Measures

Taking $X = Y$, we are able to derive the existence of invariant measures by showing that the extension $\mathcal{F}$ has fixed points on $\mathcal{P}(X)$.

Theorem 8.9.4 *Let X be a compact metric space and $F : X \rightsquigarrow X$ be a closed set-valued map with nonempty values. Then there exists an invariant probability measure μ, i.e., satisfying*

$$\forall A \in \mathcal{B}(X), \ \mu(A) \leq \mu(F^{-1}(A)) \tag{8.16}$$

Proof — Let $\mathcal{F}$ be the set-valued map from $\mathcal{P}(X)$ to itself associating with any $\mu \in \mathcal{P}(X)$ the (nonempty) set of probability measures $\nu \in \mathcal{F}(\mu)$. Then the above proposition states that $\mathcal{F}$ is a map with closed graph and convex values from the convex compact subset $\mathcal{P}(X) \subset \mathcal{C}^\star(X)$ to itself, and thus, upper semicontinuous with convex compact values. Kakutani-Fan's Fixed-Point Theorem 3.2.3[5]

[5]Recall that the proof we provided, based on the Ky Fan inequality, showed that Ky Fan's theorem remains true in locally convex Hausdorff topological vector spaces.

implies the existence of a fixed-point $\mu \in \mathcal{P}(X)$ of $\mathcal{F}$, which is a measure satisfying

$$\forall\, A \in \mathcal{B}(X), \;\; \mu(A) \;\leq\; \mu(F^{-1}(A)) \tag{8.17}$$

Therefore, μ is invariant by F. $\square$

Naturally, we can derive most of the properties of invariant measures enjoyed by single-valued maps. Let us show for instance that Poincaré's Recurrence Theorem holds true for closed set-valued dynamical systems.

Theorem 8.9.5 (Poincaré's Recurrence) *Let X be a compact metric space, $F : X \rightsquigarrow X$ be a closed set-valued map and $\mu \in \mathcal{P}(X)$ an invariant measure of F. For any Borelian subset $B \subset X$, let*

$$B_\infty \;:=\; \bigcap_{N\geq 0}\bigcup_{n\geq N} F^{-n}(B)$$

be the subset of points x such that for all N, there exists $n \geq N$ such that $F^{-n}(x) \cap B \neq \emptyset$. Then the measure of $B \cap B_\infty$ is equal to the measure of B.

Proof — The proof is a straightforward extension of the proof in the single-valued case: We introduce the subsets

$$B_N := \bigcup_{n\geq N} F^{-n}(B)$$

and we observe that $B \subset B_0$, that $\cdots \subset B_N \subset B_{N-1} \subset \cdots \subset B_0$ and that $B_N = F^{-N}(B_0)$.

Since μ is invariant, we deduce that

$$\mu(B) \;\leq\; \mu(F^{-1}(B)) \;\leq\; \cdots \;\leq\; \mu(F^{-N}(B)) \;\leq\; \cdots$$

and thus, that

$$\mu(B_0) \;\leq\; \mu(F^{-N}(B_0)) \;=\; \mu(B_N) \;\leq\; \mu(B_0)$$

Since the sequence of B_N is not increasing and since $\mu(B_N) = \mu(B_0)$, we infer that $\mu(B_\infty) = \mu(B_0)$. Therefore,

$$\mu(B \cap B_\infty) \;=\; \mu(B \cap B_0) \;=\; \mu(B) \;\; \square$$

Chapter 9

Selections and Parametrization

Introduction

One of the first questions of set-valued analysis was how to relate set-valued and single-valued maps, mainly in the hope of avoiding the necessity of dealing with set-valued maps.

This can be done in two ways:

1. — Find a *selection* $f : X \mapsto Y$ of a set-valued map $F : X \leadsto Y$, i.e., a single-valued map f satisfying

$$\forall\, x \in X, \;\; f(x) \in F(x)$$

2. — Find a *parametrization* (U, f), where U is a *control* or *parameter* space and $f : X \times U \mapsto Y$ is a single-valued map satisfying

$$\forall\, x \in X, \quad F(x) = \{f(x,u)\}_{u \in U}$$

Naturally, we require that in both cases the selection f or the parametrization (U, f) inherits the regularity properties of F, such as continuity or Lipschitzianity.

Furthermore, keeping in mind applications to control theory, we shall also study the case when both F and f are time-dependent in a measurable way.

We proved in the preceding chapter the existence of a measurable selection of a measurable map. When we ask for more regularity of a selection, we meet more technical difficulties. This is why continuous (or Lipschitz) selections are associated with continuous (respectively Lipschitz) set-valued maps with closed *convex* images.

These selections should be obtained in a constructive way whenever this is possible.

The most natural constructive way amounts to using the *minimal selection*, i.e., the selection f with minimal norm. Unfortunately, this procedure does not preserve semicontinuity, but only continuity for maps with closed convex images, as one can see in Section 3. Such selection is also Carathéodory for Carathéodory maps with closed convex images (see Section 5.)

We shall prove the existence of *continuous selections* of lower semicontinuous maps with closed convex images (Michael's Theorem) in Section 1, and the existence of *approximate continuous selections* (Cellina's Theorem) of upper semicontinuous maps with closed convex images in Section 2.

If the set-valued map is Lipschitz continuous, then, contrary to an intuitive wish, it may happen that its minimal selection is not Lipschitz.

There exist different ways to get Lipschitz selections for Lipschitz set-valued maps with convex images. The barycentric selection, which also comes easily to mind, is Lipschitz, but difficult to handle.

By constrast, the selection based on the *Steiner point*[1] enjoys attractive properties. For instance, it preserves measurability, continuity and Lipschitzianity of set-valued maps.

We first single out a point of a nonempty convex compact set

---

[1]The Steiner point of a convex compact set, known for over a century, has many important properties. It is defined for subsets of Euclidean space $\mathbf{R}^n$.

Jacob Steiner (1796-1863) introduced this concept in 1840. From a peasant family, he had only a village school education before coming late in life to Mathematics. He had a strong preference for geometrical proofs over analytical ones. Actually, he never provided any of them and succeeded however in keeping his audience of *Berliner Universität* awake without drawing any pictures. He was overshadowed by Diersteweg, in Mörs, who purposefully draws the curtains of the class room to darken it before his lectures (...die Stube auch noch ausdrücklich verfinsterte!.) Certainly he intended to help his listeners to dream freely of geometrical (with no doubts curvaceous) shapes...

$K \subset \mathbf{R}^n$, the *Steiner point* $s_n(K)$, and characterize it.

The Steiner selection $s_n(\cdot)$ being Lipschitz in the Hausdorff metric, the single-valued map $x \mapsto s_n(F(x))$ is a Lipschitz selection whenever the Lipschitz set-valued map $F : X \rightsquigarrow \mathbf{R}^n$ has convex compact images.

To be able to deal also with unbounded set-valued maps, we prove in Section 4 an Intersection Lemma, which enables us to associate with a closed convex set its convex compact subset in a Lipschitz way. Hence we obtain in a constructive way Lipschitz selections for Lipschitz set-valued maps with closed convex images.

The Steiner selection and the Intersection Lemma allow us to find selections (Section 5) and parametrizations (Sections 6 and 7) of continuous, Carathéodory and Lipschitz maps with closed convex images, preserving regularity of the set-valued map under consideration.

9.1 Case of lower semicontinuous maps

We shall prove the celebrated Michael's theorem stating that lower semicontinuous convex-valued maps do have *continuous selections*.

Definition 9.1.1 *Consider a set-valued map $F : X \rightsquigarrow Y$ with nonempty images. A single valued map $f : X \mapsto Y$ is called a* selection *of F if for every $x \in X$, $f(x) \in F(x)$.*

Theorem 9.1.2 (Michael's Theorem) *Let F be a lower semicontinuous set-valued map with closed convex values from a compact metric space X to a Banach space Y. It does have a continuous selection.*

Corollary 9.1.3 *Let F be as in Theorem 9.1.2. Consider a subset $K \subset X$ and a continuous single-valued map $\varphi : K \mapsto Y$ such that for every $x \in K$, $\varphi(x) \in F(x)$.*

Then φ can be extended to a continuous selection of F, i.e., F has a continuous selection f (defined on the whole space X) such that the restriction of f to K is equal to φ.

In particular this yields that for every $\overline{y} \in F(\overline{x})$ there exists a continuous selection f of F such that $f(\overline{x}) = \overline{y}$.

Proof — It is enough to consider the lower semicontinuous set-valued map

$$G(x) := \begin{cases} F(x) & \text{if } x \notin K \\ \{\varphi(x)\} & \text{otherwise} \end{cases}$$

and to apply Theorem 9.1.2. □

The proof of this Theorem requires the following Lemma.

Lemma 9.1.4 *Let F be a lower semicontinuous convex-valued map from a compact metric space X to a Banach space Y. Then for each $\varepsilon > 0$, there exists a continuous function f_ε satisfying*

$$\forall\, x \in X, \quad f_\varepsilon(x) \in B(F(x), \varepsilon)$$

Proof — Since F is lower semicontinuous, we can associate with any $x \in K$ and any $y_x \in F(x)$ an open neighborhood $\mathcal{U}_x$ such that

$$\forall\, x' \in \mathcal{U}_x, \quad B(y_x, \varepsilon) \cap F(x') \neq \emptyset$$

Since X is compact, it can be covered by n such neighborhoods $\mathcal{U}_{x_i}$. Let us consider a *continuous partition of unity*[2] $\{a_i\}_{i=1,\ldots,n}$ associated with this covering, i.e. $a_i : X \mapsto [0,1]$ are continuous functions such that a_i vanishes outside of $\mathcal{U}_{x_i}$ and

$$\forall\, x \in X, \quad \sum_{i=1}^{n} a_i(x) = 1$$

Define f_ε by:

$$f_\varepsilon(x) := \sum_{i=1}^{n} a_i(x) y_{x_i}$$

It is continuous since the functions a_i are continuous. Let

$$I(x) := \{i = 1, \ldots, n \mid a_i(x) > 0\}$$

It is not empty since $\sum_{i=1}^{n} a_i(x) = 1$. When $i \in I(x)$, then x belongs to $\mathcal{U}_{x_i}$ and thus,

$$y_{x_i} \in B(F(x), \varepsilon)$$

[2]See Chapter 3, Section 1.

The set $B(F(x), \varepsilon)$ being convex, we infer that $f_\varepsilon(x)$ belongs to $B(F(x), \varepsilon.)$ $\square$

We are now ready to prove the Michael Theorem.

Proof of Theorem 9.1.2 — We shall construct by induction a sequence of continuous maps $u_n : X \mapsto Y$, $n = 1, \ldots$ satisfying the following properties:

$$\begin{cases} i) & \forall\, x \in X, \quad d(u_n(x), F(x)) < 2^{-n} \\ ii) & \forall\, x \in X, \quad \|u_n(x) - u_{n-1}(x)\| < 2^{-n-1} \end{cases} \tag{9.1}$$

For $n = 1$, we apply the lemma with $\varepsilon = 1/4$. Assume that we have constructed the maps u_n up to n and let us construct u_{n+1}. We introduce for that purpose the set-valued map F_{n+1} defined by

$$\forall\, x \in X, \quad F_{n+1}(x) := \overset{o}{B}(u_n(x), 2^{-n}) \cap F(x)$$

By (9.1) $i)$, $F_{n+1}(x)$ is convex and nonempty. The set-valued map F_{n+1} is lower semicontinuous: Indeed, if x_p converges to x and if y belongs to $F_{n+1}(x)$, by the lower semicontinuity of F, there exist elements $y_p \in F(x_p)$ converging to y. For p large enough, we see that

$$\|y_p - u_n(x_p)\| < 2^{-n}$$

because u is continuous, and therefore, that y_p belongs to $F_{n+1}(x_p)$.

This being established, we deduce from Lemma 9.1.4 the existence of a continuous map u_{n+1} satisfying:

$$\forall\, x \in X, \quad d(u_{n+1}(x), F_{n+1}(x)) < 2^{-(n+1)}$$

which, together with the definition of F_{n+1}, implies (9.1) for $n + 1$.

We infer from inequalities (9.1) $ii)$ that the sequence of functions u_n is a Cauchy sequence in the Banach space of continuous maps from the compact space X to the Banach space Y. They converge uniformly to a continuous map $u : X \mapsto Y$. Inequalities (9.1) $i)$ imply that for all $x \in X$, $u(x)$ belongs to $F(x)$, for this subset is closed by assumption. Hence u is the continuous selection we were looking for.

9.2 Case of upper semicontinuous maps

Upper semicontinuous set-valued maps in general do not have continuous selections even when their values are closed and convex. The following example illustrates this issue:

Example Consider the upper semicontinuous set-valued map $F : \mathbf{R} \rightsquigarrow \mathbf{R}$ defined by

$$F(x) := \begin{cases} \{-1\} & \text{if} \quad x < 0 \\ [-1,1] & \text{if} \quad x = 0 \\ \{1\} & \text{if} \quad x > 0 \end{cases}$$

Clearly F does not have any continuous selections defined on $\mathbf{R}$. □

This is why it is more realistic to speak about *approximate selections*. We state the following result due to A. Cellina:[3]

Theorem 9.2.1 *Let $F : X \rightsquigarrow Y$ be an upper semicontinuous map from a compact metric space X to a Banach space Y. If the values of F are nonempty and convex, then for every $\varepsilon > 0$, there exists a locally Lipschitz single valued map $f_\varepsilon : X \mapsto Y$ such that*

$$\mathrm{Graph}(f_\varepsilon) \subset B(\mathrm{Graph}(F), \varepsilon)$$

and for every $x \in X$, $f_\varepsilon(x)$ belongs to the convex hull of the image of F.

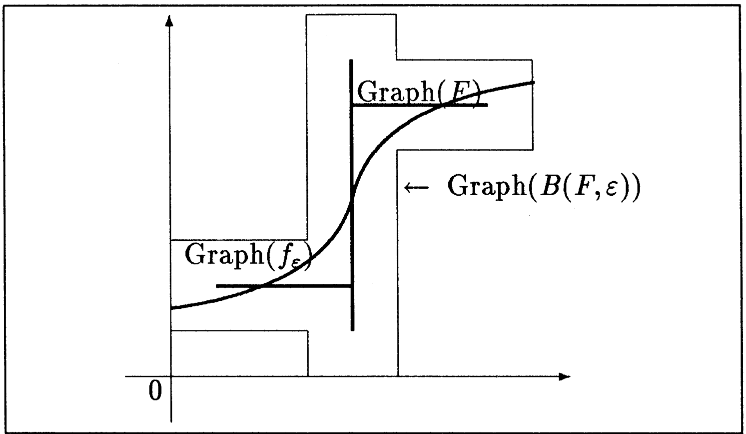
Remark — This inclusion implies that the *graphical upper limit of the approximate selections is contained in F*. □

Proof — Fix $\varepsilon > 0$. Since F is upper semicontinuous, for every $x \in X$, there exists $0 < \delta_x < 2\varepsilon$ such that

$$\forall\, y \in B(x, \delta_x), \;\; F(y) \subset F(x) + \frac{\varepsilon}{2}B$$

[3]We prove here the approximate selection result only for a compact metric space X, but the result is valid with a similar proof for any metric space X. It can be proved using paracompactness of metric spaces and the fact that a locally Lipschitz partition of unity may be subordinated to any locally finite covering of a metric space. Details can be found in [33, Section 1].

Figure 9.1: Approximate Selection of an Upper Semicontinuous Map



Approximate Selection in a Neighborhood of Graph(F)

The family of balls $\{\overset{o}{B}(x,\delta_x/4)\}_{x\in X}$ covers X. Since X is compact, there exists a finite sequence I of indices such that the family of balls $\{\overset{o}{B}(x_i,\delta_{x_i}/4)\}_{i\in I}$ covers X.

We set $\delta_i = \delta_{x_i}$ and take a locally Lipschitz partition of unity $\{a_i\}$ subordinated to this covering[4]. That is $a_i : X \mapsto [0,1]$ are locally Lipschitz functions such that for all $i \in I$, a_i vanishes outside of $B(x_i,\delta_i/4)$ and

$$\forall\, x \in X, \quad \sum_{i\in I} a_i(x) = 1$$

Let us associate with every $i \in I$ a point $y_i \in F(B(x_i,\delta_i/4))$ and define the map $f_\varepsilon : X \mapsto Y$ by

$$f_\varepsilon(x) = \sum_{i\in I} a_i(x) y_i$$

Then f_ε is locally Lipschitz and for every $x \in X$, $f_\varepsilon(x) \in coF(X)$.

To show that f_ε is the desired approximation, fix $x \in X$ and let $I(x)$ denote the subset of all $i \in I$ such that $a_i(x) \neq 0$. Thus

[4]See for instance [33, Theorem 1.1.2].

$x \in B(x_i, \delta_i/4)$ for every $i \in I(x)$. Therefore for all $i, j \in I(x)$

$$d(x_i, x_j) \leq d(x, x_i) + d(x, x_j) \leq (\delta_i + \delta_j)/4$$

Fix $k \in I(x)$ satisfying

$$\delta_k = \max_{i \in I(x)} \delta_i$$

Then the latter inequality implies that

$$\forall\, i \in I(x), \quad x_i \in B(x_k, \delta_k/2)$$

Consequently

$$y_i \in F(B(x_i, \delta_i/4)) \subset F(B(x_k, \delta_k)) \subset F(x_k) + \frac{\varepsilon}{2}B$$

Since the right hand side of the above relation is convex, from the definition of f_ε follows that

$$f_\varepsilon(x) \in F(x_k) + \frac{\varepsilon}{2}B$$

Pick $z_k \in F(x_k)$ satisfying $\|z_k - f_\varepsilon(x)\| \leq \varepsilon/2$. Thus, by the choice of δ_x and z_k,

$$d((x, f_\varepsilon(x)), (x_k, z_k)) \leq d(x, x_k) + \|z_k - f_\varepsilon(x)\| \leq \frac{\delta_k}{4} + \frac{\varepsilon}{2} \leq \varepsilon$$

which ends the proof. □

9.3 Minimal Selection

Consider a metric space X, a Banach space Y and a set-valued map $F : X \rightsquigarrow Y$.

We define the *minimal map*

$$m(F(x)) := \left\{ u \in F(x) \mid \|u\| = \min_{y \in F(x)} \|y\| \right\} \tag{9.2}$$

When Y is a Hilbert space, or, more generally, a reflexive strictly convex space[5] and F has closed convex images, the minimal map is single valued. In this case it is called the *minimal selection*.

[5] A Banach space Y is called *strictly convex* if for all $x, y \in Y$ which are not colinear, we have $\|x + y\| < \|x\| + \|y\|$.

A quite natural question arises:

What are continuity properties of the minimal selection?

We prove below that if F is continuous and takes its values in a compact subset of a strictly convex reflexive Banach space, then the minimal selection is continuous.

The upper semicontinuity (or lower semicontinuity) of F, even when it takes convex values in a compact subset, is not strong enough to imply the continuity of the minimal selection, as the following examples show:

Example — We consider the set-valued map $F : \mathbf{R} \rightsquigarrow \mathbf{R}$ defined by

$$F(x) := \begin{cases} \{2\} & \text{if} \quad x \neq 0 \\ [1,2] & \text{if} \quad x = 0 \end{cases}$$

It is upper semicontinuous with compact convex values and its minimal selection is obviously not continuous at zero. □

Example — We consider the set-valued map $F : \mathbf{R} \rightsquigarrow \mathbf{R}$ defined by

$$F(x) := \begin{cases} [0,1] & \text{if} \quad x \neq 0 \\ \{1\} & \text{if} \quad x = 0 \end{cases}$$

It is lower semicontinuous with compact convex values and its minimal selection is not continuous at zero. □

First of all we investigate continuity properties of the function $d(0, F(\cdot))$.

Lemma 9.3.1 *Let F be a set-valued map from a metric space X to a Banach space Y with nonempty images. Set*

$$\|F(x)\| = \sup_{y \in F(x)} \|y\|$$

1. — If F is lower semicontinuous (or upper semicontinuous), then the function $x \mapsto d(0, F(x))$ is upper semicontinuous (respectively lower semicontinuous.)

2. — If F has bounded images and is lower semicontinuous (or upper semicontinuous), then the function $x \mapsto \|F(x)\|$ is lower semicontinuous (respectively upper semicontinuous.)

Consequently if F is continuous, then so is $d(0, F(\cdot))$ and if in addition F has bounded images, then $\|F(\cdot)\|$ is also continuous.

Furthermore if F is Lipschitz with a Lipschitz constant c, then $d(0, F(\cdot))$ is c-Lipschitz. If moreover F has bounded images, then $\|F(\cdot)\|$ is also c-Lipschitz.

In the above *c-Lipschitz* means Lipschitz with the constant c.

Proof — If F is lower semicontinuous, then, by Corollary 1.4.17, the function $x \mapsto d(0, F(x))$ is upper semicontinuous and, in the case when F has bounded images, the function $x \mapsto \|F(x)\|$ is lower semicontinuous by Theorem 1.4.16.

If F is upper semicontinuous, then for every $\overline{x} \in X$, $\varepsilon > 0$ there exists a neighborhood $\mathcal{N}_\varepsilon$ of $\overline{x}$ such that for all $x \in \mathcal{N}_\varepsilon$, we have $F(x) \subset F(\overline{x}) + \varepsilon B$. Hence

$$d(0, F(x) \ \geq \ d(0, F(\overline{x})) - \varepsilon$$

and therefore $d(0, F(\cdot))$ is lower semicontinuous. On the other hand if $F(\overline{x})$ is bounded,

$$\|F(x)\| \ \leq \ \|F(\overline{x})\| + \varepsilon$$

This yields that $\|F(\cdot)\|$ is upper semicontinuous whenever images of F are bounded.

The latter statement follows from the very definition of Lipschitz set-valued maps. □

Proposition 9.3.2 *Let F be a lower semicontinuous set-valued map from a metric space X to a Banach space Y whose graph is closed. If the minimal map $m(F(\cdot))$ has nonempty images, then its graph is closed.*

Consequently if Y is strictly convex, reflexive, F has nonempty closed convex images and the range of $m(F(\cdot))$ is contained in a compact subset of Y, then the minimal selection is continuous.

Corollary 9.3.3 *Let F be a continuous set-valued map from a metric space X to $\mathbf{R}^n$ with nonempty closed convex images. Then the minimal selection is continuous.*

Proof — Lemma 9.3.1 implies that $\|d(0, F(\cdot))\|$ is continuous. Fix $\overline{x} \in X$. Then for some $\delta > 0$ the set

$$\{\|m(F(y))\|\}_{y \in B(F(\overline{x}), \delta)}$$

is bounded. Proposition 9.3.2 implies that $m(F(\cdot))$ is continuous at $\overline{x}$. The element $\overline{x} \in X$ being arbitrary, the proof follows. □

Proof of Proposition 9.3.2 — Indeed by Lemma 9.3.1 the map

$$x \mapsto d(0, F(x)) = \|m(F(x))\|$$

is upper semicontinuous. This yields that the set-valued map

$$x \rightsquigarrow B(0, \|m(F(x))\|)$$

has closed graph. By our assumption the graph of F is closed. Thus the intersection map

$$x \rightsquigarrow F(x) \cap B(0, \|m(F(x))\|)$$

has closed graph.

On the other hand

$$m(F(x)) = F(x) \cap B(0, \|m(F(x))\|)$$

and the proof of the first statement ensues.

To prove the second statement it is enough to observe that a single-valued map taking its values in a compact set is continuous if and only if its graph is closed. □

When Y is a Hilbert space, then the compactness assumption in the above proposition may be replaced by the upper hemicontinuity of F:

Theorem 9.3.4 *Consider a metric space X, a Hilbert space Y and a lower semicontinuous set-valued map $F : X \rightsquigarrow Y$ with closed convex values. If F is upper hemicontinuous, then the minimal selection is continuous.*

Proof — We identify Y with its dual. The projection of 0 onto the closed convex set $F(x)$ is the element

$$u := m(F(x)) \in F(x)$$

such that

$$\|u\|^2 + \sigma(-F(x), u) = \sup_{y \in F(x)} < u - 0, u - y > \leq 0 \qquad (9.3)$$

(It is actually equal to 0.) Fix $\varepsilon > 0$ and $x \in X$. Since F is upper hemicontinuous there exists a neighborhood $\mathcal{V}$ of x such that for all $x' \in \mathcal{V}$

$$\begin{cases} \langle -m(F(x')),\ m(F(x))\rangle \leq \sigma(-F(x'), m(F(x))) \\ \leq \sigma(-F(x), m(F(x))) + \varepsilon \leq -\|m(F(x))\|^2 + \varepsilon \end{cases}$$

(thanks to (9.3).)

On the other hand, by Lemma 9.3.1, the function $x \mapsto \|m(F(x))\|$ is upper semicontinuous, since F is lower semicontinuous. Then there exists another neighborhood $\mathcal{W}$ of x such that

$$\forall\, x' \in \mathcal{W}, \quad \|m(F(x'))\|^2 \leq \|m(F(x))\|^2 + \varepsilon$$

These two inequalities imply that for all $x' \in \mathcal{V} \cap \mathcal{W}$

$$\begin{cases} \|m(F(x')) - m(F(x))\|^2 \\ = \|m(F(x'))\|^2 + \|m(F(x))\|^2 + 2 < -m(F(x')), m(F(x)) > \\ \leq 2\|m(F(x))\|^2 + \varepsilon + 2(-\|F(x)\|^2 + \varepsilon) \leq 3\varepsilon \end{cases}$$

Therefore, the minimal selection is continuous. □

9.4 The Steiner Selection

To construct Lipschitz selections of set-valued maps and to parametrize Carathéodory set-valued maps, we need a selection procedure associating with a closed convex set $K \subset \mathbf{R}^n$ a point $s(K) \in K$ in a regular enough way. We start this section with such a selection procedure for convex compact sets K.

To be able to deal also with unbounded sets, we prove an Intersection Lemma which enables us to associate with a closed convex set K a convex compact subset $P(K) \subset K$ in a Lipschitz way. The last part of this section is devoted to Lipschitz selections of set-valued maps with closed convex images.

9.4.1 Steiner Points of Convex Compact Sets

For a nonempty convex compact subset K of $\mathbf{R}^n$, we define its *Steiner point* or, for the sake of simplicity, *Krümmungsschwerpunkte*, (also called sometimes the *curvature centroid*) $s_n(K)$ by:

$$\begin{cases} \text{for } n = 1, & s_1(K) = \sigma(K,+1)/2 - \sigma(K,-1)/2 \\ \text{for } n \geq 2, & s_n(K) = n \int_{\Sigma^{n-1}} p\sigma(K,p)\omega(dp) \end{cases}$$

where Σ^{n-1} denotes the unit sphere in $\mathbf{R}^n$, $\sigma(K,\cdot)$ is the support function of K, ω is the measure on Σ^{n-1} proportional to the Lebesgue measure and satisfying $\omega(\Sigma^{n-1}) = 1$.

Since $\sigma(K,p) = \sigma(-K,-p)$, it follows that $s_n(K) = -s_n(-K)$. The support function being also additive with respect to K, the map $s_n(\cdot)$ is linear:

$$\begin{cases} \text{For all convex compact } K,\ L \subset \mathbf{R}^n \text{ and all } \lambda,\ \mu \in \mathbf{R}, \\ \quad s_n(\lambda K + \mu L) = \lambda s_n(K) + \mu s_n(L) \end{cases}$$

We also observe that if K is symmetric, i.e., $K = -K$, then $s_n(K) = 0$.

We show below that s_n is a selection in the sense that $s_n(K) \in K$ and that it is Lipschitz with respect to the Hausdorff distance:

We recall that the *extended Hausdorff distance* between two closed subsets $K,\ L \subset \mathbf{R}^n$ (which may be equal to $+\infty$ when K or L is unbounded or empty) is defined by[6]

$$\mathcal{H}(K,L) := \max\left\{ \sup_{x\in K} d(x,L),\ \sup_{x\in L} d(x,K) \right\}$$

The Steiner point can be alternatively defined by using the subdifferentials of the support function of a compact convex subset K.

Recall that the subdifferential $\partial\sigma(K,p)$ of the support function $\sigma(K,\cdot)$ is given by

$$\partial\sigma(K,p) = \{x \in K \mid <p,x> = \sigma(K,p)\}$$

We denote by $m(\partial\sigma(K,p))$ the element of $\partial\sigma(K,p)$ with the minimal norm.

The function $\sigma(K,\cdot)$ being continuous, Theorem 8.2.9 implies that the subdifferential $\partial\sigma(K,\cdot)$ is measurable.

Hence $m(\partial\sigma(K,p))$ being the projection of 0 onto $\partial\sigma(K,p)$, the single-valued map $m(\partial\sigma(K,\cdot))$ is also measurable (Corollary 8.2.13.) Therefore, the formula stated in the following theorem has a meaning:

Theorem 9.4.1 *Let $\mathcal{K}$ denote the family of all nonempty convex compact subsets of $\mathbf{R}^n$. Then*

$$\forall K \in \mathcal{K},\ s_n(K) = \frac{1}{\mathrm{Vol}(B^n)} \int_{B^n} m\left(\partial\sigma(K,p)\right) dp$$

where $\mathrm{Vol}(B^n)$ is the measure of the n-dimensional unit ball $B^n \subset \mathbf{R}^n$.

Consequently,

$$\forall K \in \mathcal{K},\ s_n(K) \in K$$

Furthermore $s_n(\cdot)$ is Lipschitz with the constant n:

$$\forall K,\ L \in \mathcal{K},\ \|s_n(K) - s_n(L)\| \leq n\mathcal{H}(K,L)$$

[6] It can be verified easily that $\mathcal{H}$ is a distance. Furthermore, a set-valued map with nonempty compact images is continuous if and only if it is continuous with respect to the Hausdorff distance.

Proof — Fix $K,\ L \subset \mathbf{R}^n$.

When $n = 1$, set

$$y = \sigma(K, +1), \quad z = -\sigma(K, -1)$$

Then $y, z \in K$ and

$$s_1(K) = \frac{1}{2}y + \frac{1}{2}z \in K$$

Furthermore

$$\frac{1}{\mathrm{Vol}(B_1)} \int_{B_1} m\left(\partial\sigma(K,p)\right) dp = \frac{1}{2}\left(\int_0^1 y dp + \int_{-1}^0 z dp\right) = \frac{1}{2}y + \frac{1}{2}z$$

and

$$\begin{cases} \|s_1(K) - s_1(L)\| \\ \leq \frac{1}{2}|\sigma(K,+1) - \sigma(L,+1)| + \frac{1}{2}|\sigma(K,-1) - \sigma(L,-1)| \leq \mathcal{H}(K,L) \end{cases}$$

— We assume next that $n \geq 2$. Consider the Moreau-Yosida approximation $\sigma_\lambda(K,\cdot)$ of the support function $\sigma(K,\cdot)$ of K. Recall that it is continuously differentiable. By integrating its gradient on the unit ball B^n whose boundary is the unit sphere Σ^{n-1}, Stoke's formula implies that

$$\int_{B^n} \nabla\sigma_\lambda(K,p)dp = \int_{\Sigma^{n-1}} \sigma_\lambda(K,p)p\mu(dp)$$

where μ is the Lebesgue measure on the unit sphere Σ^{n-1}.

We also recall that $\sigma_\lambda(K,\cdot)$ satisfies

$$-\|K\| \leq \inf_{p\in\Sigma^{n-1}} \sigma(K,p) \leq \sigma_\lambda(K,p) \leq \sigma(K,p) \leq \|K\|$$

(where $\|K\| := \sup_{x\in K}\|x\|$) and converges pointwise to $\sigma(K,\cdot)$ (see Theorem 6.5.7.) Therefore

$$\lim_{\lambda\to 0+} \int_{\Sigma^{n-1}} p\sigma_\lambda(K,p)\mu(dp) = \int_{\Sigma^{n-1}} p\sigma(K,p)\mu(dp)$$

On the other hand, by Theorems 3.5.9 and 6.5.7, $\nabla\sigma_\lambda(K,\cdot)$ is the Yosida approximation of the subdifferential $\partial\sigma(K,\cdot)$. Hence

$$\nabla\sigma_\lambda(K,p) \in \partial\sigma(K,p_\lambda) \subset K$$

where p_λ is the minimizer of the function

$$q \mapsto \sigma(K,q) + \frac{1}{2\lambda}\|q-p\|^2$$

and for any $p \in B^n$, $\nabla\sigma_\lambda(K,p)$ converges to $m(\partial\sigma(K,p))$. Therefore, the maps $\nabla\sigma_\lambda(K,\cdot)$ being measurable and bounded by $\|K\|$, we infer that

$$\lim_{\lambda\to 0+}\int_{B^n}\nabla\sigma_\lambda(K,p)dp = \int_{B^n} m\left(\partial\sigma(K,p)\right)dp$$

Dividing by $\mathrm{Vol}(B^n)$, and setting $\omega(dp) := \mu(dp)/\omega(\Sigma^{n-1})$, we obtain

$$\frac{1}{\mathrm{Vol}(B^n)}\int_{B^n} m\left(\partial\sigma(K,p)\right)dp = \frac{\omega(\Sigma^{n-1})}{\mathrm{Vol}(B^n)}\int_{\Sigma^{n-1}} p\sigma(K,p)\omega(dp)$$

Since $\omega(\Sigma^{n-1}) = n\mathrm{Vol}(B^n)$, we derived the formula we were looking for.

Proposition 8.6.2 and Theorem 8.6.3 yield that

$$\int_{B^n} Kdp = \mathrm{Vol}(B^n)K$$

Consequently

$$\frac{1}{\mathrm{Vol}(B^n)}\int_{B^n} m\left(\partial\sigma(K,p)\right)dp \in \frac{1}{\mathrm{Vol}(B^n)}\int_{B^n} Kdp = K$$

To prove that $s_n(\cdot)$ is n-Lipschitz observe that for every $p \in \Sigma^{n-1}$

$$\forall\, K,\ L \in \mathcal{K},\quad |\sigma(K,p)-\sigma(L,p)| \le \mathcal{H}(K,L)$$

Hence

$$\begin{cases} \|s_n(K)-s_n(L)\| \le n\int_{\Sigma^{n-1}}|\sigma(K,p)-\sigma(L,p)|\,\|p\|\,\omega(dp) \\ \le n\mathcal{H}(K,L)\omega(\Sigma^{n-1}) = n\mathcal{H}(K,L) \end{cases}$$

The proof is complete. □

9.4.2 The Intersection Lemma

To obtain Lipschitz selections for unbounded closed convex sets we need the following technical lemma.

Lemma 9.4.2 *Let $\mathcal{K}$ denote the family of all nonempty closed convex sets in $\mathbf{R}^n$. Then the map $P : \mathbf{R}^n \times \mathcal{K} \rightsquigarrow \mathcal{K}$ defined by*

$$P(y, K) \; := \; K \cap B\,(y,\; 2d(y, K))$$

is Lipschitz: for all $K, L \in \mathcal{K}$ and $x, y \in \mathbf{R}^n$

$$\mathcal{H}(P(x, K), P(y, L)) \;\leq\; 5\,(\mathcal{H}(K, L) \;+\; \|x - y\|)$$

Proof — First we show that our statement holds true with $x = y = 0$:

$$\forall\, K, L \;\in\; \mathcal{K}, \quad \mathcal{H}(P(0, K), P(0, L)) \;\leq\; 5\mathcal{H}(K, L) \qquad (9.4)$$

Pick K, L in $\mathcal{K}$ such that $\mathcal{H}(K, L) < +\infty$ and set $\varepsilon := \mathcal{H}(K, L)$. If $\varepsilon = 0$, then $K = L$ and (9.4) holds true.

Suppose next that $\varepsilon > 0$ and let $y \in P(0, L)$ be arbitrary but fixed.

To prove inequality (9.4), we need to find a point x in $P(0, K)$ such that

$$\|x - y\| \;\leq\; 5\varepsilon \qquad (9.5)$$

Observe first that

$$\|y\| \;\leq\; 2d(0, L) \;\leq\; 2d(0, K) \;+\; 2\varepsilon \qquad (9.6)$$

Let $x_1 \in K$ be such that

$$\|y - x_1\| \;\leq\; \varepsilon$$

If $\|x_1\| \;\leq\; 2d(0, K)$ then set $x := x_1$. Otherwise let $x_2 \in K$ be such that

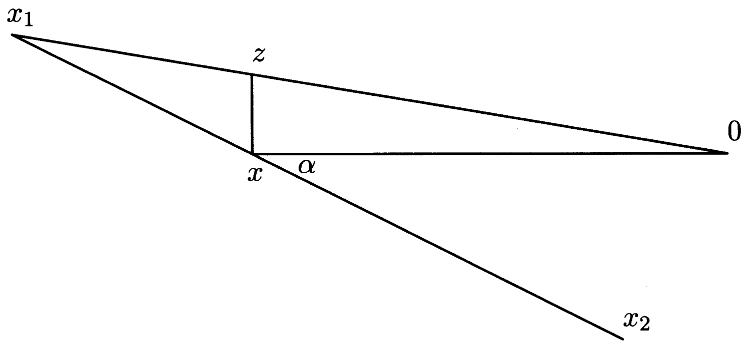
$$d(0, K) \;=\; \|x_2\|$$

If $x_2 = 0$ set $x = 0$. Then (9.6) implies (9.5.)

It remains to consider the case

$$\|x_1\| \;>\; 2d(0, K), \quad x_2 \;\neq\; 0$$

Figure 9.2: Illustration of the proof



We claim that there exists x in the interval $[x_1, x_2]$ such that

$$\|x\| = 2d(0, K)$$

Indeed define the continuous map $\varphi : [0,1] \mapsto \mathbf{R}^n$ by

$$\forall\, \lambda \in [0,1], \quad \varphi(\lambda) = \|\lambda x_1 + (1-\lambda)x_2\|$$

Since the image of φ is a connected set and

$$\varphi(0) = d(0,K) \;\; \& \;\; \varphi(1) > 2d(0,K)$$

there exists $\overline{\lambda} \in \,]0,1[$ which satisfies $\varphi(\overline{\lambda}) = 2d(0,K)$. So setting $x := \overline{\lambda} x_1 + (1-\overline{\lambda})x_2$, our claim follows. The set K being convex, $x \in P(0,K)$.

We prove next that this x is the point we are looking for.

Indeed, by the choice of x_1,

$$\|x - x_1\| \ge \|x - y\| - \|y - x_1\| \ge \|x - y\| - \varepsilon \tag{9.7}$$

In the triangle with vertices at $0, x, x_2$ denote by α the angle at x

(see Figure 9.2.) Inequality $\|x_2\| < \|x\|$ yields that $\alpha \in [0, \frac{\pi}{2}[$ and

$$\sin\alpha \leq \frac{\|x_2\|}{\|x\|} = \frac{1}{2}, \quad \cos\alpha \geq \frac{\sqrt{3}}{2} \tag{9.8}$$

If $\alpha = 0$ (that is always the case when $n = 1$), then

$$\|x_1\| = \|x\| + \|x - x_1\|$$

and (9.7) yield

$$\|y\| \geq \|x_1\| - \|y - x_1\| \geq \|x\| + \|x_1 - x\| - \varepsilon \geq 2d(0, K) + \|x - y\| - 2\varepsilon$$

Thus we deduce from (9.6) that

$$\|x - y\| \leq 4\varepsilon$$

and derive (9.5.)

Assume next that $\alpha > 0$ and consider the triangle with vertices at $0, x, x_1$. Then the angle at x is equal to $\pi - \alpha > \frac{\pi}{2}$. Let $z \in]0, x_1[$ be such that $z - x \perp x$. Then

$$\|z\| \geq \|x\| = 2d(0, K) \tag{9.9}$$

In the triangle formed by points z, x, x_1, the angle at x is equal to $\pi - \alpha - \frac{\pi}{2} = \frac{\pi}{2} - \alpha$. Therefore (9.8) implies

$$\|z - x_1\| \geq \|x - x_1\| \sin\left(\frac{\pi}{2} - \alpha\right) = \|x - x_1\| \cos\alpha \geq \|x - x_1\| \frac{\sqrt{3}}{2}$$

and from (9.9) follows that

$$\|x_1\| = \|z\| + \|z - x_1\| \geq 2d(0, K) + \frac{\sqrt{3}}{2} \|x - x_1\|$$

This and (9.7) imply

$$\|y\| \geq \|x_1\| - \|y - x_1\| \geq 2d(0, K) + \frac{\sqrt{3}}{2}(\|x - y\| - \varepsilon) - \varepsilon$$

Therefore, thanks to (9.6), we derive

$$\frac{\sqrt{3}}{2}(\|x - y\| - \varepsilon) - \varepsilon \leq 2\varepsilon$$

and thus (9.5.) So (9.4) is proved.

Pick next any $x, y \in X$ and observe that

$$\mathcal{H}(P(x,K),P(x,L)) \;=\; \mathcal{H}(P(0,K-x),P(0,L-x))$$

Hence, using (9.4), we deduce

$$\begin{cases} \mathcal{H}(P(x,K),P(y,L)) \\ \leq\; \mathcal{H}(P(x,K),P(x,L)) \;+\; \mathcal{H}(P(x,L),P(y,L)) \\ \leq\; 5\mathcal{H}(K-x,L-x) \;+\; \mathcal{H}(B(x,2d(x,L)),B(y,2d(y,L))) \\ \leq\; 5\mathcal{H}(K,L) \;+\; \|x-y\| \;+\; 2\,\|x-y\| \;=\; 5\mathcal{H}(K,L) \;+\; 3\,\|x-y\| \end{cases}$$

The proof is complete. □

9.4.3 Lipschitz Selections of Lipschitz Maps

An immediate application of the above construction is the existence of a Lipschitz selection for a Lipschitz set-valued map:

Theorem 9.4.3 *Consider a Lipschitz set-valued map F from a metric space to nonempty closed convex subsets of $\mathbf{R}^n$. Then F has a Lipschitz selection f.*

Remark — In the next section, we prove a stronger result, which in particular implies that for any $\overline{y} \in F(\overline{x})$ there exists a Lipschitz selection f of F such that $f(\overline{x}) = \overline{y}$. □

Proof — We first observe that a map $F : X \rightsquigarrow \mathbf{R}^n$ is Lipschitz if and only if there exists $c > 0$ such that

$$\forall\, x,\; y \;\in\; X,\;\; \mathcal{H}(F(x),F(y)) \;\leq\; c\,\|x-y\|$$

Define the set-valued map $G : X \rightsquigarrow \mathbf{R}^n$ with nonempty compact convex images by

$$\forall\, x \;\in\; X,\quad G(x) \;=\; F(x) \cap B(0, 2\,\|m(F(x))\|)$$

By Lemma 9.4.2 it is Lipschitz. Consider the single-valued map

$$f(x) = s_n(G(x))$$

where $s_n(\cdot)$ is the Steiner selection. The proof follows by the application of Theorem 9.4.1. □

Remark — Observe that if the set-valued map F is merely continuous in the Hausdorff metric:

$$\forall\, \varepsilon > 0,\ \exists\, \delta > 0,\ \forall\, x \in B(\overline{x}, \delta),\ \ \mathcal{H}(F(x), F(\overline{x})) \leq \varepsilon$$

then the function f defined in the above proof is a continuous selection of F. □

9.5 Selections of Carathéodory maps

From now on a measurable map always means a Lebesgue measurable map. Let X be a metric space and $a < b$ be real numbers. Consider a set-valued map

$$F : [a, b] \times X \rightsquigarrow \mathbf{R}^n$$

Definition 9.5.1 *For every $t \in [a, b]$ let $C(t) \subset X$ be a given set.*

1. — *The map F is called* Carathéodory on $\{C(t)\}_{t\in[a,b]}$ *if*

$$\begin{cases} \forall\, x \in X,\ F(\cdot, x) \text{ is measurable} \\ \\ \forall\, t \in [a, b],\ F(t, \cdot) \text{ is continuous on } C(t) \end{cases}$$

2. — *The map F is* measurable/Lipschitz *on* $\{C(t)\}_{t\in[a,b]}$ *if for every $t \in [a, b]$, there exists $k(t) \geq 0$ such that*

$$\begin{cases} \forall\, x \in X,\ F(\cdot, x) \text{ is measurable} \\ \\ \forall\, t \in [a, b],\ F(t, \cdot) \text{ is } k(t) - \text{Lipschitz on } C(t) \end{cases}$$

where $k(t)$–Lipschitz means Lipschitz with the constant $k(t)$.

Let a single valued map $z : [a,b] \mapsto X$ and a measurable single valued map $w : [a,b] \mapsto \mathbf{R}^n$ be such that

$$w(t) \in F(t, z(t)) \quad \text{almost everywhere in } [a,b]$$

We refine here the selection result proved in the preceding section:

Namely we show the existence of a measurable/Lipschitz selection $f(t,x) \in F(t,x)$ such that

$$\text{for almost every } t \in [a,b],\ w(t) = f(t, z(t))$$

whenever F is measurable/Lipschitz.

Results of this type are sometimes needed in control theory when we investigate the local behavior of solutions for nonsmooth systems.

Theorem 9.5.2 (Carathéodory Selection) *Let X be a metric space, $F : [a,b] \times X \rightsquigarrow \mathbf{R}^n$ be a set-valued map with nonempty closed convex images, $z : [a,b] \mapsto X$ be a single valued map and $w : [a,b] \mapsto \mathbf{R}^n$ be measurable.*

If F is Carathéodory on $\{C(t)\}_{t\in[a,b]}$ and

$$w(t) \in F(t, z(t)) \quad \text{almost everywhere in } [a,b]$$

then there exists a Carathéodory selection f of F on $\{C(t)\}_{t\in[a,b]}$ such that, for almost every $t \in [a,b]$, $w(t) = f(t,z(t))$.

Proof — Define the map f by

$$\forall\, (t,x) \in [a,b] \times X,\ \ f(t,x) = \Pi_{F(t,x)}(w(t))$$

where $\Pi_{F(\cdot,\cdot)}$ is the projection map from Corollary 8.2.13, i.e., for every (t,x), $\Pi_{F(t,x)}(w(t))$ is an element of $F(t,x)$ such that

$$d(w(t), F(t,x)) = \left\|\Pi_{F(t,x)}(w(t)) - w(t)\right\|$$

Corollary 8.2.13 implies that for every $x \in X$, the map $f(\cdot, x)$ is measurable. It is also clear, that $w(t) = f(t, z(t))$ almost everywhere in $[a,b]$. The last statement follows from Corollary 9.3.3. □

If $F(t,\cdot)$ is Lipschitz, then the selection f may be chosen to be Lipschitz with respect to x:

Theorem 9.5.3 (Measurable/Lipschitz Selection) *Under all as sumptions of Theorem 9.5.2 assume that the set-valued map F is measurable/Lipschitz on $\{C(t)\}_{t\in[a,b]}$ and let $k(t)$, $t \in [a,b]$ denote the corresponding Lipschitz constants.*

Then there exist $c > 0$ and a measurable/Lipschitz selection f of F on $\{C(t)\}_{t\in[a,b]}$ such that $f(t,\cdot)$ is $ck(t)$–Lipschitz on $C(t)$ and

$$\text{for almost every } t \in [a,b], \quad w(t) = f(t, z(t))$$

Furthermore c is independent of F.

Proof — Let P be the set-valued map defined in Lemma 9.4.2 and s_n be the Steiner selection.

For every $t \in [a,b]$ and $x \in \mathbf{R}^n$ set

$$f(t,x) = s_n\left(P(w(t), F(t,x))\right)$$

Since $w(t) \in F(t, z(t))$, by the very definition of the map P, we get

$$P(w(t), F(t, z(t))) = w(t) \text{ almost everywhere in } [a,b]$$

Thus $f(t, z(t)) = w(t)$ almost everywhere.

Theorem 9.4.1 implies that $f(t,x) \in F(t,x)$. We deduce from Lemma 9.4.2 and Theorem 9.4.1 that for every $t \in [a,b]$, the map $f(t,\cdot)$ is Lipschitz on $C(t)$ with Lipschitz constant $5nk(t)$.

It remains to show that f is measurable with respect to t. Fix $x \in X$. We first show that the map

$$t \rightsquigarrow \Phi(t,x) := P(w(t),\ F(t,x)) = F(t,x) \cap B(w(t), 2d(w(t), F(t,x)))$$

is measurable.

Indeed, by Theorem 8.2.4, the intersection of two measurable maps is measurable. On the other hand $F(\cdot, x)$ is measurable by assumption and

$$t \rightsquigarrow B(w(t), 2d(w(t), F(t,x)))$$

is measurable by Corollary 8.2.13.

Hence Theorem 8.2.14 implies that for every $p \in S^{n-1}$, the support function $\sigma(\Phi(\cdot,x), p)$ is measurable. Thus, by the definition of

Steiner point, $f(\cdot, x)$ is measurable when $n = 1$. When $n \geq 2$, it is enough to observe that the map $p \mapsto p\sigma(\Phi(t,x),p)$ is continuous and therefore the map

$$t \mapsto \int_{S^{n-1}} p\sigma(\Phi(t,x),p)\omega(dp)$$

is measurable. □

9.6 Carathéodory Parametrization

We prove in this section that a Carathéodory set-valued map F with closed convex images admits a Carathéodory parametrization.

Consider a metric space X, reals $a < b$ and a set-valued map

$$F : [a,b] \times X \rightsquigarrow \mathbf{R}^n$$

Definition 9.6.1 *Let U be a metric space and $C(t) \subset X$, $t \in [a,b]$ be given subsets of X. We say that a single-valued map*

$$f : [a,b] \times X \times U \mapsto \mathbf{R}^n$$

is a Carathéodory parametrization *of F on $\{C(t)\}_{t\in[a,b]}$ if*

$$\begin{cases} i) & \forall\, (t,x) \in [a,b] \times X, \;\; F(t,x) = f(t,x,U) \\ ii) & \forall\, (x,u) \in X \times U, \;\; f(\cdot,x,u) \text{ is measurable} \\ iii) & \forall\, (t,u) \in [a,b] \times U, \;\; f(t,\cdot,u) \text{ is continuous on } C(t) \\ iv) & \forall\, (t,x) \in [a,b] \times X, \;\; f(t,x,\cdot) \text{ is continuous} \end{cases}$$

We recall that $\|F(t,x)\|$ is defined by

$$\|F(t,x)\| := \max_{y\in F(t,x)} \|y\|$$

Let B denote the closed unit ball in $\mathbf{R}^n$.

Theorem 9.6.2 (Carathéodory Parametrization) *Consider a metric space X and a set-valued map $F : [a,b] \times X \rightsquigarrow \mathbf{R}^n$ with nonempty convex compact images.*

If F is Carathéodory on $\{C(t)\}_{t\in[a,b]}$, then there exists a Carathéodory parametrization f of F on $\{C(t)\}_{t\in[a,b]}$ such that $U = B$ and for all $t \in [a,b]$, $x \in X$, $u, v \in B$ we have

$$\|f(t,x,u) - f(t,x,v)\| \;\leq\; c\,\|F(t,x)\|\,\|u-v\|$$

where c is independent of F.

Furthermore if F is continuous, so is f.

Finally, if images of F are merely closed (instead of compact), then the same statement holds true with $U = \mathbf{R}^n$ and $\|F(t,x)\|$ replaced by 1.

Proof — To simplify the notations, we set $M_x(t) := \|F(t,x)\|$. Let $(t,x,u) \in [a,b] \times X \times \mathbf{R}^n$. Consider the closed ball $G(t,x,u)$ of radius $2d(M_x(t)u, F(t,x))$ and center $M_x(t)u$, i.e.,

$$G(t,x,u) \;=\; B(M_x(t)u, 2d(M_x(t)u, F(t,x)))$$

We claim that G is measurable with respect to t.

Indeed, fix $x \in X, u \in \mathbf{R}^n$. From Theorem 8.2.11 it follows that $M_x(\cdot) = \|F(\cdot,x)\|$ is measurable and from Corollary 8.2.13 that the function $d(M_x(\cdot)u, F(\cdot,x))$ is measurable. Thus, using again Corollary 8.2.13, we deduce that $G(\cdot,x,u)$ is measurable.

Let P be the map defined in Lemma 9.4.2 and set

$$\Phi(t,x,u) \;:=\; P(M_x(t)u, F(t,x)) \;=\; G(t,x,u) \cap F(t,x) \;\subset\; F(t,x)$$

Since $G(\cdot,x,u)$ is measurable, Theorem 8.2.4 and our assumptions on F imply that $\Phi(\cdot,x,u)$ is measurable.

Define the single-valued map f from $[a,b] \times X \times \mathbf{R}^n$ into $\mathbf{R}^n$ by

$$\forall\,(t,x,u) \;\in\; [a,b] \times X \times \mathbf{R}^n, \;\; f(t,x,u) \;:=\; s_n(\Phi(t,x,u))$$

where s_n is the Steiner selection.

Since Φ is measurable with respect to t, using the definition of s_n, we deduce that f is also measurable with respect to t.

Lipschitz properties of s_n and P proved in Theorem 9.4.1 and Lemma 9.4.2 imply:

$$\left\{\begin{array}{l} \|f(t,x,u) - f(t,y,v)| \;\leq\; n\mathcal{H}(\Phi(t,x,u), \Phi(t,y,v)) \\ \\ \leq\; 5n\,(\mathcal{H}(F(t,x), F(t,y)) \;+\; \|M_x(t)u - M_y(t)v\|) \end{array}\right. \tag{9.10}$$

Since F has compact images, $F(t,\cdot)$ is continuous on $C(t)$ in the Hausdorff metric. This and (9.10) imply that for every $u \in \mathbf{R}^n$, $f(\cdot,\cdot,u)$ is Carathéodory on $\{C(t)\}_{t\in[a,b]}$.

In order to prove that f is a parametrization, it remains to check $i)$ of Definition 9.6.1 with $U = B$. For this aim fix $t \in [a,b]$, $x \in X$ and $y \in F(t,x)$. Setting

$$u := \begin{cases} y/M_x(t) & \text{if } M_x(t) \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

we derive

$$u \in B, \ \ M_x(t)u = y, \ \ \Phi(t,x,u) = y$$

Hence

$$f(t,x,u) = s_n(\Phi(t,x,u)) = y$$

This means that $F(t,x) \subset f(t,x,B)$. The opposite inclusion resulting from Theorem 9.4.1, we have proved $i)$.

Assume next that F is continuous. Then, by Lemma 9.3.1, so is $\|F(\cdot,\cdot)\|$. Thus the set-valued map G with convex compact images is continuous. Consequently its graph, Graph(G), is closed and therefore the intersection map Φ has closed graph. Since for all (t,x), $F(t,x) \subset M_x(t)B$, we deduce that Φ is upper semicontinuous.

We claim that Φ is actually continuous. Indeed in order to prove this claim, it is enough to show that it is lower semicontinuous. Fix (t,x,u) and an open set $\mathcal{O} \subset \mathbf{R}^n$ such that

$$\Phi(t,x,u) \cap \mathcal{O} \neq \emptyset$$

Let $z_1 \in \Phi(t,x,u) \cap \mathcal{O}$. From the very definition of G we deduce that there exists

$$z_2 \in F(t,x) \cap \operatorname{Int}(G(t,x,u))$$

Thus the interval

$$]z_1, z_2] \subset \operatorname{Int}(G(t,x,u)) \cap F(t,x)$$

Consequently we can find an element $z \in \mathbf{R}^n$ satisfying

$$z \in \operatorname{Int}(G(t,x,u)) \cap F(t,x) \cap \mathcal{O}$$

Hence for some $\varepsilon > 0$,

$$B(z, \varepsilon) \subset G(t, x, u) \cap \mathcal{O}$$

This and continuity of G imply that for every (t', x', u') sufficiently close to (t, x, u),

$$B(z, \varepsilon/2) \subset G(t', x', u')$$

On the other hand, F being continuous, for all (t', x') near (t, x),

$$B(z, \varepsilon/2) \cap F(t', x') \neq \emptyset$$

Thus Φ is lower semicontinuous, and thus, is continuous.

Since Φ has compact images, it is also continuous in the Hausdorff metric. From the very definition of s_n, we deduce that f is continuous.

To prove the last statement, it is enough to apply the same proof replacing everywhere $M_x(t)$ by 1. □

9.7 Measurable/Lipschitz Parametrization

In this section we show that the Carathéodory parametrization defined in the preceding section is measurable/Lipschitz on $\{C(t)\}_{t\in[a,b]}$ whenever F is measurable/Lipschitz on $\{C(t)\}_{t\in[a,b]}$.

Theorem 9.7.1 (Parametrization of Unbounded Maps) *Consider a metric space X and a set-valued map $F : [a, b] \times X \rightsquigarrow \mathbf{R}^n$ with nonempty closed convex images.*

Assume that F is measurable/Lipschitz on $\{C(t)\}_{t\in[a,b]}$ and let $k(t)$, $t \in [a, b]$ denote the corresponding Lipschitz constants.

Then there exists a Carathéodory parametrization f of F with $U = \mathbf{R}^n$ on $\{C(t)\}_{t\in[a,b]}$ such that:

$$\begin{cases} \forall\, (t, u) \in [a, b] \times \mathbf{R}^n,\ f(t, \cdot, u) \text{ is } ck(t) - \text{Lipschitz on } C(t) \\ \\ \forall\, (t, x) \in [a, b] \times X,\ f(t, x, \cdot) \text{ is } c - \text{Lipschitz on } \mathbf{R}^n \end{cases}$$

where c is independent of F.

Furthermore if F is continuous, so is f.

Proof — It is enough to apply the same proof as the one of Theorem 9.6.2 replacing everywhere $M_x(t)$ by 1 and to observe that (9.10) implies that for all $(t,u) \in [a,b] \times \mathbf{R}^n$, $f(t,\cdot,u)$ is $5nk(t)$-Lipschitz on $C(t)$. □

Theorem 9.7.2 (Parametrization of Bounded Maps) *We posit the assumptions of Theorem 9.7.1 and we assume that the values of F are compact.*

Then there exists a Carathéodory parametrization f of F on the family of sets $\{C(t)\}_{t\in[a,b]}$ with $U = B$ such that:

$$\begin{cases} i) & \forall\,(t,u) \in [a,b] \times B,\ f(t,\cdot,u) \text{ is } ck(t) - \text{Lipschitz on } C(t) \\ ii) & \forall\, t \in [a,b],\ \forall\, x \in X,\ \forall\, u,v \in B \\ & \|f(t,x,u) - f(t,x,v)\| \le c\,\|F(t,x)\|\,\|u-v\| \end{cases}$$

where c is independent of F.

Furthermore if F is continuous, so is f.

Proof — We consider the same parametrization as in the proof of Theorem 9.6.2. We already know that f is a Carathéodory parametrization of F on $\{C(t)\}_{t\in[a,b]}$.

Then (9.10) and Lemma 9.3.1 imply that for all $(t,u) \in [a,b] \times B$,

$$\begin{cases} \|f(t,x,u) - f(t,y,u)\| \\ \le 5n\,(\mathcal{H}(F(t,x),F(t,y)) + \|M_x(t)u - M_y(t)u\|) \\ \le 5n\,(k(t)\,\|x-y\| + k(t)\,\|x-y\|) \end{cases}$$

Consequently, $f(t,\cdot,u)$ is $10nk(t)$-Lipschitz on $C(t)$. □

When we allow the set of parameters to range in a Banach space, then actually the parametrization may be taken linear with respect to the parameter.

Theorem 9.7.3 (Linear Parametrization) *If all the assumptions of Theorem 9.7.1 are satisfied and X is compact, then there exists a map*

$$\varphi : X \times \mathcal{C}(X;\mathbf{R}^n) \mapsto \mathbf{R}^n$$

and a measurable set-valued map $U : [a,b] \rightsquigarrow \mathcal{C}(X;\mathbf{R}^n)$ *with closed convex images such that :*

$$\begin{cases} i) & \forall\,(t,x) \in [a,b]\times X,\ \ F(t,x) \ = \ \varphi(x,U(t)) \\ ii) & \varphi(x,\cdot) \ \text{ is linear nonexpansive} \\ iii) & \forall\, t \in [a,b],\ \ \forall\, u \in U(t),\ \ \forall\, x,y \in X, \\ & \|\varphi(x,u)-\varphi(y,u)\| \ \le \ ck(t)d(x,y) \end{cases}$$

where c *is independent of* F.

If moreover F *has compact images, then the values of* U *are compact.*

Proof — Let f be a parametrization map as in the conclusion of Theorem 9.7.1. Then

$$\forall\,(t,u) \in [a,b]\times\mathbf{R}^n,\ \ f(t,\cdot,u) \ \text{ is } \ ck(t)-\text{Lipschitz}$$

Hence for all $t \in [a,b]$, $\{f(t,\cdot,u)\}_{u\in U}$ is a family of equicontinuous maps.

Define the map U with closed convex values by :

$$\forall\, t \in [a,b],\ \ \ U(t) \ := \ \overline{co}\left(\{f(t,\cdot,u)\}_{u\in\mathbf{R}^n}\right) \ \subset \ \mathcal{C}(X;\mathbf{R}^n)$$

If F has bounded images, then Ascoli's Theorem yields that for all t, $U(t)$ is compact.

To show that it is measurable, consider a dense sequence $(u_k)_{k\ge1}$ in $\mathbf{R}^n$. Then for every k the map

$$t \ \mapsto \ f(t,\cdot,u_k) \ \in \ \mathcal{C}(X;\mathbf{R}^n)$$

is measurable. Set

$$V(t) \ = \ \overline{\bigcup_{k\ge1} f(t,\cdot,u_k)}$$

Theorem 8.2.4 implies that V is measurable and Theorem 8.2.2 yields that U is measurable.

Consider the linear map

$$\varphi : X\times\mathcal{C}(X;\mathbf{R}^n) \ \mapsto \ \mathbf{R}^n,\ \ \ \varphi(x,u) \ := \ u(x)$$

Clearly it is nonexpansive.

Then equality *i*) holds true by the very definition of φ and $U(t)$. Notice that *iii*) is verified for every $u \in V(t)$. Hence, by definition of φ, it is verified for every $u \in coV(t)$. The set $U(t)$ being equal to the closure of $coV(t)$ the proof ensues. $\square$

Chapter 10

Differential Inclusions

Introduction

This chapter is just an introductory survey of differential inclusions in finite dimensional vector-spaces. Many results below are stated without proofs. Their proofs and further developments can be found in [33], [50], [195].

Differential inclusions

$$x'(t) \in F(t, x(t)) \tag{10.1}$$

where F is a set-valued map from $\mathbf{R} \times X$ to a finite dimensional vector-space X have been studied since the thirties. Their consideration was initiated by Zaremba[1] in 1934 and Marchaud in 1938. These investigators were mostly interested in the existence results[2] and also examined some qualitative properties of solutions sets.

We have first to agree on what we shall call a solution to such a differential inclusion.

[1] who kept close relations with Painlevé and other French mathematicians after his study in Paris.

[2] While Zaremba investigated the paratingent solutions, Marchaud was mainly concerned with the contingent ones, i.e., the derivatives were taken in the contingent and paratingent sense. Since 1961, when Ważewski showed that one can use more "classical" solutions than the contingent and paratingent ones, solutions have been understood in the Carathéodory sense, i.e., absolutely continuous functions verifying (10.1) almost everywhere.

In the case of differential equations, there is no ambiguity since the derivative $x'(\cdot)$ of a solution $x(\cdot)$ to the differential equation

$$x'(t) = f(t, x(t))$$

inherits the regularity properties of the map f and of the function $x(\cdot)$. This is no longer the case with differential inclusions and is one of the reasons why their study brings more difficulties than that of the ordinary differential equations.

Solutions to differential inclusion (10.1) are understood in the Carathéodory sense, i.e., absolutely continuous functions verifying (10.1) almost everywhere.

Interest in differential inclusions (10.1) was revived in the early sixties, when mathematicians became attracted by a new domain: control theory[3]. Filippov in 1959 and Ważewski in 1961 proved that under very mild assumptions the control system

$$x' = f(t, x, u(t)), \quad u(t) \in U \text{ is measurable} \tag{10.2}$$

may be reduced to the differential inclusion (10.1.)

This considerably simplified the study of the closure of solutions to (10.2) and led to the celebrated Filippov-Ważewski relaxation theorem, which we state in Section 4.

But differential inclusions also encompass much more sophisticated control systems:

1. — Closed loop control systems

$$x'(t) = f(t, x(t), u(t)), \quad u(t) \in U(t, x(t)) \tag{10.3}$$

2. — Implicit control systems

$$f(t, x(t), x'(t), u(t)) = 0, \quad u(t) \in U(t, x(t)) \tag{10.4}$$

3. — Systems with uncertainties

$$x'(t) \in f(t, x(t), u(t)) + \varepsilon(t, x)B, \quad u(t) \in U(t, x(t)) \tag{10.5}$$

[3]Set-valued maps have been carefully — maybe unconsciously — hidden in control theory, despite the early contributions of Ważewski and Filippov. They were thought of as being useless. This is often the justification of something new one is reluctant to learn... "Everybody wants to teach, nobody wants to learn..." bitterly complained Abel.

where $\varepsilon(t,x)$ is a function incorporating the disturbances of the model.

Setting

$$F(t,x) := f(t,x,U(t,x))$$

in the first case,

$$F(t,x) := \{v \mid 0 \in f(t,x,v,U(t,x))\}$$

in the second one and

$$F(t,x) := f(t,x(t),U(t,x)) + \varepsilon(t,x)B$$

in the third one, we replace the above control systems by differential inclusions of the type (10.1.)

We shall survey three main classes of differential inclusions.

First, the case when the right-hand side is upper semicontinuous with closed convex values: in this instance, we can extend Peano's Theorem on the existence of solutions to differential equations with continuous right-hand side and show that the set of solutions depends upper semicontinuously upon the initial state. This is the topic of Section 1.

In Section 4, we trade convexity of the images with more regularity on F, which is assumed to be Lipschitz with respect to the state and measurable with respect to the time. Filippov's Theorem provides not only the extension of the Cauchy-Lipschitz Theorem, but also very useful estimates *à la Gronwall*, implying in particular the Lipschitz dependence on the initial state. Under this condition, we can regard F as a kind of infinitesimal generator of a semigroup of set-valued maps, which are the *reachable maps*.

Furthermore, it allows us to estimate the contingent, adjacent and circatangent derivatives of the solution map $\mathcal{S} := \mathcal{S}_{[0,T]}$ associated with the differential inclusion

$$\text{for almost all } t \in [0,T], \quad x'(t) \in F(t,x(t))$$

where $F : [0,T] \times X \rightsquigarrow X$.

We shall express these estimates in terms of the solution maps of adequate linearizations of differential inclusion of the form

$$w'(t) \in F'(t,x(t),x'(t))(w(t))$$

where for almost all t, $F'(t,x,y)(u)$ denotes one of the (contingent, adjacent or circatangent) derivatives of the set-valued map $F(t,\cdot)$ at a point (x,y) of its graph (the set-valued map F is regarded as a family of set-valued maps $x \rightsquigarrow F(t,x)$ and *the derivatives are taken with respect to the state variable only.)*

These linearized differential inclusions can be called the *variational equations*, since they extend (in various ways) the classical variational equations of ordinary differential equations.

The Relaxation Theorem, stated in Section 4, connects these two classes of differential inclusions: when F is Lipschitz and compact-valued, the set of solutions to differential inclusion (10.1) is dense in the set of solutions to the *relaxed or convexified* differential inclusion

$$x'(t) \in coF(t,x(t))$$

We also mention the third case when

$$F(t,x) := -A(x)$$

where A is a maximal monotone map. (This covers *gradient inclusions* of the form

$$\text{for almost all } t \in [0,T], \quad x'(t) \in -\partial V(x(t))$$

where the "potential" V is a lower semicontinuous convex function, describing the continuous version of the steepest descent method.)

Monotonicity implies the uniqueness of the solutions. Existence is guaranteed if A is maximal monotone, so that A can be regarded (and is regarded) as the infinitesimal generator of a semigroup of (nonlinear) operators. This is elucidated in Section 3.

We shall also devote Section 1 to the Viability Theorem and Section 2 to the description of some of its applications.

If K is a closed subset of the domain of F (assumed to be independent of the time for simplicity), we say that a solution $x(\cdot)$ to a differential inclusion is *viable* if $x(t)$ remains in K for all t.

The purpose of the Viability Theorem is to characterize the viability property: for any initial state $x_0 \in K$, there exists at least one viable solution.

When F is upper semicontinuous with closed convex images, K enjoys the viability property if and only if K is a *viability domain* of F, i.e., if and only if

$$\forall\, x \in K, \;\; F(x) \cap T_K(x) \neq \emptyset$$

When $F \equiv f$ is single-valued, this theorem has been proved by Nagumo in 1942 and rediscovered 14 times since[4]. The above set-valued version has been proved by G. Haddad in 1981.

One can also show that if K is not a viability domain of F, it still contains the largest closed viability domain, called the *viability kernel*. This result provides a quite useful tool.

The second section is devoted to applications of the Viability Theorem. For instance, differential inequalities of the type

$$\forall\, t \in [0,T], \;\; V(x(t)) \leq w(t)$$

can be regarded as a viability constraint of the form

$$\forall\, t \in [0,T], \;\; (x(t), w(t)) \in \mathcal{E}\mathrm{p}(V)$$

so that the Lyapunov method can be adapted to differential inclusions and to functions V which are merely lower semicontinuous. This allows us to deduce the existence of the *smallest lower semicontinuous Lyapunov function larger than or equal to a given function.*

10.1 The Viability Theorem

In all this section X, Y denote finite dimensional vector-spaces, except an explicit mention to the contrary.

We describe the (non deterministic) dynamics of the system by a set-valued map F from the state space X to itself.

Consider initial value problems (or Cauchy problems) associated with differential inclusion

$$\text{for almost all } t \in [0,T], \;\; x'(t) \in F(x(t)) \tag{10.6}$$

satisfying the initial condition $x(0) = x_0$.

[4]This does not imply necessarily that it is true.....

10.1.1 Solutions to Differential Inclusions

We have first to agree on what we shall call a solution to such a differential inclusion.

Denote by $L^1(0,\infty;X,e^{-bt}dt)$ the space of integrable functions for the measure $e^{-bt}dt$ (*weighted Lebesgue spaces*) and by

$$\begin{cases} i) & W := W_0 := W^{1,1}(0,T;X) & \text{if } T < +\infty \\ ii) & W_{-b} := W^{1,1}(0,\infty;X,e^{-bt}dt) & \text{if } T = +\infty \end{cases}$$

(for some $b \geq 0$) the *weighted Sobolev spaces* defined by

$$\begin{cases} W := \{x(\cdot) \in L^1(0,T;X) \mid x'(\cdot) \in L^1(0,T;X)\} \\ W_{-b} := \{x(\cdot) \in L^1(0,\infty;X,e^{-bt}dt) \mid x'(\cdot) \in L^1(0,\infty;X,e^{-bt}dt)\} \end{cases}$$

We shall supply them with the topology for which a sequence $x_n(\cdot)$ converges to $x(\cdot)$ if and only if

$$\begin{cases} i) & x_n(\cdot) \text{ converges } \textit{uniformly} \text{ to } x(\cdot) \\ & (\text{on compact intervals if } T = +\infty) \\ ii) & x_n'(\cdot) \text{ converges } \textit{weakly} \text{ to } x'(\cdot) \text{ in } L^1(0,T;X) \\ & \text{or in } L^1(0,\infty;X,e^{-bt}dt) \text{ if } T = +\infty \end{cases}$$

Definition 10.1.1 (Viability and Invariance Properties) *Let K be a subset of* Dom(F)*.*

— *A function $x(\cdot)$ from $[0,T]$ to X is called* viable *in K if*

$$\forall\, t \in [0,T],\ \ x(t) \in K$$

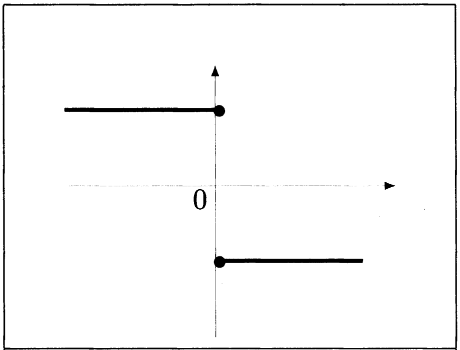
— *We shall say that K enjoys the local* viability property or controlled invariance *(for the set-valued map F) if for any initial state x_0 in K,* there exist *$T > 0$ and a solution on $[0,T]$ to differential inclusion (10.6) starting at x_0 and* viable *in K.*

It enjoys the global viability property (or, simply, the viability property*) if we can take $T = \infty$.*

— *We shall say that a subset $K \subset$* Dom(F) *is a* viability domain *of F if and only if*

$$\forall\, x \in K,\quad F(x) \cap T_K(x) \neq \emptyset \tag{10.7}$$

Figure 10.1: Example of a Map with Nonconvex Values



There is no solution starting at 0

Since the contingent cone to a singleton is obviously reduced to 0, we observe that a singleton $\{\overline{x}\}$ is a viability domain if and only if $\overline{x}$ is an equilibrium of F, i.e., a stationary solution to the inclusion $0 \in F(\overline{x})$.

In other words, the equilibria of a set-valued map provide the first examples of viability domains, actually, the *minimal viability domains*.

Open subsets of the domain of F are viability domains of F, since contingent cones to open subsets are equal to the whole space.

10.1.2 Statements of the Viability Theorems

Viability theorems hold true for the class of nontrivial upper semicontinuous set-valued maps with nonempty compact convex images. We observe that the only truly restrictive condition is the *convexity* of the images of these set-valued maps, since the continuity requirements are minimal.

But we cannot dispense with it, as the following counter example shows:

Let us consider $X := \mathbf{R}$, $K := [-1, +1]$ and the set-valued map

$F : K \rightsquigarrow \mathbf{R}$ defined by

$$F(x) := \begin{cases} \{-1\} & \text{if} \quad x > 0 \\ \{-1, 1\} & \text{if} \quad x = 0 \\ \{+1\} & \text{if} \quad x < 0 \end{cases}$$

Obviously, no solution to the differential inclusion

$$x'(t) \in F(x(t))$$

can start from 0, since 0 is not an equilibrium of this set-valued map!

Theorem 10.1.2 (Haddad's Theorem) *Let us assume that*

$$\begin{cases} i) & F : X \rightsquigarrow X \text{ is upper semicontinuous} \\ ii) & \text{the images of } F \text{ are convex and compact} \\ iii) & K \text{ is locally compact} \end{cases}$$

Then K enjoys the local viability property if and only if K is a viability domain of F.

In other words, K is a viability domain if and only if for any initial state x_0 in K, there exist $T > 0$ and a solution on $[0, T]$ to differential inclusion (10.6) starting at x_0 and viable in K.

Since open subsets of finite dimensional vector spaces are locally compact viability domains, we obtain the extension of Peano's Theorem to differential inclusions:

Theorem 10.1.3 *Let Ω be an open subset of a finite dimensional vector space X and $F : \Omega \rightsquigarrow X$ be an upper semicontinuous set-valued map with convex compact images.*

Then, for all $x_0 \in \Omega$, there exists $T > 0$ such that differential inclusion (10.6) has a solution on the interval $[0, T]$ starting at x_0.

The interesting case from the global point of view is the one when the viability subset is *closed.* In this case, we derive from Theorem 10.1.2 a more precise statement.

Theorem 10.1.4 (Local Viability Theorem) *Let us consider an* upper semicontinuous set-valued map $F : X \rightsquigarrow X$ with compact convex images *and a* closed *subset* $K \subset \mathrm{Dom}(F)$.

Then K is a viability domain *if and only if it enjoys the local viability property.*

Actually, for any $x_0 \in K$ there exist a positive T and a solution on $[0,T]$ to differential inclusion (10.6) starting at x_0 and viable in *K such that either $T = \infty$ or $T < \infty$ and* $\limsup_{t \to T-} \|x(t)\| = \infty$.

Further adequate information —a priori estimates on the growth of F — allow exclusion of the case when $\limsup_{t \to T-} \|x(t)\| = \infty$. This is the case for instance when F is bounded on K, and, in particular, when K is bounded. More generally, we can take $T = \infty$ when F enjoys linear growth:

Definition 10.1.5 *We say that a set-valued map $F : X \rightsquigarrow X$ has a linear growth if there exists $c > 0$ such that*

$$\forall\, x \in K, \quad \|F(x)\| := \sup_{v \in F(x)} \|v\| \leq c(\|x\| + 1)$$

We shall call Peano maps *nontrivial upper semicontinuous set-valued maps with nonempty compact convex images and with linear growth, or equivalently, nontrivial closed set-valued maps with convex values and linear growth.*

Theorem 10.1.6 (Viability Theorem) *Let us consider a Peano map F from X to X and a* closed *subset $K \subset$* Dom(F).

If K is a viability domain, *then for $x_0 \in K$, there exists a solution on $[0, \infty[$ to differential inclusion (10.6) starting at x_0, which is* viable *in K and belongs to the space $W^{1,1}(0, \infty; X, e^{-bt}dt)$ for some $b > 0$.*

Let us consider now a sequence of closed viability domains of a set-valued map F. We would like to address the following stability question:

Is the upper limit of these closed viability domains still a closed viability domain?

Theorem 10.1.7 (Stability of Viability Domains) *Let us consider a Peano map $F : X \rightsquigarrow X$. Then the upper limit of a sequence of closed viability domains of F is also a closed viability domain of F.*

10.1.3 Viability Kernels

Definition 10.1.8 *Denote by $\mathcal{S}(x_0)$ the (possibly empty) set of solutions to differential inclusion (10.6) on $[0,\infty[$ starting at $x_0 \in X$. We call the set-valued map*

$$\mathrm{Dom}(F) \ni x \mapsto \mathcal{S}(x)$$

the solution map *of F (or of differential inclusion (10.6).)*

Theorem 10.1.9 (Continuity of the Solution Map) *Let us assume that $F : X \rightsquigarrow X$ is a Peano map. The solution map $\mathcal{S}$ is upper semicontinuous with compact images from its domain to the space $\mathcal{C}(0,\infty;X)$.*

Definition 10.1.10 (Viability Kernel) *Let K be a subset of the domain of a set-valued map $F : X \rightsquigarrow X$. We shall say that the largest closed viability domain contained in K (which may be empty) is the* viability kernel *of K.*

Theorem 10.1.11 *Let us consider a Peano map $F : X \rightsquigarrow X$ and a closed subset $K \subset \mathrm{Dom}(F)$. Then the viability kernel of K exists (possibly empty) and is the subset of initial states such that at least one solution starting from them is viable in K.*

The viability kernels may inherit properties of both F and K. For instance, if the graph of F and the subset K are convex, so is the viability kernel of K. If F is a closed convex process (i.e., its graph is a closed convex cone) and if K is a closed convex cone, the viability kernel is a closed convex cone.

In general, viability kernels are not necessarily connected.

The viability kernel contains for instance all the *limit sets* $L(x)$ of the solutions $x(\cdot)$ to differential inclusion (10.6) defined by

$$L(x(\cdot)) := \bigcap_{T>0} cl(x([T,\infty[))$$

Indeed, they are closed viability domains:

Theorem 10.1.12 (Limit Sets are Viability Domains) *Let us consider a Peano map $F : X \rightsquigarrow X$. Then the limit sets of the solutions to differential inclusion (10.6) are* closed viability domains.

In particular, the limit of a solution $x(\cdot)$ to differential inclusion (10.6) when $t \to +\infty$, when it exists, is an equilibrium *of F.*

The trajectories of periodic solutions to the differential inclusion (10.6) are also closed viability domains.

10.1.4 Viability and Equilibria

Equilibrium Theorem 3.2.1 states that any compact convex viability domain of a Peano map contains an equilibrium. When K is a compact viability domain, the convexity of the image $F(K)$ implies also the existence of a viable equilibrium:

Theorem 10.1.13 *Let F be a Peano map. If $K \subset \mathrm{Dom}(F)$ is a compact viability domain and if $F(K)$ convex, then there exists an equilibrium of F in K.*

Proof — Assume that there is no equilibrium. This means that 0 does not belong to the closed convex subset $F(K)$, so that the Separation Theorem implies the existence of some $p \in X^\star$ and $\varepsilon > 0$ such that

$$\sup_{v \in F(K)} < v, -p > = \sigma(F(K), -p) < -\varepsilon$$

Let us take any solution $x(\cdot)$ to differential inclusion (10.6) viable in K, which exists by the Viability Theorem. We deduce that

$$\forall\, t \geq 0, \ \ \langle -p, x'(t) \rangle \leq -\varepsilon$$

so that, integrating from 0 to t, we infer that

$$\forall\, t \geq 0, \ \ \varepsilon\, t \leq < p, x(t) - x(0) >$$

But K being bounded, we thus derive a contradiction. □

With the same kind of techniques, we derive the following criterion on the existence of an equilibrium:

Theorem 10.1.14 *Let us assume that F is upper hemicontinuous with closed convex images and that $K \subset \mathrm{Dom}(F)$ is compact. If there exists a solution $x(\cdot)$ of (10.6) viable in K such that*

$$\inf_{t>0} \frac{1}{t} \int_0^t \|x'(\tau)\| d\tau \; = \; 0$$

then there exists an equilibrium of F in K.

Proof — Let us assume that there is no viable equilibrium, i.e., that for any $x \in K$, 0 does not belong to $F(x)$. Since these subsets are closed and convex, the Separation Theorem implies that there exists p in the unit sphere of X and $\varepsilon_p > 0$ such that

$$\sigma(F(x), -p) \; < \; -\varepsilon_p$$

In other words, we can cover the compact subset K by the subsets

$$\mathcal{V}_p \; := \; \{x \in K \mid \sigma(F(x), -p) \; < \; -\varepsilon_p\}$$

when p ranges over Σ. Since they are open, thanks to the upper hemicontinuity of F, K can be covered by q open subsets $\mathcal{V}_{p_i}$. Set

$$\varepsilon \; := \; \min_{j=1,\ldots,q} \varepsilon_{p_j} > 0$$

Consider now any solution to differential inclusion (10.6) viable in K. Hence, for any $t \geq 0$, $x(t)$ belongs to some $\mathcal{V}_{p_j}$, so that

$$-\|x'(t)\| \; \leq \; \langle -p_j, x'(t) \rangle \; \leq \; \sigma(F(x(t)), -p_j) \; < \; -\varepsilon$$

and thus, by integrating from 0 to t, we have proved that there exists $\varepsilon > 0$ such that for all $t > 0$,

$$\varepsilon \; < \; \frac{1}{t} \int_0^t \|x'(\tau)\| d\tau$$

a contradiction of the assumption of the theorem. □

10.2 Applications of the Viability Theorem

In this section, we shall posit the linear growth conditions which guarantee the existence of solutions to the differential inclusions on the half-line $[0, \infty[$ ($T = +\infty$.)

10.2.1 Linear Differential Inclusions

We now take for the right-hand side of F a closed convex process and for viability domain a closed convex cone K. We recall that by Lemma 4.2.5,

$$\forall\, x \in K, \; T_K(x) = \overline{K + \mathbf{R}x}$$

Corollary 10.2.1 *Let $F : X \leadsto X$ be a closed convex process and $K \subset X$ be a closed convex cone. We posit the following assumptions:*

$$\begin{cases} i) & \forall\, x \in K, \; R(x) := F(x) \cap T_K(x) \neq \emptyset \\ ii) & \text{the norm } \|R\| \text{ of } R \text{ is finite} \end{cases}$$

Then, for any initial state $x_0 \in K$, there exists a solution $x(\cdot)$ to the "linear differential inclusion"

$$\text{for almost all } t \geq 0, \quad x'(t) \in F(x(t)) \tag{10.8}$$

starting at x_0 and viable in the cone K.

10.2.2 Lyapunov Functions

We consider differential inclusion (10.6) and a time-dependent function $w(\cdot)$ defined as a solution to the differential equation

$$w'(t) = -\varphi(w(t)) \tag{10.9}$$

where $\varphi : \mathbf{R}_+ \mapsto \mathbf{R}$ is a given continuous function with linear growth.

Our problem is to characterize functions enjoying the φ-Lyapunov property, i.e., nonnegative extended functions $V : X \mapsto \mathbf{R}_+ \cup \{+\infty\}$ (such that $\mathrm{Dom}(V) \subset \mathrm{Dom}(F)$) satisfying

$$\forall\, t \geq 0, \; V(x(t)) \leq w(t), \quad w(0) = V(x_0) \tag{10.10}$$

along at least one solution $x(\cdot)$ to differential inclusion (10.6) starting at x_0 where w is a solution to differential equation (10.9.) Since this condition amounts to saying that the epigraph of V enjoys the viability property for the differential inclusion

$$(x'(t), w'(t)) \in G(x(t), w(t)) \text{ where } G(x,w) := F(x) \times \{-\varphi(w)\}$$

we can apply the Viability Theorem on the epigraph of V.

This allows us *to use lower semicontinuous instead of differentiable functions among the candidates to satisfy this Lyapunov property.*

We have then to translate the fact that $\mathcal{E}p(V)$ is a viability domain of G, and for that purpose, to use Proposition 6.1.4 which states that the contingent cones to the epigraph are the epigraphs of the contingent epiderivatives: This leads us to extend the concept of Lyapunov functions to lower semicontinuous functions

$$V : X \mapsto \mathbf{R} \cup \{+\infty\}$$

We recall that V is said to be *contingently epidifferentiable* at x if its contingent epiderivative never takes the value $-\infty$ (see Definition 6.1.2.)

Definition 10.2.2 (Lyapunov Function) *Let*

$$V : X \mapsto \mathbf{R}_+ \cup \{+\infty\}$$

be a contingently epidifferentiable function.

We shall say that V is a Lyapunov function *of F associated with φ if and only if V is a solution to the* contingent Hamilton-Jacobi inequalities

$$\forall\, x \in \mathrm{Dom}(V), \quad \inf_{v \in F(x)} D_\uparrow V(x)(v) + \varphi(V(x)) \leq 0 \qquad (10.11)$$

Theorem 10.2.3 *Let $V : X \mapsto \mathbf{R}_+ \cup \{+\infty\}$ be a contingently epidifferentiable lower semicontinuous function and $F : X \rightsquigarrow X$ be a Peano map. Then V is a Lyapunov function of F associated with $\varphi(\cdot)$ if and only if for any $x_0 \in \mathrm{Dom}(V)$,* there exist solutions *$x(\cdot)$ to (10.6) starting at x_0 and $w(\cdot)$ to (10.9) satisfying property (10.10.)*

Example W-Monotone Set-Valued Maps

Let $W : X \mapsto \mathbf{R}_+ \cup \{+\infty\}$ be a nonnegative extended function. We say that a set-valued map F is W-monotone (with respect to φ) if for every $x,\ y \in X$, we have

$$\forall\, u \in F(x),\ v \in F(y),\ \ D_\uparrow W(x-y)(v-u) + \varphi(W(x-y)) \leq 0 \qquad (10.12)$$

By setting $V(x) = W(x - \overline{x})$ we obtain the following consequence of Theorem 10.2.3:

Corollary 10.2.4 *Let $W : X \mapsto \mathbf{R}_+ \cup \{+\infty\}$ be a contingently epidifferentiable, lower semicontinuous extended function and $F : X \rightsquigarrow X$ be a Peano map such that $-F$ is W-monotone with respect to φ.*

Let $\overline{x}$ be an equilibrium of F (i.e., $0 \in F(\overline{x})$). Then, for any $x_0 \in \mathrm{Dom}(F)$, there exist solutions $x(\cdot)$ and $w(\cdot)$ to (10.6), (10.9) satisfying

$$x(0) = x_0, \; w(0) = W(x_0-\overline{x}) \text{ and } \forall\, t \geq 0, \; W(x(t)-\overline{x}) \leq w(t)$$

In particular, for $W(z) := \frac{1}{2}\|z\|^2$, we find the usual concept of monotonicity (with respect to φ):

$$\forall\, x,y \in X, u \in F(x), v \in F(y), \; \langle u-v, x-y\rangle \geq \varphi\left(\frac{1}{2}\|x-y\|^2\right) \quad \square$$

The function $U : X \mapsto \mathbf{R}_+ \cup \{+\infty\}$ being given, we can construct the smallest lower semicontinuous Lyapunov function larger than or equal to U, i.e., the smallest nonnegative lower semicontinuous solution U_φ to the contingent Hamilton-Jacobi inequalities (10.11) larger than or equal to U. Its epigraph is the viability kernel of the epigraph of U:

Theorem 10.2.5 *Let us consider a Peano map $F : X \rightsquigarrow X$, a continuous function $\varphi : \mathbf{R}_+ \mapsto \mathbf{R}$ with linear growth and*

$$U : X \mapsto \mathbf{R}_+ \cup \{+\infty\}$$

such that $\mathrm{Dom}(U) \subset \mathrm{Dom}(F)$.

Then there exists a smallest nonnegative lower semicontinuous solution

$$U_\varphi : \mathrm{Dom}(F) \mapsto \mathbf{R} \cup \{+\infty\}$$

to the contingent Hamilton-Jacobi inequalities (10.11) larger than or equal to U (which can be the constant $+\infty$), which enjoys the property:

$$\begin{cases} \forall\, x_0 \in \mathrm{Dom}(U_\varphi) \textit{ there exist solutions} \text{ to (10.6), (10.9) satisfying} \\ x(0) = x_0, \; U(x_0) = w(0), \; \forall\, t \geq 0, \; U(x(t)) \leq U_\varphi(x(t)) \leq w(t) \end{cases}$$

Let us single out the following consequence when we take for function $U := d_M$ the distance to a closed set $M \subset X$:

Corollary 10.2.6 *Let us consider a Peano map $F : X \rightsquigarrow X$. Then for all $a \geq 0$, there exists a smallest lower semicontinuous function*

$$d_{M_a} : X \mapsto \mathbf{R} \cup \{+\infty\}$$

larger than or equal to d_M such that

$\forall\, x_0 \in \mathrm{Dom}(d_{M_a})$, there exists a solution $x(\cdot)$ to (10.6) such that

$$x(0) = x_0, \ \forall\, t \geq 0, \quad d_M(x(t)) \leq d_{M_a}(x_0)e^{-at}$$

Therefore, *we can regard the subsets* $\mathrm{Dom}(d_{M_a})$ *as the* basins of exponential attraction *of M.*

10.2.3 Tracking a Differential Inclusion

Let us consider a differential inclusion (10.6) where $F : X \rightsquigarrow X$ is a Peano map and an observation map $H : X \rightsquigarrow Y$ from X to another finite dimensional vector-space Y.

We shall in some sense "project" the differential inclusion (10.6) to a differential inclusion on the observation space Y described by a set-valued map G

$$\text{for almost all } t \geq 0, \ y'(t) \in G(y(t)) \tag{10.13}$$

in order to "track" (or "filter") a solution $x(\cdot)$ to differential inclusion (10.6) in the following sense:

$$\left\{ \begin{array}{l} \forall\, x_0 \in \mathrm{Dom}(F) \text{ and } y_0 \in H(x_0), \ \text{ there exist} \\ \text{solutions } x(\cdot) \ \& \ y(\cdot) \text{ to } (10.6), (10.13) \text{ such that} \\ x(0) = x_0, \ y(0) = y_0 \text{ and } \forall\, t \geq 0, \ y(t) \in H(x(t)) \end{array} \right. \tag{10.14}$$

This property may be called the *tracking property.*

Proposition 10.2.7 *Let us consider a closed set-valued map H from X to Y and Peano maps $F : X \rightsquigarrow X$ and $G : Y \rightsquigarrow Y$. Then tracking property (10.14) holds true if and only if for every $y \in$* Im(H), *we have*

$$\forall\, x \in H^{-1}(y), \ F(x) \cap DH(x,y)^{-1}(G(y)) \neq \emptyset \tag{10.15}$$

It follows obviously from the Viability Theorem 10.1.6, because the above condition amounts to saying that

$$\forall\,(x,y) \in \mathrm{Graph}(H), \;\; (F(x)\times G(y)) \cap T_{\mathrm{Graph}(H)}(x,y) \neq \emptyset$$

i.e., that the graph of H is a viability domain[5] of the set-valued map $F \times G$.

10.3 Nonlinear Semi-Groups

We provide now an existence and uniqueness theorem for differential inclusions the right-hand side of which is minus a maximal monotone set-valued map:

Theorem 10.3.1 (Crandall-Pazy) *Let A be a maximal monotone set-valued map from a Hilbert space X to X. Consider the initial value problem for the differential inclusion*

$$x'(t) \in -A(x(t)), \;\; x(0) = x_0$$

where the initial state x_0 is given in $\mathrm{Dom}(A)$. *Then it has one unique solution $x(\cdot)$ defined on $[0,\infty[$.*

Furthermore if $m(A(x))$ denotes the element of $A(x)$ with the minimal norm, then $x(\cdot)$ is actually the solution to the differential equation

$$\text{for almost all } t \geq 0, \;\; x'(t) = -m(A(x(t)))$$

(*called the* slow solution). *Moreover, the function $t \to \|x'(t)\|$ is nonincreasing.*

Let $x(\cdot) := \mathcal{S}(x_0)$ and $y(\cdot) := \mathcal{S}(y_0)$ be the solutions starting at x_0 and y_0. Then

$$\forall\, t \geq 0, \;\; \|x(t)-y(t)\| \leq \|x_0 - y_0\|$$

Finally,

$$\begin{cases} \forall\, t \geq 0, \;\; x'(t) = \lim_{h\to 0^+}(x(t+h)-x(t))/h \\ \text{and } x'(\cdot) \text{ is continuous from the right} \end{cases}$$

In this case, the solution map $\mathcal{S}$ is single-valued and is called in the literature the *nonlinear semigroup* of the monotone map A.

[5] When $G = 0$, $y_0 = 0$ and H is single-valued, the viability kernel of $H^{-1}(0)$ is closely related to the "zero dynamics" introduced by Byrnes and Isidori.

10.4 Filippov's Theorem

Filippov's theorem for differential inclusions is as important as the Gronwall lemma for ordinary differential equations.

Consider a finite dimensional space X, reals $a < b$ and let F be a set-valued map from $[a, b] \times X$ into closed nonempty subsets of X. We associate with it the differential inclusion

$$\text{for almost all } t \in [a,b], \; x'(t) \in F(t, x(t)) \tag{10.16}$$

Consider the solution map of (10.16), associating with each initial point $x_0 \in X$ the subset

$$\mathcal{S}_{[a,b]}(x_0) := \{x \mid x \text{ is a solution to (10.16) on } [a,b], \; x(a) = x_0\}$$

of solutions to differential inclusion (10.16) on $[a, b]$.

Let $y \in W^{1,1}(a, b; X)$ be a given absolutely continuous map.

Filippov's theorem provides an estimation of the distance between y and the set $\mathcal{S}_{[a,b]}(x_0) \subset W^{1,1}(a, b; X)$ under the following assumptions on F:

$$\begin{cases} i) & \forall\, x \in X, \; F(\cdot, x) \text{ is measurable} \\ ii) & \exists\, \beta > 0, \; k \in L^1(a, b; \mathbf{R}_+) \text{ such that for a.a.} \\ & t \in [a,b], \; F(t, \cdot) \text{ is } k(t)\text{-Lipschitz on } y(t) + \beta B \\ iii) & \text{The map } t \mapsto d(y'(t), F(t, y(t))) \text{ is integrable} \end{cases} \tag{10.17}$$

where B denotes the closed unit ball in X.

Remark — Observe that under the above assumptions, the set-valued map $t \rightsquigarrow F(t, y(t))$ is measurable: Indeed, let y_n be a sequence of measurable simple maps converging pointwise to y. Since the map

$$t \rightsquigarrow F(t, y_n(t))$$

is measurable, Theorem 8.2.5 and assumptions (10.17) imply the measurability of $F(\cdot, y(\cdot))$.

Therefore, Corollary 8.2.13 implies that the function

$$t \mapsto d(y'(t), F(t, y(t)))$$

is always measurable. $\square$

Theorem 10.4.1 (Filippov) *Let us consider an absolutely continuous function $y : [a,b] \mapsto X$, $\delta \geq 0$ and assume that conditions (10.17) hold true. Set*

$$\gamma(t) \;=\; d(y'(t), F(t, y(t))), \quad m(t) \;=\; exp\left(\int_a^t k(s)ds\right)$$

and

$$\eta(t) \;=\; m(t)\left(\delta \;+\; \int_a^t \gamma(s)ds\right)$$

If $\eta(b) \leq \beta$, then for every $x_0 \in B(y(a), \delta)$, there exists $x \in \mathcal{S}_{[a,b]}(x_0)$ such that

$$\forall\, t \;\in\; [a,\, b],\;\; \|x(t) - y(t)\| \;\leq\; \eta(t)$$

and

$$\|x'(t) \;-\; y'(t)\| \;\leq\; k(t)\eta(t) \;+\; \gamma(t) \;\text{ a.e. in }\; [a,b]$$

Corollary 10.4.2 (Lipschitz Dependence) *Let $y \in \mathcal{S}_{[a,b]}(y_0)$ and assume that F, y satisfy assumptions (10.17) i) and ii).*

Then there exists $l > 0$ such that for all $x_0 \in X$ satisfying $m(b)\,\|x_0 - y_0\| \leq \beta$ we have

$$d_{W^{1,1}}\left(y,\; \mathcal{S}_{[a,b]}(x_0)\right) \;\leq\; l\,\|x_0 - y_0\|$$

We state next a "nonlinear version" of Aumann's Theorem on the convexity of the integral of a set-valued map.

This theorem compares solutions to (10.16) and to the convexified (relaxed) differential inclusion:

$$\text{for almost all } t \;\in\; [a,b],\;\; x'(t) \;\in\; \overline{co}\, F(t, x(t)) \qquad (10.18)$$

Theorem 10.4.3 (Filippov-Ważewski) *Let y be a solution to the relaxed inclusion (10.18) on $[a,b]$. Assume that F and y satisfy assumptions (10.17) and that the map $t \rightsquigarrow F(t, y(t))$ is integrably bounded on $[a,b]$. Let $\eta(\cdot)$ be defined as in Theorem 10.4.1.*

If $\eta(b) < \beta$, then for every $\delta > 0$ there exists a solution x to (10.16) on $[a, b]$ satisfying

$$x(0) = y(0) \ \& \ \|x - y\|_C \leq \delta$$

The next theorem is a simple consequence of Theorem 10.4.3 and the Convergence Theorem 7.2.1.

Theorem 10.4.4 (Relaxation Theorem) *Let $F : [a, b] \times X \mapsto X$ be a set-valued map with closed nonempty images. Assume that F is measurable with respect to t and that there exists $k \in L^1(a, b; \mathbf{R}_+)$ such that for almost every $t \in [a, b]$,*

$$F(t, \cdot) \text{ is } k(t)\text{-Lipschitz} \ \& \ \forall\, x \in X, \ F(t, x) \subset k(t)B$$

Then the solutions to (10.16) are dense in the set of solutions to the relaxed inclusion (10.18) for the metric of uniform convergence.

Corollary 10.4.5 *Let $\mathcal{S}^{co}_{[a,b]}(x_0)$ denote the set of solutions to (10.18) on $[a, b]$ with $x(a) = x_0$. We posit all assumptions of Theorem 10.4.4.*

Then for every $x_0 \in X$, the closure of $\mathcal{S}_{[a,b]}(x_0)$ in the metric of uniform convergence is equal to $\mathcal{S}^{co}_{[a,b]}(x_0)$.

Filippov's theorem allows us to investigate infinitesimal generators of reachable maps of the nonlinear dynamical system (10.16.)

For all $a \leq t \leq T \leq b$ and $x_0 \in X$, define

$$R(T, t)x_0 = \{x(T) \mid x \in \mathcal{S}_{[t,T]}(x_0)\}$$

It is called the *reachable set* of (10.16) (or of F) from (t, x_0) at time T.

Reachable sets enjoy the semigroup property

$$\forall\, t \geq s, \ R(T, s)x_0 = R(T, t)R(t, s)x_0$$

When F is sufficiently regular, the set $\overline{co}F(t, x_0)$ is the *infinitesimal generator* of the semigroup $R(\cdot, t)x_0$ in the sense that the "differential quotients" $(R(t + h, t)x_0 - x_0)/h$ converge to $\overline{co}F(t, x_0)$.

For that purpose, we posit the following assumptions:

$$\left\{\begin{array}{ll} i) & \forall\,(t,x)\in[a,b]\times X,\ F(t,x) \text{ is nonempty and compact} \\[2ex] ii) & \forall\, x\ \in\ X,\ \ F(\cdot,x)\ \text{ is continuous on }\ [a,b] \\[2ex] iii) & \forall\,(t_0,x_0)\in[a,b]\times X,\ \text{there exists a neighborhood} \\ & \mathcal{N}\subset[a,b]\times X,\ L>0\ \text{ such that }\ \forall\,(t,x),\ (t,y)\in\mathcal{N}, \\ & F(t,x)\ \subset\ F(t,y)+L\,\|x-y\|\,B \end{array}\right. \tag{10.19}$$

Theorem 10.4.6 *Assume that conditions (10.19) hold true. Then for every* $(t_0,x_0)\in[a,b[\times X$, *for all* (t,x) *near* (t_0,x_0) *and all small* $h>0$

$$R(t+h,t)x\ =\ x\ +\ h\ \mathrm{co}F(t_0,x_0)\ +\ o(t,x,h) \tag{10.20}$$

where

$$\lim_{(t,x)\to(t_0,x_0),\ h\to 0+}\ \frac{o(t,x,h)}{h}\ =\ 0$$

Remark — Equation (10.20) has to be understood in the following way

$$R(t+h,t)x\ \subset\ x+h\ coF(t_0,x_0)+\|o(t,x,h)\|\,B$$

and

$$x+h\ \mathrm{co}F(t_0,x_0)\ \subset\ R(t+h,t)x+\|o(t,x,h)\|\,B\ \ \square$$

10.5 Derivatives of the Solution Map

We now provide estimates of the contingent, adjacent and circatangent derivatives of the solution map $\mathcal{S}:=\mathcal{S}_{[0,T]}$ associated with the differential inclusion

$$\text{for almost all }\ t\ \in\ [0,T],\ \ x'(t)\ \in\ F(t,x(t)) \tag{10.21}$$

where $F:[0,T]\times X\rightsquigarrow X$.

In this section the set-valued map F is regarded as a family of set-valued maps $x \rightsquigarrow F(t,x)$ and *the derivatives are taken with respect to the state variable only.*

Let $\overline{x}$ be a solution of the differential inclusion (10.21.) We assume that F satisfies the following assumptions:

$$\left\{\begin{array}{ll} i) & \forall\, x \in X,\ \ F(\cdot,x) \text{ is measurable} \\ \\ ii) & \forall\, t \in [0,T],\ \ \forall\, x \in X,\ \ F(t,x) \text{ is a closed set} \\ \\ iii) & \exists\, \beta > 0,\ k(\cdot) \in \mathrm{L}^1(0,T) \text{ such that for almost all} \\ \\ & t \in [0,T],\ \ F(t,\cdot) \text{ is } k(t) - \text{Lipschitz on } \overline{x}(t) + \beta B \end{array}\right. \tag{10.22}$$

Consider the *adjacent variational inclusion*, which is the "linearized" along the trajectory $\overline{x}$ inclusion

$$\left\{\begin{array}{lcl} w'(t) & \in & D^{\flat}F(t,\overline{x}(t),\overline{x}'(t))(w(t)) \ \text{ a.e. in } [0,T] \\ w(0) & = & u \end{array}\right. \tag{10.23}$$

where $u \in X$. In Theorems 10.5.1, 10.5.2 below we consider the solution map $\mathcal{S}$ as the set-valued map from $\mathbf{R}^n$ to the Sobolev space $W^{1,1}(0,T;\mathbf{R}^n)$.

Theorem 10.5.1 (Adjacent Variational Inclusion) *If the assumptions (10.22) hold true, then for all $u \in X$, every solution $w \in W^{1,1}(0,T;X)$ to the linearized inclusion* (10.23) *satisfies*

$$w \in D^{\flat}\mathcal{S}(\overline{x}(0),\overline{x})(u)$$

In other words,

$$\{w(\cdot)|\ w'(t) \in D^{\flat}F(t,\overline{x}(t),\overline{x}'(t))(w(t)),\ w(0) = u\} \subset D^{\flat}\mathcal{S}(\overline{x}(0),\overline{x})(u)$$

Proof — Let $h_n > 0$ be a sequence converging to 0. Then, by the very definition of the adjacent derivative and Lipschitz continuity of $F(t,\cdot)$, for almost all $t \in [0,T]$,

$$\lim_{n\to\infty} d\left(w'(t), \frac{F(t,\overline{x}(t) + h_n w(t)) - \overline{x}'(t)}{h_n}\right) = 0 \tag{10.24}$$

Moreover, since

$$\overline{x}'(t) \in F(t, \overline{x}(t)) \text{ a.e. in } [0,T]$$

by (10.22), for all sufficiently large n and almost all $t \in [0,T]$

$$d\left(\overline{x}'(t) + h_n w'(t), F(t, \overline{x}(t) + h_n w(t))\right) \leq h_n \left(\|w'(t)\| + k(t)\,\|w(t)\|\right)$$

This, (10.24) and the Lebesgue dominated convergence theorem yield

$$\int_0^T d\left(\overline{x}'(t) + h_n w'(t), F(t, \overline{x}(t) + h_n w(t))\right) dt = o(h_n) \quad (10.25)$$

where $\lim_{n\to\infty} o(h_n)/h_n = 0$. By the Filippov Theorem and by (10.25) there exist $M \geq 0$ and solutions $y_n \in \mathcal{S}(\overline{x}(0) + h_n u)$ satisfying for all large n

$$\|y_n' - \overline{x}' - h_n w'\|_{L^1(0,T;X)} \leq M o(h_n)$$

Since

$$(y_n(0) - \overline{x}(0))/h_n = u = w(0)$$

this implies that

$$\lim_{n\to\infty} \frac{y_n - \overline{x}}{h_n} = w \text{ in } \mathcal{C}(0,T;X); \quad \lim_{n\to\infty} \frac{y_n' - \overline{x}'}{h_n} = w' \text{ in } \mathrm{L}^1(0,T;X)$$

Hence

$$\lim_{n\to\infty} d\left(w, \frac{\mathcal{S}(\overline{x}(0) + h_n u) - \overline{x}}{h_n}\right) = 0$$

Since u and w are arbitrary the proof is complete. □

Consider next the *circatangent variational inclusion*, which is the linearization involving circatangent derivatives:

$$\begin{cases} w'(t) \in CF(t, \overline{x}(t), \overline{x}'(t))(w(t)) \text{ a.e. in } [0,T] \\ w(0) = u \end{cases} \quad (10.26)$$

where $u \in X$.

Theorem 10.5.2 (Circatangent Variational Inclusion) *Let us assume that conditions (10.22) hold true. Then for all $u \in X$, every solution $w \in W^{1,1}(0,T;X)$ to the linearized inclusion* (10.26) *satisfies $w \in C\mathcal{S}(\overline{x}(0), \overline{x})(u)$.*

In other words,

$$\{w(\cdot) \mid w'(t) \in CF(t, \overline{x}(t), \overline{x}'(t))(w(t)),\ w(0) = u\} \subset C\mathcal{S}(\overline{x}(0), \overline{x})(u)$$

Proof — Consider a sequence x_n of solutions to (10.21) converging to $\overline{x}$ in $W^{1,1}(0,T;X)$ and let $h_n \to 0+$. Then there exists a subsequence $x_j := x_{n_j}$ such that

$$\lim_{j\to\infty} x_j'(t) \ = \ \overline{x}'(t) \text{ a.e. in } [0,T] \tag{10.27}$$

Set $\lambda_j = h_{n_j}$. Then, by the definition of circatangent derivative and by (10.27), for almost all $t \in [0,T]$

$$\lim_{j\to\infty} d\left(w'(t), \frac{F(t, x_j(t) + \lambda_j w(t)) - x_j'(t)}{\lambda_j}\right) \ = \ 0 \tag{10.28}$$

Moreover, using the fact that $x_j'(t) \in F(t, x_j(t))$ a.e. in $[0,T]$, we obtain that for almost all $t \in [0,T]$ and all large j

$$d\left(x_j'(t) + \lambda_j w'(t), F(t, x_j(t) + \lambda_j w(t))\right) \le \lambda_j \left(\|w'(t)\| + k(t)\,\|w(t)\|\right)$$

This, (10.28) and the Lebesgue dominated convergence theorem yield

$$\int_0^T d\left(x_j'(t) + \lambda_j w'(t), F(t, x_j(t) + \lambda_j w(t))\right) dt \ = \ o(\lambda_j) \tag{10.29}$$

where $\lim_{j\to\infty} o(\lambda_j)/\lambda_j = 0$. By the Filippov Theorem and (10.29), there exist $M \ge 0$ and solutions $y_j \in \mathcal{S}(x_j(0) + \lambda_j u)$ satisfying

$$\left\|y_j' - x_j' - \lambda_j w'\right\| \ \le \ Mo(h_j)$$

Since

$$(y_j(0) - x_j(0))/\lambda_j \ = \ u \ = \ w(0)$$

this implies that

$$\lim_{j\to\infty} \frac{y_j - x_j}{\lambda_j} \ = \ w \text{ in } C(0,T;X); \ \lim_{j\to\infty} \frac{y_j' - x_j'}{\lambda_j} \ = \ w' \text{ in } \mathrm{L}^1(0,T;X)$$

Hence

$$\lim_{j\to\infty} d\left(w, \frac{\mathcal{S}(x_j(0) + h_{n_j} u) - x_j}{h_{n_j}}\right) \ = \ 0 \tag{10.30}$$

Therefore we have proved that for every sequence of solutions x_n to (10.21) converging to $\overline{x}$ and every sequence $h_n \to 0+$, there exists a subsequence $x_j := x_{n_j}$ which satisfies (10.30.) This yields that for every sequence of solutions x_n converging to $\overline{x}$ and $h_n \to 0+$

$$\lim_{n\to\infty} d\left(w, \frac{\mathcal{S}(x_n(0)+h_n u) - x_n}{h_n}\right) = 0$$

Since u and w are arbitrary the proof is complete. □

We consider now the *contingent variational inclusion*

$$\left\{ \begin{array}{lcl} w'(t) & \in & \overline{co}DF(t,\overline{x}(t),\overline{x}'(t))(w(t)) \text{ a.e. in } [0,T] \\ w(0) & = & u \end{array} \right. \tag{10.31}$$

Theorem 10.5.3 (Contingent Variational Inclusion) *Consider the solution map $\mathcal{S}$ as a set-valued map from $\mathbf{R}^n$ to $W^{1,\infty}(0,T;\mathbf{R}^n)$ supplied with the weak-$*$ topology and let $\overline{x}(\cdot)$ be a solution to differential inclusion (10.21) starting at x_0.*

Then the contingent derivative $D\mathcal{S}(x_0,\overline{x}(\cdot))$ of the solution map is contained in the solution map of the contingent variational inclusion (10.31), in the sense that

$$D\mathcal{S}(x_0,\overline{x}(\cdot))(u) \subset \{w(\cdot) \mid w'(t) \in \overline{co}DF(t,\overline{x}(t),\overline{x}'(t))(w(t)), w(0) = u\}$$

Proof — Fix a direction $u \in \mathbf{R}^n$ and let $w(\cdot)$ be an element of $D\mathcal{S}(x_0,\overline{x}(\cdot))(u)$. By definition of the contingent derivative, there exist sequences of elements $h_n \to 0+$, $u_n \to u$ and $w_n(\cdot) \to w(\cdot)$ in the weak-$\star$ topology of $W^{1,\infty}(0,T;\mathbf{R}^n)$ and $c > 0$ satisfying

$$\left\{ \begin{array}{ll} i) & \|w_n'(t)\| \le c \text{ a.e. in } [0,T] \\ ii) & \overline{x}'(t) + h_n w_n'(t) \in F(t,\overline{x}(t) + h_n w_n(t)) \text{ a.e.} \\ iii) & w_n(0) = u_n \end{array} \right. \tag{10.32}$$

Hence

$$\left\{ \begin{array}{ll} i) & w_n(\cdot) \text{ converges pointwise to } w(\cdot) \\ ii) & w_n'(\cdot) \text{ converges weakly in } L^1(0,T;\mathbf{R}^n) \text{ to } w'(\cdot) \end{array} \right. \tag{10.33}$$

By Mazur's Theorem and (10.33) *ii*), a sequence of convex combinations

$$v_m(t) := \sum_{p=m}^{\infty} a_m^p w_p'(t)$$

converges strongly to $w'(\cdot)$ in $L^1(0,T;X)$, where $a_m^p \geq 0$,

$$\sum_{p=m}^{\infty} a_m^p = 1$$

and for every m, $a_m^p \neq 0$ only for a finite number of p.

Therefore a subsequence (again denoted) $v_m(\cdot)$ converges to $w'(\cdot)$ almost everywhere. By (10.32) *i*), *ii*) for all p and almost all $t \in [0,T]$

$$w_p'(t) \in \frac{1}{h_p}\left(F(t,\overline{x}(t) + h_p w_p(t)) - \overline{x}'(t)\right) \cap cB$$

Let $t \in [0,T]$ be a point where $v_m(t)$ converges to $w'(t)$ and $x'(t) \in F(t,x(t))$. Fix an integer $n \geq 1$ and $\varepsilon > 0$. By (10.33) *i*), there exists m such that $h_p \leq 1/n$ and $\|w_p(t) - w(t)\| \leq 1/n$ for all $p \geq m$.

Then, by setting

$$\Phi(y,h) := \frac{1}{h}\left(F(t,\overline{x}(t) + hy) - \overline{x}'(t)\right) \cap cB$$

we obtain that

$$v_m(t) \in co\left(\bigcup_{h\in]0,\frac{1}{n}],\ y\in w(t)+\frac{1}{n}B} \Phi(y,h)\right)$$

and therefore, by letting m go to ∞, that

$$w'(t) \in \overline{co}\left(\bigcup_{h\in]0,\frac{1}{n}],\ y\in w(t)+\frac{1}{n}B} \Phi(y,h)\right)$$

Since the subsets $\Phi(y,h)$ are contained in the ball of radius c, we infer that $w'(t)$ belongs to the closed convex hull of the upper limit by Lemma 1.1.9. Hence

$$w'(t) \in \overline{co} \bigcap_{\varepsilon>0,\ n\geq 1}\left(\bigcup_{h\in]0,\frac{1}{n}], y\in w(t)+\frac{1}{n}B} \Phi(y,h) + \varepsilon B\right)$$

We observe now that

$$\bigcap_{\varepsilon>0,n\geq 1}\left(\bigcup_{h\in]0,\frac{1}{n}],y\in w(t)+\frac{1}{n}B}\Phi(y,h)+\varepsilon B\right)\subset DF(t,\overline{x}(t),\overline{x}'(t))(w(t))$$

to conclude that $w(\cdot)$ is a solution to the contingent differential inclusion

$$\begin{cases} w'(t) & \in \ \overline{co}DF(t,\overline{x}(t),\overline{x}'(t))(w(t)) \ \text{ a.e. in } [0,T] \\ w(0) & = \ u \end{cases}$$

Since $w\in D\mathcal{S}(x_0,\overline{x}(\cdot))(u)$ is arbitrary we end the proof. □

Bibliographical Comments

Chapter 1

We already have mentioned that the concepts of upper and lower limits of sets are due to Painlevé in 1902 [318] as it is reported by his student Zoretti in [467].

The compactness result (Theorem 1.1.7) was obtained by Zarankie wicz [462]. Many other proofs were provided since. The Duality Theorem 1.1.8 was proved first by Walkup & Wets [427].

Some basic calculus of upper and lower limits of sets and set-valued maps appeared before and after the second World War in the two volumes of Kuratowski's basic monograph [255]. The continuous version of Proposition 1.2.6 on the lower limits (or lower semicontinuity) of intersections of convex-valued maps can be found in [35] for instance.

The square product of set-valued maps was introduced and used in [40] for studying local observability of control systems under uncertainty.

Upper and lower semicontinuous maps were introduced by Bouligand [80] and Kuratowski [254] and [255]. Berge's book [66] played an important role in disseminating these results among mathematical economists, as one can see by reading Debreu's basic monograph [132]. The convergence issues were then reexamined by Choquet in [100] in a more general case. Pseudo-Lipschitz set-valued maps have been introduced in [47,49] in the framework of the Inverse Function Theorem in finite-dimensional spaces, and later, in [39] for Banach spaces, and have also been studied by Rockafellar in [370].

The Genericity Theorem is due to Kuratowski [254,255]. It was extended later by Choquet [100] and Shi Shuzhong [388].

The Maximum Theorem appeared in Berge's monograph [66].

The lower semicontinuity criterion of an infinite intersection of lower semicontinuous set-valued maps was proved in the context of differential games [51].

Chapter 2

Closed convex processes and their transposes have been introduced by Rockafellar [356] and further studied in [358,360]. Robinson-Ursecu's Theorem has been proved in [347,348,349,417] and the Uniform Boundedness and the Crossed Convergence Theorem in [42]. The Closed Image Theorem is taken from [35]. Closed convex process are related to Ioffe's fans ([233,238]) for which results of linear functional analysis can also be extended.

Linear processes and their invariance properties were studied in [197].

Chapter 3

The literature on nonlinear analysis is too broad to give any fair account in this short review.

Ky Fan's Inequality has been proved in [160] and the proof we gave is derived from Browder in finite dimensional vector-spaces and from [45,35] in the general case, where extensions are provided and further bibliographical comments can be found. The Equilibrium Theorem was derived from Ky Fan's papers [159,161] in [45]. See also recent contributions of Simons [395]. One can find in [35] further developments on the existence of solutions to inclusions. Kakutani's Theorem appeared in [243] and the set-valued version of the Leray-Schauder Theorem in [45].

Ekeland's variational principle has been proved in [154] and used very often since then. For more recent views, see [35, Chapter 5], [156] and their bibliography. The proof given here is due to Siegel [393].

One can find in [2] the history of the concept of derivative of maps between Banach spaces, from Peano, Volterra, Gâteaux and Fréchet to the eve of the nonsmooth era. We mention only the following [323,201,199,200].

The Inverse Function Theorem for surjective single-valued maps in Banach spaces is due to Graves [205,206]. See also the earlier paper [220] in collaboration with Hildenbrandt. Ljusternik proved in [273] that the tangent space to $f^{-1}(0)$ is the kernel of $f'(x)$ when the Fréchet derivative f' is continuous and surjective (See [272] for instance.) This has been extended by Leach in [265] to strong differentiable maps.

The extension of this theorem to problems with constraints and set-valued maps in finite dimensional spaces appeared in [266,46,47, 49] and [35, Chapter 7]. The use of the localization property of Ekeland's Variational Principle is due to Lebourg [266]. It has been adapted to the case of infinite dimensional spaces in [39,178,182,189, 194].

This theorem has been extended to the case when the surjectivity of the first derivative is lacking, and also, when the set-valued map is defined on metric spaces, thanks to the concept of *high order variations* introduced in [177,182,194,189]. Such results were motivated by control theory, in particular by local controllability of control systems in [180,190]. See [194] and the forthcoming monograph [195] for an exhaustive presentation of the theory and other applications.

Monotone maps have been introduced by Zarantonello [463]. For a comprehensive account of monotone maps, see for instance the books of Brézis [83], J.-L. Lions [270] and Browder [84]. Maximal monotone maps were characterized by Minty [290] and the fact that the subdifferential is maximal monotone is due to Rockafellar [354]. See for instance [35, Chapter 6] for developments of this important domain of nonlinear analysis and further bibliography.

The existence of eigenvectors of a closed convex process on cones with compact soles appeared in [31] for characterizing controllability of closed convex processes. Existence of positive eigenvectors and extension of the Perron-Frobenius Theorem appeared in [43] and can be found in [35, Chapter 3].

Chapter 4

It is impossible to give an exhaustive account of the many papers in which the various concepts of tangent cones appeared in the literature. They have been introduced again and again in a manifold of contexts, many of them unknown to the authors.

The need to introduce *contingent and paratingent* directions was felt by Bouligand in the thirties in order to differentiate non differentiable functions (see [78,80,79].) This was taken up by Zaremba [465] and Marchaud [280] to define solutions to differential inclusions. See also the paper [100] by Choquet.

The contingent cone was used under the names of *tangent cone* and *sequential tangent cone* in optimization and control theory in the fifties (see for instance the books by Neustadt [300] and Hestenes [218]) overshadowed during the sixties when convex analysis was expending, and was resurrected in this field as a *parent pauvre* of the tangent cone introduced by Clarke in the middle of the seventies in his thesis [106]. Clarke and Rockafellar's papers extended the main results of convex analysis to the case of locally Lipschitz functions and lower semicontinuous functions respectively. See also Clarke's book [113].

Meanwhile, Russian and Eastern European mathematicians did use many of these cones. Let us mention the books by Dubovitskij & Miljutin [147], by Ioffe & Tikhomirov [229] and by Pchenitchny [321] for instance, just to cite a few.

This period saw an explosion of notions of tangent cones. We chose to emphasize only three of them, the contingent cone since it appears that it is the one which is at the origin of most theorems, the adjacent cone which is needed in Lebesgue and Sobolev spaces (see [172,173,183,181,186,196]), and the tangent cone introduced by Clarke, because of its regularity and convexity properties.

The fact that it coincides with the contingent cone for sleek subsets was first proved in [34] in finite dimensional vector-spaces. The proof that the lower limit of the contingent cones is contained in the Clarke tangent cone is due to Cornet in [118,117] in the case of Hilbert spaces (see also [325]) and the extension to any Banach space is due to Treiman [411]. We provided here a different proof based on Ekeland's Variational Principle. See also [72] on this topic.

The characterization of the polar of a viability domain of a closed convex process appeared in [31,32] to study the duality relations between controllability and observability of closed convex processes.

The calculus of tangent cones to intersections and inverse images appeared in [47] (finite dimensional case, generalizing formulas due to Rockafellar in [361] under stronger transversality conditions) and [39] (Banach spaces.) Theorem 4.3.9 was shown to us by Rockafellar.

A discussion of results on upper limits of normal cones, which we deduce by polarity from the relations between the lower limit of the contingent cones and the Clarke tangent cones can also be found in Rockafellar and Wets' book [355]. These authors also proved that the polar cones to the contingent cones coincide with the sets of proximal normals in finite dimensional vector-spaces. See also [294,232,233,237] and [365,72,70,104] among other contributions to these issues, which were not treated in this book.

Convex kernels (or asymptotic cones) of adjacent (or intermediate) cones have been used in [172,173,181] in the study of optimal control of differential inclusions because the use of the mere convexity of Clarke tangent cones required too strong properties in the infinite dimensional spaces of solutions.

The results on paratingent cones and hypertangent cones are taken from Shi Shuzhong's papers [391,392]. Ioffe's ménagerie appeared in [236]. See also [122], [283] among other contributions.

The results presented in Sections 4.6 and 4.7 appear here for the first time.

Chapter 5

There were many definitions of *pointwise derivatives* of set-valued maps proposed in the literature. Let us mention only Banks- Jacobs' [58], Martelli-Vignoli's [282], De Blasi's [125] and Petcherskaja's [326] to name a few.

The concepts of *graphical derivatives* and their use in the extension of the Inverse Function Theorem have been initiated in [46] for the contingent derivative and in [47,49] for the circatangent derivatives. Adjacent derivatives have been introduced and used in control theory in [172,173,183,181,186,196]. See also Polovinkin [332] who also used the graphical approach to define derivatives of set-valued maps, as well as [378,373]. Proposition 5.2.6 was proved in [181].

Circatangent derivatives of the subdifferential of a convex function, regarded as generalized Hessians, have been introduced in [47, 49] to study the stability of solutions to convex minimization problems. More generally, derivatives of maximal monotone maps have been extensively studied by Rockafellar in [371].

The idea to use Ekeland's Variational Principle for obtaining local surjectivity criteria for single-valued maps is due to Ekeland and was taken up in [46] in the case of set-valued maps by involving the surjectivity of the contingent derivative. This result was improved by Lebourg in [266], who used the localization property of Ekeland's Variational Principle. This appoach has been further exploited in [47,49,39,177,178,182] to provide more and more general versions of this useful theorem. Analogous approaches have been used by Ioffe, in [238] for instance.

For nondifferentiable single valued maps, an inverse function theorem using generalized Jacobian can be found in the papers [109,113] of Clarke. See also the contributions [434] of Warga, [214] of Halkin, [302], and [232,237,238] of Ioffe, who uses the concept of fans, and the references of these papers to Russian literature.

The inverse mapping theorem has been extended to the case when the set-valued map is defined on a metric space, thanks to the concept of *high order variations* introduced in [182,178,189].

Local inverse univocity and injectivity of set-valued maps has been proved in [39,38,40] and used in the study of local observability of control systems and differential inclusions. The localization results are new.

The new field of "Qualitative Physics" (see for instance [67,130]), an active branch of Artificial Intelligence, as well as of mathematical economics (see [259,260]), provide many new motivations for using the differential calculus of set-valued maps and set-valued analysis. The results on qualitative analysis appear here for the first time.

Chapter 6

The characterization of the subdifferential of a convex function through the normal cone to its epigraph goes back to Moreau and Rockafellar when they introduced the concept of subdifferential. Clarke used this characterization in [106] to define the generalized gradient of any lower semicontinuous function through the normal cone to the epigraph. The dual relation between subdifferential of locally Lipschitz functions and the epiderivatives (regarded as the support function of the subdifferential) was established by Clarke. Rockafellar extended it to lower semicontinuous functions in [362,363, 364,366]. Contingent epiderivatives have been introduced in [46] in the study of Lyapunov functions for differential inclusions, where they play a natural and crucial role. See also [33, Chapter 6] and [50]. They are also fundamentally involved in the study of Hamilton-Jacobi equations and closely related to viscosity solutions: see [179, 185,188]. The proof of Theorem 6.1.6 is derived from an idea of Quincampoix.

The epidifferential calculus has been developed in [39]. It has been established by Clarke [113] in the case of locally Lipschitz functions and by Rockafellar [362,363,364,366] under stronger transversality conditions than the ones used in [39].

Relations between the subdifferential and the local subdifferential used by Crandall, Evans and P.-L. Lions were elucidated in [191]. Local sub and super differential were extensively exploited for studying viscosity solutions of Hamilton-Jacobi equations: see [119,120] and the bibliography of P.-L. Lions' book [271].

Other generalizations of the gradients or Jacobians using approximations procedures of functions or maps have been proposed. Let us mention the Warga's derivate containers [432], Halkin's [215] and [174] among other contributions using this approach.

Demianov and Rubinov and their collaborators introduced pairs of generalized gradients to break down the lack of symmetry due to the epigraphical approach. See [140,141] for instance, and their bibliography.

We refer to the books of Rockafellar [358], Moreau [296], Ekeland-Temam [149], [35, Chapter 4], and the forthcoming book [355] by Rockafellar and Wets for an account of convex analysis. For a short introduction to convex analysis, one can also consult [48].

Results of Section 6.5 appear here for the first time.

Many suggestions have been made to define second-order approximations of functions. The one we propose fits in the unified approach of this book. Applied to the epigraph of a function, it yields Rockafellar's extension (parabolic derivatives) given in [377,369] of analogous concept introduced by Ben-Tal and Zowe in [64]. The formula on the second order contingent epiderivative of the Moreau-Yosida approximation of a lower semicontinuous function is new.

Another concept of generalized Hessian was introduced in [47, 49] by taking circatangent derivatives of subdifferentials of convex functions.

In the forthcoming book [355] by Rockafellar and Wets, generalized Hessians are defined as co-differentials of the generalized gradients. See also the contributions of Hiriart-Urruty on this topic.

Chapter 7

This chapter is just a very short introduction to epigraphical (and graphical) convergence.

The first results on this subject are due to Wijsman [454,454] and Walkup & Wets [427] in the convex case.

It has been extended to infinite dimensional spaces and applied to the approximation of variational inequalities by Mosco [297], and then extended by Joly [242] and R. Robert [345].

Extension to the non convex case has known a considerable expansion since the resurrection of this topic with the concepts of Γ and G-convergence by de Giorgi and his collaborators (see for instance [127,126,128,128]) in the framework of calculus of variations, with the concepts of epiconvergence by R. Wets and his collaborators (see [447,379,380,381,382,448,450,451,146,25]) in the framework of convergence of random sets and functions and with the papers of Attouch motivated by the theory of homogenization in [27,28] and above all in his book [30]. We refer to this book and the forthcoming monograph by Rockafellar and Wets [355] for further references and an exhaustive presentation of this topic, for which this chapter can be used as an introduction.

The results on graphical convergence of maximal monotone maps are due to Attouch [29] and the relations between lower graphical and pointwise convergence appeared in [42].

The Convergence Theorem has been used by several authors studying differential inclusions with upper semicontinuous convex-valued right-hand side (see [33, Chapter 1] for instance.) We state here a more general version of this theorem.

The results on the limits of infima are taken from Wets [449].

The theorem on conjugate functions of epilimits was proved by Wijsman [453,454] in the finite dimensional case and by Mosco [297], Joly [242] and R. Robert [345] in the infinite dimensional case.

The theorem on the lower limits of gradients of differentiable functions converging uniformly is new. It extends to the infinite-dimensional case a theorem of M. Crandall, C. Evans & P.-L. Lions [119] which allows one to study the stability of viscosity solutions to Hamilton-Jacobi equations. See P.-L. Lions' book [271] for further comments on subdifferentials of viscosity solutions to Hamilton-Jacobi equations. Their relations with generalized gradients have been investigated in [188,89].

The theorem on convergence of subdifferentials of convex functions is due to Attouch [28,26]. See also his monograph [30].

Chapter 8

The measurable selection theorem, which, as many other important mathematical innovations of this century, is due to J. von Neumann. This result was given in a stronger form by Kuratowski and Ryll-Nardzewski [253]. In [92], Castaing has proved that measurable selections are dense, the result which is often referred to now as the Castaing representation theorem.

The interested reader can learn more about the history of this subject from Wagner's survey [426] of measurable selection theorems and also from Ioffe's survey [231] of Soviet literature on this subject.

The book [91] by Castaing & Valadier remains the reference for measurable set-valued maps. In particular, the proofs of Theorem 8.1.4 and of some its applications to the calculus of measurable set-valued maps are borrowed from this book. Using the Mackey topology it is possible to prove Theorem 8.2.14 without the assumption that $X^\star$ is separable (see [91].)

The characterization of tangent cones to subsets of Lebesgue spaces extends the results obtained in the convex case by Hess.

The concept of integral has been introduced by Aumann [52] and the convexification result is due to Aumann [52], Debreu [133] and Olech [307,312,305].

The estimation Lemma and the Theorem on the convergence of selections to extremal points are taken from Olech [312]. Its more precise version was proved in [171] and used in [170] by Frankowska & Olech to study boundary trajectories of nonlinear control systems.

The extension of the Convexity Theorem on the convexity of the closure of the integral to Banach spaces is due to Uhl [415] and Hiai-Umegaki [219].

Other concepts of integral of set-valued maps have been proposed. Let us mention for instance *Pontriagin's alternate integrals* introduced in [333] and further studied by M. S. Nikolskii in [303,304].

The bang-bang principle is due to LaSalle and Olech. The proof we used is taken from Olech [312].

Existence of invariant measures of set-valued maps on compact spaces and the extension of Poincaré Recurrence Theorem can be found in [41]. A very pedagogical account of ergodic theory is given in the monograph [261] by Lasota & Mackey.

Chapter 9

The Continuous Selection Theorem is due to Michael [286,287, 288] and the Approximate Selection Theorem of an upper semicontinuous set-valued map to Cellina [96,97].

The idea to use the Steiner points of the images of a set-valued map as selection and parametrization procedures is due to Lojasiewicz

Jr. in 1979 (Olech's seminar, see also [276]).

The Steiner point was introduced by Steiner for convex polytopes of $\mathbf{R}^2$ in [402] and extended to convex polytopes of $\mathbf{R}^n$ by Grünbaum: see [207] and Grünbaum's book [208]. It has been generalized to convex bodies of finite dimensional vector spaces by Shepard in [384].

Rademacher's theorem implies that in Theorem 9.4.1 the subdifferential of the support function is almost everywhere single-valued. Hence the minimal selection in the formula of the Steiner point as an average of the faces on the unit ball can be replaced by any selection.

Vitale proved that Steiner points cannot be extended continuously to compact convex subsets of a Hilbert space in [424]. The appendix of this paper provides a sharper estimate of the Lipschitz constant of the Steiner map than the simple one we gave.

Similar Intersection Lemmas were proved by Le Donne-Marchi in [264] and by Lojasiewicz [276] and Ornelas [314].

The Parametrization Theorem for continuous maps is due to Ekeland-Valadier [151] and, for Lipschitz maps, has been proved by Lojasiewicz Jr. [276] and Ornelas [314] (with better constants, but more technical proofs.)

Chapter 10

One can find an extensive bibliography on differential inclusions and their applications to control and viability theory in the books [33], [195] and [50] for instance.

The Viability Theorem is due to Haddad [212,213].

For Fixed Point Theorems using properties of differential inclusions, see among other contributions Deimling's book [135] and also [136].

The Theorem on nonlinear semigroups is due to Crandall-Pazy, [164].

The original proof of Filippov's Theorem appeared in [169]. It can be also found in [33] for continuous with respect to the time set-valued maps. See also [196] for an infinite dimensional version of this theorem.

The proof of Filippov-Ważewski's Theorem can be found in [169], [446] and [113] and for continuous set-valued maps in [33]. Its proof involves the Aumann and the Filippov's theorems. An infinite dimensional analogue was given in [196]. Derivatives of the solution map have been studied in [181,183,40,38,333,204].

Bibliography

[1] ANICHINI G., CONTI G. & ZECCA P. (to appear) *Approximation and selection for nonconvex multifunctions in infinite dimensional spaces*, Preprint

[2] ANNARATONE S. (1989) *Breve storia del concetto di differenziale negli spazi vettoriali normati*, Thesis Università di Milano

[3] ARTSTEIN Z. & PRIKRY K. (1987) *Carthéodory Selections and the Scorza Dragoni Property*, J. Math. Anal. Appl., 127,540-547

[4] ARTSTEIN Z. & WETS R. (1988) *Approximating the integral of a multifunction*, J. Multivariate Analysis, 24, 285-308

[5] ARTSTEIN Z. (1974) *On the calculus of closed set-valued functions*, Indiana Univ. Math. J., 24, 433-441

[6] ARTSTEIN Z. (1989) *A Variational Convergence that Yields Chattering Systems*, in ANALYSE NON LINÉAIRE, Attouch, Aubin, Clarke, Ekeland Eds., Gauthier-Villars & C.R.M. Université de Montréal, 49-72

[7] ARTSTEIN Z. (1989) *Piecewise Linear Approximations of Set-Valued Maps*, J. Approx. Theory, 56, 41-47

[8] ARTSTEIN Z. (1989) *Yet another proof of Lyapunov convexity theorem*, Proceedings of the AMS

[9] ARTSTEIN Z. (to appear) *Parametrized Integration of Multifunctions with Applications to Control and Optimization*, Preprint

[10] ASEEV S.M. (1982) *Approximation of semicontinuous multivalued mappings by continuous ones*, Izvestiya Acad. Nauk SSSR, Vol.46. (English Translation. Math. USSR Izvestiya, [1983], Vol.20, p.435-448)

[11] ASEEV S.M. (1983) *Approximation of Semicontinuous Multi-valued Mappings by Continuous Ones*, Math. USSR Izvestija, 20, 435-448

[12] ASEEV S.M. (1985) *Opérateurs quasi-linéaires et leurs applications à la théorie des correspondances*, Trudy Mat. Inst. AN SSSR, Vol.167, 25-52

[13] ASPLUND E. & ROCKAFELLAR R.T. (1969) *Gradients of convex functions*, Trans. Am. Math. Soc., 139, 443-467

[14] ASPLUND E. (1966) *Farthest points in reflexive locally uniformly rotund Banach spaces*, Isr. J. Math., 4, 213-216

[15] ASPLUND E. (1968) *Fréchet differentiability of convex functions*, Acta Math., 121, 31-47

[16] ATTOUCH H., AUBIN J.-P., CLARKE F. H. & EKELAND I. Eds. (1989) ANALYSE NON LINÉAIRE, Gauthier-Villars & C.R.M. Université de Montréal

[17] ATTOUCH H. & BREZIS H. (1986) *Duality for the sum of convex functions in general Banach spaces*, in ASPECTS OF MATHEMATICS AND ITS APPLICATIONS, J. Barroso, Ed. North-Holland, 125-133

[18] ATTOUCH H. & WETS R. (1983) *A convergence theory for saddle functions*, Trans. A.M.S., 280, 1-41

[19] ATTOUCH H. & WETS R. (1983) *Convergence des points Min-Sup et des points fixes*, Comptes-Rendus de l'Académie des Sciences, PARIS, Vol.296, 657-660

[20] ATTOUCH H. & WETS R. (1987) *A quantitative approach via epigraphic distance to stability of strong local minimizers*, Publications AVACAM, 1 (1) - Université de Perpignan

[21] ATTOUCH H. & WETS R. (1987) *Lipschitzian stability of ϵ-approxiamate solutions in convex optimization*, IIASA WP 87-25

[22] ATTOUCH H. & WETS R. (1988) *Quantitative Stability of Variational Systems: I. The Epigraphical Distance*, Preprint IIASA 88-8

[23] ATTOUCH H. & WETS R. (1988) *Quantitative Stability of Variational Systems: II. A Framework for Nonlinear Conditioning*, Preprint IIASA 88-9

[24] ATTOUCH H. & WETS R. (1988) *Another Isometry for the Legendre-Fenchel transform*, J. Math. Anal. Appl.,131, 404-411

[25] ATTOUCH H. & WETS R. (1989) *Epigraphical Analysis*, in ANALYSE NON LINÉAIRE, Attouch, Aubin, Clarke, Ekeland Eds., Gauthier-Villars & C.R.M. Université de Montréal, 73-100

[26] ATTOUCH H. (1977) *Convergence de Fonctionnelles Convexes*, In PROC. JOURN. ANAL. NON LINÉAIRE, Lecture Notes in Math. # 665 Springer-Verlag, 1-40

[27] ATTOUCH H. (1977) *Convergence de fonctions convexes, de sous-différentiels et semi-groupes associés*, C.R.A.S., 284, 539-542

[28] ATTOUCH H. (1978) *Sur la* Γ*-convergence*, In SEMINAR ON PARTIAL DIFFERENTIAL EQUATIONS, H. Brézis & J.-L. Lions, Eds., Pitman,

[29] ATTOUCH H. (1979) *Famille d'opérateurs maximaux monotones et mesurabilité*, Ann. Mat, Pura Appli., 120, 35-111

[30] ATTOUCH H. (1984) VARIATIONAL CONVERGENCE FOR FUNCTIONS AND OPERATORS, Research Notes in Mathematics. Pitman, London

[31] AUBIN J.-P., FRANKOWSKA H. & OLECH C. (1986) *Controllability of convex processes*, SIAM J. on Control and Optimization, 24, 1192-1211

[32] AUBIN J.-P., FRANKOWSKA H. & OLECH C. (1986) *Contrôlabilité des processus convexes*, Comptes-Rendus de l'Académie des Sciences, Paris, 301, 153-156

[33] AUBIN J.-P. & CELLINA A. (1984) DIFFERENTIAL INCLUSIONS, Springer-Verlag, Grundlehren der math. Wiss. # 264

[34] AUBIN J.-P. & CLARKE F. H. (1977) *Monotone invariant solutions to differential inclusions*, J. London Math. Soc., 16, 357-366

[35] AUBIN J.-P. & EKELAND I. (1984) APPLIED NONLINEAR ANALYSIS, Wiley-Interscience

[36] AUBIN J.-P. & FRANKOWSKA H. (1984) *Trajectoires lourdes de systèmes contrôlés*, Comptes-Rendus de l'Académie des Sciences, PARIS, Série 1, 298, 521-524

[37] AUBIN J.-P. & FRANKOWSKA H. (1985) *Heavy viable trajectories of controlled systems*, Annales de l'Institut Henri-Poincaré, Analyse Non Linéaire, 2, 371-395

[38] AUBIN J.-P. & FRANKOWSKA H. (1987) *Observabilité des systèmes sous incertitude*, Comptes-Rendus de l'Académie des Sciences, PARIS, 306, 413-416

[39] AUBIN J.-P. & FRANKOWSKA H. (1987) *On the Inverse Function Theorem*, J. Math. Pures Appliquées, 66, 71-89

[40] AUBIN J.-P. & FRANKOWSKA H. (1989) *Observability of Systems under Uncertainty*, SIAM J. on Control and Optimization, Special Issue dedicated to Mac Shain, 27,

[41] AUBIN J.-P., FRANKOWSKA H. & LASOTA A. (to appear) *Poincaré's Recurrence Theorem for Set-Valued Dynamical Systems*, Bull. Polish Acad. Sciences

[42] AUBIN J.-P. & WETS R. (1988) *Stable approximations of set-valued maps*, Ann.Inst. Henri Poincaré, Analyse non linéaire, 5, 519-535

[43] AUBIN J.-P. (1978) *Propriété de Perron-Frobenius pour des correspondances*, C.R.A.S., 286, 911-914

[44] AUBIN J.-P. (1979) APPLIED FUNCTIONAL ANALYSIS, Wiley-Interscience

[45] AUBIN J.-P. (1979) MATHEMATICAL METHODS OF GAME AND ECONOMIC THEORY, North-Holland (Studies in Mathematics and its Applications, Vol. 7, 1-619)

[46] AUBIN J.-P. (1981) *Contingent derivatives of set-valued maps and existence of solutions to nonlinear inclusions and differential inclusions*, Advances in Mathematics, Supplementary studies, Ed. Nachbin L., 160-232

[47] AUBIN J.-P. (1982) *Comportement lipschitzien des solutions de problèmes de minimisation convexes*, Comptes-Rendus de l'Académie des Sciences, PARIS, 295, 235-238

[48] AUBIN J.-P. (1983) L'ANALYSE NON LINÉAIRE ET SES MOTIVATIONS ÉCONOMIQUES, Masson

[49] AUBIN J.-P. (1984) *Lipchitz Behavior of Solutions to Convex Minimization Problems*, Mathematics of Operations Research, 8, 87-111

[50] AUBIN J.-P. (to appear) VIABILITY THEORY

[51] AUBIN J.-P. (1990) *Differential Games: a Viability Approach*, SIAM J. on Control and Optimization

[52] AUMANN R.J. (1965) *Integrals of set-valued maps*, J. Math. Anal. Appl., 12, 1-12

[53] AUMANN R.J. (1967) *Measurable utility and the measurable choice theorem*, La Décision, Actes Coll. Int. du CNRS, AIX en Provence, 15-26

[54] AUSLENDER A. (1978) *Differentiable stability in nonconvex and nondifferentiable programming*, Mathematical Programming Study, 10, P. Huard (Ed.)

[55] AZE D. (1988) *An Inversion Theorem for Set-Valued Maps*, Bull. Aust. Math. Soc., 37, 411-444

[56] BACCIOTTI A. (1988) *The Local Stabilizability Problem for Nonlinear Systems*, IMA Journal of Mathematical Control & Information, 5, 27-39

[57] BANACH S. (1932) THÉORIE DES OPÉRATIONS LINÉAIRES, Monographie Matematyczne, P. W. N., Warszawa

[58] BANKS H.T. & JACOBS M.Q. (1970) *A differential calculus for multifunctions*, J. Math. Anal. Appl., 29, 246-272

[59] BARBU V. & CELLINA A. (1970) *On the surjectivity of multivalued dissipative mappings*, Boll. Un. Mat. Ital., 3, 817-826

[60] BEAUZAMY P. (1985) INTRODUCTION TO BANACH SPACES AND THEIR GEOMETRY, North Holland, Mathematics study, 68

[61] BEER G. (1988) *On the Young-Fenchel transform for convex functions*, Proc. Amer. Math. Soc., 104, 1115-1123

[62] BEN-TAL A. & ZOWE J. (1982) *A Unified Theory of First and Second Order Conditions for Extremum Problems in Topological Vector Spaces*, Mathematical Programming Study, 19, 39-76

[63] BEN-TAL A. & ZOWE J. (1982) *Necessary and sufficient optimality conditions for a class of nonsmooth minimization problems*, Mathematical Programming, 24, 70-91

[64] BEN-TAL A. & ZOWE J. (1985) *Second Order Optimality Conditions for the L^1-Minimization Problem*, Appl. Math. Optim, 13, 45-58

[65] BERGE C. (1957) *Théorie générale des jeux à n personnes*, Mémorial Sciences. Math. 138

[66] BERGE C. (1959) ESPACES TOPOLOGIQUES ET FONCTIONS MULTIVOQUES, Dunod, Paris

[67] BOBROW D. G. (Ed.) (1984) QUALITATIVE REASONING ABOUT PHYSICAL SYSTEMS, Artificial Intelligence Volume, 24 (North Holland)

[68] BORISOVICH YU. G., GELMAN B.D. & MYSHKIS A.D. (1987) *Multivalued analysis and operator inclusions*, Translated from Itogi Nauki i Tekhniki, Seriya Sovremennye Problemy Matematiki, Noveishie Dostizhenia, 29, 151-211

[69] BORWEIN J.M., FITZPATRICK S.P. & GILES J.R. (1987) *The Differentiability of Real Functions on Normed Linear Space Using Generalized Subgradients*, Journal of Mathematical Analisis and Applications, 28 (2), 512-534

[70] BORWEIN J.M & GILES J.R. (1987) *The Proximal Normal Formula in Banach Space*, Transactions of the Americal Mathematical Society, 302 (1), 371-381

[71] BORWEIN J.M & PREISS D. (1987) *A Smooth Variational Principle with Applications to Subdifferentiability and to Differentiability of Convex Functions*, Transactions of the American Mathematical Society, 303 (2), 517-527

[72] BORWEIN J.M. & STROJWAS H.M. (1986) *Proximal Analysis and Boundaries of Closed Sets in Banach Space, Part I: Theory*, Can. J. Math., 38, 431-452

[73] BORWEIN J.M. & ZHUANG D.M. (1989) *Verifiable Necessary and Sufficient Conditions for Openness and Regularity of Set-Valued and Single-Valued Maps*, J. Math. Anal. Appl.

[74] BORWEIN J.M (1978) *Weak tangent cones and optimization in Banach spaces*, SIAM J. Cont. Opt., 16, 512-522

[75] BORWEIN J.M (1983) *Adjoint Process Duality*, Mathematics of Operations Research, 8, 403-434

[76] BOUCHITTE G. & VALADIER M. (1989) *Multifonctions s.c.i. et régularisée s.c.i. essentielle*, in ANALYSE NON LINÉAIRE, Attouch, Aubin, Clarke, Ekeland Eds., Gauthier-Villars & C.R.M. Université de Montréal, 123-150

[77] BOUCHITTE G. & VALADIER M. (1988) *Integral representation of convex functionals on a space of measures*, J. Funct. An. 80, 398-420

[78] BOULIGAND G. (1930) *Sur les surfaces dépourvues de points hyperlimites*, Ann. Soc. Polon. Math., 9,32-41

[79] BOULIGAND G. (1932) INTRODUCTION À LA GÉOMÉTRIE INFINITÉSIMALE DIRECTE, Gauthier-Villars

[80] BOULIGAND G. (1932) *Sur la semi-continuité d'inclusions et quelques sujets connexes*, Enseignement Mathématique, 31, 14-22

[81] BRESSAN A., CELLINA A. & COLOMBO G. (1988) *Upper semicontinuous differential inclusions without convexity*, Preprint SISSA

[82] BRESSAN A. & COLOMBO G. (1986) *Extensions and Selections of Maps with Decomposable Values*, Preprint SISSA

[83] BREZIS H. (1973) OPÉRATEURS MAXIMAUX MONOTONES ET SEMI-GROUPES DE CONTRACTION DANS LES ESPACES DE HILBERT, North-Holland

[84] BROWDER F. (1976) NONLINEAR OPERATORS AND NONLINEAR EQUATIONS OF EVOLUTION IN BANACH SPACES, Am. Math. Soc. Proc. Symp., 18, 2

[85] BYRNES C.I. & ISIDORI A. (to appear) *Feedback Design From the Zero Dynamics Point of View*, Preprint

[86] BYRNES C.I. & ISIDORI A. (to appear) *Output Regulation of Nonlinear Systems*, Preprint

[87] BYRNES C.I. & ISIDORI A. (to appear) *The Analysis and Design of Nonlinear Feddback Systems. I. Zero Dynamics and Global Normal Forms*, Preprint

[88] BYRNES C.I. & ISIDORI A. (to appear) *The Analysis and Design of Nonlinear Feddback Systems. II. Global Stabilization of Minimum Phase Systems*, Preprint

[89] CANNARSA P.M. & FRANKOWSKA H. (to appear) *Some characterizations of optimal trajectories in control theory*, Preprint Scuola Normale di Pisa

[90] CAPUZZO DOLCETTA I. & FALCONE M. (1989) *Discrete Dynamic Programming and Viscosity Solutions of the Bellman Equation*, in ANALYSE NON LINÉAIRE, Attouch, Aubin, Clarke, Ekeland Eds., Gauthier-Villars & C.R.M. Université de Montréal, 161-184

[91] CASTAING CH. & VALADIER M. (1977) CONVEX ANALYSIS AND MEASURABLE MULTIFUNCTIONS, Springer-Verlag, Lecture Notes in Math. # 580

[92] CASTAING CH. (1967) *Sur les multi-application mesurables*, Inf. Rech. Op., 1, 91-126

[93] CASTAING CH. (1972) *Intégrales convexes duales*, C.R.A.S., 275, 1331-1334

[94] CASTAING CH. (1988) *Convergence faible et selections extremales*, Seminaire d'analyse convexe, Exposé #2, Montpellier

[95] CELLINA A. & COLOMBO R.M. (1989) *On the representation of measurable set valued maps through selections*, Preprint SISSA

[96] CELLINA A. (1969) *A theorem on the approximation of set-valued mappings*, Rend. Ac, Na. Lincei, 47, 429-433

[97] CELLINA A. (1969) *Approximation of set-valued functions and fixed point theorems*, Annali Mat. Pura. Appl., 82, 17-24

[98] CESARI L. (1983) OPTIMIZATION-THEORY AND APPLICATIONS, Springer-Verlag

[99] CHANEY R.W. (1982) *Second-order sufficiency conditions for nondifferentiable programming problems*, SIAM J. Control Optim., 20, 20-33

[100] CHOQUET G. (1947) *Convergences*, Annales de l' Univ. de Grenoble, 23, 55-112

[101] CHOU C.C. (to appear) *Sur l'inversion locale d'une multiapplication*, Preprint

[102] CHOW S. N. & LASOTA A. (1972) *An implicit function theorem for nondifferentiable mappings*, Proc. Amer. Math. Soc., 34, 141-146

[103] CLARKE F.H. & LOEWEN P.D. (1989) *Variational Problems with Lipschitzian Minimizers*, in ANALYSE NON LINÉAIRE, Attouch, Aubin, Clarke, Ekeland Eds., Gauthier-Villars & C.R.M. Université de Montréal, 185-210

[104] CLARKE F.H. & LOEWEN P.D. (to appear) *The Value Function in Optimal Control: Sensitivity, Controllability, and Time-Optimality*, SIAM J. Control and Optimization,

[105] CLARKE F.H. & VINTER R. (1983) *Local optimality conditions and Lipschitzian solutions to the Hamilton-Jacobi equation*, SIAM J. Control and Optimization, 21 (6), 856-871

[106] CLARKE F.H. (1975) *Generalized gradients and applications*, Trans. Am. Math. Soc., 205, 247-262

[107] CLARKE F.H. (1975) *The Euler-Lagrange differential inclusion*, J. Differ. Eq., 19, 80-90

[108] CLARKE F.H. (1976) *A new approach to Lagrange multipliers*, Math. Oper. Res., 1, 165-174

[109] CLARKE F.H. (1976) *On the inverse function theorem*, Pac. J. Math., 64, 97-102

[110] CLARKE F.H. (1977) *Inequality constraints in the calculus of variations*, Can. J. Math., 3, 528-540

[111] CLARKE F.H. (1981) *A variational proof of Aumann's theorem*, Appl. Math. Optim., 7, 373-378

[112] CLARKE F.H. (1981) *Generalized gradients of Lipschitz functionals*, Adv. Math., 40, 52-67. Appeared earlier as MRC Tech. Rep., 1687, Aug. 1976 (Madison, Wisconsin)

[113] CLARKE F.H. (1983) OPTIMIZATION AND NONSMOOTH ANALYSIS, Wiley-Interscience

[114] COLOMBO G. (1988) *Non-convex problems for multifunctions*, PhD. Thesis, SISSA

[115] CORNET B. & LAROQUE G. (1986) *Lipschitz properties of solutions in mathematical programming*, J. Optim. Theory and Appl.

[116] CORNET B. & VIAL J.-Ph. (1987) *Lipschitzian solutions of programming problems*, SIAM J. Control and Optimization, -1137

[117] CORNET B. (1981) *Regular Properties of Tangent and Normal Cones*, Cahiers de Maths. de la Décision #8130

[118] CORNET B. (1984) *Regular Properties of Open Tangent Cones*, Center for Operations Research & Econometrics - Université Catholique de Louvain

[119] CRANDALL M.G., EVANS L.C. & LIONS P.L. (1984) *Some properties of viscosity solutions of Hamilton-Jacobi equation*, Trans. Amer. Math. Soc., 282(2), 487-502

[120] CRANDALL M.G. & LIONS P.L. (1983) *Viscosity solutions of Hamilton-Jaccbi equations*, Trans. Amer. Math. Soc., 277, 1-42

[121] DA PRATO G. & LUNARDI A. (1988) *Stability, Instability and Center Manifold Theorem for Fully Nonlinear Autonomous Parabolic Equations in Banach Space*, Archive for Rational Mechanics and Analysis, 101, 2, 115-141

[122] DANCS S. (to appear) *Tangent and generalized tangent cones in finite dimensional spaces. A unified approach*, Preprint

[123] DE BLASI F. & MIJAK J. (1983) *On continuous approximations for multifunctions*, Preprint Universita Degli Studi di Firenze

[124] DE BLASI F. & MIJAK J. (to appear) *A remark on Continuous Selections*, Preprint

[125] DE BLASI F.S. (1976) *On differentiability of multifunctions*, Pac. J. Math., 66, 67-81

[126] DE GIORGI E. & FRANZONI T. (1975) *Sul un tipo di convergenza variazionale*, Atti Acc. Naz. Lincei, 58, 842-850

[127] DE GIORGI E. (1975) *Sulla convergenza di alcune successioni di integrali del tipo del area*, Rend. Mat. Univ. Roma, 8, 277-294

[128] DE GIORGI E. (1977) *Γ-convergenza e G-convergenza*, Bolletino di Unione Matematica Italiana, 14-A, 213-220

[129] DE GIORGI E. (1981) *Generalized limits in calculus of variations*, in TOPICS IN FUCNCTIONAL ANALYSIS 1980-1981, F. Strocchi, E. Zarantanello, De Giorgi, G. dal Maso & L. Modica Eds, Scuola Norm. Sup. Pisa, 117-148

[130] DE KLEER J. & BROWN J.S. (1984) *A Qualitative Physics Based on Confluences*, Artificial Intelligence, 24

[131] DE PASCALE D. & PAPINI P.L. (1984) *Cones and Projections onto Subspaces in Banach Spaces*, Bollettino U.M.I., 6, 411-420

[132] DEBREU G. (1959) THEORY OF VALUE, Wiley

[133] DEBREU G. (1967) *Integration of correspondances*, 5th Berkeley Symposium on Math. Stat. Prob., II, Part I, 351-372

[134] DEIMLING K. & HU SHOU CHUAN (1988) *Fixed Point of Weakly Inward Maps in Conical Shells*, Nonlinear Analysis, Theory, Methods 1 Applications, 12, 3, 227-230

[135] DEIMLING K. (1985) NONLINEAR FUNCTIONAL ANALYSIS, Springer-Verlag

[136] DEIMLING K. (1986) *Fixed Points of Weakly Inward Multifunctions*, Nonlinear Analysis, Theory, Methods and Applications, 10, 11, 1261-1262

[137] DEIMLING K. (1988) *Positive Fixed Points of Weakly Inward Maps*, Nonlinear Analysis, Theory, Methods and Applications, 12 #3, 223-226

[138] DEIMLING K. (to appear) *The old fixed point problem (differentiable case)*, Preprint

[139] DELFOUR M. & ZOLESIO J.-P. (1989) *Analyse des problèmes de forme par la dérivation des minimax*, in ANALYSE NON LINÉAIRE, Attouch, Aubin, Clarke, Ekeland Eds., Gauthier-Villars & C.R.M. Université de Montréal, 211-228

[140] DEMIANOV V.F. & RUBINOV A.M. (1980) *On quasidifferentiable functionals*, Sov. Math. Dokl., 21, 14-17

[141] DEMIANOV V.F. (1982) NON SMOOTH ANALYSIS (IN RUSSIAN), Leningrad University Press

[142] DEMIANOV V.F., LEMARECHAL C. & ZOWE J. (1986) *Approximation to a Set-valued mapping I: A proposal*, J. Appl. Math. Opt., 14, 203-214

[143] DIESTEL J. (1975) GEOMETRY OF BANACH SPACES, Lecture Notes in Mathematics, vol. 485, Springer Verlag, Berlin

[144] DIEUDONNE J. (1981) HISTORY OF FUNCTIONAL ANALYSIS, Mathematical Studies # 49, North-Holland

[145] DMITRUK A.V., MILYUTIN A.A & OSMOLOVSKII N.P (1980) *Ljusternik's theorem and the theory of extrema*, Uspekhi Mat. Nauk, 35:6, 11-46 / Russian Math Surveys, 35:6, 11-51

[146] DOLECKI S., SALINETTI G. & WETS R. (1983) *Convergence of functions: equi-semicontinuity*, Transactions of the American Mathematical Society, 276, 409-429

[147] DUBOVITSKIJ D. & MILJUTIN A.A. (1971) NECESSARY CONDITIONS FOR A WEAK EXTREMUM IN THE GENERAL PROBLEM OF OPTIMAL CONTROL, Nauka (in Russian)

[148] DUNFORD N. & SCHWARTZ J.T. (1957) LINEAR OPERATORS I, Wiley, N.Y.

[149] EKELAND I & TEMAM R. (1976) CONVEX ANALYSIS AND VARIATIONAL PROBLEMS, North Holland, Amsterdam

[150] EKELAND I. & LEBOURG G. (1976) *Generic Fréchet-differentiability and perturbed optimization problems in Banach spaces*, Trans. Am. Math. Soc., 224, 193-216

[151] EKELAND I. & VALADIER M. (1971) *Representation of set-valued mappings*, J. Mat. Anal. Appl., 35, 621-629

[152] EKELAND I. (1972) *Remarques sur les problèmes variationnels 1*, CRAS, 275, 1057-1059

[153] EKELAND I. (1973) *Remarques sur les problèmes variationnels 2*, CRAS, 276, 1347-1348

[154] EKELAND I. (1974) *On the variational principle*, J. Mathematical Analysis and Applications, 47, 324-353

[155] EKELAND I. (1979) *Nonconvex minimization problems*, Bull. Amer. Math. Soc., 1, 443-474

[156] EKELAND I. (to appear) *The ε-variational principle revisited*, CIME, Varena

[157] EVANS L.C. & SOUGANIDIS P.E. (1989) *A P.D.E. Approach to certain large deviation problems for systems of parabolic equations*, in ANALYSE NON LINÉAIRE, Attouch,

Aubin, Clarke, Ekeland Eds., Gauthier-Villars & C.R.M. Université de Montréal, 229-258

[158] FAN KY (1964) *Sur un théoreme minimax*, C.R. Acad. Sci., 259, 3925-3928

[159] FAN KY (1969) *Extension of two fixed-point theorems of F.E. Browder*, Math. Z., 112, 234-240

[160] FAN KY (1972) *A Minimax Inequality and Applications*, Inequalities III, 103-113, Shisha Ed., Academic Press

[161] FAN KY (1979) *Fixed point theorems and related theorems for non-compact convex sets*, In GAME THEORY AND RELATED TOPICS, 151-156, North-Holland

[162] FAN KY (1981) *A Further Generalization of Shapley's Generalization of the Knaster-Kuratowski-Mazurkiewicz Theorem*, Game Theory and Mathematical Economics, 275-279

[163] FAN KY (1984) *Some Properties of Convex Sets Related to Fixed Point Theorems*, Math. Ann., 266, 519-537

[164] FATTORINI H. & FRANKOWSKA H. (1988) *Necessary conditions for infinite dimensional control problems*, Proceedings of the 8th International Conference on Analysis and Optimization of Systems, Lecture Notes in Control and Information Sciences, Springer Verlag, 517-532

[165] FENCHEL W. (1949) *On conjugate convex functions*, Mimeographed Lecture Notes, Princeton University Press

[166] FENCHEL W. (1952) *A remark on convex sets and polarity*, Medd. Lunds Univ. Mat. Sem. (Supplement Band), 82-89

[167] FILIPPOV A.F. (1958) *On some problems of optimal control theory*, Vestnik Moskowskogo Universiteta, Math. no.2, 25-32

[168] FILIPPOV A.F. (1962) *On certain questions in the theory of optimal control*, SIAM J. Control, 1, 76-84

[169] FILIPPOV A.F. (1967) *Classical solutions of differential equations with multivalued right hand side*, SIAM J. on Control, 5, 609-621

[170] FRANKOWSKA H. & OLECH C. (1982) *Boundary solutions to differential inclusions*, J. Diff. Eqs. 44, 156-165

[171] FRANKOWSKA H. & OLECH C. (1983) *R-convexity of integral of set-valued function*, in CONTRIBUTIONS TO ANALYSIS AND GEOMETRY, John Hopkins University Press, Baltimore, Md., 117-129

[172] FRANKOWSKA H. (1983) *Inclusions adjointes associées aux trajectoires minimales d'inclusions différentielles*, Comptes-Rendus de l'Académie des Sciences, PARIS, Série 1, 297, 461-464

[173] FRANKOWSKA H. (1984) *The maximum principle for a differential inclusion problem, in* ANALYSIS AND OPTIMIZATION OF SYSTEMS, Proceedings of the 6th International Conference on Analysis and Optimization of Systems, Lecture Notes in Control and Information Sciences, Springer Verlag, 517-532

[174] FRANKOWSKA H. (1984) *The first order necessary conditions for nonsmooth variational and control problems*, SIAM J. on Control and Optimization, 22, 1-12

[175] FRANKOWSKA H. (1984) *Contrôlabilité locale et propriétés des semi-groupes de correspondances*, Comptes-Rendus de l'Académie des Sciences, PARIS, Série 1, 299, 165-168

[176] FRANKOWSKA H. (1986) *Les cônes contingents aux ensembles atteignables des systèmes contrôlés*, Comptes-Rendus de l'Académie des Sciences, PARIS 303, Série 1, 733-736)

[177] FRANKOWSKA H. (1986) *Theorème d'application ouverte pour des correspondances*, Comptes-Rendus de l'Académie des Sciences, PARIS, Série 1, 302, 559-562

[178] FRANKOWSKA H. (1987) *An open mapping principle for set-valued maps*, J. of Math. Analysis and Appl., 127, 172-180)

[179] FRANKOWSKA H. (1987) *L'équation d'Hamilton-Jacobi contingente*, Comptes-Rendus de l'Académie des Sciences, PARIS, Série 1, 304, 295-298

[180] FRANKOWSKA H. (1987) *Local controllability of control systems with feedbacks*, Proceedings of 25th CDC Conference, IEEE, Athens, December 10-12

[181] FRANKOWSKA H. (1987) *The maximum principle for an optimal solution to a differential inclusion with end point constraints* SIAM J. on Control and Optimization, 25, 145-157

[182] FRANKOWSKA H. (1987) *Théorèmes d'application ouverte et de fonction inverse*, Comptes Rendus de l'Académie des Sciences, PARIS, Série 1, 305, 773-776

[183] FRANKOWSKA H. (1987) *Local controllability and infinitesimal generators of semigroups of set-valued maps*, SIAM J. on Control and Optimization, 25, 412-432

[184] FRANKOWSKA H. (1988) *A general multiplier rule for infinite dimensional optimization problems with constraints*, Proceedings of the FIRST WORKSHOP ON STABILIZATION OF FLEXIBLE STRUCTURES, Montpellier, December 11-15, 1987, J.P. Zolesiio Editor, Comcon Conference Inc

[185] FRANKOWSKA H. (1988) *Non smooth solutions to Hamilton-Jacobi Equations*, Proceedings of the International Conference Bellman Continuum, Antibes, France, June 13-14, Lecture Notes in Control and Information Sciences, Springer Verlag

[186] FRANKOWSKA H. (1989) *Contingent cones to reachable sets of control systems*, SIAM J. on Control and Optimization, 27, 170-198

[187] FRANKOWSKA H. (1989) *Estimations a priori pour les inclusions différentielles opérationelles*, Comptes-Rendus de l'Académie des Sciences, PARIS, Série 1, 308, 47-55

[188] FRANKOWSKA H. (1989) *Hamilton-Jacobi equation: viscosity solutions and generalized gradients*, J. of Math. Analysis and Appl. 141, 21-26

[189] FRANKOWSKA H. (1989) *High order inverse function theorems*, in ANALYSE NON LINÉAIRE, Attouch, Aubin, Clarke, Ekeland Eds., Gauthier-Villars & C.R.M. Université de Montréal, 283-304

[190] FRANKOWSKA H. (1989) *Local controllability of control systems with feeedbacks*, Journal of Optimization Theory and Applications, 60, 277-296

[191] FRANKOWSKA H. (1989) *Optimal trajectories associated to a solution of contingent Hamilton-Jacobi equations*, Applied Mathematics and Optimization, 19, 291-311

[192] FRANKOWSKA H. (1989) *Set-valued analysis and some control problems*, Proceedings of the International Conference 30 YEARS OF MODERN CONTROL THEORY, Kingston, June 3-6, 1988 E. Roxin Editor, Marcel Dekker

[193] FRANKOWSKA H. (1989) *On the linearization of nonlinear control systems and exact reachability*, Proceedings of INTERNATIONAL CONFERENCE ON CONTROL OF DISTRIBUTED PARAMETER SYSTEMS, Santiago di Compostella (July 1987), Lecture Notes in Control and Information Sciences, Springer Verlag

[194] FRANKOWSKA H. (1990) *Some inverse mapping theorems*, Ann. Inst. Henri Poincaré, Analyse Non Linéaire,

[195] FRANKOWSKA H. (in preparation) SET-VALUED ANALYSIS AND CONTROL THEORY (MONOGRAPH), Birkhäuser

[196] FRANKOWSKA H. (1990) *A priori estimates for operational differential inclusions*, J. Diff. Eqs.

[197] FRANKOWSKA H. (to appear) *On controllability and observability of implicit systems*,

[198] FRECHET M. (1906) *Sur quelques points du calcul fonctionnel*, Rend. Circ. Mat. di Palermo, 22, 51-

[199] FRECHET M. (1911) *Sur la notion de différentielle*, C.R. Acad. Sc., 152, 845-847

[200] FRECHET M. (1925) *La notion de différentielle dans l'analyse générale*, Ann. Ecole Norm. Sup., 42, 293-323

[201] GATEAUX G. (1919) *Fonctions d'une infinité de variables indépendantes*, Bull. Soc. Math. France, 47, 163-177

[202] GAUVIN J. & JANIN R. (1989) *Directional lipschitzian optimal solutions and directional derivatives for the optimal value in nonlinera mathematical programming*, in ANALYSE NON LINÉAIRE, Attouch, Aubin, Clarke, Ekeland Eds., Gauthier-Villars & C.R.M. Université de Montréal, 305-324

[203] GAUVIN J. (1979) *The generalized gradient of a marginal function in mathematical programming*, Math. Opt. Res., 4, 458-463

[204] GINER E. (1986) *Thesis*, Université de Pau

[205] GRAVES L.M. (1935) *Topics in Functional Calculus*, Bull. Amer. Math. Soc.,41, 641-662

[206] GRAVES L.M. (1950) *Some mapping theorem*, Duke Math. J., 17, 111-114

[207] GRUNBAUM B. (1963) *Measures of Symmetry for Convex Sets*, Proc. Symposia Pure Maths. VII, 233-270

[208] GRUNBAUM B. (1967) CONVEX POLYTOPES, J. Wiley-Intersciences

[209] GUSEINOV H. G., SUBBOTIN A. I. & USHAKOV V. N. (1985) *Derivatives for multivalued mappings with applications to game theoretical problems of control*, Problems of Control and Information Theory, Vol.14, 155-167

[210] GWINNER J. (1981) *Bibliography on non-differentiable optimization and non-smooth analysis*, J. Comp. Appl. Math., 7, 277-285

[211] GWINNER J. (1981) *On Fixed Points and Variational Inequalities - A Circular Tour*, Nonlinear Analysis Theory, Methods and Applications, 5, 565-583

[212] HADDAD G. (1981) *Monotone trajectories of differential inclusions with memory*, Isr. J. Math., 39, 83-100

[213] HADDAD G. (1981) *Monotone viable trajectories for functional differential inclusions*, J. Diff. Eq., 42, 1-24

[214] HALKIN H. (1976) *Interior mapping theorem with set-valued derivative*, J. Analyse Math., 30, 200-207

[215] HALKIN H. (1976) *Mathematical Programming without Differentiability*, in CALCULUS OF VARIATIONS AND CONTROL THEORY, D.L. Russel, Ed., Academic Press, 279-288

[216] HAMILTON R. (1982) *The inverse function theorem of Nash and Moser*, Bull. Am. Math. Soc., 7, 1-64

[217] HESS C. (1986) *Thèse Doctorat es-Sciences*, Université de Montpellier

[218] HESTENES M. R. (1966) CALCULUS OF VARIATIONS AND OPTIMAL CONTROL THEORY, Wiley

[219] HIAI F. & UMEGAKI H. (1977) *Integrals, Conditional Expectations, and Martingales of Multivalued Functions*, Journal of Multivariate Analysis, 7, 149-182

[220] HILDEBRANDT & GRAVES L.M. (1927) *Implicit Functions and their Differentials in General Analysis*, Trans. Amer. Math. Soc.,29, 127-153

[221] HIMMELBERG C.J., JACOBS M.Q. & VAN VLECK F.S. (1969) *Measurable multifunctions, selectors and Filippov's implicit function lemma*, J. Math. An. Appl., 25(2), 276-284

[222] HIMMELBERG C.J. & VAN VLECK F.S. (1969) *Some selection theorems for measurable functions*, Can. J. Math. XXI (2), 394-399

[223] HIMMELBERG C.J. & VAN VLECK F.S. (1971) *Selection and implicit function theorems for multifunctions with Souslin graph*, Bull. Acad. Pol. Sc.; 19, 911-916

[224] HIMMELBERG C.J. (1975) *Measurable relations*, Fund. Math., 87, 53-72

[225] HIRIART-URRUTY J.-B. & PLAZANET Ph. 1989 (1989) *Moreau's Decomposition Theorem Revisited*, in ANALYSE NON LINÉAIRE, Attouch, Aubin, Clarke, Ekeland Eds., Gauthier-Villars & C.R.M. Université de Montréal, 325-338

[226] HIRIART-URRUTY J.-B. (1978) *Gradients généralisés de fonctions marginales*, SIAM J. Cont. Opt., 16, 301-316

[227] HIRIART-URRUTY J.-B. (1984) *Miscellanies on nonsmooth analysis and optimization*, Written version of a talk in Sopron

[228] HIRIART-URRUTY J.-B. (1985) *Refinements of necessary optimality conditions in nondifferentiable programming 1*, Lecture Notes in Mathematics, Springer Verlag

[229] IOFFE A.D. & TIHOMIROV V.M. (1974) THEORY OF EXTREMAL PROBLEMS, Nauka Moskow (English Translation, 1878, North-Holland)

[230] IOFFE A.D. (1977) *On lower semicontinuity of integral functionals 1*, SIAM J. Cont. Opt., 15, 521-538

[231] IOFFE A.D. (1978) *Survey of measurable selection theorems: Russian literature supplement*, SIAM J. Contr. Opt., 16, 728-723

[232] IOFFE A.D. (1981) *Nonsmooth analysis: differential calculus of nondifferentiable mappings*, Trans. Am. Math. Soc., 266, 1-56

[233] IOFFE A.D. (1982) *Nonsmooth analysis and the theory of fans*, in CONVEX ANALYSIS AND OPTIMIZATION, Aubin J.-P. & Vinter R., edts, Pitman, 93-118

[234] IOFFE A.D. (1984) *Necessary conditions in nonsmooth optimization*, Mathematics of Operations Research, 9

[235] IOFFE A.D. (1983) *Single-valued representation of set-valued mappings II ; Application to differential inclusions*, SIAM J. Control and Optimization, 21 (4), 641-651

[236] IOFFE A.D. (1986) *On the theory of subdifferential*, Mathematics for Optimization/ Mathematical studies series, North-Holland

[237] IOFFE A.D. (1986) *Approximate subdifferentials and applications II*, Mathematika, 33, 111-128

[238] IOFFE A.D. (1987) *On the local surjection property*, J. of Nonlinear Analysis, 11, 565-592

[239] IOFFE A.D. (to appear) *Approximate subdifferential of nonconvex functions*, Trans. Am. Math. Soc.

[240] ISIDORI A. (1985) NONLINEAR CONTROL SYSTEMS: AN INTRODUCTION, Springer-Verlag Lecture Notes in Control and Information Sciences, Vol.72

[241] JACOBS M.Q. (1968) *Measurable multivalued mappings and Lusin's theorem*, Trans. Am. Math. Soc., 134, 471-481

[242] JOLY J.-L. (1973) *Une famille de topologies sur l'ensemble des fonctions convexes pour lesquelles la bipolatité est continue*, J. Math. Pures et Appl., 52, 421-441

[243] KAKUTANI S. (1941) *A generalization of Brouwer's fixed point theorem*, Duke Math J., 8, 457-459

[244] KING A.J. & ROCKAFELLAR R.T. (to appear) *Asymptotic Theory for Solutions in Stochastic Programming and Generalized M-estimation*, Preprint

[245] KING A.J. & ROCKAFELLAR R.T. (to appear) *Sensitivity Analysis for Nonsmooth Generalized Equations*, Preprint

[246] KING A.J. & WETS R.J-B (1988) *Epi-consistency of convex stochastic programs*, IIASA

[247] KING A.J. (1986) *Asymptotic Behavior of Solutions in Stochastic Optimization*, PhD. Thesis, University of Washington

[248] KLEE V. & OLECH C. (1967) *Characterizations of a class of convex sets*, Math. Scand., 20 (2), 290-296

[249] KLEIN F. (1936) VORLESUNGEN ÜBER DIE ENTEWICKLUNG DER MATHEMATIK IM 19. JAHRHUNDERT, Springer-Verlag

[250] KRASNOSELSKI M.A. & ZABREIKO P.P (1975) GEOMETRICAL METHODS OF NONLINEAR ANALYSIS (IN RUSSIAN), Nauka, Moscow

[251] KRASNOSELSKI M.A. (1964) POSITIVE SOLUTIONS OF OPERATOR EQUATIONS, P. Noordhoff Ltd., The Netherlands

[252] KRIVAN V. & VRKOC I. (to appear) *Continuous and Lipschitzean Selections from Measurable Set-Valued Maps. Application to Viability Theory*, Preprint

[253] KURATOWSKI K. & RYLL-NARDZEWSKI C. (1965) *A general theorem on selectors*, Bull. Acad. Pol. Sc., 13, 397-403

[254] KURATOWSKI K. (1932) *Les fonctions semi-continues dans l'espace des ensembles fermés*, Fund. Math. 18, 148-159

[255] KURATOWSKI K. (1958) TOPOLOGIE, VOLS. 1 AND 2 4TH. ED. CORRECTED, Panstowowe Wyd Nauk, Warszawa. (Academic Press, New York, 1966)

[256] KURATOWSKI K. (1980) A HALF CENTURY OF POLISH MATHEMATICS, Pergamon Press

[257] KURZHANSKI A.B. & FILIPPOVA T.F. (1989) *On the Set-Valued Calculus in Problems of Viability and Control for Dynamic Processes: the Evolution Equation*, in ANALYSE NON LINÉAIRE, Attouch, Aubin, Clarke, Ekeland Eds., Gauthier-Villars & C.R.M. Université de Montréal, 339-364

[258] KURZHANSKI A.B. (1977) CONTROL AND OBSERVATION UNDER CONDITIONS OF UNCERTAINTY, Nauka

[259] LANCASTER K.J. (1965) *The theory of Qualitative Linear Systems*, Econometrica, 33, 395-408

[260] LANCASTER K.J. (1966) *The Solution of Qualitative Comparative Static Problems*, Quarterly Journal of Economics, 80, 278-295

[261] LASOTA A. & MACKEY M.C. (1985) PROBABILISTIC PROPERTIES OF DETERMINISTIC SYSTEMS, Cambridge University Press

[262] LASOTA A. & OPIAL Z. (1968) *Fixed point theorems for multivalued mappings and optimal control problems*, Bull. Acad. Pol. Sc., 16, 645-649

[263] LASOTA A. & PIANIGIANI G. (1977) *Invariant Measures on Topological Spaces*, Bollettino U.M.I, (5) 14B, 592-603

[264] LE DONNE A. & MARCHI M.V. (1980) *Representation of Lipschitzean compact convex valued mappings*, Rend. Acc. Naz. Lincei, 68, 278-280

[265] LEACH S.B. (1961) *A note on inverse function theorems*, Proc. Am. Math. Soc., 12, 694-697

[266] LEBOURG G. (in preparation) *Thesis*, Université de Paris-Dauphine

[267] LECHICKI A., LEVI S. (1986) *Wijsman convergence in the hyperspace of a metric space*, Preprint

[268] LEESE S.J. (1974) *Measurable selections in normed spaces*, Proc. Edinburgh Math. Soc., 19, 147-150

[269] LEESE S.J. (1974) *Multifunctions of Suslin type*, Bull. Austr. Math. Soc. II, 395-411

[270] LIONS J.-L. (1969) QUELQUES MÉTHODES DE RÉSOLUTION DE PROBLÈMES NONLINEAIRES, Dunod, Paris

[271] LIONS P.-L. (1982) GENERALIZED SOLUTIONS OF HAMILTON-JACOBI EQUATIONS, Pitman, Boston

[272] LIUSTERNIK L.A. & SOBOLEV V.I. (1965) ELEMENTY FUNKTSIONALNOGO ANALIZA,, Second ed. Nauka Moskow

[273] LIUSTERNIK L.A. (1934) *Conditional extrema of functionals*, Mat. Sb., 41, 390-401

[274] LIUSTERNIK L.A. (1947) *The topology of the Calculus of Variations in the Large*, (English trans. Vol. 16, American Mathematical Society, 1966)

[275] LOJASIEWICZ S. Jr (1985) *Some theorems of Scorza-Dragoni type for multifunctions with applications to the problem of existence of solutions for differential multivalued equations*, Mathematical Control Theory, 625-643

[276] LOJASIEWICZ S.Jr (to appear) *Parametrizations of convex sets*, J. Approx. Theory

[277] MAHANEY M. (1973) THE MATHEMATICAL CAREER OF PIERRE DE FERMAT, 1601-1665, Princeton University Press

[278] MANGASARIAN O.L. (1969) *Nonlinear Programming*, McGraw-Hill, New York

[279] MANGASARIAN O.L., & FROMOVITZ S. (1967) *The Fritz John necessary optimality conditions in the presence of equality and inequality constraints*, J. of Math. Analysis and Appl., 17, 37-47

[280] MARCHAUD H. (1934) *Sur les champs de demi-cônes et les équations différentielles du premier ordre*, Bull. Sc. Math., 62, 1-38

[281] MARCHAUD H. (1936) *Sur les champs continus de demi-cônes convexes et leurs intégrales*, Compos. Math. 1, 89-127

[282] MARTELLI M. V & VIGNOLI A. (1974) *On differentiability of multivalued maps*, Boll. Un. Mat. Stat., 10, 701-712

[283] MARTIN D.H., GARDNER J.R. & WATKINS G.G. (1981) *Indicating Cones and the Intersection Principle for Tangential Approximants in Abstract Multiplier Rules*, Journal of Optimization Theory and Applications, 33, 515-537

[284] McLEOD R.M. (1965) *Mean value theorems for vector-valued functions*, Proc. Edinburgh Math Soc., 14 (2), 197-209

[285] MICHAEL E. (1951) *Topologies on spaces of subsets*, Trans. AMS, 71, 151-182

[286] MICHAEL E. (1956) *Continuous selections I*, Annals of Math., 63, 361-381

[287] MICHAEL E. (1956) *Continuous selections II*, Annals of Math., 64, 562-580

[288] MICHAEL E. (1957) *Continuous selections III*, Annals of Math., 65, 375-390

[289] MICHAEL E. (1959) *A theorem on semicontinuous set valued functions*, Duke Math. J., 26, 647-651

[290] MINTY G. (1967) *A Theorem on Monotone Sets in Hilbert Spaces*, J. Math. Anal. Appl. 11, 434-439

[291] MIRICA S., STAICU V. & ANGELESCU N. (1987) *Fréchet Semi-Differentials of Real-Valued Functions*, Preprint Università di Firenze

[292] MIRICA S., STAICU V. & ANGELESCU N. (1988) *Equivalent definition and basic properties of Fréchet semi-differentials*, SISSA

[293] MIRICA S. (1980) *A note on the generalized differentiability of mappings*, Nonlinear Anal., 4, 567-575

[294] MORDUKHOVITCH B. (1988) APPROXIMATION METHODS IN PROBLEMS OF OPTIMIZATION AND CONTROL, Nauka

[295] MOREAU J.-J. & VALADIER M. (1987) *A Chain Rule Involving Vector Functions of Bounded Variation*, Journal of Functional Analysis, 74 (2), 333-345

[296] MOREAU J.-J. (1966) FONCTIONNELLES CONVEXES, Séminaire du Collège de France

[297] MOSCO U. (1969) *Convergence of convex sets and of solutions to variational inequalitites*, Advances in Math. 3, 510-585

[298] MOSCO U. (1971) *On the continuity of the Young-Fenchel Transformation*, J. Math. Anal. Appl., 35, 518-535

[299] NDOUTOUME J.L (1987) *Analyse Variationnelle et Applications en Mécanique et Contrôle*, Publication AVAMAC, 1,

[300] NEUSTADT L.W. (1976) OPTIMIZATION, Princeton Univ. Press, Princeton N.J.

[301] NEVEU J. (1964) BASES MATHÉMATIQUES DU CACUL DES PROBABILITÉS, Masson

[302] NG K. F. (to appear) *An inverse function theorem and its application in optimization theory*, Preprint, University of Hong Kong

[303] NIKOLSKII M. S. (1983) *On the Alternate Integral of Pontriagin*, Math. USSR Sbornik, 44, 125-132

[304] NIKOLSKII M. S. (1987) *On the Lower Alternating Integral of Pontriagin in Linear Differential Games of Pusrsuit*, Math. USSR Sbornik, 56, 33-47

[305] OLECH C. (1966) *Extremal Solutions of a Control System*, J. Diff. Eq., 2, 74-101

[306] OLECH C. (1965) *A note concerning set-valued measurable functions*, Bull. Acad. Pol. Sc., 13, 317-321

[307] OLECH C. (1967) *Lexicographical order, range of integral and bang-bang principle*, Mathematical Theory of Control, Academic Press, New York, 35-45

[308] OLECH C. (1968) *Approximation of set-valued functions by continuous functions*, Colloq. Math., 19, 285-293

[309] OLECH C. (1969) *Existence theorems for optimal control problems involving multiple integrals*, J. Diff. Eq., 6, 512-526

[310] OLECH C. (1969) *Existence theorems for optimal control problems with vector-valued cost functions*, Trans. Am. Math. Soc., 136, 157-180

[311] OLECH C. (1975) *Existence of solutions of nonconvex orientor fields*, Boll. Un. Mat. Ital., 11 (4), 189-197

[312] OLECH C. (1976) *Existence Theory In Optimal Control*, In CONTROL THEORY AND TOPICS IN FUNCTIONAL ANALYSIS, International Atom. Energy. Agency, 291-328

[313] OLECH C. (1987) *Lectures on the integration of set-valued functions*, Preprint SISSA

[314] ORNELAS A. (1988) *Parametrization of Caratheodory Multifunctions*, SISSA

[315] ORNELAS A. (1989) *Thesis, SISSA*

[316] OUTRATA J.V. & SCHINDLER Z. (to appear) *On using of bundle methods in nondifferentiable optimal control problems*, Preprint

[317] OUTRATA J.V. (1983) *On a Class of Nonsmooth Optimal Control Problems*, Appl. Math. Optim., 10, 287-306

[318] PAINLEVE P. (1909) C.R.A.S. Paris,148, 1156

[319] PAPAGEORGIOU N.S. (1989) *Convergence and representation theorems for set valued random variables*, Stochastic analysis and applications, 7(2), 187-210

[320] PAZY A. (1983) *Semi-groups of Linear Operators and Applications*, Springer Verlag, Berlin

[321] PCHENITCHNY B.W. (1980) CONVEX ANALYSIS AND EXTREMAL PROBLEMS (IN RUSSIAN), Nauka Moscow

[322] PCHENITCHNY B.W. (1986) *Implicit function theorems for multi-valued mappings*, Preprint IIASA

[323] PEANO G. (1892) *Sur la définition de la dérivée*, Mathesis,II, 2, 12-14

[324] PEANO G. (1885) *Sull'integrabilita delle equazioni differenziali di primo ordine*, Acad. Sci. Torino, 21, 677-685

[325] PENOT J.-P. (1981) *A characterization of tangential regularity*, Nonlinear Anal. T.A.M., 5, 625-643

[326] PETCHERSKAJA N.A. (1980) *On differentiability of set-valued maps (in Russian)*, Vestnik of Leningrad University, 115-117

[327] PIANIGIANI G. (1975) *Existence of solutions for ordinary differential equations in Banach spaces*, Bull. Acad. Polon. Sci., Ser. sci. math. astron. phys., 23, 853-857

[328] PIANIGIANI G. (1977) *On the Fundamental Theory of Multivalued Differential Equations*, Journal of Differential Equations, 25, 30-38

[329] PLIS A. (1961) *Remark on measurable set-valued functions*, Bull. Ac. Pol. Sc., 9-12, 857-859

[330] PLIS A. (1965) *Measurable orientor fields*, Bull. Acad. Pol. Sc., 13, 565-569

[331] POLOVINKIN E. S. & SMIRNOV G.V. (1986) *Differentiation of Multivalued Mappings and Properties of Solutions of Differential Inclusions*, Soviet Math. Dokl., 33, 662-666

[332] POLOVINKIN E. S. (1986) *Thesis*

[333] PONTRIAGIN L. S. (1980) *Linear Differential Games of Pursuit*, Mat. Sb., 112, 307-380

[334] QUIRK J. & RUPPERT R. (1965) *Qualitative Economics and the Stability of Equilibrium*, Review of economic studies, 32, 311-326

[335] RADULESCU S. & RADULESCU M. (to appear) *Local inversion theorems without assuming continuous differentiability*, Preprint

[336] REICH S. (1972) *Fixed Point in Locally Convex Spaces*, Math. Z., 125, 17-31

[337] REICH S. (1978) *Approximate Selections, Best Approximations, Fixed Point, and Invariant Sets*, Journal of Mathematical Analysis and Applications, 62, 104-113

[338] REICH S. (1979) *Fixed Point Theorems for Set-Valued Mappings*, Journal of Mathematical Analysis and Applications, 69, 353-358

[339] REICH S. (1983) *Convergence, Resolvent Consistency, and the Fixed Point Property for Nonexpansive Mappings*, Contemporary Mathematics, 18, 167-174

[340] REICH S. (1983) *The almost fixed point property for nonexpansive mappings*, Proceedings of the American Mathematical society, 88 (1), 44-46

[341] RICCERI B. (1984) *Analisi Matematica - On multifunctions with convex graph*, Rendiconti della Classe di Scienze fisiche, matematiche e naturali, 127, 64-70

[342] RICCERI B. (1985) *Analyse Fonctionnelle - Un théorème d'existence pour les inéquations variationnelles*, C. R. Acad. Sc; Paris, 301, 885-888

[343] RICCERI B. (1985) *Fixed Points of Lower Semicontinuous Multifunctions and Applications: Alternative and Minimax Theorems*, Memorie di Matematica, 9, 331-338

[344] RICCERI B. (to appear) *Existence Theorems for Nonlinear Problems*, Preprint

[345] ROBERT R. (1974) *Convergence de Fonctionnelles Convexes*, Journal of Mathematical Analysis and Applications, 45 # 3, 533-555

[346] ROBINSON S.M. (1972) *Normed Convex Processes*, Transactions of the American Mathematical Society, 174

[347] ROBINSON S.M. (1974) *Perturbed Kuhn-Tucker points and rates of convergence for a class of nonlinear-programming algorithms*, Math. Programming, 7, 1-16

[348] ROBINSON S.M. (1976) *Regularity and stability for convex multivalued functions*, Math. Op. Res., 1, 130-143

[349] ROBINSON S.M. (1976) *Stability theory for systems of inequalities, part II: differentiable nonlinear systems*, SIAM J. on Numerical Analysis,13, 497-513

[350] ROBINSON S.M. (1979) *Generalized equations and their solutions, Part 1*, Math. Prog. Stud., 10, 128-141

[351] ROBINSON S.M. (1980) *Analysis and computation of fixed points*, Academic Press, New York

[352] ROBINSON S.M. (1982) *Generalized equations and their solutions, Part II: Applications to nonlinear programming*, Mathematical Programming Study, 19, 200-221

[353] ROBINSON S.M. (1988) *An implicit-function theorem for B-Differentiable functions*, IIASA WP

[354] ROCKAFELLAR R.T. & WETS R. (1983) *Variational Systems, an Introduction*, In MULTIFUNCTIONS AND INTEGRANDS:STOCHASTIC ANALYSIS, APPROXIMATION AND OPTIMIZATION ed. Salinetti G., Springer-Verlag, Lecture Notes in Math. # 1091, 1-54

[355] ROCKAFELLAR R.T. & WETS R. (in preparation) SET-VALUED ANALYSIS AND SUBDIFFERENTIAL CALCULUS (MONOGRAPH)

[356] ROCKAFELLAR R.T. (1967) MONOTONE PROCESSES OF CONVEX AND CONCAVE TYPE, Mem. of AMS # 77

[357] ROCKAFELLAR R.T. (1969) *Measurable dependance of convex sets and functions on parameters*, J. Math. Anal. Appl., 28, 4-25

[358] ROCKAFELLAR R.T. (1970) CONVEX ANALYSIS, Princeton University Press

[359] ROCKAFELLAR R.T. (1970) *On the maximal monotonicity of subdifferential mappings*, Pacific J. Math., 33, 209-216

[360] ROCKAFELLAR R.T. (1974) *Convex algebra and duality in dynamic models of production*, Mathematical Models in Economics, Los (Ed), North Holland Amsterdam

[361] ROCKAFELLAR R.T. (1979) *Clarke's tangent cones and the boundaries of closed sets in R^n*, Nonlinear Anal. Theor. Math. Appl., 3, 145-154

[362] ROCKAFELLAR R.T. (1979) *Directionally Lipschitzian functions and subdifferential calculus*, Proc. London Math. Soc., 39, 331-335

[363] ROCKAFELLAR R.T. (1979) LA THÉORIE DES SOUS-GRADIENTS ET SES APPLICATIONS À L'OPTIMISATION: FONCTIONS CONVEXES ET NON CONVEXES, Collection Chaire Aisentadt, Presses de l'Université de Montreal, 130

[364] ROCKAFELLAR R.T. (1980) *Generalized directional derivatives and subgradients of nonconvex functions*, Can. J. Math., 32, 157-180

[365] ROCKAFELLAR R.T. (1981) *Proximal subgradients, marginal values, and augmented Lagrangians in nonconvex optimization*, Math. Oper. Res., 6, 427-437

[366] ROCKAFELLAR R.T. (1981) THE THEORY OF SUBGRADIENTS AND ITS APPLICATIONS TO PROBLEMS OF OPTIMIZATION OF CONVEX AND NONCONVEX FUNCTIONS, Helderman Verlag, W. Berlin 1981

[367] ROCKAFELLAR R.T. (1982) *Directional Differentiability of the Optimal Value Function in a Nonlinear Programming Problem*, Differentiability of Optimal Values, 213-226

[368] ROCKAFELLAR R.T. (1982) *Lagrange multipliers and subderivatives of optimal value functions in nonlinear programming*, Mathematical Programming Study, 17, 28-66

[369] ROCKAFELLAR R.T. (1983) *Marginal Values and Second-Order Necessary Conditions for Optimality*, Math. Program. Studies, 26, 245-286

[370] ROCKAFELLAR R.T. (1985) *Lipschitzian Properties of Multifunctions*, J. Nonlinear Anal. T.M.A., 9, 867-885

[371] ROCKAFELLAR R.T. (1985) *Maximal Monotone Relations and the Second Derivative of Nonsmooth Functions*, Ann. Inst. H. Poincaré, Analyse Non Linéaire, 2, 167-184

[372] ROCKAFELLAR R.T. (1986) *Lipschitzian stability in optimization: the role of nonsmooth analysis*, IIASA WP-8646

[373] ROCKAFELLAR R.T. (1987) *Pseudo-differentiability of set-valued mappings and its applications in optimization*, (Preprint)

[374] ROCKAFELLAR R.T. (1988) *First and second-order pseudo-differentiability in nonlinear programming*, Trans. AMS., 307, 75-108

[375] ROCKAFELLAR R.T. (1989) *Proto-Differentiability of set-valued mappings and its applications in optimization*, in ANALYSE NON LINÉAIRE, Attouch, Aubin, Clarke, Ekeland Eds., Gauthier-Villars & C.R.M. Université de Montréal, 449-482

[376] ROCKAFELLAR R.T. (to appear) *Extensions of subgradient calculus with applications to optimization*,

[377] ROCKAFELLAR R.T. (to appear) *Second-order optimality conditions in nonlinear programming obtained by way of pseudo-derivatives*, Math. Op. Res.

[378] RUBINOV A.M. (1985) *The contingent derivative of a multivalued mapping and differentiability of the maximum under connected constraints*, Sibirsk. Math, Zh., 26, 147-155

[379] SALINETTI G. & WETS R. (1977) *On the relation between two types of convergence for convex functions*, J. Mathematical Analysis and Applications, 60, 211-226

[380] SALINETTI G. & WETS R. (1979) *Convergence of sequences of closed sets*, Topology Proceed., 4, 149-158

[381] SALINETTI G. & WETS R. (1979) *On the convergence of sequences of convex sets in finite dimensions*, SIAM Review, 21, 18-33

[382] SALINETTI G. & WETS R. (1981) *On the convergence of closed-valued measurable multifunctions*, Trans. AMS, 266, 275-289

[383] SHEPARD G. C. (1964) *Approximation Problems for Convex Polyhedra*, Mathematika, 11, 9-18

[384] SHEPARD G. C. (1966) *The Steiner Point of a Convex Polytope*, Canad. J. Math., 18, 1294-1300

[385] SHEPARD G. C. (1968) *A Uniqueness Theorem for the Steiner Point of a Convex Region*, J. London Math. Soc., 43, 439-444

[386] SHI SHUZHONG & CHANG KUNG-CHING (1987) *A local minimax theorem without compactness*, in NONLINEAR AND CONVEX ANALYSIS Proc. in Honor of Ky Fan, Ed. B. L. Lin and S. Simons, Lecture Notes in Pure and Applied Mathematics V. 107, Dekker, 211-234

[387] SHI SHUZHONG (1980) *Remarques sur le gradient généralisé*, C.R. Acad. Série Paris 291, 443-446

[388] SHI SHUZHONG (1981) *Semi-continuités génériques de multi-applications*, C.R. Acad. Sci. Paris 293, Série A, 27-29

[389] SHI SHUZHONG (1982) *Remarques sur le gradient généralisé*, J. Math. Pures et Appl., 61, 301-310

[390] SHI SHUZHONG (1985) *Ekeland's variational principle and the mountain pass lemma*, Acta Mathematica Sinica, New Series, 1:4, 348-355

[391] SHI SHUZHONG (1987) *Théorèmes de Choquet et Analyse non régulière*, C. R. Acad. Sc. Paris, 305, Série I, 41-44

[392] SHI SHUZHONG (1988) *Choquet theorem and nonsmooth analysis*, J. Math. Pures et Appl., 67, 441-432

[393] SIEGEL J. (1977) *A new proof of Caristi's fixed point theorem*, Proc. Am. Math. Soc., 66, 54-56

[394] SIMONS S. (1986) *An Existence Theorem for Quasiconcave Functions with Applications*, Nonlinear Analysis, Theory, Methods & Applications, 10, 1133-1152

[395] SIMONS S. (1986) *Cyclical coincidences of multivalued maps*, J. Math. Soc. Japan, 38, 516-525

[396] SIMONS S. (1986) *On a fixed point theorem of Cellina*, Lincei Rend. fis. 130, 9-10

[397] SIMONS S. (1986) *Two-Function Minimax Theorem and variational inequalities for functions on compact and noncompact sets*, Proceedings of Symposia in Pure Mathematics, A.M.S. 45 (Part 2), 377-392

[398] SIMONS S. (1987) *The continuity of inf sup, with applications*, Arch. Math., 48, 426-437

[399] SIMONS S. (to appear) *Abstract Kuhn-Tucker theorems*, Preprint University of California

[400] SPINGARN J.E. (1983) *Partial Inverse of a Monotone Operator*, Applied Mathematics and Optimization, 10, 247-265

[401] STAICU V. (1988) *Generalized Derivatives for C^1-- parametrized Multifunctions*, Preprint Universita degli studi di Firenze

[402] STEINER J. (1840) *Von dem Krümmungsschwerpunkte ebener Curven*, J. Reine Angew. Math., 21, 33-63, 101-122

[403] STUDNIARSKI M. (1985) *Mean Value Theorem for Functions Possessing First Order Convex Approximations. Applications in Optimization Theory*, Analysis und ihre Anwendungen, 4, 125-132

[404] STUDNIARSKI M. (1985) *Mean Value Theorems and Sufficient Optimality Conditions for Nonsmooth Functions*, Journal of Math. Anal. and Appl., 111, 313-326

[405] SUBBOTIN A.I. & SUBBOTINA N. N. (1983) *Differentiability properties of the value function of a differential game with integral terminal costs*, Problems of Control and Information Theory, 12, 153-166

[406] TALLOS P. (to appear) *Viability problems for nonautonomous differential inclusions*, SIAM J. on Control and Optimization

[407] THIBAULT L. (1982) *Subdifferentials of nonconvex vector-valued functions*, J. Math. Anal. Appl., 86, 319-344

[408] TOLAND J.F. (1978) *Duality in nonconvex optimization*, J. Math. Anal. Appl., 66, 399-315

[409] TOLAND J.F. (1979) *A duality principle for nonconvex optimization and the calculus of variations*, Arch. Rat. Mech. Anal., 71, 41-61

[410] TREIMAN J.S. (1983) *A new characterization of Clarke's tangent cone and its applications to subgradient analysis and optimization*, PhD. Thesis, University of Washington

[411] TREIMAN J.S. (1983) *Characterization of Clarke's tangent and normal cones in finite and infinite dimensions*, Nonlinear Analysis, Theory, Methods et Applications, 7, 771-783

[412] TREIMAN J.S. (1987) *Too Many Convex Tangent Cones*, Can. J. Math.

[413] TREIMAN J.S. (to appear) *Finite Dimensional Optimality Condition: B-Gradients*, Journal of Optimization Theory and Applications

[414] TREIMAN J.S. (to appear) *Shrinking Generalized Gradients*, Nonlinear Analysis, Theory, Methods et Applications

[415] UHL J. J. (1969) *The range of a Vector-Valued Measure*, Proc. A.M.S., 23, 158-163

[416] URSESCU C. (1973) *Sur une généralisation de la méthode de différentiabilité*, Ac. Naz. Lincei, 8, 199-204

[417] URSESCU C. (1975) *Multifunctions with closed convex graph*, Czechoslovakia Mathematics J.,25, 438-441

[418] URYSOHN P. (1927) Verhandtl. Akad. Amsterdam, 13, 29

[419] VAINBERG S. (1956) VARIATIONAL METHODS FOR THE STUDY OF NONLINEAR OPERATORS, Nauka

[420] VALADIER M. (1971) *Multiaplications mesurables à valeurs convexes compactes*, J. Math. Pures et Appl., 50, 265-297

[421] VALADIER M. (1987) *Fonctions et opérateurs sur les mesures*, C.R. Ac. Sc., 304, 135-137

[422] VALADIER M. (1988) *Approximation Lipschitzienne par l'interieur d'une multifonction S.C.I.*, Seminaire d'Analyse Convexe, Montpellier 1987, Expose n 11

[423] VALADIER M. (1989) *Entraînement unilateral, lignes de descente, fonctions lipschitziennes non pathologiques*, Comptes-Rendus de l'Académie des Sciences, PARIS, Série 1, 308, 241-244

[424] VITALE R.A. (1985) *The Steiner point in infinite dimensions*, Israel J. Math., 52, 245-250

[425] VLACH M. (1983) *Closures and Neighborhoods Properties Induced by Tangential Approximation*, Preprint Universität Bielefeld

[426] WAGNER D.M. (1977) *Survey of measurable selection theorems*, SIAM J. Contr. Opt., 15, 859-903

[427] WALKUP D.W. & WETS R. (1967) *Continuity of some convex-cone-valued mapping*, Proc. Amer. Math. Soc., 18, 229-235

[428] WARD D.E. & BORWEIN J.M. (1987) *Nonsmooth Calculus in Finite Dimensions*, SIAM J. Control and Optimization, 25 (5), 1312-1340

[429] WARD D.E. (to appear) *Convex Subcones of the Contingent Cones in Nonsmooth Calculus and Optimization*, Preprint

[430] WARD D.E. (to appear) *Isotone Tangent Cones and Nonsmooth Optimization*, Preprint

[431] WARGA J. (1972) OPTIMAL CONTROL OF DIFFERENTIAL AND FUNCTIONAL EQUATIONS, Academic Press, New York

[432] WARGA J. (1976) *Derivate containers, inverse functions and controllability*, Calculus of Variations and Control Theory, Academic Press, 13-46

[433] WARGA J. (1978) *A second-Order Condition that Strengthens Pontryagin's Maximum Principle*, Journal of Differential Equations, 28, 284-307

[434] WARGA J. (1978) *An implicit function theorem without differentiability*, Proc. Am. Math. Soc., 69, 65-69

[435] WARGA J. (1978) *Controllability and a multiplier rule for nondifferentiable optimization problems*, SIAM J. Control Optim., 16, 803-812

[436] WARGA J. (1983) *Optimization and controllability without differentiability assumptions*, SIAM J. Control and Optimization, 21(6), 837-855

[437] WARGA J. (1984) *Iterative Optimization with Equality Constraints*, Mathematics of Operations Research, 9, 592-605

[438] WARGA J. (1984) *Second order necessary conditions in optimization*, SIAM J. Control and Optimization, 22, 524-528

[439] WARGA J. (1986) *Higher order conditions with and without Lagrange multipliers*, SIAM J. Control and Optimization, 24, 715-730

[440] WARGA J. (to appear) *Higher Order Conditions for Conical Controllability*, Preprint Northeastern University

[441] WARGA J. (to appear) *Homeomorphisms and local C^1 approximations*, Preprint

[442] WAŻEWSKI T. (1961) *Sur la semicontinuité inférieure du "Tendeur" d'un ensemble compact variant d'une façon continue*, Bull. Acad. Pol. Sc., 9, 869-872

[443] WAŻEWSKI T. (1961) *Sur une condition équivalente à l'équation au contingent*, Bull. Acad. Pol. Sc., 9, 865-867

[444] WAŻEWSKI T. (1961) *Systèmes de commande et équations au contingent*, Bull. Acad. Pol. Sc., 9, 151-155

[445] WAŻEWSKI T. (1962) *Sur une généralization de la notion des solutions d'une équation au contingent*, Bull. Acad. Pol. Sc., 10, 11-15

[446] WAŻEWSKI T. (1963) *On an optimal control problem*, Proc. Conference DIFFERENTIAL EQUATIONS AND THEIR APPLICATIONS, 229-242 -Prague (1962)

[447] WETS R. (1977) *On the Convergence of Random Sets*, In CONVEX ANALYSIS AND ITS APPLICATIONS, Aubin Ed., Springer-Verlag, Lecture Notes in Economics and Mathematical Systems # 144, 191-206

[448] WETS R. (1980) *Convergence of convex functions, variational inequalities and convex optimization problems*, In VARIATIONAL INEQUALITIES AND COMPLEMENTARITY PROBLEMS, R.Cottle, F. Gianesi & J.L. Lions Ed., Springer-Verlag, 405-419

[449] WETS R. (1982) *A formula for the level sets of epi-limits and some applications*, In MATHEMATICAL THEORIES OF OPTIMIZATION Eds. Cecconi P. & Zolezzi, Springer-Verlag, Lecture Notes in Mathematics #979, 256-268

[450] WETS R. (1982) *Convergence of sequences of closed functions*, In PROCEED. SYMPOSIUM ON MAHTEMATICAL PROGRAMMING AND DATA PERTURBATIONS, Eds. Fiacco A., 16-27

[451] WETS R. (1984) *On a compactness theorem for epiconvergent sequences of functions*, In MATHEMATICAL PROGRAMMING, Eds. COTTLE, LEMANSON & KORTE, North-Holland, 347-351

[452] WETS R. (1986) *An elementary proof of Minkowski's theorem for system of linear inequalities*, Preprint

[453] WIJSMAN R.A. (1964) *Convergence of sequences of convex sets, cones and functions*, Bulletin of the Americam Mathematical Society, 70, 186-188

[454] WIJSMAN R.A. (1966) *Convergence of sequences of convex sets, cones and functions II*, Bulletin of the Americam Mathematical Society, 123, 32-45

[455] WILSON W. A. (1926) Amer. Journ. Math. 48, 165

[456] WILSON W. A. (1930) *On the Phragmén-Brouwer Theorem*, Bull. Amer. Math. Soc. 111-114

[457] YOUNG L.C. (1981) MATHEMATICIANS AND THEIR TIMES, Mathematical Studies # 48, North-Holland

[458] ZAJICEK L. (1981) *On Approximate Dini Derivates and One-Sided Approximate Derivatives of Arbitrary Functions*, Commentationes Matematicae Universitatis Carolinae, 549-560

[459] ZAJICEK L. (1981) *On the Symmetry of Dini Derivates of Arbitrary Functions*, Commentationes Mathematicae Universitatis Carolinae, 195-209

[460] ZAJICEK L. (1983) *Differentiability of the Distance Function and Points of Multi-Valuedness of the Metric Projection in Banach Space*, Czechoslovak Mathematical Journal, 33, 292-308

[461] ZAJICEK L. (1984) *A Generalization of an Ekeland-Lebourg theorem and the Differentiability of Distance Functions*, Rendiconti del Circolo Matematico di Palermo, 3, 403-410

[462] ZARANKIEWICZ C. (1927) Fund. Math. 9, 124

[463] ZARANTONELLO E.H. (1975) *Conical Spectral Theory*, Topics in Functional Analysis, 37-116

[464] ZARANTONELLO E.H. (1982) *Cyclical monotonicity of maximal monotone step operators*, Bol. Soc. Bras. Mat., 13, 85-91

[465] ZAREMBA S.C. (1936) *Sur les équations au paratingent*, Bull. Sc. Math., 60, 139-160

[466] ZEIDLER E. (1984) NONLINEAR FUNCTIONAL ANALYSIS AND ITS APPLICATIONS, Springer-Verlag

[467] ZORETTI L. (1905) *Sur les fonctions analytiques uniformes qui possèdent un ensemble discontinu parfait de points singuliers*, Journ. de Math. Pures et Appl., 1, 1-51

[468] ZORETTI L. (1909) *Un théorème de la théorie des ensembles*, Bull. Soc. Math. France, 37, 116-119

Index