

Université des Antilles

Faculté des Sciences Exactes et Naturelles

ED 589 - Milieu insulaire tropical à risques :
Protection, Valorisation, Santé et Développement

Laboratoire de Mathématiques Informatique et Applications (LAMIA)

Thèse de Doctorat en Mathématiques

Présentée et soutenue par

Olguine YACINTHE

le 11 décembre 2020

Processus convexes et résolution d'inclusions variationnelles

---

Sous la direction de :

M. Alain PIÉTRUS et M. Michaël GAYDU

Devant le jury composé de :

M. Julian REVALSKI,	Professeur, Académie Bulgare des Sciences, Sofia,	Rapporteur
M. Didier AUSSEL,	Professeur, Université de Perpignan,	Rapporteur
Mme Séverine BERNARD,	Maître de conférences, Université des Antilles,	Examinatrice
M. Mathieu BONNEAU,	Chargé de Recherche, INRA Antilles-Guyane,	Examinateur
M. Michaël GAYDU,	Maître de Conférences, Université des Antilles,	Co-encadrant de thèse
M. Alain PIÉTRUS,	Professeur, Université des Antilles,	Directeur de thèse

Remerciements

Dieu, je te remercie de m'avoir donné l'humilité, la patience et surtout le courage et le sérieux pour mener à bien ce travail de recherche.
Par ailleurs, ce dernier n'aurait pas pu voir le jour sans le soutien de nombreuses personnes auxquelles je tiens à témoigner mes sincères remerciements.

Ma gratitude s'adresse au Professeur Alain PIÉTRUS, mon Directeur de Thèse, pour avoir accepté de diriger et d'encadrer ce travail réalisé sans financement. Je suis également très reconnaissant de la contribution apportée à l'enseignement des mathématiques à l'École Normale Supérieure de l'Université d'État d'Haïti.

Un remerciement spécial à M. Michaël GAYDU, mon Co-encadrant de Thèse, pour son accompagnement et ses précieux conseils dont la mise en pratique permet la réalisation du manuscrit.

Je tiens également à remercier M. Julian REVALSKI et M. Didier AUSSEL pour avoir accepté d'être les rapporteurs de cette thèse, ainsi que Mme Séverine BERNARD-ANDOUZE et M. Mathieu BONNEAU pour leur participation dans le jury et le temps qu'ils ont alloué à l'examen du manuscrit.

Il serait injuste de ne pas remercier M. Silvère Paul NUIRO et M. Raphaël PASQUIER, sans lesquels je ne saurais présenter les résutats numériques associés aux résultats théoriques. J'en suis très reconnaissant.

Ma reconnaissance s'exprime ágalement aux Professeurs Pascal POULLET, Jacques H. LAMINIE et Jean VAILLANT, ainsi qu'à tous les autres membres du Laboratoire de Mathématiques Informatiques et Applications(LAMIA) notamment Madame Marylène TROUPÉ. Sans passer outre la gentillesse de la secrétaire du Département de Mathématiques et Informatique (DMI) et celle de Mme LAURENT-PEDURAND et Mme REMUS.

Ma gratitude et ma reconnaissance se renouvellent chaque jour envers un Monsieur oeu-

vrant pour l'amélioration de la qualité de l'enseignement supérieur en Haïti à travers des anciens étudiants du Département de Mathématiques de l'École Normale Supérieure (ENS). Il répond au nom de M. Bérard CÉNATUS, membre du Conseil de l'ENS, dont j'en suis un produit résultant. J'en profite également pour remercier quelques enseignants de ladite faculté qui ont contribué d'une façon ou d'une autre dont M. Yvesner MARCELIN, M. Alvarès LOUIS et M. Lesly E. DÉJEAN. La rigueur des raisonnements de ce dernier m'a beaucoup imprégné.

Mes remerciements chaleureux s'adressent également à la Fondation Connaissance et Liberté (FOKAL).

Je ne saurais terminer sans remercier mes parents et quelques proches. D'abord, Ma Grand-Mère, décédée en 2001, qui m'a initíe à l'école où j'ai appris à savourer le goût du savoir et à me connaître face aux adversités de la vie. J'aurais bien aimé que tu sois présente aux côtés de ton petit fils pour apprécier la concrétisation de ton rêve à travers la réalisation de ce travail de recherche qui me vaut le grade de Docteur.
La part inouïe de ma mère ne saurait passer sous silence. Ma mère chérie Immacula JANVIER, je n'ai pas de mots pour saluer ton courage et ta résilience dans la vie puisqu'en partant du néant, tu m'as permis d'arriver là où j'en suis. Cette thèse vous est dédiée !
Mes deux soeurs Régine YACINTHE et Kénide HYACINTHE, vous qui en avez pris part à travers vos propos de reconfort dans les moments de détresse, je ne trouve pas également de mots. Mais, je vous dis simplement Merci.
Enfin, mes sincères remerciements à tous mes proches.

Résumé

L'objectif de cette thèse est d'utiliser le concept de processus convexe pour la résolution d'inclusions variationnelles où la partie multivoque est la somme de l'opposé d'un cône convexe fermé non vide et d'une multiapplication différentiable au sens de Nachi-Penot, à graphe fermé, à valeurs compactes et convexes.
Vu la complexité du problème, nous avons d'abord consacré une bonne partie sur des rappels d'Analyse Variationnelle et Non Lisse et les principaux théorèmes associés.

Notre contribution s'appuie fortement sur la notion de processus convexes normés ainsi qu'un théorème relatif à ceux qui sont perturbés par des opérateurs linéaires continus. Grâce à ce résultat, lorsque les dérivées sont uniformément bornées, nous étendons certaines méthodes itératives classiques via un résultat estimant la distance de deux ensembles non vides par un processus convexe normé.

Dans un premier temps, nous présentons respectivement des extensions de la méthode de Newton et la méthode de Newton non lisse lorsque la partie univoque est lisse et non lisse. Dans ces cas, nous avons montré que celles-ci convergent super-linéairement. D'autre part, nous avons adapté le résultat de convergence associé à la méthode de Newton non lisse au cas où la fonction univoque est hölderienne.

Dans un deuxième temps, lorsque la partie univoque est la somme d'une fonction lisse et d'une fonction lipschitzienne, nous associons une méthode semi-locale de type Newton-sécante ainsi que la vitesse de convergence.

Nous avons également établi que ces méthodes généralisent les travaux antérieurs associés aux cas où la multiapplication se réduit à un cône fixe plus précisément à un cône convexe fermé non vide, ainsi que la préservation du taux de convergence des méthodes classiques.

Mots clés : Inclusion variationnelle, processus convexe, différentiabilité au sens de Nachi-Penot, fonction lisse, fonction semi-lisse, fonction lipschitzienne, fonction hölderienne, méthode de Newton, méthode de Newton-sécante.

Table des matières

Notations et abréviations **vi**

Introduction **1**

1 Quelques rappels **7**
1.1 Généralités sur les applications multivoques 7
1.2 Exemples d'applications multivoques 9
1.3 Processus convexes 10
1.3.1 Processus convexes normés 11
1.3.2 Exemples de processus convexes 12
1.3.3 Perturbation d'un processus convexe normé 15
1.3.4 Estimation de la distance de l'image de deux ensembles par un processus convexe normé 17
1.4 Fonctions Fréchet-différentiables 19
1.5 Différences divisées 20
1.6 Suites majorantes 24
1.7 Fonctions semi-lisses 25
1.7.1 Fonction localement lipschitzienne 25
1.7.2 Fonctions semi-lisses 27
1.7.3 Fonctions semi-lisses d'ordre p 32
1.7.4 Fonctions sous-analytiques 34
1.8 Différentiabilité des applications multivoques 35
1.8.1 Dérivée de Nachi-Penot 37
1.8.2 Quelques dérivées multivoques 40
1.8.3 Un concept de dérivée directionnelle 47
1.8.4 π-différentiabilité de multifonctions 48
1.9 Vitesse de convergence d'une suite 51

2 Résolution d'inclusions variationnelles par des méthodes itératives utilisant les processus convexes **53**

2.1 Extension de la méthode de Newton pour des fonctions non linéaires à valeurs dans un cône 53
2.2 Extension de la méthode Newton perturbée non différentiable pour des fonctions à valeurs dans un cône 56
2.3 Extension de la méthode Newton-sécante pour des fonctions à valeurs dans un cône 58
2.4 Extension de la méthode de Newton semi-lisse pour une fonction à valeurs dans un cône 61

3 Méthode de type Newton pour une multiapplication différentiable contenue dans un cône 65
3.1 Description et convergence de la méthode 65
3.2 Analyse de convergence de la méthode 67
3.2.1 Estimation de l'erreur et vitesse de convergence 71
3.3 Applications numériques 74
3.4 Perturbation de la partie univoque par une fonction non différentiable mais lipschitzienne 76
3.4.1 Description de la méthode 76
3.4.2 Convergence de la méthode 77

4 Méthode de type Newton-sécante pour une application multivoque perturbée contenue dans un cône 81
4.1 Description et convergence de la méthode 81
4.1.1 Description de notre algorithme 82
4.2 Analyse de la convergence de la méthode 84
4.3 Convergence de la méthode 87
4.3.1 Estimation de l'erreur et vitesse de convergence 91

5 Méthode de type Newton semi-lisse pour une application multivoque contenue dans un cône 93
5.1 Mise en contexte 93
5.2 Description et convergence de la méthode 95
5.3 Analyse de la convergence de notre méthode 96
5.3.1 Estimation de l'erreur et vitesse de convergence 100
5.4 Cas où la fonction f est sous-analytique 101
5.5 Cas où la fonction f est perturbée par une fonction g lipschitzienne . . 101
5.6 Remarque : f Fréchet différentiable avec dérivée hölderienne 104

Conclusion et perspectives 108

Bibliograhie 109

Notations et abréviations

int(A) intérieur pour la topologie de la norme d'une partie A d'un espace normé ;

E un $\mathbb{R} - e.v$ espace vectoriel sur le corps des réels $\mathbb{R}$;

$\mathcal{L}(E, F)$ ensemble des applications (opérateurs) linéaires et continues de l'espace normé E dans l'espace normé F ;

$:=$ "est par définition" ;

$Conv(A)$ enveloppe convexe de l'ensemble A ;

$\partial f(v)$ jacobien généralisé de f en v au sens de Clarke ;

$\| \, . \, \|_E$ norme sur l'espace vectoriel E ;

$\| \, . \, \|_{\mathcal{L}(E,F)}$ norme sur l'espace vectoriel $\mathcal{L}(E, F)$;

$I\!B_r(x_0) = \{x \ / \ \|x - x_0\| \leq r\}$ boule fermée de centre x_0 et de rayon r ;

$d(x, B) := \inf_{b \in B} \|x - b\|$ distance entre x et l'ensemble B ;

$e(A, B)$ excès de l'ensemble A à l'ensemble B ;

$\rho(A, B)$ "distance" de Pompeiu-Hausdorff entre A et B ;

$\nabla f(x)$ dérivée de Fréchet du premier ordre de f évaluée en x ;

$\mathcal{D}G(v)$ ensemble des dérivées d'une multiapplication G différentiable au sens de Nachi et Penot ;

N-P-différentiable différentiable au sens de Nachi et Penot ;

$[x_0, y_0; g]$ différence divisée du premier ordre de la fonction g aux points x_0, y_0 ;

$[x_0, y_0, z_0; g]$ différence divisée du second ordre de la fonction g aux points x_0, y_0, z_0 ;

$T : X \rightrightarrows Y$ ou $T : X \rightsquigarrow Y$ l'application T qui à tout $x \in X$ associe un sous-ensemble de y, appelée aussi "application multivoque ou multivaluée de X dans Y " ;

Tx l'image de x par l'application multivoque T ;

$dom\ T := \{x \ / \ Tx \neq \emptyset\}$ domaine de T ;

$Im\ T := \{y \in Y \ / \ \exists x, y \in Tx\}$ l'image de T ;

T^{-1} l'inverse de T ;

$T(A) := \bigcup_{x \in A} Tx$ l'image de A par T ;

$\inf\{\|y\|,\ y \in Tx\}$ borne inférieure dans $\mathbb{R}_+$ des $\|y\|$ pour $y \in Tx$;

s.c.i. semi-continuité inférieure ;

s.c.s. semi-continuité supérieure.

Introduction générale

Formellement les inclusions variationnelles sont des problèmes de la forme

$$0 \in f(x) + G(x), \tag{1}$$

où $f : X \to Y$ est une fonction, $G : X \rightrightarrows Y$ est une multiapplication(ou application multivoque), X et Y désignent deux espaces de Banach ; c'est une relation qui, à chaque $x \in X$ associe un sous-ensemble éventuellement vide de Y.
Par ailleurs, les applications multivoques s'avèrent être très utiles pour résoudre les problèmes en théorie du contrôle, en économie et gestion, en biologie, en combinatoire, etc.
Le concept d'inclusions variationnelles a été introduit par S. M. Robinson [84, 85] dans les années 70 et 80 comme outils mathématiques pour décrire, analyser et résoudre différents problèmes rencontrés en ingénierie, économie, théorie de transport, etc. Robinson a montré que les problèmes de complémentarité peuvent être transformés en termes d'inclusions variationnelles en utilisant le cône normal d'un ensemble. Par exemple, si $G : x \rightrightarrows N_C(x)$ est une multiapplication, appelée *cône normal* associé á C de $\mathbb{R}^n$ avec $N_C(x)$ l'ensemble des vecteurs v de $\mathbb{R}^n$ tels que $\langle v, x^{'} - x\rangle \leq 0$ pour tout $x^{'} \in C$, vide si $x \notin C$, l'équation généralisée (1) devient une inégalité variationnelle.
Soulignons également qu'un problème d'optimisation convexe peut être vu, aussi, comme une inégalité variationnelle. En fait, pour une *fonction-objectif* $g : \mathbb{R}^n \to \mathbb{R}$, différentiable sur un ouvert $\mathcal{O} \in \mathbb{R}^n$, et C un sous-ensemble de *contraintes*, convexe et fermé de $\mathcal{O}$, minimiser $g(x)$ pour tout $x \in C$ revient à résoudre l'inégalité variationnelle

$$\nabla g(x) + N_C(x) \ni 0,$$

où $\nabla g(x)$ est le *vecteur-gradient de* g *en* x, défini par $\nabla g(x) = \left[\frac{\partial g}{\partial x_j}(x_1, \cdots x_n)\right]_{j=1}^{n}$ (voir Théorème 2A.6 de [29]. De même, dans certains problèmes de la théorie des jeux, la recherche d'un équilibre de Nash revient à résoudre une inégalité variationnelle [voir [29] de Dontchev et Rockafellar, page 73, exercices 2A.10 et 2A.11].
Vu son vaste champ d'applications pratiques, plusieurs théories relatives à la résolution de (1) se sont accrues durant les trois dernières décennies.
Plusieurs méthodes itératives ont été présentées pour approcher une solution x^* de (1) ;

l'une des plus originales est celle de Dontchev [25, 26] où il a introduit une méthode de type Newton de la forme

$$0 \in \underbrace{f(x_k) + \nabla f(x_k)(x_{k+1} - x_k)}_{\text{Approx. d' ordre 1 de } f(x)} + G(x_{k+1}) \tag{2}$$

et établissait la convergence quadratique quand la dérivée de Fréchet ∇f était localement Lipschitz dans un voisinage de x^* et sous l'hypothèse de pseudo-Lipschitz(ou d'Aubin continuité) de l'application multivoque $(f + F)^{-1}$; dans [27], il a également prouvé la stabilité de la méthode (3). Suite à ces travaux, A. Piétrus, dans [73], a obtenu des résultats similaires quand la dérivée de Fréchet f' satisfait une condition de type Hölder ; de ce fait, il a montré que la méthode (3) était applicable à un cadre un peu plus général.
Lorsque la partie univoque f est deux fois Fréchet-différentiable, en utilisant un développement de Taylor à l'ordre 2 de f, Geoffroy, Hilout et Piétrus [31] ont associé à (1) la relation

$$0 \in \underbrace{f(x_k) + \nabla f(x_k)(x_{k+1} - x_k) + \frac{1}{2}\nabla^2 f(x_k)(x_{k+1} - x_k)^2}_{\text{Approx. d'ordre 2 de } f(x)} + G(x_{k+1}), \tag{3}$$

où $\nabla f(x_k)$ et $\nabla^2 f(x_k)$ représentent respectivement la dérivée première et la dérivée seconde de Fréchet de f au point x. Ils ont prouvé la convergence locale et cubique de la suite $(x_k)_k$ vers x^* en supposant que l'application multivoque

$$\Big[f(x) + \nabla f(x^*)(. - x^*) + \frac{1}{2}\nabla^2 f(x^*)(. - x^*)^2 + G(.)\Big]^{-1}$$

est M-pseudo-Lipschitz en $(0, x^*)$.
Le calcul de la dérivée seconde de Fréchet étant coûteux, C. Jean-Alexis, dans [46], a associé à (1) une méthode sans dérivée seconde qui convergeait aussi cubiquement.

Dans le cas où la partie univoque f de (1) n'est pas Fréchet différentiable, les méthodes précédentes ne sont plus valables. Cependant, la modélisation par l'inclusion (1) de certains problèmes rencontrés dans les domaines évoqués précédemment subsiste. Il est fondamental de recourir à des méthodes appropriées. En ce sens, plusieurs méthodes ont été donc proposées, dans la littérature, pour résoudre (1). En fait, lorsque la partie univoque du problème (1) est la somme d'une fonction différentiable f dans un voisinage d'une solution x^* et d'une fonction g différentiable en x^* mais pas nécessairement dans un voisinage de x^*, les auteurs M.H. Geoffroy et A. Piétrus [33] ont prouvé l'existence d'une suite (x_k) convergeant super-linéairement vers x^* et vérifiant la relation

$$0 \in f(x_k) + g(x_k) + (\nabla f(x_k) + [x_{k-1}, x_k; g])(x_{k+1} - x_k) + G(x_{k+1}),$$

où $\nabla f(x_k)$ est la dérivée première de Fréchet en x_k et $[x_{k-1}, x_k; g]$ est la *différence divisée* de g aux points x_{k+1} et x_k.
L'idée de considérer cette suite vient d'une méthode de Cãtinas lorsque $G = \{0\}$ (voir [20]) ; cette méthode est locale et d'ordre $p = \frac{1+\sqrt{5}}{2}$.
Inspirés de la méthode de Dontchev, ces auteurs dans [32] ont proposé la suite définie par

$$\begin{cases} x_0 \text{ est une initialisation proche d'une solution } \bar{x}, \\ 0 \in f(x_k) + \nabla f(x_k)(x_{k+1} - x_k) + g(x_k) + G(x_{k+1}), \text{ pour } k = 0, 1, 2, \ldots \end{cases}$$

où f est différentiable, g est une fonction lipschitzienne (non nécessairement différentiable) et x^* est une solution du problème

$$0 \in f(x) + g(x) + G(x).$$

Ils ont prouvé, sous une condition de régularité métrique et un théorème de point fixe, la convergence semi-locale et super-linéaire de leur méthode.
Remarquons que leur méthode itérative serait celle de Zinchenko [92] si G était identiquement égale au singleton $\{0\}$.
Quand f est sous-analytique et localement Lipschitz, Cabuzel et Piétrus [21] ont proposé une autre méthode de type Newton en remplaçant, dans la suite de Dontchev, la suite $\nabla f(x_k)$ par $\Delta f(x_k)$ où $\Delta f(x_k) \in \partial f(x_k)$ et $\partial f(x_k)$ est la (matrice) jacobienne généralisée de Clarke de f au point x_k. La suite (x_k) ainsi obtenue vérifie l'inclusion

$$0 \in f(x_k) + \Delta f(x_k)(x_{k+1} - x_k) + G(x_{k+1}) \tag{4}$$

et converge super-linéairement vers une solution x^* de (1).
Notons que toutes ces méthodes ont été étudiées sous l'hypothèse de régularité métrique ou Aubin continuité.
En revanche, cette hypothèse n'est pas l'unique condition si nous voulons obtenir des résultats intéressants dans la mesure où S. Burnet et al. [14] ont pu montrer, sans l'Aubin continuité, que la méthode proposée par ces derniers, en l'occurrence la suite définie par (4), converge super-linéairement vers une solution semi-stable x^* de (1).
La morale de l'histoire est que la régularité métrique n'est pas une condition nécessaire pour approcher avec la même vitesse une solution semi-stable de (1) par la méthode conçue par Cabuzel et Piétrus [21].
Notons que le concept de semi-stabilité a été introduit par Bonnans [11] pour résoudre les inégalités variationnelles.
Pour de plus amples détails sur la régularité métrique et les théorèmes de point fixe évoqués, le lecteur peut se référer par exemple à [3, 4, 24, 28, 67, 68, 86] et leurs références.

Ces différentes méthodes provenaient d'une approche faisant intervenir la régularité métrique ou des variantes et un théorème de point fixe. Elles se révèlent efficaces quant au calcul du taux de convergence, cependant, en pratique, elles comportent des inconvénients car la régularité métrique est un concept local.
D'où, plusieurs travaux contournant ces concepts ont été envisagés. L'idée est de proposer des méthodes à convergences semi-locales. Le principe comme pour les équations classiques est de construire une suite qui converge vers un point qui est solution du problème à résoudre.

De nouvelles méthodes issues d'une approche d'optimisation faisant appel aux processus convexes sont apparues depuis moins de dix ans. Ces méthodes prennent leur origine dans un travail de Robinson [82] proposant une extension de la méthode de Newton à une fonction Fréchet différentiable à valeurs dans un cône convexe fermé non vide. La mise en oeuvre de cette approche était fondée sur la projection de l'origine sur l'image d'un élément approprié par l'inverse d'un processus convexe normé.
En fait, sa méthode consistant à résoudre numériquement le problème

$$\text{trouver } x \in X \text{ tel que } 0 \in f(x) - K, \tag{5}$$

où K est un cône convexe, fermé et non vide de Y. C'est une classe de méthode de projection étendue au cadre multivoque puisque les approximations (x_k) résulte de la projection de l'origine de X sur un ensemble convexe fermé et non vide.
Notons que la convergence et l'ordre de cette méthode se basait en partie sur la suite majorante utilisée par Rheinboldt [81] et quelques propriétés d'une classe d'applications multivoques appelée "Processus Convexes".
La technique utilisée sera notre boussole tout au long de cette thèse. Pour en faire une idée, nous tâcherons de présenter, dès l'entame du chapitre 2, une description succincte de cette méthode ainsi que son résultat de convergence.
Les travaux de Robinson ont mis quelques temps avant de s'imposer comme techniques fondamentales à la construction des extensions de méthodes de résolution d'équations multivoques.
40 ans plus tard, Piétrus et al. [78, 79], s'inspirant des travaux de Robinson, ont présenté d'autres extensions relatives aux cas où la partie univoque f de (5) est perturbée par une application g lipschitzienne et non différentiable puis dans le cas où g admet des différences divisées.
En 2016, soit 44 ans après, S. Bernard et al. [15] ont proposé une extension de la méthode de Newton semi-lisse au problème (5) quand l'application univoque f est non différentiable au sens de Fréchet mais seulement semi-lisse.
Soulignons que la partie multivoque G de (1) était toujours fixe et plus précisément un cône convexe fermé.

La contribution principale de cette thèse consiste à étendre les résultats obtenus par Piétrus, Bernard et al. [15, 78, 79] au cas où la partie multivoque G du problème (1) est différentiable dans un certains sens. En fait, on écrira G comme somme d'un cône convexe, fermé, non vide $K \subset Y$ et d'une application multivoque F différentiable au sens de Nachi-Penot [69]. Ainsi, la thèse sera organisée en cinq chapitres repartis comme suit.

Le premier chapitre sera débuté avec quelques notions générales relatives aux applications multivoques. Nous en profitons pour présenter un exemple concret ayant rapport avec le milieu médical. Ensuite, nous rappelons la définition et quelques résultats relatifs au processus convexe, puis la notion de suite majorante, qui toutes deux vont jouer un rôle central dans les résultats à établir. Enfin nous enchaînons avec quelques notions d'analyse notamment le sous-différentiel de Clarke et la différentiabilité au sens de Nachi et Penot.
Le deuxième chapitre est consacré à un résumé de l'état de l'art sur les différentes méthodes itératives n'utilisant pas la régularité métrique et les thèorèmes de point fixe mais utilisant les processus convexes pour résoudre (5) et de ses extensions. L'originalité de ce qui suit réside dans l'utilisation des dérivées d'applications multivoques.

Dans le troisième chapitre, nous nous intéresserons à l'extension de la méthode de Newton associé au problème

$$0 \in f(x) + F(x) - K, \tag{6}$$

où la dérivée première de Fréchet ∇f de f est Lipschitz sur une partie $X_0 \subset X$ et F une multiapplication. La preuve de la convergence se fait en utilisant une méthode analogue à celle de A. Piétrus [78]. Ce résultat généralise donc les travaux réalisés par ce dernier. Puis nous terminons par un exemple numérique illustrant la sensibilité de la méthode par rapport au choix du point initial.

Le quatrième chapitre sera consacré à la méthode de type Newton-sécante relative au problème perturbé

$$0 \in f(x) + g(x) + F(x) - K, \tag{7}$$

où g admet des différences divisées. Notons que le cas où $F(x) = \{0\}$ a été étudié par Piétrus et Jean-Alexis [79] ; ils ont prouvé que la convergence est semi-locale et d'ordre $\frac{1+\sqrt{5}}{2}$.

Finalement, dans le cinquième chapitre, nous présentons une adaptation des résultats de Bernard et al. [15] au problème (6) avec f non nécessairement différentiable mais seulement semi-lisse. Ainsi, nous obtenons une convergence semi-locale et super-linéaire. Et

nous terminerons par des remarques importantes, ainsi qu'une conclusion générale et quelques perspectives pour des travaux futurs.

Chapitre 1

Quelques rappels

Dans ce chapitre, nous recueillerons certaines notions dont nous aurons besoin dans le cadre de cette thèse. Nous commencerons par quelques généralités sur les applications multivoques. Puis, nous présenterons un problème concret ayant rapport avec les multiapplications. Ensuite, nous rappelons la définition d'un processus convexe ainsi que des résultats qui lui sont relatifs. Nous verrons ensuite, pour mieux comprendre les différences divisées, la notion de dérivée de Fréchet. Nous terminerons avec les différences divisées, les fonctions semi-lisses, la différentiabilité de multiapplications, etc.

1.1 Généralités sur les applications multivoques

Avant de rappeler certaines notions générales relatives aux applications multivoques, il est commode d'exhiber quelques problèmes ayant rapport avec celles-ci. Comme nous l'avons mentionné précédemment, une application multivoque est une application à valeurs dans l'ensemble des parties de l'espace d'arrivée. Ces multiapplications ont permis de modéliser une multitude de problèmes issus de diverses disciplines dont ceux rencontrés en théorie de l'optimisation convexe ; celles-ci ont permis d'étudier les problèmes du type :

$$\begin{cases} \text{minimiser } F(x) \\ \text{tel que } x \in D, \end{cases}$$

où F est une application multivoque appelée fonction-objectif à valeurs dans un espace normé, appelé espace objectif, partiellement ordonné dans un certains sens et D est un sous-ensemble non vide du domaine de F appelé ensemble-contrainte ou domaine de faisabilité. Résoudre ce problème revient à déterminer tous les éléments $\bar{x} \in D$ pour lesquels il existe $\bar{y} \in F(\bar{x})$ tel que $\bar{y}$ soit un élément minimal dans un certains sens de l'ensemble $F(D)$. Dans [39], A. H. Hamel et al. l'ont appliqué dans le domaine de la

finance, plus précisément, dans le cadre de la théorie de la mesure de risques dans un modéle de marché avec frictions. D'autre part, certains problèmes de complémentarité, de contrôlabilité, d'équilibre, d'optimisation paramétrique, etc.

Les notations utilisées dans cette thèse sont standards, elles figurent néanmoins dans la rubrique notations.
Les définitions et propriétés concernant les applications multivoques sont extraites de l'ouvrage d'Aubin et al. [4].

Définition 1.1.1 *Étant donné deux ensembles X et Y, si à chaque élément de X, on associe un sous-ensemble éventuellement vide de Y noté $F(x)$, on définit une application multivoque F de X vers Y. On note $F : X \rightrightarrows Y$ ou bien $F : X \rightsquigarrow Y$.*

Dans la littérature, les applications multivoques sont aussi appelées multiapplications, multifonctions, relations ou correspondances.

Définitions 1.1.1 *Soient X et Y deux ensembles, $F : X \rightrightarrows Y$ une application multivoque.*

1. *On appelle domaine de F, l'ensemble $dom\ F := \{x \in X \ / \ F(x) \neq \emptyset\}$.*
2. *On appelle image de F, l'ensemble $im\ F := \{y \in Y \ / \ \exists\ x \in X \ ; \ y \in F(x)\}$.*
3. *On appelle image d'une partie A de X par F, l'ensemble $F(A) := \bigcup_{x \in A} F(x)$ soit $F(A) := \{y \in Y \ / \ x \in A \ ; \ y \in F(x)\}$.*
4. *On appelle image réciproque d'une partie D de Y par F, l'ensemble $F^{-1}(D) := \{x \in X \ / \ F(x) \cap D \neq \emptyset\}$.*
5. *Le graphe de F est le sous-ensemble de $X \times Y$ défini par*

$$gph\ F := \{(x, y) \in X \times Y \ / \ y \in F(x)\}.$$

6. *L'inverse de F est l'application multivoque $F^{-1} : Y \rightrightarrows X$ telle que $x \in F^{-1}(y)$ si et seulement si $y \in F(x)$. Autrement dit, $x \in F^{-1}(y)$ si et seulement si $(x, y) \in gph\ F$.*
7. *La restriction de F à un sous-ensemble K de X, notée $F|_K$ définie par*

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

Le domaine de F est donc l'image de F^{-1} et coincide avec la projection du graphe de F sur l'espace X et, de façon symétrique, l'image de F est égale au domaine de F^{-1} et également à la projection du graphe de F sur l'espace Y. Notons également que pour une application multivoque, l'inverse est toujours défini.

On définit également, comme pour les applications univoques, une addition et une multiplication par un scalaire pour les applications multivoques par :

$$(F_1 + F_2)(x) := F_1(x) + F_2(x) \text{ et } (\lambda F)(x) := \lambda F(x)$$

où le second signe d'addition désigne la somme des deux ensembles $F_1(x)$ et $F_2(x)$.

Remarque 1.1.2
Soient X et Y deux espaces métriques, $\mathcal{P}$ une propriété d'un sous-ensemble de $X \times Y$ par exemple fermé, convexe, compact,... On dit qu'une application multivoque vérifie la propriété $\mathcal{P}$ si et seulement si son graphe la vérifie. Une application multivoque est donc caractérisée par son graphe. Par ailleurs, on dira que F est à valeurs convexes (resp. fermés, compactes,...) si pour tout $x \in$ dom F, $F(x)$ est un sous-ensemble convexe (resp. fermés, compacts,...) de l'ensemble Y.

1.2 Exemples d'applications multivoques

Nous énumérons ici quelques exemples d'applications multivoques.

1. Soit $f : X \to Y$ une application univoque.
   Alors, pour tout $x \in Y$, l'ensemble $f^{-1}(y) = \{x \in X \ / \ f(x) = y\}$ est un sous-ensemble de X. On peut donc définir une application multivoque $F : Y \rightrightarrows X$ par $F(y) = f^{-1}(y)$.
2. Soient $g : X \to Y$ une application univoque et C un sous-ensemble de Y.
   L'application $F(x) := g(x) + C$ est une application multivoque de X dans Y.
3. Certains problèmes provenant de divers domaines conduisent naturellement à l'utilisation des applications multivoques. C'est le cas par exemple de nombreux problèmes de complémentarité, de contrôlabilité, d'équilibre, d'optimisation paramétrique, etc. Nous en présentons un exemple ayant rapport avec le milieu médical.

   On considère le cas d'un patient qui va consulter un médecin.
   Soient P l'ensemble des pathologies, D l'ensemble des diagnostics possibles du médecin et S l'ensemble des symptômes pouvant être présentés par le patient. Un diagnostic correspond à une et une seule pathologie mais toutes les pathologies n'étant pas connues à ce jour, on peut considérer que l'on a $D \subset P$.

   L'application $F : D \rightrightarrows S$ qui, à un diagnostic d, associe l'ensemble de symptômes correspondant $F(d)$, est une application multivoque ; et l'on introduit la fonction $f : P \times P \to [0, 1]$, où $f(p_1, p_2)$ est le pourcentage de la pathologie p_2 soignée par un traitement prévu pour la pathologie p_1. Pour $p \in P$ et $d_1, d_2 \in D$, on

dira que d_1 est plus proche de p que d_2 si $f(d_1, p) \leq f(d_2, p)$.
On note que pour un patient atteint d'une pathologie p et $d_1, d_2 \in D$,

$$F(d_1) \supset F(d_2) \Rightarrow f(d_1, p) \geq f(d_2, p).$$

On peut donc construire une suite dans D à partir d'une suite croissante (au sens de l'inclusion) de parties de S, ce qui pourrait nous amener à considérer, par exemple, le problème de minimisation suivant :

$$\min\{-f(d, \bar{d}); d \in D\},$$

où le diagnostic idéal $\bar{d} \in P$ correspond à la pathologie du patient.

Dans la section qui suit, nous allons définir les Processus Convexes et en exhiber quelques propriétés indispensables à nos résultats.
Juste avant, il convient de rappeler la définition d'une application multivoque convexe.

Définition 1.2.1 *Une application multivoque $F : X \rightrightarrows Y$ est dite convexe si pour tous $x, x' \in X$, et $\lambda \in [0, 1]$*

$$(1 - \lambda)F(x) + \lambda F(x') \subset F((1 - \lambda)x + \lambda x').$$

Lorsque F est convexe, il en est de même de son domaine et son image. Par ailleurs, par caractérisation de la multiapplication F par son graphe, sa convexité est équivalente à celle de son inverse F^{-1}.

1.3 Processus convexes

Vu la propriété de convexité des processus convexes, ils joueront un rôle central dans la mesure où nos différentes méthodes résulteront de la projection de l'origine sur des ensembles convexes d'un espace réflexif.
Le terme processus convexes à été introduit par Rockafellar [87] pour désigner des applications multivoques de $\mathbb{R}^n$ dans $\mathbb{R}^m$ dont le graphe est un cône convexe contenant l'origine de $\mathbb{R}^m$. Il a montré, entre autres, que cette classe contient strictement les opérateurs linéaires continues. Ce concept a été généralisé, en 1972, par S. M. Robinson [82], aux espaces de Banach quelconques. Ceux qui sont normés dans un certains sens sont appelés *processus convexes normés*. Il a également prouvé que la classe des processus convexes fermés possède des propriétés similaires à celles des opérateurs linéaires continues, par exemple les théorèmes de l'application ouverte et du graphe fermé, le théorème de Banach-Steinhaus. En cette même année, il a utilisé certaines propriétés des processus convexes pour construire un shéma itératif pour résoudre (5).
Débutons ce paragraphe en rappelant la définition et quelques propriétés d'un processus convexe extraites de l'article de Robinson [82].

Définition 1.3.1 *Soient X, Y deux $\mathbb{R}$-espaces vectoriels. Une application multivoque T de X dans Y est un processus convexe si elle vérifie :*

a) $T(x) + T(y) \subset T(x+y)$ *pour tous* $x,\ y \in X$;

b) $T(\lambda x) = \lambda T(x)$, *pour tous* $\lambda > 0,\ x \in X$;

c) $0 \in T(0)$.

Si la multiapplication T vérifie uniquement les propriétés $b)$ et $c)$, T est dite *positivement homogène.*
De façon équivalente, un processus convexe de X dans Y est une multiapplication définie sur X à valeurs dans un sous-ensemble de Y, dont le graphe est un cône convexe contenant l'origine de $X \times Y$. Si son graphe est fermé, on dira que c'est un processus convexe fermé. Remarquons qu'un processus convexe est une application multivoque à graphe convexe. Mais, la réciproque est fausse.

Comme nos différentes extensions seront basées sur la projection de l'origine sur des processus convexes normés, il convient de les rappeler.

1.3.1 Processus convexes normés

Soient X et Y deux espaces normés sur $\mathbb{R}$, on définit la norme de T par

$$\|T\| := \sup\{\inf\{\|y\| \mid y \in T(x)\} \mid \|x\| \leq 1,\ x \in domT\},$$

et nous dirons que T est un processus normé si celle-ci est finie. Cette norme-là est égale à l'infimum des κ strictement positifs tels que l'intersection des ensembles $T(x)$ et $\kappa\mathbb{B}$ soit non vide lorsque x décrit la boule unité $\mathbb{B}$. Ainsi, en particulier, un processus convexe à graphe fermé dont le domaine est égal à l'espace X est normé. D'autre part, elle permet de caractériser sa surjectivité, dans le sens où l'ensemble des images est tout l'espace Y, à partir de la finitude de la norme de son inverse ; c'est ainsi qu'un processus convexe à graphe fermé est surjectif si et seulement si son inverse est normé. Pour de plus amples informations sur ce sujet, le lecteur est invité à consulter les pages 253 à 267 de l'ouvrage de Dontchev et Rockafellar [29].

Remarques 1.3.2
Pour un processus convexe T, $Im\ T$ et $dom\ T$ sont des cônes convexes contenant l'origine, puisqu'ils sont les projections naturelles du graphe de T sur X et Y respectivement i.e $domT = P_1(gph\ T)$, $ImT = P_2(gph\ T)$ *avec* $P_1 : X \times Y \longrightarrow X,\ (x,y) \longmapsto x$ *et* $P_2 : X \times Y \longrightarrow Y,\ (x,y) \longmapsto y$.

L'appellation norme a été utilisée par abus de langage puisque l'application $\|.\|$ n'est pas une norme au sens classique sur l'ensemble des processus convexes. En fait, elle peut prendre la valeur $+\infty$, comme par exemple pour le processus convexe $T : \mathbb{R} \rightrightarrows \mathbb{R}$ défini par $T(x) = \{0\}$, si $x \geq 0$, $T(x) = \emptyset$ sinon.

Dans la littérature, on utilise souvent l'appellation *norme intérieure* en parallèle à la *norme extérieure* définie par

$$\|T\|^+ := \sup\{\sup\{\|y\| \mid y \in T(x)\} \mid \|x\| \leq 1,\ x \in domT\}.$$

Celle-ci équivaut à l'infimum des réels strictement positifs κ tels que l'image de la boule unité $\mathbb{B}$ par le processus convexe T soit contenue dans l'ensemble $\kappa\mathbb{B}$. Donc la finitude de la norme extérieure d'un processus convexe implique que l'image du vecteur nul de X par T est égale au singleton $\{0\}$; cette implication devient une équivalence quand T est à graphe fermé et l'espace X est de dimension finie.
Notons que les applications *norme intérieure* et *norme extérieure* ne sont pas des normes et qu'en général, leurs valeurs sont distinctes. Mais, elles sont égales pour un processus convexe fermé et ses adjoints suivant le Théorème 5C.10 de la page 270 de l'ouvrage de Dontchev et Rockafellar [29].

Faisons place maintenant à quelques exemples de processus convexes.

1.3.2 Exemples de processus convexes

1) Soit l'application multivoque $T : \mathbb{R}^2 \rightsquigarrow \mathbb{R}$, définie par :

$$T(x,y) := \begin{cases} \{z/z \geq 0\} & \text{si} \quad y > 0 \\ \{z/z \geq |x|\} & \text{si} \quad y = 0 \\ \emptyset & \text{si} \quad y < 0. \end{cases}$$

T est un processus convexe, en effet :

$T(0,0) := \{z \ / \ z \geq |0|\} = \{z \in \mathbb{R} \ / \ z \geq 0\} := \mathbb{R}_+$; d'où $0 \in T(0,0)$.

Soient $(x,y),(z,k) \in \mathbb{R}^2$.
. Si $y < 0$ ou $k < 0$, on a $T(x,y) + T(z,k) = \emptyset$. On a plus rien à prouver.
. Sinon, i.e $y \geq 0$ et $k \geq 0$.
Remarquons d'abord l'équivalence
$(y \geq 0 \text{ et } k \geq 0) \Leftrightarrow [(y = 0 \text{ et } k = 0) \text{ ou } (y = 0 \text{ et } k > 0) \text{ ou } (y > 0 \text{ et } k = 0) \text{ ou } (y > 0 \text{ et } k > 0)]$.

-Si $y = 0$ et $k = 0$.

On a alors : $T(x,y) = \{p \ / \ p \geq |x|\}, T(z,k) = \{t \ / \ t \geq |z|\}$.
Donc $t + p \geq |x| + |z| \geq |x + z|$.
Ainsi $T(x,y) + T(z,k) \subset T(x+z, y+k) = T[(x,y) + (z,k)]$.
-Si $y = 0$ et $k > 0$.
On a alors $T(x,y) = \{t \ / \ t \geq |x|\} \subset \mathbb{R}_+$ et $T(z,k) = \{p \ / \ p \geq 0\} = \mathbb{R}_+$.
Comme $y + k > 0 \Rightarrow T[(x,y) + (z,k)] = T(x+z, y+k) = \{w \ / \ w \geq 0\} := \mathbb{R}_+$, donc $T(x,y) + T(z,k) \subset T[(x,y) + (z,k)]$.
Le cas où $y > 0$ et $k = 0$ se traite de façon similaire puisque y et k jouent des rôles symétriques.
-Si $y > 0$ et $k > 0$.
Alors $y + k > 0$, $T(x,y) = \{p \ / \ p \geq 0\} := \mathbb{R}_+$ et $T(z,k) = \{t \ / \ t \geq 0\} := \mathbb{R}_+$.
Par suite $T(x+z, y+k) = \{t \ / \ t \geq 0\} := \mathbb{R}_+$ [car $y + z > 0$].
Donc $T(x,y) + T(z,k) \subset T[(x,y) + T(z,k)]$.
Ainsi T vérifie a) de la Définition 1.2.1.

Soit $((x,y),\ \lambda) \in \mathbb{R}^2 \times \mathbb{R}_+^*$.
Si $y < 0$, on a $T(x,y) = \emptyset$. On a plus rien à prouver. Sinon, comme $\{z/z \geq |x|\} \subset \{z/z \geq 0\} := \mathbb{R}_+$, on obtient par la stabilité de $\mathbb{R}_+$ pour l'addition des ensembles, $T(\lambda(x,y)) = \lambda T(x,y)$ [car $\lambda > 0$]. $\square$

Il serait commode de proposer quelques multiapplications qui ne sont pas des processus convexes.
Contre - exemples
1) Soit $Z(w)x := \nabla f(w)x + \frac{1}{2}\nabla^2 f(w)(x^2) + K$ n'est pas un processus convexe, car pour $\lambda > 1, x, y \in X$, on a

$$Z(w)(\lambda x) = \lambda(\nabla f(w)x + \frac{1}{2}\nabla^2 f(w)(\lambda x)^2) \neq \lambda Z(x) = \lambda(\nabla f(w)x + \frac{1}{2}\nabla^2 f(w)(x^2)).$$

2) $T := X \rightsquigarrow X$, définie par $Tx := I\!B_r(x)$ (boule de centre x et de rayon r).

Remarque 1.3.3
Si un processus convexe T est normé, alors son inverse T^{-1} n'est pas nécessairement normé.

En effet, soit le processus convexe de $\mathbb{R}$ dans $\mathbb{R}^2$ défini par :

$$Tx := \begin{cases} \{(y,z)/y^2 \leq zx \text{ et } 0 \leq z\} & \text{si} \quad x \geq 0 \\ \emptyset & \text{sinon.} \end{cases}$$

Sa norme est égale à zéro.
En effet, si $(y,z) \in Tx$ pour $x \in \mathbb{B}(0,1) \cap \mathbb{R}_+$, alors $y^2 \leq zx$ et $0 \leq z$.
Comme $0^2 \leq 0x$ et $0 \leq 0$ donc $(0,0) \in Tx$.
Puisque $0 \leq \|(y,z)\|$, pour tout $(y,z) \in Tx$, donc l'inf devient un max égale à zéro.
Donc $\|T\| = 0$.

Son inverse est

$$T^{-1}(y,z) := \begin{cases} \{x \ / \ x \geq y^2 \ / \ z\} & \text{pour} \quad z > 0 \text{ et } y > 0, \\ \mathbb{R}_+ & \text{pour} \quad z = 0 \text{ et } y = 0, \\ \emptyset & \text{sinon.} \end{cases}$$

Considérons la suite $x_n = (1, 1/n), n \in \mathbb{N}^*$. C'est une suite d'éléments de la boule unité puisqu'en prenant, par exemple, la norme $\|.\|_\infty$ dans $\mathbb{R}^2$, on obtient $\|x_n\| = \max\{1, 1/n\} = 1$, $\forall n \in \mathbb{N}$.
De plus, $T^{-1}(1,1/n) = \{x \ / \ x \leq 1^2/1/n\} = \{x \ / \ x \leq n\}$(car $1/n > 0$).
Puisque $\inf\{\|y\|, y \in T^{-1}(1,1/n)\} = n$, il s'en suit que

$$\sup_{x_n \in \mathbb{B}(0,1) \cap domT} \inf\{\|y\|, y \in T^{-1}(1,1/n)\} = +\infty.$$

D'où $\|T^{-1}\| = +\infty$.
Ainsi, on a bien une application multivoque T normée dont l'inverse T^{-1} ne l'est pas. □

Dans ce qui suit, nous allons établir l'inclusion au sens strict de l'ensemble des applications multivoques linéaires dans l'ensemble des processus convexes. Dans la foulée, nous aurons donc une multiapplication non linéaire dont l'inverse sera un processus convexe.

On considère E et F deux espaces vectoriels sur le corps des réels $\mathbb{R}$.

Définition 1.3.4 *Une application multivoque $T : E \rightsquigarrow F$ est dite linéaire si :*

a) $T(x) + T(y) \subset T(x+y)$ pour tous $x, y \in E$;

b) $T(\lambda x) = \lambda T(x)$ pour tous $x \in E$, $\lambda \in \mathbb{R}$.

Remarquons que $0 \in T(0)$ puisque $\{0\} = 0.T(x) = T(0x) = T(0)$. Par suite, une application multivoque linéaire est un processus convexe. Mais la réciproque est fausse. En fait, l'application multivoque T définie par

$$Tu = \begin{cases} \{x \ / \ x \leq Bu\} & \text{si} \quad u \geq 0 \\ \emptyset & \text{sinon,} \end{cases}$$

où B est une application linéaire de $\mathbb{R}^m$ dans $\mathbb{R}^n$ est un processus convexe dont l'inverse est

$$T^{-1}x = \{u \ / \ u \leq 0, Bu \geq x\}, \forall x.$$

Observons que la multiapplication T n'est pas une application multivoque linéaire puisque si $\lambda \leq 0$ et $y \geq 0 \Rightarrow T(\lambda y) = \emptyset$ et $T(y) \neq \emptyset$.

Nous terminons cette sous-section par une proposition qui montrera que la nature d'un processus convexe est conservée en lui ajoutant une application univoque linéaire.

Proposition 1.3.5 *Soient un processus convexe $T : E \rightsquigarrow F$ et $A : E \rightsquigarrow F$ une application linéaire. Alors*

a) La multiapplication $T + A$ est un processus convexe ;

b) si T est normé, alors $(T + A)$ est normé.

Preuve
Soit $(T + A)0 = T(0) + A(0) = T(0)$ puisque $A(0) = 0$ et T est un processus convexe.
D'où $(T + A)(0) \ni 0$.
Soient x, x' de E et soit $y \in (T + A)(x) + (T + A)(x')$.
On a $y \in (T + A)(x) + (T + A)(x') = T(x) + A(x) + T(x') + A(x') = (T + A)(x + x')$.
Ceci montre que

$$(T + A)(x + x') \supset (T + A)x + (T + A)x'.$$

Soient $x \in E$, $\lambda > 0$ et soit $(T + A)(\lambda x)$.
Alors $y \in (T+A)(\lambda x) = T(\lambda x)+A(\lambda x))$ (car T est un processus convexe et A linéaire).
Ainsi, $(T + A)(\lambda x) = \lambda(T + A)x$.

La partie b) s'obtient par inégalité triangulaire de la norme. □

Ce serait faux de dire que la partie b) de la proposition est vraie pour l'inverse $(T+A)^{-1}$ du processus convexe $T + A$ si T^{-1} est normée.
Le thèorème qui suit nous fournira une condition sous laquelle cela est vraie.

1.3.3 Perturbation d'un processus convexe normé

Nous allons présenter un théorème, dû à Robinson, qui sera constamment appliqué dans cette thèse. En fait, il est prouvé que l'inverse d'un processus convexe normé qui est perturbé par un processus convexe normé est normé, ainsi qu'une estimation de la norme de l'inverse en fonction de la norme de l'inverse du processus convexe.

Théorème 1.3.6 *Si X un espace de Banach, Y un espace normé. Soient T et Δ deux processus convexes de X dans Y ; notons dom T par K et $Im\ T$ par R. Supposons que T, T^{-1} et Δ sont normés, et que $\|T^{-1}\|\|\Delta\| < 1$. Supposons de plus que $K \subset dom\ \Delta$, $\Delta(K) \subset R$, K soit fermé, et $(T-\Delta)(x)$ soit fermé pour chaque $x \in K$. Alors le processus convexe $T - \Delta$ a les propriétés suivantes :*

a) $Im\ T \subset Im\ (T-\Delta)$;

b) $(T-\Delta)_R^{-1}$ est un processus convexe normé, et $\|(T-\Delta)_R^{-1}\| \leq \|T^{-1}\|/(1 - \|T^{-1}\|\|\Delta\|)$.

Pour la preuve de ce résultat, le lecteur est renvoyé à l'article de S. M. Robinson [82].

Il est clair que si le perturbateur Δ est linéaire, les conclusions se tiennent. Cependant, la conclusion b) ne se tient plus si le processus convexe $(T-\Delta)^{-1}$ n'est pas restreint à $Im\ T$. Par exemple, si $X = Y = \mathbb{R}^2$ avec la norme infinie, T et Δ définis par

$$Tx := \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} x,\ x \in \mathbb{R}^2 \right\}$$

et

$$\Delta x := \left\{ \begin{bmatrix} 0.5 & 0 \\ 0 & -0.1 \end{bmatrix} x,\ x \in \mathbb{R}^2 \right\}.$$

Pour chaque $x = (x_1, x_2)^t \in \mathbb{R}^2$, nous avons, pour $x \in \mathbb{B}(0,1)$, $\|x\|_\infty = \max(|x_1|, |x_2|) \leq 1$, donc $|x_1| \leq 1$. Un petit calcul nous fournit $Tx = \{(x_1, 0)^t,\ x_1 \in \mathbb{R}\}$.
On a donc $\|y\| = \max(|x_1|, 0) = |x_1| \leq 1$. Puisque pour $x = (1,0)^t \in \mathbb{B}(0,1)$, l'inf y est atteint, $\|y\| = 1$, d'où $\|T\| = 1 = \|T^{-1}\|$ avec

$$T^{-1}y := \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} y,\ y \in \mathbb{R}^2 \right\}.$$

Un calcul similaire nous donne $\|\Delta\| = 1/2$ et $\|(T-\Delta)^{-1}\| = 10$ avec

$$(T-\Delta)^{-1}y := \left\{ \begin{bmatrix} 2 & 0 \\ 0 & 10 \end{bmatrix} y,\ y \in \mathbb{R}^2 \right\}.$$

Cependant sur l'image de T, $\|(T-\Delta)_R^{-1}\| = 2$. $\square$

Dans le cas où T et Δ sont des applications linéaires continues dans un espace de Banach dans lui-même avec T et Δ inversibles, nous pouvons obtenir également une borne inférieure de $\|(T-\Delta)^{-1}\|$,

$$\|(T-\Delta)^{-1}\| \geq \|T^{-1}\|/(1 + \|T^{-1}\|\|\Delta\|). \tag{1.1}$$

Cependant cette inégalité est généralement fausse pour les processus convexes. Par exemple, soit $X = Y = \mathbb{R}$, et soit T l'application identique sur $\mathbb{R}$. Soit Δ défini par $\Delta x := \mathbb{R}$ pour chaque $x \in \mathbb{R}$. Alors $\|T\| = \|T^{-1}\| = 1$, mais $\|(T-\Delta)^{-1}\| = 0$, ainsi l'inégalité (1.1) n'est plus vraie.
Pour établir (1.1), notons d'abord que si $\|(T-\Delta)^{-1}\| = +\infty$ il n'y a rien à prouver, et que si $\|(T-\Delta)^{-1}\|\|\Delta\| \geq 1$, il s'en suit que

$$\|(T-\Delta)^{-1}\| \geq \|T^{-1}\|(1 - \|\Delta\|\|(T-\Delta)^{-1}\|) \tag{1.2}$$

qui est équivalente à (1.1).
Il reste seulement à prouver (1.1) dans le cas où $\|(T-\Delta)^{-1}\|\|\Delta\| < 1$. Nous allons donc passer par la définition du *cône de récession* 0^+S d'un ensemble convexe S de Rockafellar [87] ; qui est, le cône constitué de tous les points x vérifiant la propriété $x + S \subset S$.
Si l'on suppose que pour tout $x \in dom\ T \cap dom\ \Delta$, on a

$$(\Delta - \Delta)x \subset 0^+Tx, \tag{1.3}$$

on en déduit que, pour tout x, $(T-\Delta)x - (-\Delta)x \subset Tx$; mais puisque $0 \in (\Delta - \Delta)x$, l'inclusion inverse est triviale. De plus, on a $(T-\Delta) - (-\Delta) = T$, et maintenant en faisant des hypothèses nécessaires à l'application du Théorème 1.3.6 à $(T-\Delta)$ et $-\Delta$, nous pouvons établir (1.2), donc (1.1).
Nous remarquons que (1.3) est toujours vraie quand Δ est une fonction univoque. □

Dans la sous-section suivante, nous allons énoncer un résultat, dû à Robinson [83], à travers lequel il a décrit que la distance des images de deux ensembles non vides par un processus convexe normé, sous certaines conditions, ne dépasse pas le produit de la norme de ce dernier par la distance de ces ensembles. Ce résultat lui a permis de construire, par récurrence, la suite majorante (t_k) dont la convergence entraînera celle des itérés (x_k) générés par sa méthode. Cette idée sera également à la base de la convergence de nos différentes extensions à construire.

1.3.4 Estimation de la distance de l'image de deux ensembles par un processus convexe normé

Comme ce résultat a été établi par rapport à la notion "excès", il convient de la rappeler.

Définition 1.3.7 *Soient A, B deux sous-ensembles d'un espace normé X.*
On appelle excès de A sur B la quantité $e(A,B) = \sup_{a\in A} d(a,B)$, avec la convention

$$e(\emptyset, B) = \begin{cases} 0 & si \quad B \neq \emptyset, \\ +\infty & sinon. \end{cases}$$

$*$ Si $a \in A$, $B \neq \emptyset$, $d(a, B) := \inf\{\|a - b\|, b \in B\}$.

Remarque 1.3.8

a) La distance de Pompeiu-Hausdorff entre A et B est la quantité

$$\rho(A, B) := \max\{e(A, B), e(B, A)\}.$$

De façon équivalente, cette quantité peut être exprimée par

$$\rho(A, B) = \inf\{\varepsilon \leq 0 \,/\, A \subset B + \varepsilon\mathbb{B}_X \ \text{ et } \ B \subset A + \varepsilon\mathbb{B}_X\};$$

b) $e(A, B) \neq e(B, A)$ en général; $e(A, B) \in [0, +\infty]$;

c) ρ vérifie tous les axiomes de la distance, cependant elle ne l'est pas, car $\rho(A, B)$ peut prendre la valeur $+\infty$, par exemple si A n'est pas borné. Toutefois, elle l'est sur l'ensemble des parties bornées de X.

Théorème 1.3.9 *(Robinson [83])*
X, Y espaces normés, P, $Q \neq \emptyset$ deux parties de X, T un processus convexe normé tels que les images TP, TQ soient non vides.

a) Si $Q - P \subset domT$, alors $e(TP, TQ) \leq \|T\| e(P, Q)$.

b) Si $(Q - P) \bigcup (P - Q) \subset domT$, alors $\rho(TP, TQ) \leq \|T\| \rho(P, Q)$.

La preuve de ce théorème serait assez pédagogique pour permettre aux lecteurs de manipuler les calculs liés à la notion de distance d'images d'ensembles par un processus convexe normé. Les lecteurs peuvent se reporter à l'article de Robinson [83].

Nous allons maintenant rappeler une notion fondamentale aux différentes extensions à construire. Cette technique de construction, due à Kantorovich dans [50], n'est pas nouvelle puisqu'elle avait déjà été utilisée dans la convergence des suites générées par plusieurs méthodes itératives notamment celles de Rheinboldt [81]. Cette technique était de construire, par récurrence, une suite croissante convergente de réels positifs qui majore la suite des itérés. Par ailleurs, elle a permis de prouver sa convergence ainsi qu'une estimation de l'erreur de la méthode.

Comme la notion de différences divisées est liée à celle de la différentiabilité au sens de Fréchet, nous allons commencer par rappeler la différentiabilité au sens de Fréchet.

1.4 Fonctions Fréchet-différentiables

Considérons deux $\mathbb{R}$-espaces de Banach E et F. Notons $\mathcal{L}(E,F)$ l'espace des applications linéaires continues de E dans F, muni de la norme

$$\|l\| = \sup_{h\in E} \frac{\|lh\|}{\|h\|}$$

laquelle correspond à la norme matricielle lorsque $E = F = \mathbb{R}^n$. Observons que $\mathcal{L}(E,F)$ est également un espace de Banach.

Supposons que U soit un ouvert (non vide !) de E et considérons une fonction $f : U \to F$.

Définition 1.4.1 *Nous dirons que la fonction f est différentiable en un point $x \in U$ si elle est continue au point x et s'il existe $l(x) \in \mathcal{L}(E,F)$ tel que*

$$\lim_{h\to 0_E} \frac{\|f(x+h) - f(x) - l(x)h\|}{\|h\|_E} = 0. \tag{1.4}$$

L'application $l(x)$ dépend du point x, nous la noterons désormais $l(x) = \nabla f(x)$ ou $f'(x)$, de sorte que (1.4) se réécrit

$$\lim_{h\to 0_E} \frac{\|f(x+h) - f(x) - \nabla f(x)h\|}{\|h\|_E} = 0.$$

L'application $\nabla f(x)$ est appelée la dérivée de Fréchet de f au point x.

Définition 1.4.2 *Nous dirons que la fonction est différentiable sur U si elle est différentiable en tout point $x \in U$. Dans ce cas, nous appelons la dérivée de Fréchet de f la fonction*

$$\nabla f : x \in U \mapsto \nabla f(x) \in \mathcal{L}(E,F).$$

Si de plus ∇f est continue par rapport à x, nous dirons que f est continûment Fréchet-différentiable, ou de façon équivalente que f est de classe $\mathcal{C}^1(E,F)$.

Nous allons clôturer la sous-section par un exemple de processus convexe constitué de la dérivée de Fréchet et d'un cône K convexe fermé contenant l'origine de Y.

Soit la multiapplication S définie par

$$S(w)x := \nabla f(w)x + K.$$

S est un processus convexe.
En effet, $S(w)0 = K$. K étant un cône, donc $0 \in K$. Ainsi $0 \in S(w)0$.

Soient $x, z \in X$.
$S(w)x + S(w)z = \nabla f(w)x + \nabla f(w)z + K + K$. Le fait que K soit un cône convexe, on a $\nabla f(w)x + \nabla f(w)z + K + K \subset \nabla f(w)x + \nabla f(w)z + K$.
On a aussi $S(w)(x+z) = \nabla f(w)(x+z)+K$.Par linéarité de$\nabla f(w)$, on a $S(w)(x+z) = \nabla f(w)x + \nabla f(w)z + K$.
Par suite $S(w)(x + z) \supset S(w)x + S(w)z$.

Soit $\lambda > 0, x \in X$.

$$\begin{aligned} S(w)(\lambda x) &= \lambda \nabla f(w)x + K \\ &= \lambda[\nabla f(w)x + \tfrac{1}{\lambda}K] \subset \lambda[\nabla f(w)x + K] \\ &= \lambda S(w)x. \end{aligned}$$

Réciproquement,

$$\begin{aligned} \lambda S(w)(x) &= \lambda[\nabla f(w)x + K] = \nabla f(w)(\lambda x) + \lambda K \\ &= \nabla f(w)(\lambda x) + K (\text{car } K \text{ est un cône}) \\ &= S(w)(\lambda x). \quad \square \end{aligned}$$

Nous allons maintenant aborder la notion de *différences divisées*. L'utilisation de cette notion a déjà permis d'obtenir une méthode itérative pour résoudre numériquement des systèmes d'équations non linéaires lorsque la fonction en question n'était pas Fréchet différentiable. Elle consistait à remplacer, dans la suite de Newton, la dérivée par la pente de la droite reliant les points x_n et x_{n-1} de la fonction ; d'où l'appellation "Méthode Newton-sécante" dont une extension fera l'objet du chapitre 3.
Il sera donc question de définir cette notion, ainsi que certaines propriétés qui leur sont relatives.

1.5 Différences divisées

Dans le cadre de cette thèse, il se pourra que la différentielle de la partie univoque du problème (1) soit coûteuse à calculer ou bien qu'elle n'existe pas. Si c'est le cas, on suppose qu'elle admet des différences divisées dont nous donnerons les définitions dans ce paragraphe. Notons qu'elles sont extraites du papier de Piétrus et al. [79].

Définition 1.5.1 *Soient X et Y deux espaces de Banach. Un opérateur linéaire de X dans Y est appelé* ***différence divisée d'ordre 1*** *de l'opérateur $g : X \to Y$ aux points x_0, y_0, noté $[x_0, y_0; g]$, s'il vérifie la propriété suivante :*

$$[x_0, y_0; g](y_0 - x_0) = g(y_0) - g(x_0), \quad x_0 \neq y_0.$$

En particulier, si g is Fréchet-différentiable en $x_0 \in X$, alors $[x_0, x_0; g] = g'(x_0)$.

Remarquons que $[y_0, x_0; g](x_0 - y_0) = g(x_0) - g(y_0)$ équivaut à $-[y_0, x_0; g](y_0 - x_0) = -(g(y_0) - g(x_0))$. Ainsi $[x_0, y_0; g] = [y_0, x_0; g]$.

Définition 1.5.2 *Soient X et Y deux espaces de Banach. Un opérateur linéaire de X dans Y est appelé* ***différence divisée du second ordre*** *d'un opérateur $g : X \to Y$ aux points x_0, y_0 et z_0, noté $[x_0, y_0, z_0; g]$, s'il vérifie la propriété suivante :*

$$[x_0, y_0, z_0; g](z_0 - x_0) = [y_0, z_0; g] - [x_0, y_0; g], \quad x_0, y_0, z_0, distincts.$$

En particulier, si g est deux fois Fréchet différentiable en $x_0 \in X$ alors $[x_0, x_0, x_0; g] = \frac{1}{2}g''(x_0)$.
Notons que l'opérateur différence divisée n'est pas nécessairement unique.

Quelques propriétés
Pour toute permutation σ de l'ensemble $\mathfrak{S}(0, 1, ..., n)$ des permutations de $\{0, 1, ..., n\}$, on a $[x_{\sigma(0)}, x_{\sigma(x_1)}, ..., x_{\sigma(x_n)}] = [x_0, x_1, ..., x_n]$.

Les notions de différences divisées et de différentiabilité sont liées. Citons par exemple, les deux propriétés suivantes.
Soit une fonction $g : X \to Y$ et un ouvert de $\Omega \subset X$.
1) On montre dans [5] que si g admet des différences divisées vérifiant les inégalités suivantes :
pour tous $x, y, z \in \Omega$

$$\|[x, y; g] - [x, z; g]\| \leq C_0\|y - z\|,$$

$$\|[y, x; g] - [z, x; g]\| \leq C_1\|y - z\|,$$

où C_0 et C_1 sont des constantes positives, alors g est Fréchet différentiable sur Ω. Et si les deux inégalités sont vérifiées, alors g est lipschitzienne sur Ω de constante $C_0 + C_1$.

2) Si l'opérateur g admet des différences divisées du premier ordre et qu'il vérifie une condition de Lipschitz de constante C sur Ω, avec Ω convexe, alors g est Fréchet différentiable sur Ω. De plus, pour tous x, y, z de Ω,

$$\|g'(x) - g'(y)\| \leq 2C\|x - y\|$$

et

$$\|[x, y; g] - g'(z)\| \leq C(\|x - z\| + \|y - z\|).$$

Réciproquement, si g est différentiable sur Ω et que g' est lipschitzienne sur Ω, alors g admet des différences divisées vérifiant une condition de Lipschitz sur Ω.

Finalement, notons que l'on définit, par récurrence, les différences divisées d'ordre n par la relation

$$[x_0, x_1, ..., x_n; g](x_n - x_0) = [x_1, ...x_n; g] - [x_0, ..., x_{n-1}; g],$$

pour $x_0, x_1, ..., x_n$ deux à deux distincts, et si g est n fois Fréchet en x_0, alors $[x_0, ..., x_0; g] = \frac{1}{n!}g^n(x_0)$.

Pour plus de détails sur les différences divisées, le lecteur peut consulter les références [5, 45, 89].

Remarque 1.5.3
Pour une fonction réelle de variable réelle f de classe C^1 sur un intervalle contenant x_0 et x_1, on a alors

$$f(x_1) - f(x_0) = \int_0^1 f'(x_0 + t(x_1 - x_0))(x_1 - x_0)dt.$$

On en déduit la relation

$$[x_0, x_1; f] = \int_0^1 f'(x_0 + t(x_1 - x_0))dt.$$

Les différences divisées du premier ordre peuvent être vues comme moyenne de dérivées sur l'intervalle d'extrémités x_0 et x_1. On va généraliser ce résultat :

Théorème 1.5.4 *Supposons f de classe C^n sur l'intervalle*

$$\Big[\min_{0\leq j\leq n} x_j, \max_{0\leq j\leq n} x_j\Big].$$

Alors

$$[x_0, ..., x_n; f] = \int_0^1 \int_0^{t_1} ... \int_0^{t_{n-1}} f^{(n)}(x_0 + t_1(x_1 - x_0)) + t_2(x_2 - x_1) + ... + t_n(x_n - x_{n-1})dt_n...dt_1.$$

Ce résultat a été prouvé dans l'ouvrage de Michelle Schatzman [88].
Nous tenons à montrer le cas particulier mentionné plus haut à savoir

$$[x_0, ..., x_0; f] = \int_0^1 \int_0^{t_1} ... \int_0^{t_{n-1}} f^{(n)}(x_0)dt_n dt_{n-1}...dt_1 = f^{(n)}(x_0)V_n,$$

où V_n est le volume du n-simplexe $\Sigma_n = \{t \in \mathbb{R}^n,\ 0 \leq t_1 \leq t_2 \leq ... \leq t_n \leq 1\}$, de sommets $(0, 0, ..., 0)$, $(1, 0, ..., 0)$, et ainsi de suite en transformant successivement les 0

en 1 pour avoir à chaque fois un nouveau sommet, jusqu'à $(1, 1, ..., 1)$.
On a

$$\begin{aligned} V_n &= \int_0^1 ... \int_0^{t_{n-1}} dt_n ... dt_1 \\ &= \int_0^{t_{n-2}} t_{n-1} dt_{n-1} ... dt_1 \\ &= \int_0^1 ... \int_0^{t_{n-3}} \frac{t_{n-2}^2}{2} dt_{n-2} ... dt_1 \\ &= ... = \tfrac{1}{n!} \end{aligned}$$

On a donc obtenu la relation

$$[x_0, ..., x_0; f] = \frac{f^{(n)}(x_0)}{n!}.$$

D'autre part, le calcul pratique des différences divisées est fondé sur la relation de récurrence ci-dessus. Dans la pratique, on aura le tableau suivant, d'où on déduira aisément un algorithme automatique :

$$\begin{array}{ll} x_0 & f(x_0) \\ & \\ & \Longrightarrow \frac{f(x_0)-f(x_1)}{x_0-x_1} = [x_0, x_1; f] \\ x_1 & f(x_1) \\ & \\ & \Longrightarrow \frac{f(x_1)-f(x_2)}{x_0-x_1} = [x_1, x_2; f] \\ x_2 & f(x_2) \end{array}$$

Passer de n points à $n+1$ points demande n calculs de différences divisées : dans le tableau ci-dessus, rajouter le point x_3 demande le calcul de $(x_3), [x_2, x_3; f], [x_1, x_2, x_3; f]$ et $[x_0, x_1, x_2, x_3; f]$, soit 3 calculs de différences divisées pour passer de 3 à 4 points.

On peut construire des processus convexes à l'aide des différences divisées. Par exemple, prenons X, Y deux espaces normés, w un vecteur fixé de X, f et g deux applications respectivement Fréchet-différentiable et admettant une différence divisée du 1^{er} ordre, K un cône convexe contenant l'origine de Y.
L'application multivoque $T(w)$ définie par

$$T(w)x := (\nabla f(w) + [y, w; g])x + K \text{ où } y \text{ est un point fixe de } X$$

est un processus convexe.
En effet, on a $T(w)0 = K$. K étant un cône, donc $0 \in T(w)0$.
Soient $x, z \in X$.

$$\begin{aligned} T(w)x + T(w)z &= \nabla f(w)x + \nabla f(w)z + [y, w; g]x + [y, w; g]z + K + K, \\ T(w)x + T(w)z &= (\nabla f(w) + [y, w; g])(x + z) + K + K \text{ (par linéarité de } \nabla f(w) + [y, w; g]) \end{aligned}$$

K étant un cône convexe,
$K + K \subset K$, donc $T(w)(x+z) \supset T(w)x + T(w)z$.
Soient $\lambda > 0$, $x \in X$.

$$\begin{aligned} T(w)(\lambda x) &= (\nabla f(w) + [y, w; g])(\lambda x) + K \\ &= \lambda(\nabla f(w) + [y, w; g])x + K. \\ &= \lambda[\nabla f(w)(x) + [y, w; g](x) + \tfrac{1}{\lambda}K](\text{car, } \tfrac{1}{\lambda} > 0\ ) \\ &\subset \lambda[\nabla f(w)x + [y, w; g](x) + K](\text{car } K \text{ cône} \Rightarrow \tfrac{1}{\lambda}K \subset K\ ). \end{aligned}$$

D'où, $T(w)(\lambda x) \subset \lambda T(w)(x)$.

$$\begin{aligned} \lambda T(w)(x) &= \lambda(\nabla f(w) + [y, w; g])x + \lambda K \\ &= (\nabla f(w) + [y, w; g])(\lambda x) + K \text{ (par linéarité de } \nabla f(w) + [y, w; g]\ ). \end{aligned}$$

D'où, $\lambda T(w)(x) = T(w)(\lambda x)$. □

Nous allons maintenant rappeler une notion fondamentale aux différentes extensions à construire. Cette technique de construction, due à Kantorovich dans [50], n'est pas nouvelle puisqu'elle avait déja été concourue à la convergence des suites générées par plusieurs méthodes itératives notamment celles de Rheinboldt dans [81]. Cette technique était de construire, par récurrence, une suite croissante convergente de réels positifs qui majore la suite des itérés. Elle a permis d'ailleurs de trouver une estimation de l'erreur de la méthode.

1.6 Suites majorantes

La définition que nous allons énoncer est issue de l'article de Rheinboldt [81].

Définition 1.6.1 *Soit $(x_k)_k$ une suite dans un espace métrique (X, ρ). Alors une suite croissante de réels positifs $(t_k)_k$ est dite majorante de la suite $(x_k)_k$ si*

$$\rho(x_{k+1},\, x_k) \leq t_{k+1} - t_k, \quad k = 0, 1, ...$$

Observons que la suite (t_k) est une suite croissante et pour tout $m > k \geq 0$,

$$\rho(x_m,\, x_k) \leq \sum_{j=k}^{m-1} \rho(x_{j+1},\, x_j) \leq \sum_{j=k}^{m-1} (t_{j+1} - t_j) = t_m - t_k.$$

Ainsi, si $\lim_k t_k = t^*$ existe, alors (x_k) est une suite de Cauchy dans X. De plus, comme X est complet, il existe $x^* \in X$ tel que $\lim x_k = x^*$, et pour $m \to +\infty$, l'erreur estimée est immédiate, et $\rho(x^*, x_k) \leq t^* - t_k$ pour $k = 0, 1, ...$ □

Comme nous aurons à proposer une extension de (1) lorsque la fonction f sera semi-lisse, il est naturellement question de faire place à de telles fonctions.

1.7 Fonctions semi-lisses

Dans cette section, nous n'allons pas présenter, de manière exhaustive, les différentes notions liées aux fonctions non-lisses. Ainsi nous nous intéressons seulement aux notions dont nous aurons besoin dans le cadre de cette thèse. Pour ce faire, nous faisons appel au jacobien ou gradient généralisé introduit par Clarke. Nous en profitons pour rappeler quelques notions liées au sous-différentiel de Clarke ainsi que le théorème de Rademacher à partir duquel on définit la semi-lissité.
Nous commençons par rappeler la définition d'une fonction localement lipschitzienne. C'est une condition nécessaire à la définition d'une fonction semi-lisse.

1.7.1 Fonction localement lipschitzienne

Soient X, Y deux espaces normés, B un sous-ensemble ouvert de X et une fonction $f : X \to Y$ est dite lipschitzienne (ou Lipschitz) sur B s'il existe un nombre positive K tel que

$$\|f(y) - f(z)\| \leq K\|y - z\| \text{ pour tous } y, z \text{ de } B.$$

Si f est Lipschitz sur un sous-ensemble borné de B alors f est dite *localement lipschitzienne*.
Une fonction est dite *localement lipschitzienne* (tout court) si pour tout x, il existe un voisinage V de x sur lequel elle est lipschitzienne.
Notons qu'une fonction vectorielle à valeurs dans $\mathbb{R}^n$ est localement lipschitzienne, si chaque composante l'est.
Remarquons que toute fonction de f de $\mathbb{R}$ dans $\mathbb{R}$ de classe C^1 est localement lipschitzienne. En effet, soient f de classe C^1 et $x \in \mathbb{R}$.
La dérivée première $f^{'}$ de f est bornée par M sur le segment $[x-1, x+1]$, car $f^{'}$ est continue. Par inégalité des accroissement finis, f est M-lipschitzienne sur $[x-1, x+1]$.
Ce résultat se généralise aux fonctions définies sur $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ ou bien $\mathbb{R}^m$.
Observons aussi qu'une fonction localement lipschitzienne peut ne pas être lipschitzienne. Par exemple, la fonction $x \mapsto x^2$ est de classe C^1 et pour tout $k \geq 0$,

$$|(k+1)^2 - k^2| = 2k + 1 > k|k + 1 - k|.$$

Intéressons-nous maintenant au concept de fonctions semi-lisses qui à été introduit par Mifflin [65]. Ceci sera généralisé, plus tard, aux fonctions vectorielles par Qi et Sun [75]. Particulièrement, ils ont montré l'équivalence entre la semi-lissité et la convergence uniforme des dérivées directionnelles dans toutes les directions.
Rappelons d'abord la définition du jacobien de Clarke qui s'appuie sous un résultat d'analyse classique : théorème de Rademacher ; ceci affirme que toute fonction de $\mathbb{R}^n$ dans $\mathbb{R}^m$ localement lipschitzienne est différentiable presque partout, i.e. l'ensemble des points où elle n'est pas différentiable est de mesure de Lebesgues nulle. Ce résultat

a été déjà utilisé, par Pang, dans les travaux sur la différentiation des applications multivoques positivement homogènes [77].

Théorème 1.7.1 *(Théorème de Rademacher)*
Soient $\mathcal{O}$ un ouvert de $\mathbb{R}^n$ et $f : \mathcal{O} \to \mathbb{R}^m$ une fonction lipschitzienne. Soit D l'ensemble des points de $\mathcal{O}$ où la fonction est différentiable. Alors l'ensemble des points $\mathcal{O} \setminus D$ où la fonction n'est pas différentiable est de mesure nulle dans $\mathbb{R}^n$. En particulier, D est dense dans $\mathcal{O}$, i.e., la fermeture de D dans $\mathcal{O}$ contient $\mathcal{O}$.

Faisons place maintenant à la définition du jacobien de Clarke qui a été évidemment introduit par Clarke [19] ; ce qui est nécessaire à la notion de semi-lissité d'une fonction.

Définition 1.7.2 *(Jacobien de Clarke) Soient $f : \mathcal{O} \to \mathbb{R}^m$ une fonction lipschitzienne (i.e. localement lipschitzienne), $\mathcal{O} \subset \mathbb{R}^n$ un ensemble ouvert, et soit $D \subset \mathcal{O}$ l'ensemble des points où f est différentiable. La jacobienne généralisée au sens de Clarke de f en $\overline{x}$ est définie par*

$$\partial f(\overline{x}) := conv\{A \in \mathbb{R}^{m\times n} : \exists x_i \to \overline{x} avec \quad x_i \in D, Jf(x_i) \to A\},$$

où $Jf(x_i)$ désigne la matrice jacobienne usuelle $m \times n$ des dérivées partielles de f en $x \in D$.

Nous pouvons dire que $\partial f(\overline{x})$ est l'enveloppe convexe des limites de la forme

$$A = \lim_{x_i \to \overline{x}} J(x_i),\ x_i \in D.$$

Il est évident du théorème de Rademacher que le jacobien de Clarke en $\overline{x}$ est un ensemble convexe compact non vide de l'espace matriciel $M_{m,n}(\mathbb{R})$. L'application multivoque $x \rightrightarrows \partial f(x)$ est localement bornée et semi-continue supérieurement. Pour $m = 1$, il est bien connu que le jacobien de Clarke se réduit au gradient généralisé au sens de Clarke, c'est-à-dire

$$\partial f(\bar{x}) = conv\Big\{ \lim_{x_i \to \overline{x}} \nabla f(x_i),\ x_i \in D\Big\},$$

où $\nabla f(x_i)$ désigne le vecteur-gradient de f en x_i.
Cet objet ∂f a été énormément étudié, généralisé et utilisé depuis son introduction par Clarke en 1973.

Maintenant, nous allons aborder les fonctions semi-lisses ainsi que certaines propriétés qui leur sont relatives.

1.7.2 Fonctions semi-lisses

Définition 1.7.3 *[65] On dit qu'une fonction f est semi-lisse en $x \in \mathbb{R}^n$ si f est localement lipschitzienne en x et*

$$\lim_{\substack{V \in \partial f(x+th') \\ h' \to h,\, t \downarrow 0}} \{Vh'\}$$

existe pour tout $h \in \mathbb{R}^n$.

On dit que f est semi-lisse sur $X_0 \subset \mathbb{R}^n$ si f est semi-lisse en tout point $x \in X_0$.

Remarque 1.7.4

(a) De [75], semi-lissité f implique l'existence de la dérivée directionnelle au sens d'Hadamard, de plus

$$\lim_{\substack{h' \to h \\ t \downarrow 0}} \frac{f(x+th') - f(x)}{t} = \lim_{\substack{V \in \partial f(x+th') \\ h' \to h, t \downarrow 0}} \{Vh'\}.$$

(b) Quand $f : \mathbb{R}^n \to \mathbb{R}$ localement Lipschitz, $\partial f(x)$ désigne le gradient généralisé de f en x défini par

$$\partial f(x) = \{g \in \mathbb{R}^n : \langle g, d \rangle \leq f^o(x; d) \text{ pour tout } d \in \mathbb{R}^n\},$$

où $\langle ., . \rangle$ désigne le produit scalaire usuel sur $\mathbb{R}^n$, et $f^o(x;d) = \lim_{\substack{h \to 0 \\ t \downarrow 0}} \sup \frac{f(x+h+td)-f(x+h)}{t}$ est la dérivée directionnelle généralisée de f en x dans la direction de d. Notons que cette dérivée existe toujours pour les fonctions localement Lipschitz [66].

(c) Si $f : \mathbb{R}^n \to \mathbb{R}$ est une fonction convexe, $\partial f(x)$ est le sous-différentiel $\partial_c f(x)$ de f en x,
avec $\partial_c f(x) := \{x^* \in \mathbb{R}^n : f(z) \geq f(x) + \langle x^*, z - x \rangle, \forall z \in \mathbb{R}^n\}$. Chaque élément $x^* \in \partial f(x)$ est appelé un sous-gradient de f en x. ([66], Théorème 3.1.8).

(d) Si $f : \mathbb{R}^n \to \mathbb{R}$ est continûment différentiable en $x \in \mathbb{R}^n$, alors f est localement Lipschitz, $\partial f(x) = \{\nabla f(x)\}$, et f semi-lisse en x [65] où
$\nabla f(x)$ désigne le vecteur-gradient de f en x.

Pour plus de détails le lecteur peut, par exemple, se reporter aux articles [65, 66, 87].

Exemples de fonctions semi-lisses

Nous poursuivons avec une série d'exemples qui peut être jugés utiles aux lecteurs. Nous enchaînons avec un exemple détaillé de calcul du gradient généralisé de la norme euclidienne.

Exemple 1.7.1 *Le sous-gradient de* $P(x) = \|x\| = \sqrt{\langle x, x\rangle}$ *(*$\langle ., . \rangle$ *désigne un produit scalaire sur* $\mathbb{R}^n$*) est*

$$\partial P(x) = \begin{cases} \mathbb{B} & si\ x = 0 \\ \left\{ \frac{x}{\|x\|} \right\} & sinon \end{cases}$$

où $\mathbb{B}$ *est la boule unité fermée de* $\mathbb{R}^n$.
En effet,
Pour $x \neq 0$, $\|x + h\| = \sqrt{\|x\|^2 + 2\langle x, h\rangle + \|h\|^2} = \|x\|\sqrt{1 + 2\frac{\langle x,h\rangle}{\|x\|^2} + \frac{\|h\|^2}{\|x\|^2}} = \|x\| + \frac{\langle x,h\rangle}{\|x\|} + o(h)$,
donc la fonction P *est différentiable en* $x \neq 0$ *et de différentielle* $h \mapsto \frac{\langle x,h\rangle}{\|x\|}$.
Par convexité de l'application P, *d'après* (d) *de la remarque précédente, le vecteur-gradient en* $x \neq 0$ *est* $x/\|x\|$.
D'où $\partial P(x) = \{x/\|x\|\}$ *lorsque* $x \neq 0$.
Soit $v \in \partial P(0)$. *Par définition du sous-gradient, on a*

$$\langle v, x\rangle = \langle v, x - 0\rangle \leq P(x) - P(0) = \|x\| \quad \forall x.$$

En particulier, pour $x = v$, $\langle v, v\rangle \leq \|v\| \Longrightarrow \|v\|(\|v\| - 1) \leq 0$.
Comme $\|v\|$ *est positive, on obtient* $\|v\| - 1 \leq 0$, *soit* $\|v\| \leq 1$. *C'est-à-dire que* $v \in \mathbb{B}$.
Réciproquement, soit $v \in \mathbb{B}$.
En utilisant l'inégalité de Cauchy-Schwartz, on obtient $\langle v, x - 0\rangle = \langle v, x\rangle \leq \|v\|\|x\| \leq \|x\| = P(x) - P(0)$ *pour tout* $x \in \mathbb{R}^n$. *Ainsi,* $v \in \partial P(0)$.
D'où l'égalité $\partial P(0) = \mathbb{B}$.

Exemple 1.7.2 *Fonction localement Lipschitz et semi-lisse sur* $\mathbb{R}$.
La fonction f *de* $\mathbb{R}^2$ *dans* $\mathbb{R}^2$ *définie par*

$$f(x, y) = (|x| + y, 2x + |y|).$$

Considérons sur $\mathbb{R}^2$ *la norme uniforme :* $\|(a, b)\|_\infty = \max(|a|, |b|)$.

•*Montrons que* f *est lipschitzienne pour la norme* $\|(x, y)\|_\infty$.
Soient $X = (x_1, x_2) \in \mathbb{R}^2$, $Y = (y_1, y_2) \in \mathbb{R}^2$.
$f(X) - f(Y) = (|x_1| + x_2, 2x_1 + |x_2|) - (|y_1| + y_2, 2y_1 + |y_2|)$
$= (|x_1| - |y_1| + x_2 - y_2, 2(x_1 - y_1) + |x_2| - |y_2|) = (z_1, z_2)$

$z_1 = |x_1| - |y_1| + x_2 - y_2$
$\Rightarrow |z_1| \leq ||x_1| - |y_1|| + |x_2 - y_2| \leq |x_1 - y_1| + |x_2 - y_2| \leq 2\max(|x_1 - y_1|, |x_2 - y_2|) = 2\|X - Y\|_\infty$.
Un calcul analogue montre que $|z_2| \leq 2\|X - Y\|_\infty$. *Donc*

$$\|f(X) - f(Y)\|_\infty \leq 2\|X - Y\|_\infty.$$

La fonction f *est lipschitzienne de rapport 2, elle est donc localement Lipschitz.*

•*Le jacobien de Clarke de* f *en* $(0,0)$ *est* $\partial f(0,0) = \left\{ \begin{bmatrix} x & 1 \\ 2 & y \end{bmatrix}, -1 \leq x \leq 1, -1 \leq y \leq 1 \right\}$.
Soient $h = (h_1, h_2)^T$, $h' = (h_1', h_2')^T$ *de* $\mathbb{R}^2$, $f(0 + th') = f(th') = (th_1', th_2')^T$.
Soit $V \in \partial f(0 + th')$. *Alors, si* $h_1' = h_2' = 0$, *on a pour* $x, y \in [-1, 1]$, *on a* $Vh' = [xh_1' + h_2', 2h_1' + yh_2']^T$.
Sinon, la fonction f *est différentiable en* th' *; chaque matrice jacobienne est indépendante de* t.
Par conséquent

$$\lim_{\substack{V \in \partial f(x+th') \\ h' \to h, t \downarrow 0}} \{Vh'\}$$

existe. La fonction f *est donc semi-lisse en* $(0,0)$. *Pour* $(x,y) \neq (0,0)$, f *est différentiable, donc semi-lisse.*

Exemple 1.7.3 *Fonction semi-lisse sur* $\mathbb{R}$, *mais qui n' y est pas différentiable (i.e en* 0*)* $f(x) = \log(1 + |x|)$.
Le gradient généralisé (ou gradient de Clarke) de f *est l'ensemble défini comme suit :*

$$\partial f(x) = \begin{cases} \left\{\frac{1}{1+x}\right\} & si\ x > 0 \\ conv\{-1, 1\} & si\ x = 0 \\ \left\{\frac{-1}{1-x}\right\} & si\ x < 0. \end{cases}$$

On remarque que pour tout x *tel que* $-1 < x < 1$,

$$f(x) = \max[\log(1 + x), \log(1 - x)].$$

En effet, $-1 < x < 1 \Rightarrow -1 < x \leq 0$ *ou* $0 \leq x < 1$.
• *Si* $-1 < x \leq 0$.
$\Rightarrow |x| = -x$, *donc* $f(x) = \log(1 - x)$, *et* $(1 - x) - (1 + x) = -2x \geq 0$.
Par stricte croissance de la fonction logarithme,

$$\log(1 - x) > \log(1 + x).$$

Donc,

$$\max[\log(1+x), \log(1-x)] = \log(1-x).$$

On conclut que pour $-1 < x \leq 0, f(x) = \max[\log(1+x), \log(1-x)]$.

• *Si* $0 \leq x < 1$, *alors* $|x| = x$. *On a donc* $f(x) = \log(1+x)$.
Comme $(1+x) - (1-x) = 2x \geq 0$, *et par stricte croissance de la fonction logarithme, on obtient*

$$\log(1+x) > \log(1-x)$$

Par suite

$$\max[\log(1+x), \log(1-x)] = \log(1+x) = f(x).$$

On en déduit que si $0 < x < 1$, $f(x) = \max[\log(1+x), \log(1-x)]$.
f *est différentiable comme maximum de telles fonctions.* f *est donc semi-lisse en* 0.
En somme f *est semi-lisse sur* $\mathbb{R}$ *; ce qui achève la preuve.* □

Il a été prouvé dans [75] qu'une fonction vectorielle est semi-lisse en x si et seulement si chaque composante l'est. La classe de fonctions semi-lisses est très vaste ; en effet, selon [65], elle contient les fonctions lisses(fonctions continûment différentiables), toutes les fonctions convexes et les fonctions semi-lisses par morceaux. Par ailleurs, les sommes, différences, produits et composées de fonctions semi-lisses sont semi-lisses. En particulier, si f est une fonction semi-lisse, alors la fonction $g(x) = \min(x, f(x))$ l'est aussi [76].

Énonçons deux résultats des travaux de Mifflin [65] qui seront nécessaires aux résultats à établir dans le chapitre 5.

Lemme 1.7.5
Supposons que $F : \mathbb{R}^n \to \mathbb{R}^m$ *soit une fonction localement lipschitzienne et la dérivée directionnelle classique* $F'(x;h) = \lim_{t \downarrow 0} \frac{F(x+th)-F(x)}{t}$ *existe pour tous* h *et* x. *Alors*

(i) $F'(x;.)$ *est lipschitzienne ;*

(ii) pour tout h, *il existe un* $V \in \partial F(x)$ *tel que* $F'(x;h) = Vh$.

Le résultat suivant propose des caractérisations d'une fonction semi-lisse.

Théorème 1.7.6
On suppose que $F : \mathbb{R}^n \to \mathbb{R}^m$ *est une fonction localement lipschitzienne. Les énoncés suivants sont équivalents :*

(i) F *est semi-lisse(semismooth) en* x;

(ii) $\lim_{V\in\partial F(x+th')_{h'\to h,t\downarrow 0}}\{Vh'\}$ *est uniformément convergente pour tout vecteur unitaire* h;

(iii) $\lim_{t\downarrow 0}\{Vh\}$ *est uniformément convergente pour tout vecteur unitaire* h;

(iv) pour tout $V \in \partial F(x+h)$, $h \to 0$,

$$Vh - F'(x;h) = o(\|h\|); \tag{1.5}$$

(v)

$$\lim_{x+h\in D_F \ h\to 0} \frac{F'(x+h;h) - F'(x;h)}{\|h\|} = 0. \tag{1.6}$$

Il convient de passer à un exemple de fonction localement lipschitzienne qui n'est pas semi-lisse.

Contre-exemple

Fonction localement lipschitzienne qui n'est pas semi-lisse.
Soit la fonction de $\mathbb{R}$ dans $\mathbb{R}$ définie par

$$f(x) = \begin{cases} x^2\sin(\frac{1}{x}) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0. \end{cases}$$

La fonction dérivée de la fonction f est définie par :

$$f'(x) = \begin{cases} 2x\sin(\frac{1}{x}) - \cos(\frac{1}{x}) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0. \end{cases}$$

La fonction f' n'est pas continue en 0 puisque les fonctions $\cos(\frac{1}{x})$, $\sin(\frac{1}{x})$ n'ont pas de limite en 0. Elle n'est donc pas continûment différentiable en 0 mais différentiable en 0.
La dérivée f' étant bornée sur toute partie bornée de $\mathbb{R}$, f est localement lipschitzienne (via l'inégalité des accroissements finis).
Pour montrer la non semi-lissité de f en 0, on va montrer qu'elle ne vérifie pas la relation (iv) du théorème. Cela revient à montrer l'existence d'un $h \to 0$, $V \in \partial f(0+h)$ tels que

$$Vh - F'(x;h) \neq o(\|h\|).$$

Posons $h_n = \frac{1}{2n\pi}$, pour tout $n \in \mathbb{N}^*$.
$\partial f(0+\frac{1}{2n\pi}) = \{f'(\frac{1}{2n\pi})\} = \{-1\}$. Pour tout $n \in \mathbb{N}^*$, on a

$$f'\left(0;\frac{1}{2n\pi}\right) = \lim_{t\to 0}\frac{f(t\frac{1}{2n\pi}) - f(0)}{t} = \lim_{t\to 0}\frac{t}{4n^2\pi^2}\sin\left(\frac{2n\pi}{t}\right) = 0$$

et pour tout $h \neq 0$,

$$\frac{Vh - f^{'}(0;h)}{h} = \frac{-1}{\frac{1}{2n\pi}} = -2n\pi.$$

Par suite, $Vh - f^{'}(0;h)$ est non négligeable devant h. $\square$

Pour arriver à terme de notre passage en revue, nous présentons des résultats et définitions essentiels aux chapitre 5. Ce sont des extraits de l'article de Qi et Sun [75].

1.7.3 Fonctions semi-lisses d'ordre p

Définition 1.7.7 *(Fonction semi-lisse d'ordre p) On dit qu'une fonction $F : \mathbb{R}^n \to \mathbb{R}^m$ est **semi-lisse d'ordre** p **en** x, si pour tout $V \in \partial F(x+h)$ et $h \to 0$, on a $Vh - F^{'}(x;h) = O(\|h\|^{1+p})$, où $0 < p \leq 1$. Dans le cas où $p = 1$, on dit que F est **fortement semi-lisse**.*

Dans la définition précédente, le symbole $O(\|h\|^{1+p})$ désigne une fonction vectorielle $e : \mathbb{R}^n \to \mathbb{R}^m$ vérifiant $\|e\| \leq M\|h\|^{1+p}$ pour tout $h \in \mathbb{R}^n$ tel que $\|h\| \leq s$ avec M et s des constantes strictement positifs.

Remarque 1.7.8
On peut facilement vérifier les implications suivantes :

(a) Semi-lissité d'ordre p implique semi-lissité ;

(b) si F est semi-lisse en x, alors pour tout $h \to 0$,

$$F(x+h) - F(x) - F^{'}(x;h) = o(\|h\|);$$

(c) si F est semi-lisse d'ordre p en x, alors pour tout $h \to 0$,

$$F(x+h) - F(x) - F^{'}(x;h) = O(\|h\|^{1+p}).$$

Donnons maintenant quelques exemples de fonctions lisses, fortement semi-lisses.

Exemple 1.7.4
Les fonctions minimum $\Phi_{min}(a,b) = \min(a;b)$ et Fischer-Burmeister $\Phi_{FB}(a;b) = \sqrt{a^2+b^2} - a - b$ $((a;b) \in \mathbb{R}^2)$ ne sont pas lisses (non continûment différentiables) mais fortement semi-lisses [91].
La fonction $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = |x|\sqrt{|x|}$ est lisse en 0, mais n'est pas fortement semi-lisse en 0 [15].

La fonction $\psi : \mathbb{R}^3 \to \mathbb{R}$ *définie par* $\psi(a,b,c) = \sqrt{a^2+b^2+c^2} - (b+c)$ *a été introduite par Kanzow [54] dans l'étude des problèmes de complémentarité linéaire.*
Elle est localement Lipschitz, fortement semi-lisse sur $\mathbb{R}^3$. *De plus, si* $a^2+b^2+c^2 > 0$, *alors* ψ *est continûment différentiable sauf en* $(0,0,0)$. *Le jacobien généralisé de* ψ *en* $(0,0,0)$ *est*

$$\partial\psi(0,0,0) = \mathcal{O} \equiv \{(\alpha,\beta,\gamma)^t / \alpha^2 + (\beta+1)^2 + (\gamma+1)^2 \leq 1\}.$$

En effet, remarquons que $\sqrt{a^2+b^2+c^2}$ *est la norme Euclidienne du vecteur* $(a,b,c)^t$. *Donc* $\sqrt{a^2+b^2+c^2}$ *est localement Lipschitz et fortement semi-lisse sur* $\mathbb{R}^3$. $-(b+c)$ *est continûment différentiable, donc localement Lipschitz et fortement semi-lisse sur* $\mathbb{R}^3$. ψ *est donc fortement semi-lisse sur* $\mathbb{R}^3$ *comme somme de telles fonctions.*
Si $a^2+b^2+c^2 > 0$, $\sqrt{a^2+b^2+c^2}$ *est continûment différentiable en* (a,b,c), ψ *l'est.*
Soit $d = (d_1,d_2,d_3) \in \mathbb{R}^3$ *et* $d \neq 0$. *Alors* ψ *est continûment différentiable en* td *pour* $t > 0$.*Et*

$$\nabla\psi(td) = \Big(\frac{d_1}{\sqrt{d_1^2+d_2^2+d_3^2}}, \frac{d_2}{\sqrt{d_1^2+d_2^2+d_3^2}} - 1, \frac{d_3}{\sqrt{d_1^2+d_2^2+d_3^2}} - 1\Big)^t.$$

Pour simplifier, $\nabla\psi(td)$ *est noté par* $(\alpha,\beta,\gamma)^t$. *Clairement* $\alpha^2+(\beta+1)^2+(\gamma+1)^2 = 1$.
Par la propriété de semi-continuité du jacobien de Clarke, en faisant tendre t *vers* 0, *on obtient*

$$(\alpha,\beta,\gamma) \in \partial\psi(0,0,0).$$

Il résulte de la convexité du jacobien généralisé que

$$\mathcal{O} \in \partial\psi(0,0,0).$$

D'autre part, pour tout $(a,b,c) \neq 0$

$$(\nabla_a\psi(a,b,c))^2 + (\nabla_b\psi(a,b,c))^2 + (\nabla_c\psi(a,b,c))^2 = 1.$$

Par définition du jacobien généralisé de Clarke, on peut conclure que

$$\partial\psi(0,0,0) \in \mathcal{O}.$$

Ceci montre que $\partial\psi(0,0,0) = \mathcal{O}$.

On peut aisément trouver dans la littérature d'autres exemples de fonctions semi-lisses, fonctions semi-lisses d'ordre p et fortement semi-lisses.

Notons également qu'une fonction F de $\mathbb{R}^n$ dans lui-même, localement Lipschitz et sous-analytique est semi-lisse. Ce résultat est dû à Bolte dans [13] et lui a permis de prouver que la méthode de Newton appliquée à une fonction sous-analytique converge

super-linéairement. Par ailleurs, celui-ci est fondé sur un concept lié aux fonctions appelé *Tame functions*.
Les *Tame* vérifient une propriété dont nous aurons besoin dans le cadre d'une remarque à faire sur notre extension relative au cas semi-lisse de la partie univoque. Notons qu'elle a déjà été utilisé par J. P. Dedieu dans [23].

Les définitions et la propriété suivantes sont extraites de l'article de Dedieu [23].

1.7.4 Fonctions sous-analytiques

Commençons par rappeler la notion de fonction semi-analytique.

Définition 1.7.9 *Une partie X de $\mathbb{R}^n$ est semi-analytique si pour tout point a de $\mathbb{R}$, il existe un voisinage U de a et des fonctions analytiques $f_{i,j}$ sur U telles que*

$$X \cap U = \cap_{i=1}^{r} \cup_{j=1}^{s} \{x \in U \ / \ f_{i,j} \varepsilon_{i,j} 0\}$$

où $\varepsilon_{i,j} \in \{<, >, =\}$.

On dit que X est semi-algébrique quand $U = \mathbb{R}^n$ et $f_{i,j}$ sont polynômiales.

Définition 1.7.10 *Un partie X de $\mathbb{R}^n$ est sous-analytique si tout $a \in \mathbb{R}^n$ admet un voisinage U tel que $X \cap U$ est une projection d'un ensemble relativement compact semi-analytique : il existe une partie A borné semi-analytique de $\mathbb{R}^{n+p}$ tels que $X \cap U = \Pi(A)$ où $\Pi : \mathbb{R}^{n+p} \to \mathbb{R}^n$ est la projection.*

Définition 1.7.11 *Soit X une partie de $\mathbb{R}^n$. Une fonction $f : X \to \mathbb{R}^n$ est sous-analytique (resp. semi-analytique) si son graphe est sous-analytique (resp. semi-analytique).*

Observons que l'intersection et la réunion d'une famille finie d'ensembles sous-analytiques est sous-analytique, et la fermeture d'un ensemble sous-analytique est sous-analytique.

Nous allons énoncer la proposition mentionnée précédemment. Elle a été à la base de la convergence super-linéaire de la méthode de type Newton proposée par Cabuzel et al. [21].

Proposition 1.7.12 *Soit f une fonction localement Lipschitz et sous-analytique définie sur $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^n$. Il existe un nombre rationnel γ et une constante C_x tels que*

$$\|f(y) - f(x) - \Delta f(y)(y - x)\| \leq C_x \|y - x\|^{1+\gamma}. \tag{1.7}$$

Nous achevons cette section avec une série d'exemples relatifs aux différentes notions que nous venons de voir.
Nous débutons avec des exemples intéressants de fonctions sous-analytiques ayant rapport avec l'optimization.

Exemple 1.7.5 *Si X est une partie sous-analytique de $\mathbb{R}^n$, alors la fonction-distance $d(x,X) := \min_{z\in\bar{X}} |x-z|$ est sous-analytique.*
En effet, supposons que X soit relativement compact.
Soit $A := \{(x,z,y) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n : z \in \bar{X},\ y \geq |x-z|\}$. Alors A est sous-analytique. Soit Π la projection $\Pi(x,z,y) = (x,y)$. Donc l'ensemble $A = \{(x,y) \in \mathbb{R}^n \times \mathbb{R} : y \geq d(x,X)\} = \Pi(A)$, dont le complémentaire est le graphe de la fonction-distance, est sous-analytique.
Ainsi, la fonction-distance est sous-analytique par complémentarité d'une partie sous-analytique (voir Théorème 3.10 de [10]).

Exemple 1.7.6 *Soient X et T des sous-ensembles sous-analytique de $\mathbb{R}^n$ et $\mathbb{R}^m$ où T est compact, si $f : X \times T \to \mathbb{R}$ est sous-analytique et continue, alors $g(x) = \min_{t\in T} f(x,t)$ est sous-analytique.*

Pour d'autres exemples et propriétés de fonctions sous-analytiques et semi-analytiques, le lecteur peut se référer à [10] et [23] et leurs références.

Nous savons que dans l'étude des fonctions de variables réelles, la notion de dérivée joue un rôle clé, notamment dans la recherche d'extrema. Lorsque l'on passe dans un cadre plus général des fonctions agissant dans des espaces vectoriels normés par exemple, les notions de différentiabilité généralisent bien la dérivabilité.
De même, dans l'étude d'extrema d'applications multivoques, la notion de différentiabilité est essentielle, mais celle-ci a été approchée différemment. Nous ne disposons pas de techniques générales pour calculer la différentielle. Elle est souvent définie en fonction de la théorie que l'auteur veut établir. Nous sommes donc tenus à le justifier à travers une série d'exemples de dérivées multivoques déjà rencontrées dans la littérature.

1.8 Différentiabilité des applications multivoques

Les applications multivoques ont de nombreuses applications intéressantes dans les problèmes de contrôle et en théorie d'équations contingentes (par exemple, voir [6]-[7], [16] - [17], [30], [34]-[36], [51]-[52], [57], [58] et [71]), en économie[[1], [22]], et dans diverses branches de l'analyse (par exemple dans l' étude des sous-différentiels des fonctions convexes [62]). Beaucoup de théories relatives aux intégrations d'applications multivoques ont été assez développées et leurs applications en théorie de contrôle et en économie ont été traitées [1, 16, 22, 36, 35, 34, 43, 44, 70, 71]. Plusieurs théories ont été donc développées autour de La différentiation généralisée dans le cadre multivoque. En effet, H.T. Banks et M.Q. Jacobs ont introduit dans[8] la notion de $\pi-$différentiabilité. La lettre π désignant une isométrie de l'ensemble $\mathcal{B}$ des parties non-vides, convexes,

fermées et bornées de l'espace de Banach réflexif F dans $\mathbb{B}(F)$ l'espace quotient $\mathcal{B}^2/\mathcal{R}$ où $\mathcal{R}$ est la relation déquivalence définie sur $\mathcal{B}^2 = \mathcal{B} \times \mathcal{B}$ par

$$(A, B)\mathcal{R}(C, D) \quad \Leftrightarrow \quad A + D = B + C.$$

Notons que l'existence de cette isométrie est due à Rådström [80]. Cette dérivée est une multiapplication à valeurs dans l'espace quotient $\mathbb{B}(F)$; elle est aussi stable par composition de multiapplications. Par ailleurs, Banks et al. ont montré ce concept est une condition nécessaire et non suffisante pour qu'une application multivoque $\Omega : I \subset \mathbb{R} \to \mathcal{B}(\mathbb{R}^n)$ soit différentiable au sens de Hukuhara [36] ; puis, ils ont appliqué cette notion aux multiapplications définies par une intégrale au sens de Debreu [22]. Soulignons que la différentielle au sens de Hakuhara D_h de $\int_a^t F(s)ds$ est égale à $F(t)$.

Il est bien connu qu'en théorie d'optimisation univoque, les concepts de dérivée et de dérivée directionnelle d'une fonction sont des outils pratiques pour établir des conditions nécessaires et suffisantes d'optimalité du premier ordre.
Ces dernières années, plusieurs recherches ont été effectuées en optimization multivoque. Celles-ci, entre autres, visent à étendre ces concepts aux multiapplications. En tout premier lieu, on prend un point du graphe de la multiapplication et on associe une autre multiapplication dont le graphe est en quelque sorte le cône tangent au graphe de la multiapplication originale au point en question ; c'est la *dérivée contingente* introduite par Jean-Pierre Aubin [2, 4]. Notons que la *codérivée* d'une application multivoque est définie de façon similaire en remplaçant cône tangent par cône normal, voir [4, 68]. Le concept de cône tangent est dû à Pierre de Fermat dès la première moitié du $XVII^e$ siècle [4].
Soulignons également que cette dérivée est un processus convexe fermé, puis la dérivée contingente de la restriction d'une application Fréchet-différentiable à une partie K est la restriction de la dérivée au cône tangent de K.
Plus tard, différents concepts de dérivée directionnelle associés aux relations d'ordre sur des ensembles ont été introduits ; ces relations servent également à définir les solutions optimales. Par exemple, Kuroiwa dans [56], Hoheisel Kanzow, Mordukhovich et Phan [40], Jahn [49].
Plus près de nous, en 2017, Troung Xuan Duc Ha [41] définit une dérivée directionnelle basée sur une distance de type Hausdorff par rapport à l'ordre partiel induit par un cône pointé. Cette distance combinée avec des relations d'ordre d'ensembles définies dans [47, 55] lui a permis d'établir des conditions nécessaires et suffisantes d'optimalité du premier ordre ainsi que des propriétés similaires à celles des applications (univoques).
Dans [42] Ioffe définit une classe d'applications multivoques, positivement homogènes et à valeurs fermées pour approximer localement une application non lisse (non différentiable), ce sont les *prédérivées*. Notons aussi que les dérivées strictes et usuelles de Fréchet sont respectivement les dérivées et prédérivées strictes au sens de Ioffe. De plus, les dérivées

de Warga [90] et fans de Halkin [37] appartiennent également à cette classe. En outre, il a mentionné, entre autres, que la dérivée contingente (ou dérivée graphique) introduite par Jean-Pierre Aubin (voir [2, 4]) et la dérivée faible coincident (si elle existe). Pour une étude plus approfondie sur ce sujet, nous pouvons nous référer à [42] où les relations entre les concepts précédents et plusieurs autres notions de dérivées généralisées, telles que les gradients généralisées de Clarke [18] et des dérivées (codérivées) introduites par Mordukhovich [64] sont explorés en détail.
Inspirés des travaux de Ioffe, Pascaline et al.[74] ont utilisé un nouveau concept de différentia-tion : la H- différentiabilité dont la stabilité a été étudiée en utilisant la convergence de Fisher ; celle-ci est une notion relative aux applications univoques mais adapté par ces derniers au cadre multivoque.
Notons également le nouveau concept introduit par F. S. De Blasi dans [9] relative aux applications multivoques à valeurs bornées F ; la différentielle D_x de F en x est une application multivoque semi-continue supérieurement et homogène, à valeurs convexes, fermées et bornées telle que

$$d(F(x+h)), F(x) + D_x(h)) = o(h) \text{ quand } \|h\| < \delta.$$

Il a démontré l'unicité d'une telle multiapplication (si elle existe) ainsi que sa différentielle est une condition nécessaire et non suffisante de celle de Martelli et Vignoli dans [63]. Plus près de nous, C. H. Jeffrey Pang dans [77] a défini les notions de $T-$ différentiabilité intérieure, $T-$ différentiabilité extérieure ainsi que la stricte $T-$ différentiabilité avec T une application multivoque positivement homogène. Pour une application univoque les trois notions coincident.
En 2005, Khadra Nachi et Jean-Paul Penot ont introduit divers concepts de différentiabilité notamment celle qui est similaire à celle définie par Fréchet ; c'est une application univoque linéaire et continue. Ils ont construit cette dérivée dans le but d'établir un théorème d'inversion relative aux applications multivoques ainsi que son application à la résolution d'une inclusion différentielle. Soulignons que ce type de différentielle n'est pas unique et généralise évidemment la notion de différentiabilité au sens de Fréchet pour une application univoque.
Cette liste est loin d'être exhaustive !
Nous allons adopter celle définie par Nachi et Penot qui jouera un rôle fondamental tout au long de cette thèse. Nous verrons que l'intérêt de cette dérivée est qu'elle possède des propriétés similaires à la dérivée de Fréchet, à savoir la linéarité et la continuité ; ce sont des propriétés qui seront utiles à la construction de nos différentes extensions.

1.8.1 Dérivée de Nachi-Penot

Définition 1.8.1 *Considérons X_0 un ouvert de X, une application multivoque $F : X_0 \rightrightarrows Y$ avec $X_0 \subset \mathrm{dom}(F)$, et $x_0 \in X_0$. F est dite différentiable au sens de Nachi et*

Penot en x_0 si F est semi continue inférieurement en x_0 et s'il existe une application linéaire continue (univoque) $A \in \mathcal{L}(X,Y)$ telle que la fonction o définie

$$e(F(x_0+x), F(x_0)+A(x)) = o(x) \text{ est négligeable par rapport à } \|x\|. \tag{1.8}$$

Alors, $A \in \mathcal{L}(X,Y)$ est appelée une N-P dérivée de F en x_0 et l'ensemble des N-P dérivées de F en x_0 est noté $\mathcal{D}F(x_0)$.

Rappelons qu'une fonction $o : X \to \mathbb{R}$ est *négligeable* si $\lim_{x\to 0, x\neq 0} \frac{o(x)}{\|x\|} = 0$, ou de manière équivalente s'il existe une fonction module $\mu : \mathbb{R}^+ \to \mathbb{R}^+ \cup \{+\infty\}$, continue en 0 avec $\mu(0) = 0$ et telle que $\|o(x)\| \leq \mu(\|x\|)\|x\|$; sans perte de généralité, on peut supposer que μ est croissante. Donc, une formulation équivalente à (1.8) est

$$F(x_0+x) \subset F(x_0) + A(x) + \mu(\|x\|)\|x\|\mathbb{B}_Y.$$

Nous disons que F est N-P-différentiable sur X_0 si et seulement si elle est N-P-différentiable en tout point $x_0 \in X_0$.
Si F était univoque, la définition précédente serait exactement celle définie pour la différentiabilité au sens de Fréchet.
Nous pouvons en déduire que la notion dérivée au sens de Nachi-Penot généralise celle au sens de Fréchet.

Dans la suite, nous abrégerons "différentiable au sens de Nachi et Penot" en "N-P-différentiable" ou plus simplement "différentiable". Notons que cette abréviation ne s'emploiera que pour une application multivoque.
Par contre, pour une application univoque "différentiable", nous utiliserons "Fréchet différentiable".

Nous poursuivons avec une séries d'exemples relatifs aux calculs de dérivée de Nachi-Penot.

Exemple 1.8.1 *Soit $F : \mathbb{R} \rightrightarrows \mathbb{R}$ définie par $F(x) = [a-x^2, b+x^2]$ pour tout $x \in \mathbb{R}$, avec $a \leq b$ dans $\mathbb{R}$. Alors F est différentiable en 0 ; pour $x_0 \neq 0$, elle n'est pas différentiable en x_0.*

Preuve
Observons d'abord que pour tout $x \in \mathbb{R}$, $F(x) = [a,b] + x^2[-1,1]$. En effet, pour $y \in [a-x^2, b+x^2]$, il existe un réel $t \in [0,1]$ tel que $y = (1-t)(a-x^2) + t(b+x^2)$, soit $y = (1-t)a + tb + x^2[(1-t)(-1) + t(1)]$.
D'où $y \in [a,b] + x^2[-1,1]$.
Ceci prouve que $[a-x^2, b+x^2] \subset [a,b] + x^2[-1,1]$. $\square$

L'autre inclusion s'obtient aisément en rebroussant chemin.

Montrons que F est semi-continue inférieurement(s.c.i.) en 0.
Soient $\overline{y} \in F(0)$ et une suite (x_n) convergeant vers 0. Pour tout x, $F(0) = [a,b] \subset [a,b] + x^2[-1,1] = F(x)$. On a donc $\overline{y} \in [a,b] + x^2[-1,1]$ pour tout x. En particulier, $\overline{y} \in [a,b] + x_n^2[-1,1] = F(x_n)$ pour tout n. Prenons la suite constante $y_n = \overline{y}, \forall n$. Ainsi $y_n \in F(x_n)$ et $\lim\limits_{n\to\infty} y_n = \overline{y}$.
Ceci montre que pour tout $\overline{y} \in F(x_0)$ et quelle que soit la suite (x_n) convergeant vers 0, il existe une suite d'éléments $y_n \in F(x_n)$ qui converge vers $\overline{y}$. L'application multivoque est donc s.c.i. en 0.
Nous abordons le deuxième volet de la définition de la N-P-différentiabilité.
D'après ce qui précède, pour tout $x \in \mathbb{R}$, on a

$$F(x) = [a - 0^2, b + 0^2] + x^2[-1,1] = F(0) + x^2[-1,1] = F(0) + |x||x|[-1,1].$$

On a donc l'inclusion

$$F(0+x) \subset F(0) + 0(x) + \mu(|x|)|x|\bar{\mathbb{B}}_{\mathbb{R}},$$

avec $\mu : \mathbb{R}_+ \to \mathbb{R}_+, x \mapsto \mu(x) = |x|$, $\bar{\mathbb{B}}_{\mathbb{R}} = \{x \in \mathbb{R} \ / \ |x| \leq 1\} = [-1,1]$.
D'où la N-P-différentiabilité de F en 0 avec $DF(0) = 0$.
Pour tous $x_0 \neq 0$, $h \in \mathbb{R}$,

$$F(x_0 + h) = [a - (x_0 + h)^2, b + (x_0 + h)^2]$$

$$F(x_0 + h) = [a,b] + x_0^2[-1,1] + 2x_0h[-1,1] + h^2[-1,1].$$

S'il existait une telle application linéaire $A(x_0)$, on aurait $A(x_0)h = 2x_0ht$ avec $t \in [-1,1]$. Ainsi, pour $h = t$, $A(x_0)h = 2x_0t^2$ est non linéaire (contradiction). □

Exemple 1.8.2 *Plus généralement, soit $g : X \to Y$, $r : X \to \mathbb{R}$ différentiables en x_0 et r à valeurs positives. Alors l'application multivoque $F : X \rightrightarrows Y$ définie par $F(x) := g(x) + r(x)\overline{\mathbb{B}}_Y$ est N-P-différentiable en x_0.*

Exemple 1.8.3 *Soit $g : X \to Y$ une application différentiable en x_0. Alors pour tout sous-ensemble C de Y l'application multivoque $F : X \rightrightarrows Y$ définie par $F(x) := g(x)+C$ est différentiable en x_0.*

Ces trois exemples précédents sont extraits de l'article de Nachi et Penot [69].

Exemple 1.8.4 *Soit $f : \mathbb{R} \to \mathbb{R}$ Fréchet-différentiable, de dérivée ∇f,*
l'application multivoque $F : \mathbb{R} \rightrightarrows \mathbb{R}$ définie par $F(x) = [f(x),\ f(x) + 1]$ est N-P-différentiable en tout x_0 de $\mathbb{R}$.
En effet, comme $F(x) = [0,1] + \{f(x)\}$, s'inspirant de l'un des exemples précédents, F est donc N-P-différentiable avec $A(x_0) = \nabla f(x_0)$.

Nous savons que dans le cadre univoque, la dérivée (si elle existe) est toujours une application univoque, mais peut être multivoque pour une application multivoque. Nous allons justifier cela à travers les dérivées mentionnées dans l'introduction de cette sous-section. Nous tenons à préciser que cette liste sera loin d'être exhaustive !

1.8.2 Quelques dérivées multivoques

Nous commençons par présenter des approches géométriques ou graphiques de la dérivée d'une application multivoque. Considérons dans un premier temps celle d'Aubin. Ce dernier a utilisé le concept de cône tangent au graphe. Rappelons donc la définition du cône tangent.

Définition 1.8.2 *Soit C un sous-ensemble d'un espace normé X. On considère une suite (x_k) dans C qui converge vers x et une suite de réels strictement positifs (τ_k) qui décroît vers 0. Un vecteur v est dit tangent à C en x si*

$$\frac{1}{\tau_k}(x_k - x) \to v \ \textit{lorsque}\ k \to +\infty.$$

L'ensemble des vecteurs tangents à C en x est appelé le cône tangent(ou cône tangent de Bouligand) à C et est noté $T_C(x)$. Lorsque $x \notin C$, $T_C(x)$ est l'ensemble vide.
Une description équivalente à la définition est que $v \in T_C(x)$ si et seulement si il existe des suites de vecteurs v_k convergeant vers v et $\tau_k \searrow 0$ avec $x + \tau_k v_k \in C$, ou de manière équivalente, il existe des suites $x_k \in C$, $x_k \to x$ telles que $v_k := (x_k - x)/\tau_k \to v$ quand $k \to +\infty$.

On peut vérifier que si v est dans C ouvert, alors $T_C(v) = X$; si v est adhérent à C, alors $T_C(v) = \emptyset$.
Vérifions que si v est dans C ouvert, $T_C(v) = X$.
Soit $h \in X$. Montrons que $h \in T_C(v)$.
$v \in C$ ouvert, il existe $r > 0$ tel que $\mathbb{B}_r(v) \subset C$.
Choisissons $v_n = v + \lambda_n h$, pour tout n, et une suite (λ_n) de réels strictement positifs telle que $\lim\limits_{n\to\infty} \lambda_n = 0$.
La suite $(v_n)_n$ étant convergente vers v, il existe un rang N_0 tel qu'on ait $v_n \in \mathbb{B}_r(v)$ dès que $n \geq N_0$.
Considérons la suite $y_n := v_n,\ n \geq N_0$, et $t_n := \frac{1}{\lambda_n},\ n \in \mathbb{N}$.
Or, la suite (y_n) est une suite d'éléments de C, et $\lim\limits_{n\to\infty} t_n(y_n - v) = \lim\limits_{n\to\infty} \dfrac{1}{\lambda_n}(v + \lambda_n - v) = \lim\limits_{n\to\infty} h = h$.
D'où $h \in T_C(v)$; ce qui achève la vérification.
Notons également que $T_C(x)$ n'est pas nécessairement convexe (mais il est convexe si C est convexe). En voici un contre-exemple (Exemple (1.8.5)).

Exemple 1.8.5 *L'ensemble non convexe* $C := \{(x, y) : x^2 \leq y^2\}$ *a son cône tangent en* $(0, 0)$ *qui s'écrit* $T_C(0, 0) = \{(d_1, d_2) \in \mathbb{R}^2 : |d_1| \leq |d_2|\}$ *et qui n'est pas un ensemble convexe.*
Par contre, le cône tangent défini par Clarke est nécessairement un cône convexe fermé contenant l'origine [66].
Remarquons également que $T_C(0, 0)$ *est la réunion des graphes des applications multivoques* $F(x) =]-\infty, -|x|] \cup [|x|, +\infty[$ *et*

$$G(x) = \begin{cases} \{x\} \cup \{-x\} & si\ x \leq 0, \\ [-x, x] & si\ x \geq 0. \end{cases}$$

Exemple 1.8.6 $C = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 \leq x_2 \leq 2x_1^2\}$, *alors* $T_C(0, 0) = \mathbb{R} \times \{0\}$.

Nous pouvons maintenant définir la dérivée contingente introduite par Pang [48].

Définition 1.8.3 *(Dérivée graphique) Soient* X *et* Y *deux* $\mathbb{R}$*-espaces normés. Pour une application multivoque* $S : X \rightrightarrows Y$ *telle que* $(\bar{x}, \bar{y}) \in gph(S)$*, la dérivée graphique, appelée aussi dérivée contingente, est notée par* $DS(\bar{x}\ \bar{y}) : X \rightrightarrows Y$ *et définie par*

$$gph(DS(\bar{x},\ \bar{y})) = T_{gph(S)}(\bar{x}, \bar{y}).$$

Donc, $v \in DF(\bar{x}, \bar{y})(u)$ si et seulement si il existe des suites $u_k \to u$, $v_k \to v$ and $\tau_k \searrow 0$ telles que $\bar{y} + \tau_k v_k \in F(x + \tau_k v_k)$ pour tout k.

Exemple 1.8.7 *Soit l'application multivoque* $S : \mathbb{R} \rightrightarrows \mathbb{R}$ *définie par*

$$S(x) :=]-\infty, -|x|] \cup [|x|, +\infty[.$$

la dérivée graphique (ou dérivée contingente) $D(S(0, 0))$ en $(0; 0)$ est l'application multivoque définie par $gph(DS(0, 0)) = T_{gph(S)}(0, 0)$ avec

$$T_{gph(S)}(0, 0) = \{(x, y) : |x| \leq |y|\}.$$

Exemple 1.8.8 *Soit* $F : \mathbb{R} \rightrightarrows \mathbb{R}$ *définie par*

$$F(x) := \begin{cases} [\sqrt{1 - (x-1)^2}, +\infty[, & pour\ 0 \leq x \leq 1\ , \\ [\sqrt{1 - (x+1)^2}, +\infty[, & pour\ -1 \leq x \leq 0\ , \\ \emptyset & sinon\ . \end{cases}$$

La dérivée contingente en $(0, 0)$ *est donc*

$$DF(0, 0)(u) = \begin{cases} [0, +\infty[& pour\ u = 0 \\ \emptyset & pour\ u \neq 0\ . \end{cases}$$

Une autre approche géométrique très présente dans la littérature est celle de la codérivée de Mordukhovich. Par soucis de simplicité, nous allons nous restreindre aux applications multivoques de $\mathbb{R}^n$ dans $\mathbb{R}^m$. Pour cette approche, nous rappelons d'abord le concept de cône normal.

Définition 1.8.4 *Pour un ensemble $C \subset \mathbb{R}^n$, le cône normal régulier en $\bar{x}$ est défini par*

$$\hat{N}_C(\bar{x}) := \{y \ / \ \langle y, x - \bar{x}\rangle \leq o(\|x - \bar{x}\|)\| \ \textit{pour tout} \ x \in C\}.$$

La limite du cône normal $N_C(\bar{x})$ est définie comme $\limsup_{x \underset{C}{\to} \bar{x}} \hat{N}_C(x)$ *ou comme*

$$N_C(\bar{x}) = \Big\{y \ / \ \ \textit{il existe} \ {}^{x_i \underset{C}{\to} \bar{x}}, y_i \in \hat{N}_C(x_i) \ \textit{telle que} \ y_i \to y\Big\}.$$

Le symbole ${}^{x_i \underset{C}{\to} \bar{x}}$ signifie que $x_i \to x$ avec $x_i \in C$.
On déduit immédiatement de la définition que $\hat{N}_C(\bar{x})$ avec C convexe est un cône convexe fermé, qui est réduit à $\{0\}$ si $x \in int(\Omega)$. Une propriété remarquable d'un cône normal d'un ensemble convexe non vide en dimension finie est que $\hat{N}_C(\bar{x}) \neq \{0\}$ si et seulement si $\bar{x}$ est un point du bord de Ω. Par ailleurs,

$$v \in \hat{N}_C(\bar{x}) \quad \Leftrightarrow \quad \langle v, w\rangle \leq 0 \ \text{pour tout} \ w \in T_C(\bar{x}).$$

Exemple 1.8.9 *Soient les ensembles de $\mathbb{R}^2$, $\Omega_1 := \{(x_1, x_2) \in \mathbb{R}^2 : x_2^2 \geq -|x_1|\}$ et $\Omega_2 := \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \leq x_2\}$.*

$N_{\Omega_1}((0,0)) = \{(v, v) \ / \ v \leq 0\} \cup \{(v, -v) \ / \ v \geq 0\}$, $\hat{N}_{\Omega_1}(0;0) = \{0\}$ et $N_{\Omega_2}(0,0) = \{(v, -v) \ / \ v \geq 0\}$.

Exemple 1.8.10 *Soit l'ensemble $\Omega := gph|x|$.*
$N_\Omega(0,0) := \{(v_1, v_2) \mid v_2 \leq -|v_1|\} \cup \{(v_1, v_2)| \ v_2 = |v_1|\}$.

Définition 1.8.5 *(Codérivée) Pour une application multivoque, la codérivée de Mordukhovich en $(\bar{x}, \bar{y}) \in gphS$ notée $D^*S(\bar{x} \mid \bar{y}) : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ et définie par*

$$v \in D^*S(\bar{x} \mid \bar{y})(u) \quad \Leftrightarrow \quad (v, -u) \in N_{gph(S)}(\bar{x}, \bar{y}).$$

Ceux qui sont intéressés aux cas généraux (i.e dans un espace de Banach quelconque) pourront consulter l'ouvrage de Mordukhovich [68].

Exemple 1.8.11 *Soit F l'application multivoque de $\mathbb{R}$ dans $\mathbb{R}$ définie par*

$$F(x) := [\max\{x, 0\}, +\infty[.$$

Si $x > 0$, $F(x) = [x, +\infty[$, *alors* $gph(F) = \{(u, v) \ / \ v \geq u\}$.

$$N_{gph(F)}(x, y) = \begin{cases} \{(t, -t) \mid t \geq 0\} & si\ y = x \\ \{(0, 0)\} & si\ y > x. \end{cases}$$

Comme $v \in D^*(x \mid y)u \quad \Leftrightarrow \quad (v, -u) \in N_{gph(F)}(x, y)$, *le graphe de* $D^*(x \mid y)$ *est* $\{(u, u),\ u \geq 0\}$ *si* $y = x$ *et* $\{(0, 0)\}$ *si* $y > x$.

Nous poursuivons avec la notion de différentiation généralisée de Pang, la T-différentiabilité. Cette définition nécessite la notion de multiapplications positivement homogènes.
Ces dernières jouent un rôle très important en analyse variationnelle, plus particulièrement dans la théorie de la différentiation généralisée dans le cadre multivoque où la plupart de dérivées introduites le sont. De telles applications possèdent des propriétés intéressantes qui permettent à plusieurs auteurs de généraliser nombreux résultats de l'analyse classique dont les théorèmes des fonctions inverses et implicites.

Définition 1.8.6 *Soient* X *et* Y *deux espaces de Banach.*
Une multiapplication $T : X \rightrightarrows Y$ *est positivement homogène si* $0 \in T(0)$ *et* $T(\lambda x) = \lambda T(x)$ *pour tout* $\lambda > 0$, *ou de manière équivalente* $gph(T)$ *est un cône.*

Remarquons qu'un processus convexe est positivement homogène mais l'inverse est fausse.

Définition 1.8.7 *(*T*-différentiabilité) Soit* $T : X \rightrightarrows Y$ *une multiapplication positivement homogène.*

(a) On dit que $S : X \rightrightarrows Y$ *est extérieurement* T*-différentiable en* $\bar{x}$ *si pour tout* $\delta > 0$, *il existe un voisinage* V *de* $\bar{x}$ *tel que*

$$S(x) \subset S(\bar{x}) + T(x - x^{'}) + \delta|x - x^{'}|\mathbb{B} \ \textit{pour tous} \ \ x,\ x' \in V.$$

S est dite intérieurement T*-différentiable en* $\bar{x}$ *si la formule précédente est remplacée par*

$$S(\bar{x}) \subset S(x) - T(x - \bar{x}) + \delta|x - \bar{x}|\mathbb{B} \ \textit{pour tout}\ x \in V.$$

Si S est à la fois extérieurement T*-différentiable et intérieurement* T*-différentiable en* $\bar{x}$, *alors S est dite* T*-différentiable.*

(b) On dit que $S : X \rightrightarrows Y$ *est strictement* T*-différentiable en* $\bar{x}$ *si pour tout* $\delta > 0$, *il existe un voisinage* V *de* $\bar{x}$ *tel que*

$$S(x) \subset S(x^{'}) + T(x - x^{'}) + \delta|x - x^{'}|\mathbb{B} \ \textit{pour tous}\ x, x^{'} \in V.$$

(c) On dit que $S : X \rightrightarrows Y$ est pseudo strictement T-différentiable en $(\bar{x}, \bar{y}) \in gph(S)$ si pour tout $\delta > 0$, il existe des voisinages U_δ de $\bar{x}$ et V_δ de $\bar{y}$ tels que

$$S(x) \cap V_\delta \subset S(x^{'}) + T(x - x^{'}) + \delta|x - x^{'}|\mathbb{B}, \quad \forall x, x^{'} \in U_\delta.$$

Remarque 1.8.8 *Lorsque S est univoque, les définitions T-différentiabilité extérieure, T-différentiabilité intérieure sont équivalentes.*

Remarque 1.8.9 *Si S est strictement T-différentiable avec T une application positivement homogène vérifiant $H(x) \subset \kappa\|x\|\mathbb{B}$ pour tout $x \in X$, alors S est lipschitzienne de constante κ.*

Exemple 1.8.12 *Soit $S : \mathbb{R} \rightrightarrows \mathbb{R}$ définie par $S(x) :=]-\infty, x]$, et soit $T : \mathbb{R} \rightrightarrows \mathbb{R}$ définie par*

$$T(w) = \begin{cases} \{0\} & si\ w \leq 0 \\ \{(0,0)\} & si\ w \geq 0. \end{cases}$$

S est strictement T-différentiable en x pour tout $x \in \mathbb{R}$, et elle est strictement T-différentiable en x pour tout $x \in \mathbb{R}$.

Un exemple où la dérivée généralisée T n'est pas unique.

Exemple 1.8.13 *Soit $S : \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ définie par $S(x, y) = \{x\} \times \mathbb{R}^2$. Soient $T_1, T_2 : \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ définie par $T_1(x, y) = \{(x, y)\}$ et $T_2(x, y) = \{(x, 0)\}$. Il est clair que S est pseudo T_1-différentiable strictement et pseudo T_2-différentiable strictement en tout point, mais qu'elle n'est pseudo $(T_1 \cap T_2)$-différentiable nulle part.*

La section sera achevée par la dérivée directionnelle de Ha [41]. Celle-ci est basée sur une distance de type Hausdorff.
Il a utilisé quelques résultats auxiliaires que nous allons rappeler à nouveau.
Soient X et Y deux espaces de Banach. $K \subset Y$ un cône convexe fermé pointé (pointé signifie que $K \cap (-K) = \{0\}$). Le cône K induit une relation d'ordre partiel dans Y : pour tout $y_1, y_2 \in Y$

$$y_1 \leq_K y_2 \Longleftrightarrow y_2 - y_1 \in K.$$

Par soucis de simplicité, nous omettrons l'indice K dans la notation $\leq_K$. Il existe plusieurs relations d'ordre d'ensembles, voir [49, 55] mais nous utiliserons principalement ce qui suit.

Définition 1.8.10 *([60]) Soit $A \subset Y$ un sous-ensemble non-vide et $a \in A$. On dit que (i) A est $K-$borné s'il existe un sous-ensemble non-vide $M \subset Y$ tel que $A \subset M + K$; (ii) A est $K-$compact si tout recouvrement de la forme $\{U_\alpha + K \ / \ \alpha \in I, U_\alpha \ \ ouvert\ \}$ admet un sous-recouvrement fini.*

Définissons maintenant la distance de type-Hausdorff évoquée plus haut. Cette définition passe par la fonction distance Δ_U d'Hiriart-Urruty associée à un sous-ensemble non-vide $U \subset Y$ (voir [38]). Rappelons la définition dans le cas $U = -K$.

$$\Delta_{-K}(y) := d_{-K}(y) - d_{Y\setminus(-K)}(y) = \begin{cases} -d_{Y\setminus(-K)}(y) & \text{si } y \in -K, \\ d_{-K}(y) & \text{sinon.} \end{cases}$$

En voici quelques exemples illustratifs.

Exemple 1.8.14 *(i) Si $K = \{0\}$, $\Delta_{-K}(y) = \|y\|$ pour tout $y \in Y$.*

(ii) Si $Y = \mathbb{R}$ et $K = \mathbb{R}_+$, alors $\Delta_{-K}(y) = y$ et son sous-différentiel $\partial\Delta_{-K}(y) = 1$ pour tout $y \in \mathbb{R}$. Rappelons que la fonction Δ_{-K} est une fonction convexe, positivement homogène et Lipschitz de rapport 1.

(iii) Soit $Y = \mathbb{R}^n$ muni de la norme de la norme euclidienne $\|.\|_2$ et $K = \mathbb{R}^n_+$. On a $d_{-K}(y) = \|y^+\|$, où $y^+ = (y_i^+)$ avec $y_i^+ = \max(y_i, 0), i = 1, ..., n$ et

$$d_{Y\setminus(-K)}(y) = \begin{cases} 0 & \text{si } y_i \geq 0 \text{ pour un certain } i, \\ -\max\limits_i y_i & \text{si } y_i < 0, \ \forall i = 1, ..., n \ . \end{cases}$$

Donc

$$\Delta_{-K}(y) = \begin{cases} \|y^+\| & \text{si } y \notin -K, \\ \max\limits_i y_i & \text{si } y \in -K. \end{cases}$$

(iv) La fonction Δ_{-K} prend une forme similaire si on considère $Y = \mathbb{R}^n$ muni de la norme $\|y\|_\infty = \max\limits_i |y_i|$; dans ce cas on a

$$d_{-K}(y) = \begin{cases} \max\limits_i y_i & \text{si } y \notin -K \\ 0 & \text{si } y \in -K \end{cases}$$

et

$$d_{Y\setminus(-K)}(y) = \begin{cases} 0 & \text{si } y \notin -K, \\ -\max\limits_i y_i & \text{si } y \in -K, \end{cases}$$

et donc,
pour tout $y \in Y$,

$$\Delta_{-K}(y) = \max_i y_i.$$

Soient A, B deux sous-ensembles non-vides de Y. On note

$$h_K(A,B) := \sup_{b\in B}\inf_{a\in A}\Delta_{-K}(a-b).$$

Maintenant nous pouvons définir une distance spéciale dans la famille des sous-ensembles non-vides $K-$bornés.

Définition 1.8.11 *Soient A, B deux sous-ensembles $K-$borné et non-vides de Y. Une distance type-Hausdorff relative à la relation d'ordre induit par K entre A et B, notée $d_K(A,B)$, est définie par*

$$d_K(A,B) := \max\{h_K(A,B),\ h_K(B,A)\}.$$

Remarque 1.8.12
L'appellation "distance type-Hausdorff" vient de la coincidence avec la distance classique de Hausdorff définie par

$$d(A,B) := \max\{\sup_{a\in A} d(a,B), \sup_{b\in B} d(b,A)\}$$

lorsque $K = \{0\}$ puisque dans ce cas $\Delta_{-K}(y) = \|y\|$ pour tout $y \in Y$.

Nous listons quelques propriétés de la fonction $d_K(.,.)$; celle-ci s'abrège en $d(.,.)$.

Proposition 1.8.13 *Supposons que A et B soient des parties non-vides $K-$bornés. Alors*

(i) $d(A,B) = d(B,A)$ *;*

(ii) $d(A,B) = d(A+K, B+K)$ *;*

(iii) $d(\lambda A, \lambda B) = \lambda d(A,B)$ *pour tout* $\lambda \geq 0$ *;*

(iv) l'inégalité triangulaire : pour tout sous-ensemble non-vide $K-$ borné, on a

$$d(A,B) \leq d(A,C) + d(C,B);$$

(v) supposons que A et B soient $K-$compacts. Alors $d(A,B) = 0$ ssi $A + K = B + K$.

Les preuves résultent de l'article de Ha [41].
Le corollaire suivant est nécessaire pour la définition de la dérivée directionnelle.

Corollaire 1.8.14 *Soient A et B deux parties non-vides $K-$compacts de Y. Alors*

$$d(A,B) = d(Min(A), Min(B)).$$

1.8.3 Un concept de dérivée directionnelle

Dans cette sous-section, nous allons énoncer la dérivée directionnelle pour une application multivoque introduite par Ha[41], puis nous énumérons quelques exemples où l'on calcule cette dérivée.
On suppose que F est à **valeurs compactes**. Rappelons que $d \in Y$ est une direction admissible de F en $x \in domF$ si $x + td \in domF$ pour tout $t > 0$ suffisamment petit.

Définition 1.8.15 *Soit $x \in dom\ F$ et d une direction admissible de F en x. On note*

$$W(x,d) := \Big\{A \subset Y/A \ est\ K-compact\ et\ \lim_{t\to 0^+} d\Big(\frac{F(x+td)-F(t)}{t}, A\Big) = 0\Big\}.$$

La dérivée directionnelle $DF(x,d)$ de F en x dans la direction d est définie par

$$DF(x,d) := \begin{cases} Min(A) & pour\ A \in W(x,d)\ si\ W(x,d) \neq \emptyset \\ \emptyset & sinon. \end{cases}$$

On dit que F est la dérivée directionnelle $DF(x,d)$ en x dans la direction d si $DF(x,d) \neq \emptyset$.

Exemple 1.8.15 *Soit $X = \mathbb{R},\ Y = \mathbb{R}^2$ et $K = \mathbb{R}^2_+$.*
(i) Soient

$$F(x) := \begin{cases} \{(x,1),(x,2)\}\ si\ x \neq 0 \\ \{(0,0),(0,1)\}\ si\ x = 0. \end{cases}$$

Nous calculons $DF(x,d)$ en $x = 0$.
Pour $d = 1$ et $t > 0$, on a $F(x+td) = F(t) = \{(t,1),(t,2)\},\ F(t) - F(0) = \{(t,1),(t,2),(t,0)\}$ et

$$A_t := \frac{(F(t) - F(0))}{t} = \{(1,1/t),(1,2/t),(1,0)\}.$$

Il est clair que $Min(A_t) = \{(1,0)\}$. Posons $A := Min(A_t)$. Le corollaire précédent implique que $d(A_t,A) = d(Min(A_t),Min(A_t))) = 0$. Par suite, $DF(0,1) = \{(1,0)\}$.
Fixons $d = -1$ et $t > 0$. Donc $F(x+td) = F(-t) = \{(-t,1),(-t,2)\}, F(t) - F(0) = \{(-t,1),(-t,2),(-t,0)\}$ et

$$A_t := (F(t) - F(0))/t = \{(-1,1/t),(-1,2/t),(-1,0)\}.$$

Il s'en suit que $Min(A_t) = \{(-1,0)\}$. En posant $A := Min(A_t)$, puis en vertu du corollaire précédent $d(A_t,A) = d(Min(A_t),Min(A_t)) = 0$. On en déduit que $DF(0,-1) = \{(-1,0)\}$.

ii) Considérons l'application multivoque

$$F(x) := \begin{cases} \{(2x^2, 1), (3x, 2)\} & si\ x > 0 \\ \{(0,0), (0,1)\} & si\ x = 0 \\ \{(x^4, 1), (-x^3, 2)\} & si\ x < 0. \end{cases}$$

Calculons $DF(x, d)$ *en* $x = 0$ *dans la direction* $d = 1$.
Pour $t > 0$, *on a* $F(x+td) = F(t) = \{(2t^2, 1), (3t, 2)\}$, $F(t)-F(0) = \{(2t^2, 0), (3t, 2), (2t^2, 0), (3t, 1)\}$ *et*

$$A_t := (F(t) - F(0))/t = \{(2t, 1/t), (3, 2/t), (2t, 0), (3, 1/t)\}.$$

Évidemment $Min(A_t) = \{(2t, 0)\}$ *pour* $t < 1$.
Posons $A := \{(0,0)\}$. *Un raisonnement similaire nous permet d'écrire*

$$d(A_t, A) = d(Min(A_t), Min(A)) = 2t.$$

Ainsi $DF(0, 1) = \{(0,0)\}$.
Examinons le cas où $d = -1$ *et* $t > 0$.
Comme $F(x+td) = F(-t) = \{(t^4, 1), (-t^3, 2)\}$, $F(-t)-F(0) = \{(t^4, 1), (t^3, 2), (t^4, 0), (t^3, 1)\}$ *et*

$$A_t := (F(t) - F(0))/t = \{(t^3, 1/t), (t^2, 2/t), (t^3, 0), (t^2, 1/t)\},$$

on en déduit que $Min(A_t) = \{(t^3, 0)\}$ *pour* $t < 1$. *Posons* $A := \{(0,0)\}$. *En vertu du corollaire précédent,* $d(A_t, A) = d(Min(A_t), A)) = t^3$. *Par conséquent,* $DF(0, -1) = \{(0,0)\}$. □

1.8.4 π-différentiabilité de multifonctions

Nous achevons notre périple sur la différentiabilité de multifonctions avec la $\pi-$différentiablité. Elle se basait sur un théorème établi par Rådström [80] selon lequel si F est un espace de Banach réflexif (sur le corps des réels $\mathbb{R}$), il existe un espace normé $\mathbb{B}(F)$ et une isométrie $\pi : \mathcal{B} \to \mathbb{B}(F)$, où $\mathcal{B}$ est une métrique pour la distance de Hausdorff d_H, tels que $\pi((\mathcal{B}))$ soit un cône convexe de $\mathbb{B}(F)$.
Soulignons par ailleurs que $\mathcal{B}$ désigne l'ensemble des parties non-vides, convexes, fermées et bornées de l'espace réflexif F.

Comme précédemment, nous tenons à rappeler en détail certains préliminaires nécessaires à la description de l'espace $\mathbb{B}(F)$, nous passons ensuite à la définition de la $\pi-$ différentiabilité et d'en exhiber quelques exemples.

Une relation d'équivalence $\mathcal{R}$ est définie sur $\mathcal{B}^2 = \mathcal{B} \times \mathcal{B}$ par

$$(A, B)\mathcal{R}(C, D) \Longleftrightarrow A + D = B + C.$$

La classe d'équivalence du couple (A, B) est notée $\langle A, B\rangle$. L'espace quotient $\mathcal{B}^2/\mathcal{R}$, noté $\mathbb{B}(F)$, est muni d'une addition définie par

$$\langle A, B\rangle + \langle C, D\rangle \equiv \langle A + C, B + D\rangle,$$

et si $\alpha \geq 0$, alors $\alpha\langle A, B\rangle \equiv \langle \alpha A, \alpha B\rangle$ tandis que si $\alpha < 0$, alors

$$\alpha\langle A, B\rangle \equiv \langle |\alpha|B, |\alpha|A\rangle.$$

L'addition et la multiplication par scalaires ainsi définies confèrent à $\mathbb{B}(F)$ une structure d'espace vectoriel sur le corps des réels. L'application $\pi : \mathcal{B} \to \mathbb{B}(F)$ est définie par $\pi(A) \equiv \langle A, 0\rangle$, $A \in \mathcal{B}$, i.e., $\langle A, 0\rangle$ est la classe d'équivalence $\{(A+D, D)/D \in \mathcal{B}\}$. Nous adopterons la convention d'écriture $\pi(A)$ par $\hat{A}$ quand A est un élément de $\mathcal{B}$, et donc le cône convexe $\pi(\mathcal{B}) = \hat{\mathcal{B}}$. Une métrique $\mathfrak{d}_H$ sur $\mathbb{B}(F)$ est définie par

$$\mathfrak{d}_H(\langle A, B\rangle), \langle C, D\rangle) \equiv d_H(A + D, B + C).$$

L'élément neutre de $\mathbb{B}(F)$ est la classe d'équivalence $\{(D, D)/D \in \mathcal{B}\}$, notée $\langle 0, 0\rangle$. Comme la distance de Hausdorff est invariante par translation et positivement homogène, la relation $\|\langle A, B\rangle\| = \mathfrak{d}(\langle A, B\rangle), \langle 0, 0\rangle)$ définie une norme sur $\mathbb{B}(F)$ telle que

$$\mathfrak{d}(\langle A, B\rangle), \langle C, D\rangle) = \|\langle A, B\rangle - \langle C, D\rangle\|.$$

Rappelons qu'une fonction $f : E \to F$, où E et F sont deux espaces normés sur $\mathbb{R}$ est dite égale $o(\|h\|)$ si $\|f(x)\|/\|h\|$ quand $\|h\| \to 0$.

Définition 1.8.16 *(π-différentiabilité) Une multifonction, $\Omega : G \to \mathcal{B}(F)$, où G est un sous-ensemble ouvert d'un $\mathbb{R}$-espace vectoriel normé E, est dite π-différentiable en $x_0 \in G$ si la fonction, $\hat{\Omega} : G \to \mathbb{B}(F), x \to \hat{\Omega}(x)$, $x \in G$ est différentiable en $x_0 \in G$. La multiapplication Ω est π-différentiable sur G si elle l'est en tout point de G.*
Donc Ω est π-différentiable en $x_0 \in G$ signifie qu'il existe une fonction linéaire continue $\hat{\Omega}'(x_0) : E \to \mathbb{B}(F)$ telle que

$$\hat{\Omega}(x) - \hat{\Omega}(x_0) - \hat{\Omega}'(x_0)(x - x_0) = o(\|x - x_0\|). \tag{1.9}$$

Remarquons que si $\hat{\Omega}'(x_0)(\Delta x) = \langle A_{\Delta x}, B_{\Delta x}\rangle, \Delta x = x - x_0 \in E$, et $A_{\Delta x}, B_{\Delta x} \in \mathcal{B}$, alors en termes de la métrique de Hausdorff, (1.9) équivaut à

$$d_H(\Omega(x) + B_{x-x_0}, \Omega(x_0) + A_{x-x_0}) = o(\|x - x_0\|). \tag{1.10}$$

Si E est de dimension finie de base $\xi_1, \xi_2, ..., \xi_n$, alors $\Delta x = \sum \Delta x^i \xi_i, \Delta x \in E$. Si $\hat{\Omega}'(x_0)(\xi_i) = \langle A_{\xi_i}, 0\rangle, i = 1, 2, ..., n$, on dira que Ω est *canoniquement différentiable* en $x_0 \in G$ et on a

$$\hat{\Omega}'(x_0)(\Delta x) = \sum \Delta x^i \langle A_{\xi_i}, 0\rangle.$$

Nous exhibons maintenant quelques exemples de multifonctions π-différentiables. Ces exemples sont simples et seront utilisés dans la preuve de la stabilité de la π-différentiabilité par composition de multiapplications.

Exemple 1.8.16 *Soit E un espace normé sur $\mathbb{R}$, F un espace de Banach réflexif. Soit A un élément fixé de $\mathcal{B}(F)$ et r une application différentiable, $r : G \to F$, où G est un sous-ensemble de E. Considérons la multiapplication $\Omega : G \to \mathcal{B}(F)$ définie par $\Omega(x) = \{r(x)\} + A, x \in G$. Il est facile de voir que Ω est π-différentiable avec $\Omega'(x_0)(\Delta x) = \langle\{r'(x_0)(\Delta x)\}, 0\rangle, x_0 \in G,\ \Delta x \in E$, puisque*
$\|\hat{\Omega}(x) - \hat{\Omega}(x_0) - \langle\{r'(x_0)(x - x_0)\}, 0\rangle\|$

$$\begin{aligned}
&= \|\langle\{r(x)\} + A, 0\rangle - \langle\{r(x_0)\} + A, 0\rangle - \langle\{r'(x_0)(x - x_0)\}, 0\rangle\| \\
&= d_H(\{r(x)\} + A, \{r(x_0) + r'(x_0)(x - x_0)\} + A) \\
&= d_H(\{r(x)\}, \{r(x_0) + r^{'}(x_0)(x - x_0)\}) \\
&= \|r(x) - r(x_0) - r'(x_0)(x - x_0)\| \\
&= o(\|x - x_0\|.
\end{aligned}$$

Notons que l'application $\Delta x \to \langle\{r'(x_0)(\Delta x)\}, 0\rangle$ est un opérateur linéaire borné. L'additivité par rapport à $\Delta(x)$ est claire, et l'homogénéité résulte de $\langle \alpha A, 0\rangle = \alpha\langle A, 0\rangle$ pour tout $\alpha \in \mathbb{R}$ quand $A = \{a\}$ est un singleton dans F. Finalement on a

$$\|\langle\{r'(x_0)(\Delta x)\}, 0\rangle\| = \|r'(x_0)(\Delta x)\| \leq \|r'(x_0)\|\|\Delta x\|, \qquad \Delta x \in E.$$

Si dans l'exemple E est de dimension finie, alors Ω est canoniquement différentiable. En fait si $\xi_1, ..., \xi_n$ est une base de E, et $\Delta x = \sum \Delta x^i \xi_i$,

$$\hat{\Omega}'(x_0)(\Delta x) = \sum \Delta x^i \langle\{r'(x_0)(\xi_i)\}, 0\rangle.$$

Exemple 1.8.17 *Soient E, F, G, A définis comme dans l'exemple précédent. Soit g une fonction différentiable de G dans $\mathbb{R}$ de signe constant sur G, i.e. $g(x) \geq 0, x \in G$ ou bien $g(x) \leq 0, x \in G$. Soit $\Omega : G \to \mathcal{B}(F)$ la multifonction définie par $\Omega(x) = g(x)A,\ x \in G$.*
Considérons le premier cas où $g(x) \geq 0,\ x \in G$. Dans ce cas, on a $\pi((g(x))A) = \langle g(x)A, 0\rangle = g(x)\langle A, 0\rangle$, ceci implique que $\hat{\Omega}'(x_0)(\Delta x) = g'(x_0)(\Delta x)\langle A, 0\rangle,\ x_0 \in G,\ \Delta x \in E$. D'autre part pour $\alpha \leq 0$, on a $\langle \alpha A, 0\rangle = \langle -\alpha(-A), 0\rangle = \alpha\langle 0, -A\rangle$. Par conséquent, si $g(x) \leq 0, x \in G$, alors

$$\pi(g(x)A) = \langle g(x)A, 0\rangle = g(x)\langle 0, -A\rangle.$$

On en déduit que $\hat{\Omega}'(x_0)(\Delta x) = g'(x_0)(\Delta x)\langle 0, -A\rangle$. *Si* E *est de dimension finie de base* $\xi_1, \xi_2, ..., \xi_n$, Ω *est canoniquement différentiable en* $x_0 \in G$, *quand* $g \geq 0$ *et* $g'(x_0)(\xi_i) \geq 0, i = 1, 2, ..., n$ *(respectivement,* $g \leq 0$ *et* $g'(x_0)(\xi_i), i = 1, 2, ..., n$*), et en fait dans chacun des cas on a la relation*

$$\hat{\Omega}'(x_0)(\Delta x) = \sum \Delta x^i \langle g'(x_0)(\xi_i)A, 0\rangle \text{ pour } \Delta x = \sum \Delta x^i \xi_i \in E.$$

Observons que dans l'exemple précédent l'hypothèse de signe constant est crucial. En effet, prenons $E = G = \mathbb{R}, F = \mathbb{R}^2$ et $\Omega(t) = t\sigma$ où σ est le disque unité dans $\mathbb{R}^2$. Cette multifonction n'est pas différentiable en $t = 0$ puisque ces deux dérivées $\hat{\Omega}'(t_0)(\Delta t) = \Delta t\langle \sigma, 0\rangle$ pour $t_0 > 0$, et $\hat{\Omega}'(t_0) = \langle 0, -\sigma\rangle$, pour $t_0 < 0$ existent en $t_0 = 0$ et sont différentes. □

1.9 Vitesse de convergence d'une suite

La vitesse de convergence d'une suite représente la vitesse à laquelle les termes de la suite se rapprochent de sa limite. Ce concept a une importance capitale pratique lorsque nous travaillons avec une suite d'approximations successives obtenues à partir d'une méthode itérative, car en général peu d'itérations sont nécessaires pour donner une valeur approchée intéressante lorsque la vitesse de convergence est grande. Il est intéressant également de comparer les vitesses de convergence de suites ; on peut ainsi obtenir un "classement" des méthodes itératives.

Définition 1.9.1
Soit (x_n) *une suite qui converge vers* $\bar{x} \in X$. *On suppose que* $x_k \neq x^*$.
Si $K_p = \lim_{k\to\infty} \dfrac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^p}$ *existe et* $K_p > 0$, *alors* (x_n) *est dite* ***convergente d'ordre*** p *vers* x^*.

- *Lorsque* $p = 1$, (x_n) *est dite* ***linéairement convergente***.
- *Lorsque* $p = 2$, (x_n) *est dite* ***quadratiquement convergente***.
- *Lorsque* $p = 3$, (x_n) *est dite* ***cubiquement convergente***.

Si $K_1 = 0$, *alors* (x_n) *est dite* ***super-linéairement convergente*** *et* $K_2 = 0$, *alors* (x_n) *est dite* ***super-quadratiquement*** *convergente.*

S'il existe une constante C *et un réel positif* p *tel que*

$$\|x_{k+1} - x^*\| \leq C\|x_k - x^*\|^p$$

pour tout n *assez grand, on dit que la convergence est au* ***moins d'ordre*** p.

Pour plus de détails concernant la vitesse de convergence, le lecteur pourra consulter [12].

Chapitre 2

Résolution d'inclusions variationnelles par des méthodes itératives utilisant les processus convexes

Dans ce chapitre, nous allons passer en revue les méthodes itératives utilisant les processus convexes. Elles sont toutes issues d'une approche d'optimisation faisant intervenir les processus convexes. Leur utilisation a permis à travers des travaux, dont certains sont récents, d'obtenir une convergence semi-locale tandis que celles utilisant la régularité métrique ou des variantes ne fournissent qu'une convergence locale.
Ces méthodes prennent leur origine dans un travail de Robinson [83], proposant une extension de la méthode de Newton associée au problème ci-dessous que nous tenons à présenter.

2.1 Extension de la méthode de Newton pour des fonctions non linéaires à valeurs dans un cône

En 1972, Stephen M. Robinson [83] a proposé une extension de la méthode de Newton dédiée à la résolution de l'inclusion

$$0 \in f(x) - K, \tag{2.1}$$

où $f : X \to Y$ est une fonction continûment Fréchet-différentiable sur un sous-ensemble $X_0 \subset X$, K est un cône convexe fermé non vide, X et Y désignent respectivement un espace de Banach réflexif et un espace de Banach.

En fait, pour approcher une solution de (2.1), il a construit une méthode itérative basée sur la projection de l'origine de X sur l'image d'un élément approprié par l'inverse d'un processus convexe normé.
Il a donc considéré pour $\bar{w}$ fixé dans X_0, le processus convexe

$$T_{\bar{w}}x := \nabla f(x) - K, \quad x \in X$$

d'inverse $T_{\bar{w}}^{-1}$, qui est aussi un processus convexe, défini par

$$T_{\bar{w}}^{-1}y := \{z \ / \ \nabla f(\bar{w})z \in y + K\}, \quad y \in Y.$$

L'existence de la suite générée se justifie par la convexité et la fermeture de l'inverse dans l'espace de Banach réflexif X. En fait, l'itéré initial x_0 est choisi dans $X_0 \subset X$ tel que l'ensemble convexe

$$T_{x_0}[-f(x_0)] \text{ soit non vide .}$$

Par convexité et la fermeture de l'ensemble $T_{x_0}^{-1}[-f(x_0)]$, il existe un élément de celui-ci se trouvant à une distance minimale par rapport à l'origine 0 de X_0. L'itéré suivant x_1 est tel que la différence $x_1 - x_0$ soit une projection de l'origine 0 sur l'ensemble convexe non vide $T_{x_0}^{-1}$. Le point x_1 est alors utilisé pour répéter la procédure. Ainsi, à la k^{eme} étape nous avons x_k, et on définit l'itéré suivant x_{k+1} comme la somme de x_k et d'une projection de l'origine 0 sur l'ensemble $T_{x_k}^{-1}[-f(x_k)]$. Une façon équivalente de décrire l'algorithme est, si x_k est calculé, l'itéré suivant x_{k+1} est une solution du problème de minimisation convexe

$$\min\{\|x - x_k\| \ / \ f(x_k) + \nabla f(x_k)(x - x_k) \in K\}, \tag{2.2}$$

c'est-à-dire que l'itéré x_{k+1} est un point de l'ensemble des x tels que le vecteur $f(x_k) + \nabla f(x_k)(x - x_k)$ soit dans le cône K et qui soit le plus proche de x_k.

Observons que si $K = \{0\}$, le point x_{k+1} est un/le point le plus proche de x_k qui vérifie $f(x_k) + \nabla f(x_k)(x_{k+1} - x_k) = 0$; géométriquement, c'est donc l'abscisse d'un point de la tangente à la courbe de f au point d'abscisse x_k qui soit le plus proche de x_k. On a ainsi retrouvé la méthode de Newton classique.

On devrait s'attendre à ce qu'en général le problème (2.2) est plus facile à résoudre que le problème (2.1) à travers des commentaires sur des cas spécifiques que nous allons faire.
En fait, nous remarquons que si (2.2) est faisable, c'est-à-dire vérifie l'équation caractéristique de l'ensemble des contraintes dans (2.2), il doit être solvable (la solution existe) pour les raisons suivantes. Le fait que l'opérateur $\nabla f(x_k)$ soit continue (c'est la dérivée de Fréchet) et l'hypothèse de convexité de K, implique que l'ensemble des

points de faisabilité est un ensemble convexe fermé. L'existence d'un point de faisabilité $\tilde{x}$ implique que n'importe quelle solution de (2.2) doit appartenir à l'intersection de l'ensemble des point de faisabilité de (2.2) et de la boule fermée de rayon $\|\tilde{x} - x_k\|$ et centrée en x_k ; cette intersection est un ensemble convexe fermé borné. Comme X est un espace de Banach réflexif et la fonction $\|x - x_k\|$ est semi-continue inférieurement par rapport à la topologie faible, une solution de (2.2) existe. Ainsi, puisque (2.2) est un problème de minimisation convexe, toute solution locale sera une solution globale.

Pour la convergence, Robinson initialise la méthode en un point x_0 tel que processus T_{x_0} soit défini sur X et surjectif, puis il a établi un résultat sur la norme de l'inverse du processus résultant d'une perturbation de T_{x_0} par une transformation linéaire continue. Ce résultat s'énonce ainsi.

Théorème 2.1.1 *Soient X, X_0, Y, K, f et T définis comme précédemment. Supposons l'existence d'un $v \in X_0$ tel que T_v soit défini sur X, surjectif et à images dans Y. Alors*

a) T_v^{-1} est normé ;

b) Si Δ est une transformation linéaire continue de X dans Y telle que $\|T_v^{-1}\|\|\Delta\| < 1$, alors le processus convexe $\tilde{T}$ défini par

$$\tilde{T}x := [\nabla f(v) + \Delta]x - K$$

est défini sur X et à valeurs dans Y. De plus, $\tilde{T}^{-1}$ est normé et

$$\|\tilde{T}^{-1}\| \leq \frac{\|T_v^{-1}\|}{1 - \|T_v^{-1}\|\|\Delta\|}.$$

La preuve résulte du Théorème 1.3.6 du chapitre précédent.

Rappelons que ce résultat est indispensable au théorème de convergence qui suit.

Théorème 2.1.2 *Soient $f : X_0 \subset X \to Y$, K et T définis comme précédemment. Supposons qu'il existe un point $x_0 \in X_0$ tel que T_{x_0} soit défini sur X, surjectif et à images dans Y, et des nombres réels strictement positifs B, L, η vérifiant les propriétés suivantes :*

a) $\|T_{x_0}^{-1}\| \leq B$,

b) Pour tous $x, w \in X_0$, on a

$$\|\nabla f(x) - \nabla f(w)\| \leq L\|x - w\|,$$

c) $\|x_1 - x_0\| \leq \eta$, *où* x_1 *est un point obtenu à partir de* x_0 *par la construction de l'algorithme (un tel point existe puisque* T_{x_0} *est surjectif),*

d) $< h \leq \frac{1}{2}$, *où* $h := BL\eta$.

Soient

$$\bar{t} := \eta h^{-1}(1 - \sqrt{1 - 2h})$$

et

$$\Omega := \{x/ \ \|x - x_0\| \leq \bar{t}\}.$$

Si $\Omega \subset X_0$, *alors l'algorithme ci-dessus génère une suite* $(x_k)_k$ *de* Ω *qui converge vers un élément* x^* *de* Ω *qui est une solution du problème* (2.1).

Notons que la convergence de la suite résultante $(x_k)_k$ a été établie à partir de celle de la suite majorante $(t_k)_k$ due à Ortega [72]. Par ailleurs, il a prouvé que la convergence est quadratique lorsque $K := \{0\}$.

Les détails concernant la preuve du théorème précédent se trouvent dans l'article de Robinson [83].

Nous enchaînons maintenant avec deux cas où l'on perturbe le problème (1) par une fonction g dans un premier temps seulement lipschitzienne sur X_0, puis dans un second temps non différentiable dans un voisinage d'une solution x^* du problème perturbé.

2.2 Extension de la méthode Newton perturbée non différentiable pour des fonctions à valeurs dans un cône

En 2012, Alain Piétrus [78] a proposé une méthode itérative dédiée à la résolution d'une inclusion dont la partie univoque est perturbée par une fonction lipschitzienne non différentiable. En effet, il a proposé d'approcher une solution de l'inclusion perturbée du problème (2.1) à savoir

$$0 \in f(x) + g(x) - K, \tag{2.3}$$

où la fonction f est Fréchet-différentiable et g est seulement lipschitzienne sur X_0.

Pour la construction du processus itératif associé, il a donc fait suite au processus de construction utilisé par Robinson [83] comme suit :
il démarre le processus avec un point initial x_0 telle que l'image $-f(x_0) - g(x_0)$ par

le processus convexe défini ci-dessus $T_{x_0}^{-1}[-f(x_0) - g(x_0)]$ soit non vide ; connaissant l'itéré x_k, le nouvel itéré x_{k+1} est solution du problème de minimisation

$$\min\{\|x - x_k\| \ / \ f(x_k) + g(x_k) + \nabla f(x_k)(x - x_k) \in K\}.$$

Observons par exemple que le fait d'utiliser le même processus convexe que Robinson ramène à perturber l'équation caractéristique du membre de gauche de l'ensemble des contraintes, figurant dans l'itération de Robinson, par l'image de la fonction g au point x_k.

Rappelons ci-dessous le lemme suivant, dû à Rheinboldt [81], donnant la convergence de la suite majorante $(t_k)_k$ sur laquelle repose la convergence de la suite précédente.

Lemme 2.2.1
Considérons la suite $(t_k)_k$ définie par

i) $t_0 = 0, \ t_1 = \alpha,$

ii) $t_{k+1} - t_k = \frac{1}{1-p_4 t_k}\Big[p_1(t_k - t_{k-1})^2 + (p_2 + p_3 t_{k-1})(t_k - t_{k-1})\Big]$, *pour* $k = 1, 2, ...$
avec

$0 \leq p_4, \ \alpha < 1, \ p_i \geq 0, i = 1, ..., 4.$

Si pour certains paramètres (p_i^0, α^0) la solution (t_k^0) satisfait $t_k^0 \leq t_{k+1}^0, k \geq 0$, et $\lim t_k^0 = t^{0*} < \frac{1}{p_4}$, alors pour tous $0 \leq p_i \leq p_i^0, i = 1, ..., 4, 0 \leq \alpha \leq \alpha^0$, la suite (t_k) est croissante et $\lim t_k = t^* \leq t^{0*}$. Si $p_1 > 0, 0 \leq p_2 < 1, p_3 + p_4 = 2p_1$ et $0 < \alpha \leq \frac{(1-p_2)^2}{4p_1}$, alors la suite (t_k) est strictement croissante et

$$\lim t_k = t^* = \frac{1}{2p_1}\Big[(1 - p_2) - \sqrt{(1 - p_2)^2 - 4p_1\alpha}\Big].$$

Notons que la suite majorante utilisée par A. Piétrus [78] est totalement différente de celle utilisée par Robinson.

La convergence de sa méthode suit le procédé utilisé par Robinson dans la mesure où il a initié l'itération en un point x_0 où le processus convexe $T(x_0)$, défini par ce dernier, soit défini sur l'espace X et surjectif. Ceci a été fait aux fins d'établir, sous des hypothèses similaires, la convergence super-linéaire de sa méthode lorsque g est lipschitzienne.
La convergence a été établie selon le théorème suivant.

Théorème 2.2.2 *(Théorème de convergence) Soient X, X_0, Y, K, f et T définis comme précédemment. Supposons qu'il existe un point $x_0 \in X_0$ tel que $T(x_0)$ soit défini sur X, dont l'image est Y, des nombres strictement positifs B, L, η, l, M vérifiant les propriétés suivantes :*

a) $||T^{-1}(x_0)|| \leq B$;

b) Pour tous $x, w \in X_0$, *on a*

$$\|\nabla f(x) - \nabla f(w)\| \leq L\|x - w\|,$$
$$\|g(x) - g(y)\| \leq l\|x - y\|;$$

c) $\|x_1 - x_0\| \leq \eta \leq \min(\frac{1}{BL}, \frac{(1-Bl)^2}{2BL})$, *où* x_1 *est obtenu à partir de* x_0 *par l'algorithme*

(ce point existe puisque T_{x_0} *est surjectif) ;*

d) Soit $t^* = \frac{1}{BL}\Big((1-Bl) - \sqrt{(1-Bl)^2 - 2BL\eta}\Big)$ *et* $\mathbb{B}_{t^*}(x_0) = \{x \mid \|x - x_0\| \leq t^*\}$.

Si $\mathbb{B}_{t^*}(x_0) \subset X_0$, *alors l'algorithme génère une suite* $(x_k)_k$ *telle que pour tout* $k \geq 0, x_k \in \mathbb{B}_{t^*}(x_0)$ *et converge vers* x^* *vérifiant* $f(x^*) \in K - g(x^*)$. *C'est-à-dire* x^* *est solution du problème* (2.3).
Notons que le nombre t^* *est la limite de la suite* $(t_k)_k$ *définie dans le lemme précédent avec*

$$p_1 = \frac{BL}{2};\ p_2 = Bl;\ p_3 = 0;\ p_4 = BL;\ \alpha = \eta.$$

2.3 Extension de la méthode Newton-sécante pour des fonctions à valeurs dans un cône

Quelques années plus tard, faisant suite aux travaux sur les inclusions perturbées, A. Piétrus et C. Jean-Alexis [79] ont initiés des travaux portant sur la résolution de l'inclusion (2.1) perturbée par une fonction g non différentiable dans un voisinage d'une solution x^*.
Ils ont proposé de résoudre le problème suivant

$$0 \in f(x) + g(x) - K, \tag{2.4}$$

où g est non différentiable sur un voisinage de x^*.

Ici, afin de pallier au manque de régularité de g, ils ont combiné une méthode de type Newton à f et une méthode de type sécante à g pour approcher x^* par une méthode décrite comme suit :
On démarre avec x_0 et $x_1 \in X_0$ tels que l'ensemble $T^{-1}(x_0, x_1)[-f(x_1)-g(x_1)]$ soit non vide.
On pose $x_{k+1}-x_k$ égale à une projection de l'origine sur l'ensemble convexe $T^{-1}(x_0, x_1)[-f(x_1)-g(x_1)]$,
où pour tous v et w fixés dans X_0,

$$T^{-1}(v, w)y := \{z \in X, (\nabla f(w) + [v, w; g])z \in y + K\}, \forall y \in Y,$$

$[v, w; g]$ est une différence divisée du premier ordre de la fonction g.
D'une manière équivalente, x_{k+1} est une solution du problème

$$\min \Big\{ \|x - x_k\| \mid f(x_k) + (\nabla f(x_k) + [x_{k-1}, x_k])(x - x_k) \in K - F(x_k) \Big\}. \tag{2.5}$$

Remarquons qu'ils ont procédé de façon similaire en utilisant le processus convexe, défini par Robinson, auquel ils ont ajouté la différence divisée du premier ordre de g.

La convergence de la méthode a été établie via une suite majorante. Ils ont fait appel à la suite de Fibonacci ; l'idée vient du fait qu'il y a une relation entre l'ordre de la méthode Newton-sécante classique et celle-ci.

Proposition 2.3.1 *Considérons la suite $(t_n)_n$ définie par*

i) $t_0 = 0,\ t_1 = \alpha,\ t_2 = \beta$

ii) $t_{n+1} - t_n = M\Big(\frac{L}{2}(t_n - t_{n-1})^2 + K(t_n - t_{n-1})(t_n - t_{n-2})\Big)$

où α, β, M et K sont des constantes strictement positives.
Si les conditions

iii) $\alpha \leq M,\ \beta \leq 2\alpha$;

iv) $q = M^2(\frac{L}{2} + 2K) < 1$

sont toutes vérifiées, alors tous les termes de la suite $(t_n)_n$ appartiennent à la boule $\mathbb{B}_s(t_1)$, où $s = \frac{M}{q}\Sigma_{l=1}^{\infty} q^{u_l}$, (u_l) est la suite de Fibonacci définie par $u_0 = u_1 = 1$ et $u_{l+1} = u_l + u_{l-1}$ pour tout $l \geq 1$. De plus, la suite $(t_n)_n$ est convergente, et on a une estimation de l'erreur

$$t^* - t_n \leq \frac{M}{q\Big(1 - q^{\frac{p^n(p-1)}{\sqrt{5}}}\Big)} q^{\frac{p^n}{\sqrt{5}}}$$

avec $t^* = \lim t_n$ et $p = \frac{1+\sqrt{5}}{2}$.

La méthode sera d'ordre p au moins, évidemment plus rapide que celle présentée précédemment dans [78].

La convergence de leur méthode a été prouvée par le théorème suivant :

Théorème 2.3.2 *Considérons $X_0 \subset X$, un sous-ensemble ouvert. Soit $f : X \to Y$ et $g : X \rightrightarrows Y$ deux fonctions respectivement différentiables de dérivée $\nabla f(x)$ et admettant des différences divisées du premier et du second ordre sur X_0. De plus, supposons qu'il existe des nombres réels strictement positifs B, L, K, M_1, α et β vérifiant*

(i) $\|\tilde{T}^{-1}_{(x_0,x_1)}\| \leq B$;

(ii) $\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$, *pour tous* $x, y \in X_0$;

(iii) g *est continue sur* X_0.

(iv) $BM_1 < 1$ *et* $\alpha \leq \beta \leq 2\alpha$;

(v) $\|x_1 - x_0\| \leq \alpha \leq \frac{B}{1-BM_1}$, $\|x_2 - x_1\| \leq \beta$, *où* x_2 *est un point obtenu à partir de* x_0 *et* x_1 *par notre algorithme ;*

(vii) pour tous points distincts $x, y,$ *et* z *de* X_0, $\|[x, y, z; g]\| \leq K_1$, $\|\nabla f(y) + [x, y; g] - \nabla f(x_1) - [x_0, x_1; g]\| \leq M_1$;

(viii) $q = \frac{B^2}{(1-BM_1)^2}(\frac{L}{2} + 2K) < 1$.

Alors il existe t^* tel que pour tout point initial x_0 satisfait $\mathbb{B}_{t^*}(x_0) \subset X_0$, il existe une suite $(x_k)_k$ générée par l'algorithme (2.5) telle que la suite $(x_k)_k$ est une suite de la boule $\mathbb{B}_{t^*}(x_0)$ et converge vers une solution x^* du problème (1).

On va clore l'état de l'art par des nouveaux travaux s'inspirant de l'extension de la méthode de Newton due à Robinson [83]. Ces résultats sont dus à S. Bernard et al. [15] à travers lesquels ils ont proposé une méthode itérative pour résoudre (2.1) mais sous l'hypothèse de la non différentiabilité de la fonction f.
Ils ont montré que quand la fonction f est une fonction de $\mathbb{R}^n$ dans $\mathbb{R}^m$ et localement lipschitzienne, il existe un processus itératif basé sur les processus convexes qui converge super-linéairement vers une solution du problème. Ainsi, leur méthode généralise une version non lisse de la méthode de Newton proposée par L. Qi et J. Sun [75] ; sa convergence a été établie en combinant les procédés de ces derniers et ceux de Robinson.

2.4 Extension de la méthode de Newton semi-lisse pour une fonction à valeurs dans un cône

En 2016, S. Bernard et al. [15] ont proposé de résoudre l'inclusion

$$0 \in f(x) - K, \tag{2.6}$$

où la fonction $f : \mathbb{R}^n \to \mathbb{R}^m$ est non différentiable mais localement lipschitzienne.

Ils ont associé à (2.6) une méthode de la forme

$$\begin{cases} T^{-1}(x_0)[-f(x_0] := \{z \in X \ / \ f(x_0) + \Delta f(x_0) \in K\} \neq \emptyset \\ x_{k+1} \text{ solution du problème de minimisation } \min\{\|x - x_k\| \ / \ f(x_k) + \Delta f(x_k)(x - x_k) \in K\} \end{cases}$$

avec $\Delta f(x_k) \in \partial f(x_k)$ et x_0 le point initial.

Remarquons qu'on substitue la dérivée de Fréchet de f au point x_k à un jacobien généralisé de Clarke $\Delta f(x_k)$ de f au point x_k.

Voyons maintenant le lemme sur lequel a été basé le théorème de convergence.

Lemme 2.4.1 *Supposons qu'une fonction f est semi-lisse d'ordre p sur un sous-ensemble X_0, alors il existe des constantes $r > 0$ et $L > 0$ telles que pour tous $x, y \in X_0$ vérifiant $\|x - y\| \leq r$ et pour tout $\nabla f(y) \in \partial f(y)$,*

$$\|f(x) - f(y) - \Delta f(y)(x - y)\| \leq L\|x - y\|^{1+p}.$$

Ce lemme leur a permis de prouver le théorème suivant.

Théorème 2.4.2 *Soient X_0, T définis comme précédemment et f une fonction semi-lisse d'ordre p sur un sous-ensemble X_0. Supposons qu'il existe un point $x_0 \in X_0 \subset \mathbb{R}^n$ tel que $T(x_0)$ soit défini sur $\mathbb{R}^n$ à valeurs dans l'ensemble des parties de $\mathbb{R}^m$ et des réels strictement positifs B, M_1, M_2 et α vérifiant propriétés suivantes :*

(a) $2BM_1 < 1$ *et* $\|T^{-1}(x_0)\| \leq B$;

(b) $\|\Delta f(y)\| \leq M_1$ *et* $\Delta f(y) \in \partial f(y)$, $y \in X_0$;

(c) $\alpha < \frac{r}{2}, 0 < h = ML\alpha^p < 1 - \frac{2\alpha}{r}$, *où L et r sont donnés par le lemme précédent,* $M = \frac{B}{1-2BM_1}$;

(d) $\|x_1 - x_0\| \leq \alpha$ *avec* x_1 *est obtenu à partir de* x_0 *via l'algorithme précédent.*

Il existe t^* *tel que si* $\mathbb{B}_{t^*}(x_0) \subset X_0$, *alors l'algorithme génère une suite* $(x_k)_k$ *telle que pour tout* $k \geq 0, x_k \in \mathbb{B}_{\frac{r}{2}}(x_0)$ *et converge vers* x^* *tel que* $f(x^*) \in K$. *De plus, on a une estimation de l'erreur*

$$\|x^* - x_k\| \leq \alpha \frac{h^{\frac{(1+p)^k-1}{p}}}{1-h^{(1+p)^k}}, \quad \forall k > 0.$$

La suite converge super-linéairement vers x^*.

Dans les chapitres qui suivent, nous allons étendre les résultats à des problèmes plus généraux ne se limitant plus à un cône fixe mais à des multiapplications différentiables. Ceci conduit à la construction de nouvelles méthodes pour résoudre chacun des problèmes suivants :

1) trouver $x \in X$ tel que $0 \in f(x) + F(x) - K$ où f, F différentiables ;

2) trouver $x \in X$ tel que $0 \in f(x) + g(x) + F(x) - K$ où f, F différentiables et g Lipschitz ;

3) trouver $x \in X$ tel que $0 \in f(x) + F(x) - K$, où f semi-lisse, F différentiable.

Nous achevons ce chapitre avec deux propositions qui seront essentielles à nos résultats. La première est inspirée du théorème préliminaire au théorème de convergence de la méthode établie par Robinson dans [82]. Elle montrera que, sous certaines conditions, les résultats de ce théorème restent inchangés en perturbant le processus convexe $T_v(x) := \nabla f(v)x - K$ par une dérivée de Nachi-Penot.
La seconde nous donne une condition suffisante de continuité d'une application multivoque à partir d'une suite convergente.

Dans ce qui suit, on note $\nabla f(v)$ et $A(v)$ respectivement la dérivée de Fréchet et une dérivée de Nachi-Penot évaluées en $v \in X_0$.

Proposition 2.4.3 *Soit* X_0 *un sous-ensemble ouvert de* X. *Considérons* $f : X \to Y$ *et* $F : X \rightrightarrows Y$ *respectivement continûment Fréchet-différentiable et N-P-différentiable sur* X_0. *Soit* $v \in X_0$ *tel que le processus convexe défini par* $T_v(x) := \nabla f(v)(x) - K$ *soit surjectif. Alors pour le processus convexe* $\tilde{T}_v(x) := (\nabla f(v) + A(v))x - K$, *on a*

(a) $\tilde{T}_v$ *et* $\tilde{T}_v^{-1}$ *sont normés ;*

(b) Si $\|T_v^{-1}\|\|A(v)\| < 1$*, une estimation de la norme de l'inverse* $\tilde{T}_v^{-1}$ *de* $\tilde{T}_v$ *est*

$$\|\tilde{T}_v^{-1}\| \leq \frac{\|T_v^{-1}\|}{1 - \|T_v^{-1}\|\|A(v)\|}. \tag{2.7}$$

Preuve
Remarquons d'abord que $\tilde{T}_v$ est un processus convexe puisque c'est la somme du processus convexe $T_v(x) := \nabla f(v)x - K$ et de l'application linéaire $A(v)$ (voir Proposition (1.3.5)).
Soit $x \in \operatorname{dom} \tilde{T}_v \cap I\!B_X$, et remarquons que

$$(\nabla f(v) + A(v))x \in (\nabla f(v) + A(v))x - K = \tilde{T}_v(x).$$

Considérons y un élément quelconque de l'ensemble $\tilde{T}_v(x)$.
Donc, par définition de la borne inférieure d'un ensemble,

$$\inf_{y \in \tilde{T}_v(x)} \|y\| \leq \|(\nabla f(v) + A(v))x\| \leq \|\nabla f(v) + A(v)\|.$$

La continuité de l'application $\nabla f(x) + A(x)$ implique que

$$\|\nabla f(v) + A(v)\| = \sup_{x \in B_X} \|(\nabla f(v) + A(v))x\| < \infty.$$

D'où,

$$\|\tilde{T}_v\| := \sup_{x \in \operatorname{dom} \tilde{T}_v \cap B_X} \inf_{y \in \tilde{T}_v(x)} \|y\| \leq \|\nabla f(v) + A(v)\| < \infty.$$

Ceci montre que $\tilde{T}_v$ est normé.

Le processus convexe T_v étant surjectif et à graphe fermé, on en déduit que le processus convexe $\tilde{T}_v$ l'est également.
Par suite, $\tilde{T}_v^{-1}$ est à graphe fermé et défini sur l'espace Y tout entier et par conséquent, $\tilde{T}_v^{-1}$ est normé.
Rappelons que cette affirmation résulte du fait qu'un processus convexe normé et à graphe fermé est surjectif.

Pour montrer la partie (b), remarquons d'abord que $\tilde{T}_v(x) = T_v(x) + A(v)x$. Par conséquent en utilisant la partie $b)$ du Théorème 1.3.6, on obtient l'inégalité (2.7) et ce qui achève la preuve. □

Nous allons maintenant énoncer et démontrer notre seconde proposition, indispensable aux principaux résultats, selon laquelle la suite d'images d'une suite convergente par une application multivoque semi-continue inférieurement et à graphe fermé converge vers l'image de la limite.

Proposition 2.4.4
Soit $G : X \rightrightarrows Y$ *une application multivoque semi-continue inférieurement sur une partie fermée* $X_0 \subset X$. *Supposons que* $\operatorname{gph} G$ *soit fermé et à valeurs compactes. Alors, pour toute suite* $(x_n) \subset X_0$ *qui converge vers* x^*, *on a* $\lim_{n\to\infty} G(x_n) = G(x^*)$.

Preuve
Comme $\operatorname{gph} G$ est fermé, la fermeture de l'ensemble $G(x^*)$ est encore $G(x^*)$.
Donc $\liminf_{x_n\to x^*} G(x_n) \subset G(x^*) \subset \limsup_{x_n\to x^*} G(x_n)$. La multiapplication G étant s.c.i. sur X_0, donc elle l'est en $x^* \in X_0$.
Par suite, grâce à la Proposition 1.4.7 dans [4], $G(x^*) \subset \liminf_{x_n\to x^*} G(x_n)$.
Il en résulte que $G(x^*) = \liminf_{x_n\to x^*} G(x_n)$.
La multiapplication de G étant à valeurs compactes, par une remarque faite sur la Proposition 1.4.7, la fermeture et la s.c.s. de G coincident.
Puisque la suite $(x_n)_n$ converge vers x^*, tout voisinage de x^* contient une infinité de termes de la suite à partir d'un certain rang. Par compacité et la s.c.s. de G, les ensembles-images $G(x_n)$ sont tous contenus dans une boule centrée en $G(x^*)$ et de rayon quelconque à partir de ce rang.
Il s'en suit qu'il existe un voisinage de l'ensemble $G(x^*)$ qui contient tous les ensembles $G(x_n)$ (à partir de ce rang).
On en déduit que

$$\limsup_{x_n\to x^*} G(x_n) \subset \overline{G(x^*)} = G(x^*).$$

Ainsi, la suite d'ensembles $G(x_n)$ converge et $\lim_{n\to\infty} G(x_n) = G(x^*)$. $\square$

Dans les chapitres qui suivent, on construit des nouvelles méthodes itératives permettant de résoudre des inclusions variationnelles résultant d'une perturbation des inclusions précédentes par une multiapplication F qui est N-P-différentiable et à dérivée uniformément bornée, à graphes fermés et à valeurs compactes et convexes.
La fin de chacun d'eux comportera des remarques sur des cas spéciaux notamment celui où la multiapplication est identiquement égale au singleton $\{0\}$.

Chapitre 3

Méthode de type Newton pour une multiapplication différentiable contenue dans un cône

Dans ce chapitre, nous allons proposer une méthode itérative pour résoudre numériquement le problème

$$\text{trouver } x \in X \text{ tel que } 0 \in f(x) + F(x) - K, \tag{3.1}$$

où f est une fonction Fréchet-différentiable sur un sous-ensemble $X_0 \subset X$; F est une fonction N-P-différentiable, K est un cône convexe fermé et contenant l'origine de Y.

3.1 Description et convergence de la méthode

Pour tout $v \in X_0$, on définit une application multivoque $\tilde{T}_v : X \rightrightarrows Y$ par

$$\tilde{T}_v(x) := (\nabla f(v) + A(v))x - K, \quad x \in X,$$

où $\nabla f(v)$ et $A(v)$ désigne respectivement la dérivée de Fréchet de f et une N-P-dérivée de F en v.
De la Proposition 2.4.3, la multiapplication $\tilde{T}_v$ est un processus convexe. Son inverse est aussi un processus convexe, défini par

$$\tilde{T}_v^{-1}(y) := \{z \in \ / \ (\nabla f(v) + A(v))z \in y + K\}, \quad y \in Y.$$

Choisissons un point initial $x_0 \in X_0$ tel que l'ensemble $\tilde{T}_{x_0}^{-1}[-f(x_0) - F(x_0)] \neq \emptyset$. On définit x_1 comme la somme de x_0 et d'une projection de l'origine de X sur l'ensemble $\tilde{T}_{x_0}^{-1}[-f(x_0) - F(x_0)]$. Pour construire x_2 nous répétons le processus en partant de x_1. Alors, à la $k-$ème pas, nous avons x_k, et on définit x_{k+1} comme la somme de x_k et

d'une projection de l'origine de X sur $\tilde{T}_{x_k}^{-1}[-f(x_k) - F(x_k)]$.

Cet algorithme peut s'écrire de façon équivalente : si x_k est calculé, alors le nouvel itéré x_{k+1} est une solution du problème de minimisation

$$\min\{\|x - x_k\|/\ |f(x_k) + (\nabla f(x_k) + A(x_k))(x - x_k) \in K - F(x_k)\}. \qquad (3.2)$$

Observons que le processus convexe utilisé est similaire à celui utilisé par Robinson dans [83] ; c'est le processus convexe utilisé par ce dernier auquel on ajoute une N-P-dérivée de la multiapplication F. Ainsi, l'itération (3.2) peut être vue comme la recherche d'un / du minimum de la fonction $\|x - x_k\|$ sur l'ensemble des contraintes où la dérivée de Fréchet de f au point x_k est perturbée par une N-P-dérivée de F en ce même point x_k et le cône K par l'opposé de l'image de F au point x_k.
Par suite, lorsque $F(x) = \{0\}$, la méthode est exactement la même que celle introduite par Robinson dans [83].

L'itération (3.2) a un sens selon la remarque suivante.

Remarque 3.1.1
Comme nous l'avons mentionné dans la remarque du chapitre précédent, l'existence d'un point de faisabilité $\tilde{x}$ implique que toute solution de (3.2) *est contenue dans l'intersection de l'ensemble des points de faisabilité de* (3.2) *et de la boule de centre x_k et de rayon $\|x - x_k\|$; cette intersection est un ensemble convexe fermé borné. Puisque X est réflexif et la fonction est semi-continue inférieurement par rapport à la topologie faible, une solution de* (3.2) *existe, (voir [59]).*

Nous allons maintenant nous intéresser à la convergence de notre processus itératif (3.2).
Nous rappelons ci-dessous une proposition, due à Rheinboldt, donnant une suite convergente de laquelle on déduira la convergence.
Rappelons que cette proposition avait déjà été utilisée par Piétrus dans [78].

Proposition 3.1.2 *Considérons la suite $(t_k)_k$ définie par*

i) $t_0 = 0,\ t_1 = \alpha,$

ii) $t_{k+1} - t_k = \frac{1}{1-p_4 t_k}\Big[p_1(t_k - t_{k-1})^2 + (p_2 + p_3 t_{k-1})(t_k - t_{k-1})\Big],$ *pour* $k = 1, 2, ...$
avec

$0 \leq p_4\alpha < 1, p_i \geq 0, i = 1, ..., 4.$

Si pour certains paramètres (p_i^0, α^0) *la solution* (t_k^0) *satisfait* $t_k^0 \leq t_{k+1}^0, k \geq 0$, *et* $\lim t_k^0 = t^{0*} < \frac{1}{p_4}$, *alors pour tous* $0 \leq p_i \leq p_i^0, i = 1, ..., 4, 0 \leq \alpha \leq \alpha^0$, *la suite* (t_k) *est croissante et* $\lim t_k = t^* \leq t^{0*}$. *Si* $p_1 > 0, 0 \leq p_2 < 1, p_3 + p_4 = 2p_1$ *et* $0 < \alpha \leq \frac{(1-p_2)^2}{4p_1}$, *alors la suite* (t_k) *est strictement croissante et*

$$\lim t_k = t^* = \frac{1}{2p_1}\Big[(1-p_2) - \sqrt{(1-p_2)^2 - 4p_1\alpha}\Big].$$

Nous avons vu que la lipschitziannité du perturbateur univoque g était essentielle à la convergence de la méthode de Newton perturbée non différentiable proposée par Piétrus [78] dans la mesure où la convergence de la suite majorante des itérés (x_k) dépendait de la constante de Lipschitz. Cela nous inspirera à supposer que la somme de la dérivée de Fréchet et la dérivée de Nachi-Penot bornée du perturbateur multivoque F est lipschitzienne sur la partie X_0.

3.2 Analyse de convergence de la méthode

Théorème 3.2.1
Étant donnés un sous-ensemble ouvert $X_0 \subset X$, $f : X \to Y$ *et* $F : X \rightrightarrows Y$ *respectivement Fréchet-différentiable et N-P-différentiable sur* X_0, *et dont les dérivées sont respectivement* $\nabla f(x)$ *et* $A(x)$ *pour tout* $x \in X_0$. F *soit à graphe fermé, à valeurs compactes et convexes. Supposons qu'il existe un point* $x_0 \in X_0$ *tel que le processus convexe* $\tilde{T}_{x_0}$ *soit défini sur* X *et surjectif. De plus, supposons qu'il il existe des nombres strictement positifs* B, L, l, M *et* η *vérifiant*

(i) $\|\tilde{T}_{x_0}^{-1}\| \leq B$;

(ii) $\sup_{x \in X_0} \|A(x)\| \leq M$;

(iii) $\|\nabla f(x) - \nabla f(x')\| \leq l\|x - x'\|$, *pour tous* $x, x' \in X_0$;

(iv) $\|(\nabla f(x) + A(x)) - (\nabla f(x') + A(x'))\| \leq L\|x - x'\|$, *pour tous* $x, x' \in X_0$;

(v) Pour tout x_{k+1} *solution de* (3.2), $F(x_k) \subset F(x_{k+1}) - K$, $k = 1, 2, \ldots$;

(vi) $l < L$, $BM < 1$;

(vii) $\|x_1 - x_0\| \leq \eta \leq \min\left\{\frac{1}{BL}, \frac{(1-BM)^2}{2BL}\right\}$, *où* x_1 *est un point obtenu à partir de* x_0 *par notre algorithme (un tel point existe via la Remarque* (3.1.1)*).*

Alors il existe $t^* = \frac{1}{BL}((1-BM) - \sqrt{(1-BM)^2 - 2BL\eta}) > 0$ *de la Proposition 3.1.2 tel que pour tout point initial* x_0 *satisfait* $\mathbb{B}_{t^*}(x_0) \subset X_0$*, il existe une suite* $(x_k)_k$ *générée par l'algorithme* (3.2) *telle que* $\forall k \geq 0$*,* x_k *appartient à la boule* $\mathbb{B}_{t^*}(x_0)$ *et converge vers* x^* *solution du problème* (3.1)*.*

Preuve

Dans un premier temps, nous construisons par récurrence une suite $(x_k)_k$ qui satisfait (3.2) dont tous les termes x_k restent dans la boule $\mathbb{B}_{t^*}(x_0)$ et une suite croissante $(t_k)_k$ qui vérifie

$$\|x_{k+1} - x_k\| \leq t_{k+1} - t_k, \text{ pour tout } k \in \mathbb{N}.$$

Prenons $t_0 = 0$ et $t_1 = \eta$.

L'hypothèse (vii) nous permet d'écrire

$$\|x_1 - x_0\| \leq \eta = t_1 - t_0,$$

et $x_1 \in \mathbb{B}_{t^*}(x_0)$(via un élément de la preuve du Lemme 2.2.1 de Rheinboldt [81]).

Maintenant, supposons que l'algorithme (3.2) génère $x_1, x_2, \ldots, x_j$ tels que

$$\|x_k - x_{k-1}\| \leq t_k - t_{k-1} \text{ et } x_k \in \mathbb{B}_{t^*}(x_0) \text{ avec } t_k \geq t_{k-1}, \ \forall k = 1, \ldots, j.$$

En combinant (i) et (iv), nous obtenons

$$\|\tilde{T}_{x_0}^{-1}\|\|(\nabla f(x_j) + A(x_j)) - (\nabla f(x_0) + A(x_0))\| \leq BL\|x_j - x_0\| \leq BLt^*$$

car, par hypothèse, l'itéré x_k est dans la boule $\mathbb{B}_{t*}(x_0)$.

Puisque

$$t^* = \frac{1}{BL}\Big(1 - \Big(BM + \sqrt{(1-BM)^2 - 2BL\eta}\Big)\Big) < \frac{1}{BL}$$

et le produit BL étant strictement positif,

$$\|\tilde{T}_{x_0}^{-1}\|\|(\nabla f(x_j) + A(x_j)) - (\nabla f(x_0) + A(x_0))\| \leq BL\|x_j - x_0\| < 1.$$

Remarquons que

$$\begin{aligned}\tilde{T}_{x_j}(x) &= \Big(\nabla f(x_0) + A(x_0)\Big)x + \Big(\nabla f(x_j) + A(x_j) - (\nabla f(x_0) + A(x_0))\Big)x - K \\ &= T_{x_0}(x) + [\nabla f(x_j) + A(x_j) - (\nabla f(x_0) + A(x_0)]x.\end{aligned}$$

Alors, en utilisant le Théorème 1.3.6 et les hypothèses (i) - (iv), il s'en suit que $\tilde{T}_{x_j}$ est défini sur X à images dans Y ; $\tilde{T}_{x_j}^{-1}$ est normé et sa norme satisfait

$$\begin{aligned}\|\tilde{T}_{x_j}^{-1}\| &\leq \frac{\|\tilde{T}_{x_0}^{-1}\|}{1 - \|\tilde{T}_{x_0}^{-1}\|\|(\nabla f(x_j) + A(x_j)) - (\nabla f(x_0) + A(x_0))\|} \\ &\leq \frac{B}{1 - BL\|x_j - x_0\|}.\end{aligned}$$

Comme $\tilde{T}_{x_j}$ est défini sur X et surjectif, nous déduisons que (3.2) est faisable et donc solvable pour $k = j$ et nous obtenons l'existence de x_{j+1} solution du problème (3.2).

À présent, considérons le nouveau problème consistant à trouver x solution de

$$f(x_j)+(\nabla f(x_j)+A(x_j))(x-x_j) \in f(x_{j-1})+(\nabla f(x_{j-1})+A(x_{j-1}))(x_j-x_{j-1})+K. \quad (3.3)$$

Puisque x_j est solution de (3.2) (pour $k = j-1$), l'hypothèse (v) nous permet d'écrire l'inclusion sous la forme

$$f(x_j) + (\nabla f(x_j) + A(x_j))(x - x_j) \in K - F(x_{j-1}) \subset K - F(x_j),$$

et nous en déduisons que tout x satisfait (3.3) est nécessairement faisable pour (3.2). Nous pouvons réécrire (3.3) sous la forme

$$x - x_j \in \tilde{T}_{x_j}^{-1}\Big(- f(x_j) + f(x_{j-1}) + (\nabla f(x_{j-1}) + A(x_{j-1}))(x_j - x_{j-1})\Big). \quad (3.4)$$

Donc, son ensemble des solutions est l'ensemble

$$T_{x_j}^{-1}\Big(- f(x_j) + f(x_{j-1}) + (\nabla f(x_{j-1}) - A(x_{j-1}))(x_j - x_{j-1})\Big) + x_j.$$

Par suite

$$d\Big(0, T_{x_j}^{-1}\Big(- f(x_j) + f(x_{j-1}) + (\nabla f(x_{j-1}) + A(x_{j-1}))(x_j - x_{j-1})\Big) + x_j\Big)$$

$$= \inf\{\|t - x_j - 0\|,\ t \in T_{x_j}^{-1}(- f(x_j) + f(x_{j-1}) + (\nabla f(x_{j-1}) + A(x_{j-1}))(x_j - x_{j-1})\}.$$

Par ailleurs, en appliquant le Théorème 1.3.9, nous avons l'inégalité

$$d\Big(\tilde{T}_{x_j}^{-1}\Big(- f(x_j) + f(x_{j-1}) + (\nabla f(x_{j-1}) + A(x_{j-1}))(x_j - x_{j-1})\Big),\ \tilde{T}_{x_j}^{-1}(0)\Big)$$

$$\leq \|T_{x_j}^{-1}\| \times \| - f(x_j) + f(x_{j-1}) + (\nabla f(x_{j-1}) + A(x_{j-1})(x_j - x_{j-1})\|.$$

Par définition de la distance de deux ensembles et puisque $0 \in \tilde{T}_{x_j}^{-1}(0)$,

$$d\Big(0, T_{x_j}^{-1}\Big(- f(x_j) + f(x_{j-1}) + (\nabla f(x_{j-1}) + A(x_{j-1})(x_j - x_{j-1}))\Big)\Big)$$

$$\leq \|T_{x_j}^{-1}\| \times \| - f(x_j) + f(x_{j-1}) + (\nabla f(x_{j-1}) + A(x_{j-1})(x_j - x_{j-1})\|.$$

Le membre de droite de (3.4) étant un ensemble convexe fermé non vide inclus dans l'espace de Banach réflexif X, il existe un élément $t := \tilde{x}$ solution de (3.4) tel que

$$\begin{aligned}\|\tilde{x}-x_j\| &\leq \|\tilde{T}_{x_j}^{-1}\|\|f(x_{j-1})-f(x_j)+(\nabla f(x_{j-1})+A(x_{j-1}))(x_j-x_{j-1})\| \\ &\leq \|\tilde{T}_{x_j}^{-1}\|\Big(\|f(x_{j-1})-f(x_j)+\nabla f(x_{j-1})(x_j-x_{j-1})\|+\|A(x_{j-1})\|\|x_j-x_{j-1}\|\Big) \\ &\leq \|\tilde{T}_{x_j}^{-1}\|(\tfrac{l}{2}\|x_j-x_{j-1}\|^2+M\|x_j-x_{j-1}\|) \\ &\leq \tfrac{B}{1-BL\|x_j-x_0\|}\Big(\tfrac{L}{2}\|x_j-x_{j-1}\|^2+M\|x_j-x_{j-1}\|\Big).\end{aligned}$$

Le membre de droite de la dernière inégalité nous conduit à considérer la suite $(t_k)_k$ définie par

$$\begin{cases} t_0=0,\ t_1=\eta. \\ t_{k+1}-t_k=\dfrac{1}{1-BLt_k}\Big[\dfrac{BL}{2}(t_k-t_{k-1})^2+BM(t_k-t_{k-1})\Big],\ \text{pour } k=0,1,2,\cdots \end{cases}$$

Ainsi, en appliquant le Lemme 2.2.1 avec $\alpha:=\eta$, $p_1:=\frac{BL}{2}$, $p_2:=BM$, $p_3:=0$ et $p_4:=BL$, nous en déduisons que la suite $(t_k)_k$ est croissante et converge vers t^*, et vérifie l'inégalité

$$\|\tilde{x}-x_j\|\leq t_{j+1}-t_j.$$

Comme $\tilde{x}$ est un point de faisabilité de (3.2) pour $k=j$, nous obtenons

$$\|x_{j+1}-x_j\|\leq\|\tilde{x}-x_j\|\leq t_{j+1}-t_j,$$

donc $x_{j+1}\in I\!B_{t^*}(x_0)$. Ceci complète la récurrence.

En outre, la suite $(t_k)_k$ étant une suite majorante de $(x_k)_k$, par conséquent $(x_k)_k$ est une suite de Cauchy de l'espace de Banach réflexif X, donc converge vers un élément $x^*\in X$.

Dans un second temps, nous remarquons que la suite $(x_k)_k$ vérifie

$$f(x_k)+(\nabla f(x_k)+A(x_k))(x_{k+1}-x_k)\in K-F(x_k),\ \forall k\in\mathbb{N}.$$

De plus, il existe $w\in K$, tel que

$$(x_k,f(x_k)+(\nabla f(x_k)+A(x_k))(x_{k+1}-x_k)-w)\in \mathrm{gph}(-F). \tag{3.5}$$

La fermeture de $\mathrm{gph}(F)$, la continuité des applications f, ∇f, A et $\lim_{k\to\infty}x_k=x^*$ impliquant que

$$(x^*,f(x^*)-w)\in\mathrm{gph}(-F)\Rightarrow f(x^*)\in K-F(x^*).$$

Ceci montre que x^* est solution de (3.3) et qui achève la preuve. □

Il est intéressant de remarquer que l'énoncé du théorème comporte des conditions similaires à celles du résultat de convergence établi dans [83] à savoir la finitude de la norme de l'inverse d'un processus convexe surjectif associé au point initial, la lipschitziannité de la dérivée de f sur X_0 et un nombre strictement positif η pour lequel la distance entre le premier itéré x_1 et le point initial x_0 est inférieure à $1/2$. Cette dernière résulte du fait que η inférieur à $\frac{(1-BM)^2}{2BL}$ entraîne que $\eta \leq \frac{1}{2BL}$, soit $h = BL\eta \leq \frac{1}{2}$.
Remarquons également que le théorème est similaire au Théorème 3.2.1 de Piétrus où la lipschitziannité de la fonction perturbée $\nabla f(x)+A(x)$ de la dérivée de Fréchet $\nabla f(x)$ par une N-P-dérivée uniformément bornée $A(x)$ à la lipschitziannité de la fonction g. Cela justifie pourquoi la convergence de notre processus itératif repose sur celle de la suite majorante utilisée par ce dernier.

La proposition qui suit montre que notre méthode converge super-linéairement. La démontra-tion de laquelle permettra de déduire la convergence quadratique lorsque la multiapplication F est identiquement égale au singleton $\{0\}$.

3.2.1 Estimation de l'erreur et vitesse de convergence

Proposition 3.2.2 *Soit une suite $(x_k)_k$ générée par la méthode et majorée par la suite $(t_k)_k$ définie par*

$$\begin{cases} t_0 = 0,\ t_1 = \eta. \\ t_{k+1} - t_k = \dfrac{1}{1 - BLt_k}\Big[\dfrac{BL}{2}(t_k - t_{k-1})^2 + BM(t_k - t_{k-1})\Big],\ \textit{pour}\ k = 0, 1, 2, \cdots \end{cases}$$

avec $\eta,\ B,\ L, M$ vérifient les hypothèses du théorème précédent. Alors s'il existe une constante $C > 0$ telle que $h = C\frac{BL}{2}\eta + BM < 1$, alors une estimation de l'erreur de la méthode à l'ordre k

$$\|x^* - x_k\| \leq \eta\frac{h^k}{1-h}$$

où x^* est la limite de la suite (x_k).

Preuve
On remarque d'abord qu'on a , en utilisant le Lemme 2.2.1, la stricte croissance de la suite $(t_k)_k$ vers le réel positif t^* (limite de la suite $(t_k)_k$) figurant dans l'hypothèse de notre théorème de convergence.
On a par suite, l'existence d'une constante strictement positive $C > 0$ telle que $\frac{1}{1-BLt_k} \leq C$.
On en déduit que

$$t_{k+1} - t_k \leq C\Big(\frac{BL}{2}(t_k - t_{k-1}) + BM\Big)(t_k - t_{k-1}). \tag{3.6}$$

Remarquons maintenant que si $t_k - t_{k-1} \leq \eta$, alors

$$t_{k+1} - t_k \leq C\Big(\frac{BL}{2}\eta + BM\Big)(t_k - t_{k-1}) \leq h\eta \leq \eta.$$

Ainsi, $t_{k+1} - t_k \leq \eta,\ k = 0, 1, ...$
Donc,

$$t_{k+1} - t_k \leq h(t_k - t_{k-1}),\ k = 0, 1, ...$$

En appliquant k fois l'inégalité précédente, on obtient

$$t_{k+1} - t_k \leq h^2(t_{k-1} - t_{k-2}) \leq .... \leq h^k(t_1 - t_0) = h^k\eta.$$

Il en résulte que pour tous entiers naturels $m \geq 1$ et $k \geq 0$,

$$\begin{aligned}\|x_{k+m} - x_k\| &\leq \|(x_{k+m} - x_{k+m-1})\| + \|(x_{k+m-1} - x_{k+m-2})\| + ... + \|(x_{k+1} - x_k)\| \\ &\leq (t_{k+m} - t_{k+m-1}) + (t_{k+m-1} - t_{k+m-2}) + ... + (t_{k+1} - t_k) \\ &\leq \eta h^{k+m-1} + \eta h^{k+m-2} + ... + \eta h^k = \eta h^k(h^{m-1} + h^{m-2} + ... + 1) = \eta h^k \sum_{j=0}^{m-1} h^j.\end{aligned}$$

La constante positive h étant strictement inférieure à 1, la somme géométrique $\sum_{j=0}^{m-1} h^j$ converge vers $\frac{1}{1-h}$ quand m tend vers l'infini, et ceci indépendemment de k.
On en déduit que

$$\|x^* - x_k\| \leq \eta\frac{h^k}{1-h}.$$

D'où, une estimation de l'erreur à l'ordre k

$$\|x^* - x_k\| \leq \eta\frac{h^k}{1-h}.$$

Ceci prouve que notre méthode converge super-linéairement, ce qui achève la preuve. □

Il importe de souligner que si $F(x) := \{0\}$, la borne uniforme $M = 0$ et donc $h = \frac{1}{2}\eta CBL$. L'inégalité (3.6) devient

$$t_{k+1} - t_k \leq \frac{1}{2}CBL(t_{k-1} - t_k)^2.$$

D'autre part, si l'on suppose que $t_k - t_{k-1} \leq \eta h^{2^{k-1}-1}$, on a

$$t_{k+1} - t_k \leq \frac{1}{2}CBL(\eta h^{2^{k-1}-1})^2 = \frac{1}{2}CBL\eta^2 h^{2(2^{k-1}-1)} = \frac{1}{2}\eta CBL\eta h^{2^k-2}.$$

Or, $h = \frac{1}{2}\eta CBL$, il vient

$$t_{k+1} - t_k \leq \eta h^{2^k-1}.$$

Ce qui permet d'écrire

$$t_{k+1} - t_k \leq \eta h^{2^k-1} \text{ pour tout } k = 0, 1, ...$$

Cela implique que

$$t_{k+m} - t_k \leq \eta h^{2^{k+m-1}-1} + \eta h^{2^{k+m-2}-1} + ... + \eta h^{2^k-1}, \ \forall m \geq 1, \ k \geq 0.$$

Comme quel que soit $p \in \mathbb{N}$, $2^p \geq p$, l'inégalité précédente s'écrit

$$t_{k+m} - t_k \leq \eta h^{-1}\left[h^{2^k} + ... + h^{(m-2)2^k} + h^{(m-1)2^k}\right] = \eta h^{-1} \sum_{j=1}^{m-1} \left(h^{2^k}\right)^j.$$

Fixons k. En passant à la limite sur m et en tenant compte que cette somme est une somme géométrique de m termes et de raison $h^{2^k} < 1$, on obtient

$$t^* - t_k \leq \eta h^{-1} \frac{h^{2^k}}{1 - h^{2^k}} = \eta \frac{h^{2^k-1}}{1 - h^{2^k}}.$$

Or, $\|x^* - x_k\| \leq t^* - t_k$, une estimation de l'erreur à l'ordre k est

$$\|x^* - x_k\| \leq \eta \frac{h^{2^k-1}}{1 - h^{2^k}}.$$

Ainsi la méthode converge avec une vitesse quadratique vers x^*. □

De telles bornes d'erreur peuvent également se retrouver par exemple dans l'article de Mysovskih [61].

Notons que lorsqu'on perturbe la partie multivoque de l'extension de Robinson [83] par une multiapplication (variable) N-P-différentiable, à graphe fermé et de dérivée Nachi-Penot uniformément bornée ajoutée à une dérivée de Fréchet telle que leur somme soit Lipschitz, la convergence de la méthode s'obtient, à partir des hypothèses similaires et imposant l'hypothèse (v) aux solutions des itérés de la méthode à construire.

Soulignons également que si $F(x) := \{0\}$, la méthode est identique à celle proposée par Robinson [83]. Ainsi, la méthode généralise celle proposée par ce dernier.

La section qui suit aborde l'aspect numérique de notre méthode. Nous présentons un exemple où l'on montre que les solutions approchées associées à deux points initiaux qui sont proches sont un peu éloignées.
Dans la foulée, nous verrons que la deuxième solution approchée nécessite un temps de calcul un peu plus long. Ainsi, Celle-ci sera donc plus éloignée de la solution exacte.

3.3 Applications numériques

Considérons un exemple en dimension finie où $X = \mathbb{R}^2$, $Y = \mathbb{R}^3$ et $K = \mathbb{R}^2_- \times \{0\}$. Les traitements numériques ont été efffectués dans MAPLE Software avec un processeur "Intel®core i7 ".

Exemple. Soit à résoudre le système :

$$\begin{cases} x_1^2 + x_2^2 - 1 \leq 0, \\ x_1^2 + (x_2 - 1)^2 - 1 \leq 0, \\ (x_1 - 1)^2 + (x_2 - 1)^2 - 1 \in [0, 10^{-2}]. \end{cases} \tag{3.7}$$

Ce système est inspiré de celui de Robinson ([83]) en remplaçant le second membre de l'équation (3) par un intervalle contenant zéro.
Nous remarquons que les points

$$x^* = \left(\frac{1}{2}, 1 - \frac{1}{2}\sqrt{3}\right) = (0.5000000000, 0.1339745962)$$

et $x^{**} = (0, 1)$ sont deux solutions du système. En effet, l'ensemble des solutions (voir Fig 3.1) est un arc de couronne de centre $(1, 1)$ et de rayons $r_1 = 1$ et $r_2 = 1.01$, délimité par des arcs de cercle de rayon 1, centrés en $(0, 1)$ et $(0, 0)$ passant respectivement par les points x^* et x^{**}.

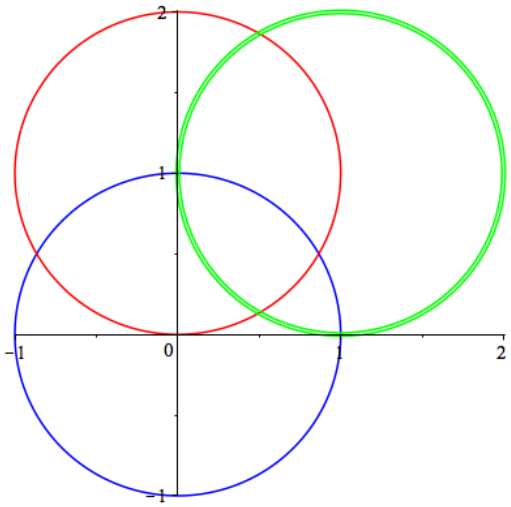


FIGURE 3.1 – L'ensemble des solutions du système

Le système (3.7) est résolu numériquement par la méthode (3.2) en posant :

$$f(x) := \left(x_1^2 + x_2^2 - 1, x_1^2 + (x_2 - 1)^2 - 1, (x_1 - 1)^2 + (x_2 - 1)^2 - 1\right),$$

et

$$F(x) := \{0\} \times \{0\} \times [-10^{-1}, 0].$$

Rappellons que notre méthode génère une suite (x_k) où l'itéré x_{k+1} est solution du problème

$$Min\{\|x - x_k\| \ / \ f(x_k) + (A(x_k) + f'(x_k))(x - x_k) \in K - F(x_k)\}, \qquad (3.8)$$

où A est une dérivée Nachi-Penot de F qui est identiquement nulle.
En utilisant l'algorithme (3.2) avec $x_0 = (0.55, 0.1)$ comme point initial, M_k la valeur de la solution du problème de minimisation précédente au pas k, et $\overline{x} = (0.50418137, 0.13640221)$ la solution approximative de la solution exacte x^*, on obtient :

step k	$\|x_k - \overline{x}\|_1$	M_k	$\|x_k - x_{k-1}\|_1$
0	0.04581863		
1	0.00222084	0.00325000	0.08000000
2	0.00000180	0.00000263	0.00221904
3	$2.08684747 \times 10^{-9}$	$1.73194458 \times 10^{-12}$	0.00000180
4	$2.08653623 \times 10^{-9}$	$7.49908004 \times 10^{-25}$	$1.18445885 \times 10^{-12}$
5	$2.08653623 \times 10^{-9}$	$1.40590503 \times 10^{-49}$	$5.12854270 \times 10^{-25}$

Table 1 *: Solutions de* (3.7) *en utilisant (3.2) avec* $x_0 = (0.55, 0.1)$ *; cpu time = 0.037 s.*

En démarant l'algorithme avec un point initial $x_0 = (0.51, 0.13)$ plus proche de la solution exacte, on a les résultats suivants :

step k	$\|x_k - \overline{\overline{x}}\|_1$	M_k	$\|x_k - x_{k-1}\|_1$
0	0.687005×10^{-3}		
1	0.4909×10^{-4}	0.7177×10^{-4}	0.1183908×10^{-1}
2	$5.45764467 \times 10^{-9}$	$1.28774923 \times 10^{-9}$	0.4909×10^{-4}
3	$6.33853714 \times 10^{-9}$	$4.14574519 \times 10^{-19}$	$8.80892473 \times 10^{-10}$
4	$6.33853714 \times 10^{-9}$	$4.29680079 \times 10^{-38}$	$2.83592151 \times 10^{-19}$
5	$6.33853714 \times 10^{-9}$	$4.61562427 \times 10^{-76}$	$2.93925198 \times 10^{-38}$

Table 2 *: Solutions de* (3.7) *en utilisant* (3.2) *avec* $x_0 = (0.51, 0.13)$ *; cpu time = 0.12 s.*

Ici, la solution approximative est $\overline{\overline{x}} = (0.50498188, 0.13687005)$.
Nous pouvons constater que ce point est plus éloigné de la solution exacte que la première solution approchée x^* pourtant l'algorithme a été initialisé avec un point plus proche de la solution exacte. Ceci montre que la méthode semble avoir une sensibilité importante par rapport au choix du point initial.

Il semble naturel de se demander : qu'adviendra-t-il si on perturbe la partie univoque f du problème (3.1) par une fonction univoque g qui est l_2-lipschitzienne (non nécessairement différentiable) sur X_0 ? Nous verrons dans quelle mesure que notre méthode sera adaptée à la résolution du problème perturbé.

3.4 Perturbation de la partie univoque par une fonction non différentiable mais lipschitzienne

Considérons une fonction g qui est l_2-lipschitzienne (non nécessairement différentiable) sur X_0.
Nous nous recherchons donc une méthode itérative pour approcher une solution x^* de l'inclusion perturbée

$$0 \in f(x) + g(x) + F(x) - K. \tag{3.9}$$

3.4.1 Description de la méthode

Considérons le processus convexe

$$T_w(x) = (\nabla f(w)v + A(w))x - K, \quad x \in X,\ w \in X_0,$$

d'inverse T_w^{-1} qui est aussi un processus convexe, défini par

$$T_w^{-1}(y) := \{z \in X \ / \ (\nabla f(w) + A(w))z \in y + K\}, \quad y \in Y.$$

De façon similaire à la description de la méthode relative au problème (3.1), en projetant successivement l'origine 0 de X sur les ensembles convexes non vides $T_{x_k}^{-1}[-f(x_k) - g(x_k) - F(x_k)]$ où l'itéré initial $x_0 \in X_0$ vérifie

$$T_{x_0}^{-1}[-f(x_0) - g(x_0) - F(x_0)] \neq \emptyset,$$

nous obtenons une suite (x_k) dont le terme x_{k+1} est solution du problème

$$\min\{\|x - x_k\| \ / \ f(x_k) + g(x_k) + (\nabla f(x_k)) + A(x_k))(x - x_k) \in K - F(x_k)\}. \tag{3.10}$$

Observons que nous avons utilisé l'inverse du processus convexe de l'itération précédente dans la construction de la méthode (3.10). Ici, au lieu de projeter sur l'image de l'opposé de la somme $f(x_k) + F(x_k)$ par l'inverse de ce processus convexe, nous projetons plutôt sur l'image de l'opposé de cette somme perturbée par $g(x_k)$ par l'inverse de ce même processus convexe. Par suite, l'itération (3.10) revient à minimiser la fonction $x \mapsto \|x - x_k\|$ sur l'ensemble des contraintes où le membre de gauche de l'équation caractéristique est celui de l'ensemble des contraintes de (3.2) perturbée par $g(x_k)$ et de prendre pour l'itéré suivant l'un ou le minimum.
Notons par ailleurs que l'existence d'une telle suite est justifiée par la Remarque 3.1.1.

3.4.2 Convergence de la méthode

Le théorème qui suit témoignera que l'hypothèse de Lipschitz du perturbateur g influencera seulement l'hypothèse $BM < 1$ du Théorème 3.2.1. Nous illustrerons que la convergence nécessitera le changement $BM < 1$ en $B(M + l_2) < 1$ où l_2 sera précisé ci-dessous.

Théorème 3.4.1
Considérons un sous-ensemble ouvert $X_0 \subset X$. Soient $g : X \to Y$ une fonction l_2-lipschitzienne sur $X_0 \subset X$, $f : X \to Y$ et $F : X \rightrightarrows Y$ respectivement Fréchet-différentiable et N-P-différentiable sur X_0, et dont les dérivées sont respectivement $\nabla f(x)$ et $A(x)$ pour tout $x \in X_0$. Supposons que F soit à graphe fermé, à valeurs compactes et convexes.
S'il existe des nombres strictement positifs B, L, l, l_2, M et η vérifiant

(i) $\|\tilde{T}_{x_0}^{-1}\| \leq B$;

(ii) $\sup_{x \in X_0} \|A(x)\| \leq M$;

(iii) $\|\nabla f(x) - \nabla f(x')\| \leq l\|x - x'\|$, *pour tous* $x, x' \in X_0$;

(iv) $\|g(x) - g(x^{'}\| \leq l_2\|x - x^{'}\|$, *quels que soient* $x, x' \in X_0$;

(v) $\|(\nabla f(x) + A(x)) - (\nabla f(x') + A(x'))\| \leq L\|x - x'\|$, *pour tous* $x, x' \in X_0$;

(vi) Pour tout x_{k+1} solution de (3.10), $F(x_k) \subset F(x_{k+1}) - K$, $k = 1, 2, \ldots$;

(vii) $l < L$, $B(M + l_2) < 1$;

(viii) $\|x_1 - x_0\| \leq \eta \leq \min\left\{\dfrac{1}{BL}, \dfrac{(1 - B(M + l_2))^2}{2BL}\right\}$, *où x_1 est un point obtenu à partir de x_0 par notre méthode,*

alors il existe un nombre strictement positif

$$t^* = \frac{1}{BL}\Big(1 - B(M + l_2) - \sqrt{(1 - B(M + l_2))^2 - 2BL\eta}\Big)$$

tel que pour tout point initial x_0 satisfait $\mathbb{B}_{t^*}(x_0) \subset X_0$, il existe une suite $(x_k)_k$ générée par l'algorithme (3.2) *telle que $\forall k \geq 0$, x_k appartienne à la boule $\mathbb{B}_{t^*}(x_0)$ et converge vers x^* solution du problème* (3.9).

Preuve
Encore une fois, des arguments similaires à ceux utilisés dans la preuve du théorème précédent nous assure que si les itérés $x_0, x_1, ..., x_j$ générés par notre méthode vérifient

$$\|x_{k+1} - x_k\| \leq t_{k+1} - t_k, \ k = 1, ..., j,$$

alors ils sont tous dans la boule $I\!B_{t^*}(x_0)$.
Pour montrer l'existence de x_{j+1} solution du problème de minimisation convexe (3.2), nous ferons appel à la surjectivité du processus convexe T_{x_j}.
Comme le processus convexe T_{x_j} peut s'écrire sous la forme

$$\begin{aligned} T_{x_j}(x) &= \Big(\nabla f(x_0) + A(x_0)\Big)x + \Big(\nabla f(x_j) + A(x_j) - (\nabla f(x_0) + A(x_0))\Big)x - K \\ &= T_{x_0}(x) + (\nabla f(x_j) + A(x_j) - (\nabla f(x_0) + A(x_0))x, \end{aligned}$$

en combinant la conclusion du Théorème 1.3.6 et des hypothèses (i) et (v), le processus convexe T_{x_j} est défini sur X et à ensemble-images défini sur tout l'espace Y ; on conclut donc $T_{x_j}^{-1}$ est normé et sa norme satisfait

$$\begin{aligned} \|T_{x_j}^{-1}\| &\leq \frac{\|T_{x_0}^{-1}\|}{1 - \|T_{x_0}^{-1}\|\|(\nabla f(x_j) + A(x_j)) - (\nabla f(x_0) + A(x_0))\|} \\ &\leq \frac{B}{1 - BL\|x_j - x_0\|}. \end{aligned}$$

Il résulte de la surjectivité et la définition sur X du processus convexe T_{x_j} que le problème (3.2) est faisable et donc solvable pour $k = j$ et nous obtenons l'existence de x_{j+1} solution du problème (3.2).

Notons que jusqu'ici les mêmes arguments de la preuve du théorème précédent ont été utilisés.
Maintenant, l'hypothèse de lipschitziannité de g sera prise en compte pour justifier l'appartenance de x_{j+1} à la boule $I\!B_{t*}(x_0)$.
Comme x_j est solution du problème (3.2), l'inclusion

$$f(x_j)+g(x_j)+(\nabla f(x_j)+A(x_j))(x-x_j) \in f(x_{j-1})+g(x_{j-1})+(\nabla f(x_{j-1})+A(x_{j-1})(x_j-x_{j-1})$$

peut s'écrire

$$f(x_j) + g(x_j) + (\nabla f(x_j) + A(x_j))(x - x_j) \in K - F(x_{j-1}) \subset K - F(x_j).$$

Des raisonnements similaires à la preuve de notre résultat précédent nous conduisent à l'existence de x_{j+1} tel que

$$\|x_{j+1}-x_j\| \leq \|T_{x_j}^{-1}\| \, \|f(x_{j-1})-f(x_j)+g(x_j)+g(x_{j-1})+(\nabla f(x_{j-1})+A(x_{j-1}))(x_j-x_{j-1})\|.$$

Par inégalité triangulaire, le second facteur du membre de droite de l'inégalité précédente est inférieur à

$$\frac{l}{2}(\|x_k - x_{k-1}\|^2 + (M + l_2)\|x_k - x_{k-1}\|.$$

D'où

$$\|x_{j+1} - x_j\| \leq \frac{B}{1 - BL\|x_j - x_0\|}\Big(\frac{L}{2}\|x_j - x_{j-1}\|^2 + (M + l_2)\|x_j - x_{j-1}\|\Big)(\text{ car } l < L).$$

Ainsi, la suite $(t_k)_k$ définie par

$$\begin{cases} t_0 = 0,\ t_1 = \eta. \\ t_{k+1} - t_k = \dfrac{1}{1 - BLt_k}\Big[\dfrac{BL}{2}(t_k - t_{k-1})^2 + (M + l_2)(t_k - t_{k-1})\Big],\ \text{pour } k = 0, 1, 2, \cdots \end{cases}$$

est donc une suite majorante de la suite $(x_k)_k$ généré par notre méthode.
En appliquant donc le Lemme 2.2.1 avec $\alpha := \eta$; $p_1 := \frac{BL}{2}$; $p_2 := B(M + l_2)$; $p_3 := 0$, $p_4 = BL$, la suite $(t_k)_k$ est croissante et converge vers t^*.
Il en résulte que la suite $(x_k)_k$ est de Cauchy dans l'espace réflexif X ; elle est donc convergente vers x^* qui est solution du problème (3.9) (via des raisonnements similaires figurant dans la deuxième partie de la preuve du théorème précédent).

Par ailleurs, l'inégalité $\|x_{j+1} - x_j\| \leq t_{j+1} - t_j$ implique que

$$\begin{aligned} \|x_{j+1} - x_0\| &\leq \|x_{j+1} - x_j\| + \|x_j - x_{j-1}\| + ... + \|x_1 - x_0\| \\ &\leq t_{j+1} - t_j + t_j - t_{j-1} + ... + t_1 - t_0 = t_{j+1} - t_0 \leq t^*. \end{aligned}$$

Donc $x_{j+1} \in I\!B_{t^*}(x_0)$. $\square$

Nous venons de confirmer que si la partie multivoque du problème (3.1) est perturbée par une application l_2- lipschitzienne (non nécessairement différentiable), la suite des itérés $(x_k)_k$ s'obtient, sous certaines hypothèses, en rajoutant l'image $g(x_k)$ à l'ensemble des contraintes du problème (3.2). De plus, si le produit de facteurs la somme de constante de Lipschitz l_2 et la borne uniforme M d'une dérivée Nachi-Penot de F par un nombre strictement positif B majorant la norme de l'inverse du processus convexe surjectif T_{x_0} est strictement inférieur à 1, la suite des itérés converge vers une solution du problème perturbé.
D'autre part, notons également que toutes les hypothèses du Théorème 3.2.1 demeurent inchangés sauf $BM < 1$ qui change en $B(M + l_2) < 1$.
Ainsi, pour avoir la convergence de la méthode, il suffit de garder toutes les hypothèses du Théorème 3.4.1 et de remplacer M par $M + l_2$. Par suite, la borne uniforme M de

la dérivée Nachi-Penot figurant dans la suite majorante se changera en $M + l_2$.

Il convient de faire des remarques sur des cas spéciaux.

Remarques 3.4.2 *1) Si $F(x) := \{0\}$, on peut prendre pour N-P-dérivée $A(x) = 0$ et donc la borne uniforme $M = 0$. Ce qui entraîne que les hypothèses (iii) et (v) coincident, le nombre positif η est inférieur ou égal au minimum des deux nombres $\frac{1}{BL}$ et $\frac{(1-Bl_2)^2}{2BL}$ et le nombre*

$$t^* = \frac{1}{BL}\Big((1 - Bl_2) - \sqrt{(1 - Bl_2)^2 - 2BL\eta}\Big).$$

On obtient par ailleurs la chute de l'hypothèse (v) puisqu'on a l'appartenance de l'origine 0 à $-K$ et dès le départ on a supposé que l'origine était élément du cône K, donc de $-K$.
Notre méthode est alors similaire à celle proposée par Piétrus dans [78] ;

2) L'estimation de l'erreur est préservée en remplaçant M par $M + l_2$;

3) Si $K := \{0\}$ et $F(x) := \{0\}$ notre méthode est exactement celle construite en 1963 par Zinchenko [92].
Ce qui est important de dire, c'est que la méthode généralise celle introduite par Piétrus [78].

Le chapitre suivant fera l'objet de conception et construction d'une méthode itérative pour résoudre le problème où l'on perturbe la partie univoque du problème précédent par une fonction g non différentiable mais admettant des différences divisées.
Notons que le même processus convexe sera considéré ; la convergence de notre méthode sera basée sur une suite majorante que nous construirons à partir de celle de Piétrus et Jean-Alexis [79]. Donc, la preuve de sa convergence sera inspirée de celle de ces derniers. Par ailleurs, nous constaterons que certaines hypothèses du théorème de Piétrus et Jean-Alexis seront nécessaires à la convergence de notre théorème.
Notons également que le problème que nous aurons à résoudre pourra être vu comme un problème où l'on perturbe la partie multivoque du problème résolu par Piétrus et Jean-Alexis [79] par une multiapplication N-P-différentiable avec dérivées uniformément bornées.

Chapitre 4

Méthode de type Newton-sécante pour une application multivoque perturbée contenue dans un cône

Suite aux travaux portant sur la résolution des inclusions perturbées par une multiapplication N-P-différentiable et à dérivées uniformément bornées, nous nous intéressons au cas où la partie univoque de l'inclusion (3.1) est perturbée par une telle multiapplication, c'est-à-dire, la résolution de l'inclusion perturbée

$$0 \in f(x) + g(x) + G(x) - K. \tag{4.1}$$

À cause du manque de régularité de g, nous ne pouvons pas appliquer les méthodes précédentes. Alors pour atteindre notre objectif, nous nous inspirons des travaux de Piétrus et Jean-Alexis [79]. Nous allons donc résoudre le problème précédent par une combinaison de deux méthodes : les méthodes de Newton et sécante. Nous obtiendrons une méthode semi-locale avec une vitesse de convergence égale à $\frac{1+\sqrt{5}}{2}$.

Dans la suite de ce chapitre, rappelons qu'on suppose encore que le perturbateur g admette des différences divisées du premier et du second ordre.

4.1 Description et convergence de la méthode

Nous faisons d'abord appel à une proposition qui permettra de décrire notre méthode. Elle établira que la somme du processus convexe fermé $\tilde{T}_w$ de la proposition du chapitre précédent et la différence divisée du premier ordre est un processus convexe. Ainsi, sous certaines hypothèses, son inverse sera normé.

Proposition 4.1.1 *Soient v, w fixés de X_0 et $A(w)$ une application linéaire continue, on définit l' application multivoque $T(v,w)$ de X dans Y par*

$$T(v,w)x := (\nabla f(w) + A(w) + [v,w;g])x - K, \quad x \in X.$$

a) L'application multivoque $T(v,w)$ est un processus convexe ;

b) $T(v,w)$ est fermé et normé lorsque l'inverse $\tilde{T}_w^{-1}$ du processus convexe normé $\tilde{T}_w x := (\nabla f(w) + A(w))x - K$ vérifie $\|\tilde{T}_w^{-1}\|\|[v,w;g]\| < 1$. De plus, $T^{-1}(v,w)$ est normé et

$$\|T^{-1}(v,w)\| \leq \frac{\|\tilde{T}_w^{-1}\|}{1 - \|\tilde{T}_(w)^{-1}\|\|[v,w;g]\|}.$$

Preuve
Remarquons d'abord que

$$T(v,w)x = (\nabla f(w) + A(w))x - K + [v,w;g]x = \tilde{T}_w x + [v,w;g]x,$$

avec $\tilde{T}_w x := (\nabla f(w) + A(w))x - K$ qui est un processus convexe (via la Proposition 2.4.3).
La linéarité de la différence divisée $[v,w;g]$ et la Proposition 1.3.5 impliquant que la multiapplication $T(v,w)$ est un processus convexe.
Par suite, l'inverse $T^{-1}(v,w)$ de $T(v,w)$ est aussi un processus convexe puisque l'inverse d'un processus convexe est un processus convexe.

La continuité de $\nabla f(w) + A(w) + [v,w;g]$ et la fermeture du cône K nous permet d'affirmer que $T(v,w)$ est fermé.

Observons que l'hypothèse faite dans (b) nous permet d'appliquer le Théorème 1.3.6. Par conséquent, le processus convexe $T(v,w)$ est normé et surjectif. Son inverse $T^{-1}(v,w)$ est donc normé et une majoration de la norme de l'inverse

$$\|T^{-1}(v,w)\| \leq \frac{\|\tilde{T}_w^{-1}\|}{1 - \|\tilde{T}_(w)^{-1}\|\|[v,w;g]\|}.$$

Faisons place maintenant à la description de notre méthode.

4.1.1 Description de notre algorithme

Soient v, w fixés de X_0, on définit l'application multivoque $T(v,w)$ de X dans Y

$$T(v,w)x := (\nabla f(w) + A(w) + [v,w:g])x - K, \quad x \in X.$$

La proposition précédente implique que l'application multivoque $T(v,w)$ est un processus convexe de X dans Y. L'inverse de ce dernier est aussi un processus convexe, défini pour tout $y \in Y$ par

$$T^{-1}(v,w)y := \{z \in X \ / \ (\nabla f(w) + A(w) + [v,w;g])z \in y + K\}.$$

À cause de la différence divisée, deux points sont nécessaires pour initialiser notre méthode, nous les appelons x_0 et x_1.
Partant de ces deux points, nous faisons l'hypothèse que

$$T^{-1}(x_0,x_1)[-f(x_1) - g(x_1) - G(x_1)]$$

=

$$\{z \ \in X \mid f(x_1) + g(x_1) + (f^{'}(x_1) + A(x_1) + [x_0,x_1;g])z \in K - G(x_1)\} \neq \emptyset.$$

Puisque cet ensemble est convexe, fermé et inclus dans l'espace réflexif X, on y projette l'origine 0 de X; notons d_0 une projection.
On pose $x_2 - x_1 = d_0$.
Supposons qu' à la $k-$ième étape nous obtenions x_{k-1}, x_k, donc on définit $x_{k+1} - x_k$ comme une projection de l'origine 0 de X sur l'ensemble

$$T^{-1}(x_{k-1},x_k)[-f(x_k) - g(x_k) - G(x_k)].$$

On pose $x_{k+1} = x_k + d_{k-1}$.
Ainsi, on a construit une suite (x_k) vérifiant

$$x_{k+1} = x_k + d_{k-1}, \quad k = 0,1,...$$

où d_k est une projection de l'origine de X sur $T^{-1}(x_{k-1},x_k)[-f(x_k) - g(x_k) + G(x_k)]$.
De façon équivalente, si l'itéré précédent x_k étant trouvé, le nouvel itéré x_{k+1} est une solution du problème de minimisation

$$\min\{\|x - x_k\| \ / \ f(x_k) + g(x_k) + (\nabla f(x_k) + A(x_k) + [x_{k-1},x_k;g])(x - x_k) \in K - G(x_k)\}. \tag{4.2}$$

Il n'est pas difficile de voir que l'équation caractéristique de l'ensemble des contraintes de (4.2) est celle de l'itération proposée par Piétrus et Jean-Alexis dont l'application linéaire du membre de gauche est perturbée par $g(x_k)$ et le membre de droite par l'opposé de l'image $G(x_k)$.

Nous avons vu que la convergence de la méthode proposée par Piétrus et Jean-Alexis [79] se basait sur la convergence d'une suite majorante construite par ces derniers.

Comme l'idée est de généraliser ces travaux, nous allons adopter une approche similaire à celle proposées par ces derniers.
Nous allons donc construire une suite de réels positifs $(t_n)_n$ dont la relation de récurrence $t_{n+1} - t_n$ sera fonction de $(t_n - t_{n-1})^2$, du produit $(t_n - t_{n-1})(t_n - t_{n-2})$ et de $t_n - t_{n-1}$ avec des coefficients tels que la suite $(t_n)_n$ converge vers un nombre t^*.

4.2 Analyse de la convergence de la méthode

Pour la convergence de notre algorithme, on va énoncer et prouver une proposition fondamentale à notre résultat de convergence.

Proposition 4.2.1
Considérons la suite $(t_n)_n$ définie par

1. $t_0 = 0, \quad t_1 = \alpha, \quad t_2 = \beta;$

2. $t_{n+1} - t_n = M(\frac{L}{2}(t_n - t_{n-1})^2 + K_1(t_n - t_{n-1})(t_n - t_{n-2}) + K_2(t_n - t_{n-1}))$ *où* $\alpha, \beta, M, L, K_1, K_2$

   sont des constantes strictement positives.

   Si les conditions

3. $\alpha \leq M, \beta \leq 2\alpha;$

4. $q^{u_{l-2}} + K_2 M \leq q^{u_l}$ *avec* $q = M^2(\frac{L}{2} + 2K_1)$, *où la suite* (u_l) *est celle de Fibonacci définie par* $u_0 = u_1 = 1$ *et* $u_{l+1} = u_l + u_{l-1}$ *pour tout entier naturel* $l \geq 1$

sont toutes vérifiées, alors la suite $(t_n)_n$ est une suite d'éléments de la boule $\mathbb{B}_s(t_1)$ où $s = \frac{M}{q}\sum_{l=1}^{\infty} q^{u_l}$. De plus, la suite (t_n) converge, et une estimation de l'erreur est

$$t^* - t_n \leq \frac{M}{q\Big(1 - q^{\frac{p^n(p-1)}{\sqrt{5}}}\Big)} q^{\frac{p^n}{\sqrt{5}}}$$

avec $t^* = \lim t_n$ *et* $p = \frac{1+\sqrt{5}}{2}$.

Preuve
Nous allons procéder par récurrence, comme dans la preuve faite par Cãtinas dans [20], qui consistait à montrer (pour tout $n \geq 2$) que

$$t_n \in \mathbb{B}_s(t_1), \tag{4.3}$$

$$t_n - t_{n-1} \leq t_{n-1} - t_{n-2}, \tag{4.4}$$

$$t_n - t_{n-1} \leq q^{u_{n-1}-1} M. \tag{4.5}$$

Pour $k \geq 2$, supposons que (4.3), (4.4), (4.5) soient vérifiés pour $n \leq k$.
Comme

$$t_{k+1} - t_k = M\Big[\frac{L}{2}(t_k - t_{k-1})^2 + K_1(t_k - t_{k-2})(t_k - t_{k-1}) + K_2(t_k - t_{k-1})\Big]$$

et

$$t_k - t_{k-2} = t_k - t_{k-1} + t_{k-1} - t_{k-2},$$

on a

$$t_{k+1}-t_k = M\Big[(\frac{L}{2}(t_k-t_{k-1})(t_k-t_{k-1})+K_1(t_k-t_{k-1})(t_k-t_{k-1})+K_1(t_k-t_{k-1})(t_{k-1}-t_{k-2})+K_2(t_k-t_{k-1})\Big].$$

Compte tenu de l'hypothèse de récurrence (4.4), on a

$$t_{k+1} - t_k \leq M\Big[\Big(\frac{L}{2} + 2K_1\Big)(t_{k-1} - t_{k-2})(t_k - t_{k-1}) + K_2(t_k - t_{k-1})\Big].$$

Par application de l'hypothèse de récurrence (4.5), et en observant que les indices $k-1, k-2$ sont plus petits que k, on en déduit que

$$t_{k+1} - t_k \leq M\Big[\Big(\frac{L}{2} + 2K_1\Big)q^{u_{k-2}-1} M(t_k - t_{k-1}) + K_2(t_k - t_{k-1})\Big],$$

soit

$$t_{k+1} - t_k \leq \Big(M^2\Big(\frac{L}{2} + 2K_1\Big)q^{u_{k-2}-1} + K_2 M\Big)(t_k - t_{k-1}) = (q^{u_{k-2}} + K_2 M)(t_k - t_{k-1})$$

car $q = M^2(\frac{L}{2} + 2K_1)$.
Puisque $q^{u_{k-2}} + K_2 M \leq q^{u_k}$,

$$t_{k+1} - t_k \leq q^{u_k}(t_k - t_{k-1}) \leq t_k - t_{k-1}.$$

(4.4) est vraie pour $n = k + 1$.

Maintenant, montrons que

$$\forall n \geq 1,\ t_n - t_{n-1} \leq q^{u_{n-1}-1} M.$$

Pour $k \geq 2$, supposons que (4.5) soit vraie pour $n \geq k$. Montrons qu'elle est encore vraie pour $k+1$.
En combinant l'inégalité

$$t_{k+1} - t_k \leq (q^{u_{k-2}} + K_2 M)(t_k - t_{k-1})$$

et l'hypothèse de récurrence, on déduit que

$$t_{k+1} - t_k \leq (q^{u_{k-2}} + K_2 M) q^{u_{k-1}-1} M \leq q^{u_k + u_{k-1}-1} M = q^{u_{k+1}-1} M \leq q^{u_k - 1} M$$

car la suite de Fibonacci (u_n) est croissante et $q < 1$.
Ce qui prouve (4.4) pour $n = k+1$.
De plus, on a

$$t_{k+1} - t_k = (t_2 - t_1) + (t_3 - t_2) + ... + (t_k - t_{k-1}) + (t_{k+1} - t_k).$$

En utilisant (4.5) on obtient

$$t_{k+1} - t_1 \leq q^{u_1 - 1} M + q^{u_2 - 1} M + ... + q^{u_{k-1}-1} + q^{u_k - 1} M$$

$$= \frac{M}{q}(q^{u_1} + q^{u_2} + ... + q^{u_{k-1}} + q^{u_k}).$$

Or, $s = \frac{M}{q} \sum_{l=1}^{\infty} q^{u_l}$ et $q^{u_l} > 0$, donc $t_{k+1} - t_1 < s$. (4.4) est donc vraie pour $n = k+1$.

Nous allons montrer que la suite $(t_k)_k$ est une suite de Cauchy de X.
Pour tous $k \geq 1, m \geq 1$, on a

$$t_{k+m} - t_k \leq \frac{M}{q}(q^{u_k} + q^{u_{k+1}} + ... + q^{u_{k+m-2}} + q^{u_{k+m-1}}).$$

On sait que $u_l \geq \frac{p^l}{\sqrt{5}}, \quad l \geq 1$. En utilisant $p^{k+s} - p^k = p^k(p^s - 1)$, puis l'inégalité de Bernoulli, on a

$$t_{k+m} - t_k \leq \frac{M}{q} q^{\frac{p^k}{\sqrt{5}}} (1 + q^{\frac{p^k(p-1)}{\sqrt{5}}} + q^{\frac{p^k 2(p-1)}{\sqrt{5}}} + ... + q^{\frac{p^k(m-2)(p-1)}{\sqrt{5}}} + q^{\frac{p^k(m-1)(p-1)}{\sqrt{5}}}).$$

La somme de m termes $1 + q^{\frac{p^k(p-1)}{\sqrt{5}}} + q^{\frac{p^k 2(p-1)}{\sqrt{5}}} + ... + q^{\frac{p^k(m-2)(p-1)}{\sqrt{5}}} + q^{\frac{p^k(m-1)(p-1)}{\sqrt{5}}}$ étant géométrique de raison $q^{\frac{p^k(p-1)}{\sqrt{5}}} < 1$, et égale à $\frac{1 - q^{\frac{p^k(p-1)m}{\sqrt{5}}}}{1 - q^{\frac{p^k(p-1)}{\sqrt{5}}}}$ nous permet d'écrire

$$t_{k+m} - t_k \leq \frac{M q^{\frac{p^k}{\sqrt{5}}} (1 - q^{\frac{p^k(p-1)m}{\sqrt{5}}})}{q(1 - q^{\frac{p^k(p-1)}{\sqrt{5}}})}.$$

Lorsque k tend vers l'infini, le membre de droite de la dernière inégalité tend vers zéro, indépendemment de m. La suite (t_n) est donc de Cauchy dans X complet. D'où, elle converge vers $t^* \in X$. $\square$

Ainsi, nous justifions que la relation de récurrence de la suite $(t_n)_n$ diffère de celle construite par Piétrus et Jean-Alexis [79] du troisième terme $K_2(t_n - t_{n-1})$.

La convergence de notre méthode itérative sera inspirée de celle proposée par Piétrus et Jean-Alexis dans [79]. Ainsi, les conditions de notre théorème seront constituées de celles du Théorème 2.3.2, de la Proposition 2.4.4 et de la bornitude d'une dérivée de Nachi-Penot.

4.3 Convergence de la méthode

Théorème 4.3.1 *Soient v, w fixés de X_0, on définit l'application multivoque $T(v,w)$ de X dans Y par*

$$T(v,w)x := (\nabla f(w) + A(w) + [v, w : g])x - K, \quad x \in X.$$

Supposons qu'il existe des points $x_0,\ x_1 \in X_0$ tels que $T(x_0, x_1)$ soit défini sur X et surjectif et une application multivoque $G : X \rightrightarrows Y$ à graphe fermé , à valeurs compactes et convexes, N-P-différentiable et de dérivée $A(w)$ évaluée en $w \in X_0$.
S'il existe des nombres réels strictement positifs $B, L, K_1, K, M_1,\ \alpha$ satisfaisant les conditions suivantes :

(a) $\|T^{-1}(x_0, x_1)\| \leq B$;

(b) *pour tous $x, y \in X_0$,* $\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$;

(c) *g est continue sur X_0 et admet des différences divisées du premier et second ordre sur X_0;*

(d) *pour tout $k \in \mathbb{N}$, $G(x_k) \subset G(x_{k+1}) - K$, où la suite (x_k) est générée par notre méthode.*

(e) $BM_1 < 1$; $\alpha \leq \beta \leq 2\alpha$;

(f) $\|x_1 - x_0\| \leq \alpha \leq \frac{B}{1-BM_1}$, $\|x_2 - x_1\| \leq \beta$; *où x_2 est obtenu à partir de x_0 et x_1 suivant*

l'algorithme précédent(un tel point existe via la convexité de l'inverse de $T^{-1}(x_0, x_1)$*) ;*

(g) *pour tous points distincts* x, y *et* z *dans* $X_0, \|[x, y, z; g]\| \leq K_1$,

$\|A(x)\| \leq K_2, \|\nabla f(y) + A(y) + [x, y; g] - (\nabla f(x_1) + A(x_1)) - [x_0, x_1; g]\| \leq M_1$;

(h) $q^{u_l - 2} + \frac{B}{1 - BM_1} K_2 \leq q^{u_l}$, *avec* $q = \frac{B^2}{(1-BM_1)^2}(\frac{L}{2} + 2K_1) < 1$,

où la suite (u_l) *est celle de Fibonacci définie par* $u_0 = u_1 = 1$ *et* $u_{l+1} = u_l + u_{l-1}$ *pour tout entier naturel* $l \geq 1$.

Il existe t^* *tel que si* $I\!B_{t^*}(x_0) \subset X_0$*, alors l'algorithme* (4.2) *génère une suite* $(x_k)_k$ *telle que pour tout* $k \geq 0, x_k \in I\!B_{t^*}(x_0)$ *et converge vers un* x^* *vérifiant* $f(x^*) + g(x^*) \in K - G(x^*)$.

Preuve
Si l'on montre l'existence d'une suite convergente $(t_k)_k$ et d'une suite $(x_k)_k$ générées par l'algorithme telles que

$$x_k \in I\!B_{t^*}(x_0); \quad \|x_{k+1} - x_k\| \leq t_{k+1} - t_k, \quad k = 0, 1, ... \tag{4.6}$$

La suite (x_k) sera une suite de Cauchy dans la boule $I\!B_{t^*}(x_0)$, et convergera vers une solution x^* du problème (4.1), avec $x^* \in I\!B_{t^*}(x_0)$.

On a pour tout $k \in \mathbb{N}$,

$$f(x_{k+1} + g(x_{k+1}) - f(x^*) - g(x^*) - [f(x_{k+1}) - f(x_k) - (\nabla f(x_k) + A(x_k)$$

+

$$[x_{k-1}, x_k; g])(x_{k+1}) - x_k) - g(x_{k+1} - g(x_k)] \in K - f(x^*) - g(x^*) - G(x_k).$$

L'hypothèse de continuité implique que le membre de gauche tend vers zéro. Puisque l'ensemble $K - G(x_k) - f(x^*) - g(x^*)$ est fermé, d'après la Proposition 2.4.4,

$$K - G(x_k) - f(x^*) - g(x^*) \text{ tend vers } K - G(x^*) - f(x^*) - g(x^*).$$

Ainsi x^* est solution du problème (4.1).

On va montrer l'existence d'une suite $(x_k)_k$ de la boule $I\!B_{t*}(x_0)$ et d'une suite croissante de réels strictement positifs (t_k) de limite t^* telles que

$$\|x_{k+1} - x_k\| \leq t_{k+1} - t_k, \forall k \in \mathbb{N}.$$

L'hypothèse (e) donne l'existence d'un point x_2 solution de (4.1) à partir des points initiaux x_0, x_1.
Ainsi, en posant $t_2 = \beta$, $t_1 = \alpha$ et $t_0 = 0$, on obtient $\|x_1 - x_0\| \leq \alpha = t_1 - t_0$, $\|x_2 - x_1\| \leq \beta - \alpha = t_2 - t_1 \leq \alpha$ (car $\beta \leq 2\alpha$).
Ceci prouve l'inégalité figurant dans (4.6) pour $k = 0$ et $k = 1$.
Maintenant, supposons que pour $k \geq 2, x_0, x_1, x_2, \cdots, x_k$ soient générés par l'algorithme précédent, et $t_0, t_1, \cdots, t_k$ tels que $\|x_j - x_{j-1}\| \leq t_j - t_{j-1}, \ \forall j = 3, \cdots, k$.
Par suite,

$$\|x_j - x_0\| \leq \|x_j - x_{j-1}\| + ... + \|x_1 - x_0\| \leq t_j - t_0 = t_j$$

Nous allons donc prouver l'existence de x_{k+1} et t_{k+1} tels que $\|x_{k+1} - x_k\| \leq t_{k+1} - t_k$.

Remarquons que le processus convexe

$$T(x_{k-1}, x_k) = (\nabla f(x_k) + A(x_k) + [x_{k-1}, x_k; g])x - K$$

$\Leftrightarrow T(x_{k-1}, x_k) = (\nabla f(x_1) + A(x_1) + [x_0, x_1; g]x - K + (\nabla f(x_k) - \nabla f(x_1) + A(x_k) - A(x_1) + [x_{k-1}, x_k; g] - [x_0, x_1; g])x = (T(x_0, x_1) + \Delta_k)x$ avec

$$\Delta_k(x) = -(\nabla f(x_k) + A(x_k) - (\nabla f(x_1) + A(x_1) + [x_{k-1}, x_k; g] - [x_0, x_1; g])x.$$

En combinant les hypothèses (a), (d) et (f), on obtient

$$\|T^{-1}(x_0, x_1)\|.\|\Delta_k\| < 1.$$

Dans ce cas, l'application du théorème précédent nous permettra d'affirmer que $T(x_{k-1}, x_k)$ est normé, défini sur X et surjectif, donc vérifie

$$\|T^{-1}(x_{k-1}, x_k)\| \leq \frac{\|T^{-1}(x_0,\ x_1)}{1 - \|T^{-1}(x_0, x_1)\|.\|\Delta_k\|} \leq \frac{B}{1 - BM_1}.$$

Considérons le problème consistant à trouver un point x solution de

(4.7) $\quad f(x_k) + g(x_k) + (\nabla f(x_k) + A(x_k) + [x_{k-1}, x_k; g](x - x_k) \in f(x_{k-1}) + g(x_{k-1}) + ((\nabla f(x_{k-1}) + A(x_{k-1})) + [x_{k-2}, x_{k-1}; g])x + K.$

x_k étant solution de (4.1), le membre de droite de (4.7) est contenu dans $K - G(x_{k-1}) + K \subset K - G(x_k)$ (d'après (d)). On en déduit que tout x solution de (4.7) est nécessairement un point où le problème (4.1) est faisable.

Le membre de droite étant fermé et non vide, inclus dans l'espace réflexif X, il existe $\tilde{x}$ solution de (4.7), donc de (4.1).
Par inégalité triangulaire de la norme,

$$\|\tilde{x} - x_k\| \leq \|T^{-1}(x_{k_1}, x_k)\|(\| - f(x_k) + f(x_{k-1}) + f(x_{k-1})(x_k - x_{k-1})\|$$

+

$$\|g(x_{k-1}) - g(x_k)([x_{k-1} - x_k; g]) - [x_{k-2}, x_k; g])(x_k - x_{k-1})\| + \|A(x_{k-1})\|\|(x_k - x_{k-1}))\|).$$

Un petit calcul nous fournit

$$\| - f(x_k) + f(x_{k-1}) + \nabla f(x_{k-1})(x_k - x_{k-1})\| \leq \frac{L}{2}\|x_k - x_{k-1}\|^2,$$

$$\|g(x_{k-1}) - g(x_k) + [x_{k-2}, x_k; g](x_k - x_{k-1})\|$$

=

$$\|[x_{k-1}, x_k; g](x_{k-1} - x_k) - [x_{k-2}, x_{k-1}; g](x_{k-1} - x_k)\|$$

=

$$\|x_{k-2}, x_{k-1}, x_k; g](x_k - x_{k-2})(x_k - x_{k-1}) \leq K_1\|(x_k - x_{k-2})(x_k - x_{k-1})\| \text{ et}$$

$$\|A(x_{k-1})(x_k - x_{k-1})\| \leq \|A(x_{k-1}\|\|x_k - x_{k-1}\| \leq K_2\|x_k - x_{k-1}\|.$$

On en déduit que

$$\|\tilde{x} - x_k\| \leq M\Big(\frac{L}{2}(x_k - x_{k-1})^2 + K_1\|x_k - x_{k-2}\|\|x_k - x_{k-1}\| + K_2\|(x_k - x_{k-1}\|)\Big),$$

avec $M = \frac{B}{1-BM_1}$.

Grâce à l'inégalité $\|x_{k+1} - x_k\| \leq t_{k+1} - t_k, \ \forall k \in \mathbb{N}$, on a

$$\|\tilde{x} - x_k\| \leq M\Big(\frac{L}{2}(t_k - t_{k-1})^2 + K_1(t_k - t_{k-2})(t_k - t_{k-1}) + K_2(t_k - t_{k-1})\|\Big).$$

Le membre de droite de l'inégalité précédente nous invite à considérer la suite majorante définie par

$$t_{k+1} - t_k = M\Big(\frac{L}{2}(t_k - t_{k-1})^2 + K_1(t_k - t_{k-2})(t_k - t_{k-1}) + K_2(t_k - t_{k-1})\Big)$$

avec $t_1 = \alpha$, $t_2 = \beta$ et $t_0 = 0$.
La proposition et les hypothèses du théorème nous permettront de conclure que la suite $(t_k)_k$ est strictement croissante et converge vers un nombre t^*.
Le point $\tilde{x}$ étant un point où le problème de minimisation est faisable pour $j = k$, la définition de x_{k+1} nous fournit

$$\|x_{k+1} - x_k\| \leq t_{k+1} - t_k.$$

Ce qui termine la preuve. □

Notons que ce théorème est constitué des conditions du Théorème 3.4.1 et d'hypothèses liées à la multiapplication G. En fait, toutes sont conservées sauf l'hypothèse f) qui devient petit g) en perturbant la dérivée de Fréchet par une N-P-dérivée uniformément bornée $A(x)$ et imposant une condition sur le nombre q.

Nous achevons ce chapitre par une proposition justifiant que notre méthode est au moins d'ordre p et un majorant de l'erreur $\|x^* - x_n\|$ est fonction de M, q et p.

4.3.1 Estimation de l'erreur et vitesse de convergence

Proposition 4.3.2 *Soit $(x_k)_k$ une suite générée par la méthode (4.2) et majorée par la suite convergente $(t_k)_k$ définie dans la Proposition 4.2.1. Alors*

a) la méthode est au moins d'ordre p ;

b) une estimation de l'erreur

$$\|x^* - x_n\| \leq \frac{M}{q\Big(1 - q^{\frac{p^n(p-1)}{\sqrt{5}}}\Big)} q^{\frac{p^n}{\sqrt{5}}}.$$

Preuve
Soit $(x_n)_n$ une suite générée par la méthode (4.2) et majorée par la suite convergente $(t_k)_k$ définie dans la Proposition 4.2.1.
Par suite, $\|x_{k+m} - x_k\| \leq t_{k+m} - t_k$ pour tous $m, n \in \mathbb{N}$. Puis en fixant k et faisant tendre m vers l'infini, on a

$$\|x^* - x_k\| \leq t^* - t_k(\text{ car la suite } t_k \text{ converge vers } t^* \ ).$$

Une estimation de l'erreur de la suite $(t_k)_k$ étant

$$t^* - t_k \leq \frac{M}{q\Big(1 - q^{\frac{p^n(p-1)}{\sqrt{5}}}\Big)} q^{\frac{p^n}{\sqrt{5}}},$$

on a donc

$$\|x^* - x_k\| \leq \frac{M}{q\Big(1 - q^{\frac{p^n(p-1)}{\sqrt{5}}}\Big)} q^{\frac{p^n}{\sqrt{5}}}.$$

Ceci prouve que l'itération (4.2) converge vers x^* avec une vitesse p. $\square$

Remarque 4.3.3 *Si $G(x) = \{0\}$ pour tout $x \in X$, la méthode est exactement celle de Newton-sécante proposée dans [79].*

Remarque 4.3.4 *Si $g(x) = 0$ et $G(x) = \{0\}$ pour tout $x \in X$, la méthode est exactement la méthode de type Newton proposée dans [83] et dans ce cas on obtient une convergence quadratique.*

Il est intéressant de remarquer également que le théorème de convergence est constitué des hypothèses du résultat énoncé dans [79], des hypothèses d) et i) et de la bornitude uniforme d'une N-P-dérivée de G.

Chapitre 5

Méthode de type Newton semi-lisse pour une application multivoque contenue dans un cône

5.1 Mise en contexte

La méthode de Newton

$$x_{k+1} = x_k - [F^{'}(x_k)]^{-1}F(x_k)$$

est une méthode itérative efficace pour résoudre des systèmes d'équations(linéaires ou non linéaires)

$$F(x) = 0$$

où $F : \mathbb{R}^n \to \mathbb{R}^n$ est une fonction continûment Fréchet-différentiable, $F^{'}(x_k)$ est le jacobien de F au point x_k.
Il est bien connu que cette méthode converge quadratiquement et localement vers une solution x^* du système si $F^{'}(x_k)$ est inversible à chaque itération k.
Si la fonction F n'est pas différentiable, la méthode ne fonctionne pas. Il est donc nécessaire de construire des méthodes adaptées à la résolution des systèmes d'équations non différentiables. Lorsque F est localement lipschitzienne, les auteurs Mifflin, Liqun Qi et Jie Sun ont proposé respectivement dans [65] et [75] une version non lisse de la méthode de Newton en utilisant le jacobien généralisé $\partial F(x_k)$ de F en x_k, défini par Clarke [19].
Pour établir la convergence, ils ont utilisé, entre autres, des concepts de fonctions localement lipschitzienne et la semi-lissité d'ordre p.

Ces mêmes propriétés ont permis aux auteurs de l'article [15] d'établir la convergence

super-linéaire de la méthode itérative associée au problème

$$0 \in K - f(x), \tag{5.1}$$

où f n'est pas nécessairement différentiable mais localement lipschitzienne.

Il est tout à fait naturel d'examiner le cas où l'on perturbe le cône K de (5.1) par une application multivoque variable F qui est N-P- différentiable et à dérivées uniformément bornées.

L'objet de ce chapitre est de construire une méthode itérative qui généralise celle proposée par S. Bernard et al. [15] pour résoudre l'inclusion

$$0 \in f(x) + G(x) - K. \tag{5.2}$$

En effet, ici, le cadre n'est plus une fonction univoque mais celui d'une multiapplication différentiable au sens de Nachi-Penot.

La procédure sera la même, nous allons énoncer une proposition selon laquelle l'ensemble sur lequel nous allons projeter soit un processus convexe.

Proposition 5.1.1 *Soit un sous-ensemble $X_0 \subset \mathbb{R}^n$. Pour tout sous-ensemble $v \in X_0$, on définit une application multivoque $T(v)$ de $\mathbb{R}^n$ dans $\mathbb{R}^m$ par*

$$T(v)x := (\Delta f(v) + A(v))x - K, \;\; x \in \mathbb{R}^n, \quad \Delta f(v) \in \partial f(v), \;\; A(v) \in \mathcal{D}G(v),$$

où $\mathcal{D}G(v)$ est l'ensemble des dérivées de G au sens de Nachi et Penot et $\Delta f(v)$ est un jacobien généralisé de f en v.
Alors $T(v)$ est un processus convexe.

Preuve
Comme $T(v)x := (\Delta f(v) + A(v))x - K = \Delta f(v)x + A(v)x - K$, il suffit de montrer que la multiapplication $\tilde{T}(v)x = \Delta f(v)x - K$ est un processus convexe.

En fait, tout $z \in (\tilde{T}(v)x + \tilde{T}(v)y)$ peut s'écrire sous la forme

$$z = z_1 + z_2, \;\text{ avec }\; z_1 = \Delta f(v)x - k_1, \quad z_2 = \Delta f(v)y - k_2, \quad k_1, \; k_2 \in K$$

Comme $\Delta f(v) \in \partial f(v)$, il est donc une combinaison convexe de familles de p limites de la forme

$A_j = \lim\limits_{x_i^j \to v} J(x_i^j) \in D$ (D l'ensemble des points où f est différentiable et de matrice jacobienne J).

Par suite

$$\Delta f(v)x + \Delta f(v)y = \sum_{j=1}^{p} A_j(x) + \sum_{j=1}^{p} A_j(y).$$

Par linéarité de la limite et de la matrice jacobienne, on a $A_j(x) + A_j(y) = A_j(x+y)$. Donc,

$$\Delta f(v)x + \Delta f(v)y = \Delta f(v)(x+y).$$

D'autre part , $k_1 + k_2 = 2(\frac{1}{2}(k_1 + k_2))$ puisque K est un cône convexe.

Combinant les deux égalités précédentes, et tenant compte de la définition $\tilde{T}(v)(x+y)$, il s'en suit que

$$z \in \tilde{T}(v)(x+y).$$

D'où

$$\tilde{T}(v)(x+y) \supset \tilde{T}(v)(x) + \tilde{T}(v)y.$$

Soit $\lambda > 0$, $\tilde{T}(v)(\lambda x) = \Delta f(v)(\lambda x) - K$.
Alors

$$\tilde{T}(v)(\lambda x) = \lambda(\Delta f(v)x - \frac{1}{\lambda}K).$$

K étant un cône, $\frac{1}{\lambda}K = K$. Par conséquent, $\tilde{T}(v)(\lambda x) = \lambda\tilde{T}(v)(x)$.

Il est clair que $0 \in \tilde{T}(v)(0)$; ce qui achève la preuve. □

5.2 Description et convergence de la méthode

Soit v un élément fixe d'une partie $X_0 \subset \mathbb{R}^n$. Considérons l' application multivoque $T(v)$ de $\mathbb{R}^n$ dans $\mathbb{R}^m$, défini par

$$T(v)x := (\Delta f(v) + A(v))x - K, \;\; x \in \mathbb{R}^n, \;\; \Delta f(v) \in \partial f(v), \;\; A(v) \in \mathcal{D}G(v),$$

où $\mathcal{D}G(v)$ est l'ensemble des dérivées de G au sens de Nachi et Penot et $\Delta f(v)$ est un jacobien généralisé de f en v.
De la proposition précédente, la multiapplication $T(v)$ est un processus convexe de $\mathbb{R}^n$ dans $\mathbb{R}^m$. Son inverse, défini pour tout $y \in \mathbb{R}^m$ par

$$T^{-1}(v)y := \{z \in \mathbb{R}^n \;/\; (\Delta f(v) + A(v))z \in y + K\}$$

est aussi un processus convexe.

Nous allons procéder de manière analogue à ce que nous avons fait dans les cas précédents.
Tout d'abord, nous choisissons un point initial $x_0 \in X_0$ tel que l'ensemble convexe

$$T^{-1}(x_0)[-f(x_0) - G(x_0)] = \{z \in \mathbb{R}^n \ / \ f(x_0) + (\Delta f(v) + A(v))z \in K - G(x_0)\}$$

soit non vide et projetons ensuite l'origine $0 \in X_0$ sur ce dernier, puis en itérant ce procédé, nous obtenons une suite $(x_k)_k$ telle que x_{k+1} soit solution du problème de minimisation convexe

$$\min\{\|x - x_k\| \ / \ f(x_k) + (\Delta f(v) + A(v))(x - x_k) \in K - G(x_k)\}, \quad k \geq 0. \tag{5.3}$$

Nous allons maintenant énoncer et établir notre résultat de convergence.

5.3 Analyse de la convergence de notre méthode

Nous entamons cette section par un lemme indispensable à notre résultat de convergence et qui a déjà été utilisé par S. Bernard et al. dans [15]. Par ailleurs, Nous verrons que leur méthode peut résoudre le problème perturbé (5.2) en supposant seulement l'uniforme bornitude d'une dérivée de Nachi-Penot du perturbateur G à graphe fermé.

Ce lemme sera donc fondamental pour montrer la convergence de notre algorithme.

Lemme 5.3.1
Supposons qu'une fonction f soit semi-lisse d'ordre p sur un sous-ensemble X_0, alors il existe des constantes $r > 0$ et $L > 0$ tels que pour tous $x, y \in X_0$ vérifiant $\|x - y\| \leq r$ et pour tout $\Delta f(y) \in \partial f(y)$,

$$\|f(x) - f(y) - \Delta f(y)(x - y)\| \leq L\|x - y\|^{1+p}.$$

Le théorème qui suit étendra le résultat de S. Bernard et al. [15] à notre problème perturbé. Cela montrera que leur résultat de convergence est conservé lorsqu'on perturbe leur problème par une multiapplication N-P-différentiable, à dérivées uniformément bornées et à graphe fermé.

Théorème 5.3.2
Soit X_0 un sous-ensemble ouvert de $\mathbb{R}^n$. Considérons $f : \mathbb{R}^n \to \mathbb{R}^m$ et $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ respectivement semi-lisse d'ordre p et N-P-différentiable sur X_0. On suppose que F est à graphe fermé. De plus, l'existence d'un point x_0 tel que $T(x_0)$ soit surjectif et défini sur $\mathbb{R}^n$, et des réels positifs B, M_1, M_2 et α vérifiant les propriétés suivantes :

(a) $2BM_1 < 1$ *et* $\|T^{-1}(x_0)\| \leq B$;

(b) $\|\Delta f(y)\| \leq M_1$ *et* $\|A(y)\| \leq M_2$ *pour tous* $\Delta f(y) \in \partial f(y)$ *et* $y \in X_0$, $A(y)$ *est une N-P-dérivée de l'application multivoque* G *au point* y ;

(c) Pour tout k, $G(x_k) \subset G(x_{k+1}) - K$, *où la suite* (x_k) *est générée par notre itération.*

(d) $\alpha < \frac{r}{2}, 0 < h = M(L\alpha^p + M_2) < 1 - \frac{2\alpha}{r}$, *où* L *et* r *sont donnés par le lemme précédent, avec* $M = \frac{B}{1-2BM_1}$;

(e) $\|x_1 - x_0\| \leq \alpha$ *avec* x_1 *est obtenu à partir de* x_0 *via l'algorithme précédent.*

Il existe t^* *tel que si* $I\!B_{t^*}(x_0) \subset X_0$, *alors l'algorithme précédent génère une suite* (x_k) *telle que pour tout* $k \geq 0, x_k \in I\!B_{\frac{r}{2}}(x_0)$ *et converge vers* x^* *tel que* $f(x^*) \in K - G(x^*)$.

Preuve
Remarquons que si pour tout $k \in \mathbb{N}$, l'élément

$$A_k = f(x_{k+1}) - f(x^*) - [f(x_{k+1}) - f(x_k) - (\Delta f(x_k) + A(x_k))(x_{k+1-x_k})]$$

appartient à l'ensemble $K - G(x_k) - f(x^*)$ et s'il existe une suite (x_k)de limite x^*, il s'en suit de la continuité des applications linéaires $\Delta f(x_k)$ et $A(x_k)$, la suite (A_k) converge vers 0.
La fermeture de l'ensemble $K - G(x_k) - f(x^*)$ et la Proposition 2.4.4 impliquant que la suite d'ensembles $(G(x_k))$ converge vers l'ensemble $G(x^*)$. Il en résulte que l'élément x^* est solution du problème (5.2).

On va s'inspirer de la preuve du résultat de convergence de S. Bernard et al. [15] pour prouver le nôtre. On va donc montrer par récurrence l'existence d'une suite (x_k) qui vérifie (5.3), ainsi qu'une suite strictement croissante de réels positifs (t_k) telles que

$$\|x_{k+1} - x_k\| \leq t_{k+1} - t_k, \ \ \forall k \in \mathbb{N}. \tag{5.4}$$

L'hypothèse d) nous permet d'affirmer l'existence d'un élément x_1 vérifiant (5.3) pour $k = 0$. De plus, en posant $t_1 = \alpha$ et $t_0 = 0$, on obtient $\|x_1 - x_0\| \leq \alpha = t_1 - t_0$. Cela prouve que (5.4) est vérifiée pour $k = 0$.
À présent, supposons que notre algorithme génère des points $x_1, x_2, ..., x_k$ et des réels strictement positifs $t_0, t_1, ..., t_k$ tels que

$$\|x_j - x_{j-1}\| \leq t_j - t_{j-1} \ \ , \forall j = 1, ..., k.$$

Pour montrer l'existence de x_{k+1} et de t_{k+1} vérifiant (5.4), nous allons écrire le processus convexe $T(x_k)$ en fonction de $T(x_0)$, $\Delta f(x_k)$ et $\Delta f(x_0)$ afin d'appliquer le Théorème 2.1.2.
En fait,
il résulte de l'écriture du processus convexe $T(x_k)x = (\Delta f(x_0) + A(x_0))x - K + (\Delta f(x_k) - \Delta f(x_0)x$ sous la forme $T(x_k)x = (T(x_0)) - \Gamma_k)x$ avec $\Gamma_k(x) = -(\Delta f(x_k) - \Delta f(x_0))x$ que

$$\|T^{-1}(x_0)\|.\|\Gamma_k\| < 1$$

puisque $2BM_1 < 1$.
Dans ce cas, l'application du Théorème 2.1.2 de la première section nous permettra d'affirmer que $T(x_k)$ est un processus convexe normé, défini sur $\mathbb{R}^m$ tout entier et surjectif, et son inverse $T^{-1}(x_k)$ vérifie

$$\|T^{-1}(x_k)\| \leq \frac{\|T^{-1}(x_0)\|}{1 - \|T^{-1}(x_0)\|\|\Gamma_k\|} \leq \frac{B}{1 - 2BM_1}.$$

La définition de $T^{-1}(x_k)$ sur $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^m$ implique que le problème de minimisation (5.3) est faisable pour $j = k$ et l'existence d'un x_{k+1} solution de (5.3).

Maintenant, considérons le nouveau problème qui consiste à trouver un point x solution du problème

$$f(x_k) + (\Delta f(x_k) + A(x_k))(x - x_k) \in f(x_{k-1}) + (\Delta f(x_{k-1}) + A(x_{k-1}))(x_k - x_{k-1}) + K. \tag{5.5}$$

Puisque x_k est solution du problème (5.2) pour $j = k-1$, le membre de droite de (5.5) est contenu dans $K - G(x_{k-1}) \subset K - G(x_k)$. On peut conclure que tout x satisfait (5.5) est nécessairement faisable pour (5.2) avec $j = k - 1$.

Observons que le problème (5.5) peut se réécrire

$$x - x_k \in T^{-1}(x_k)\Big(- f(x_k) + f(x_{k-1} + (\Delta f(x_{k-1}) + A(x_{k-1}))(x_k - x_{k-1}\Big). \tag{5.6}$$

Comme le membre de droite de (5.6) est un ensemble convexe, fermé et non vide de l'espace réflexif X, il existe donc $\tilde{x}$ solution de (5.6)(donc de (5.5)) qui vérifie

$$\|\tilde{x} - x_k\| \leq \|T^{-1}(x_k)\|\| - f(x_k) + f(x_{k-1}) + (\Delta f(x_{k-1}) + A(x_{k-1}))(x_k - x_{k-1})\|. \tag{5.7}$$

Par inégalité triangulaire de la norme, on a

$$\|\tilde{x}-x_k\| \leq \|T^{-1}(x_k)\|\Big(\|-f(x_k)+f(x_{k-1})+\Delta f(x_{k-1})(x_k-x_{k-1})\|+\|A(x_{k-1}))(x_k-x_{k-1})\|\Big).$$

En utilisant le lemme précédent et des hypothèses du théorème, on obtient

$$\|\tilde{x} - x_k\| \leq M(L\|x_k - x_{k-1}\|^{1+p} + M_2\|x_k - x_{k-1}\|).$$

Le membre de droite de la dernière inégalité nous conduit à considérer la suite (t_k)

$$t_{k+1} - t_k = M(L(t_k - t_{k-1})^{1+p} + M_2(t_k - t_{k-1}))$$

avec $t_0 = 0$ et $t_1 = \alpha$.
Par définition de x_{k+1},

$$\|x_{k+1} - x_k\| \leq \|\tilde{x} - x_k\| \leq t_{k+1} - t_k.$$

Cela signifie que la suite (x_k) est majorée par la suite (t_k).
Par ailleurs, il est clair que, par construction, la suite (t_k) est strictement croissante.

Étant donné que la suite (x_k) est une suite d'éléments de l'espace de Banach réflexif X, pour montrer qu'elle converge, il suffit de montrer la convergence de la suite $(t_k)_k$.

D'abord, remarquons que $t_2 - t_1 = M(L(t_1 - t_0)^p + M_2)(t_1 - t_0)) = h\alpha < \alpha$ car $h < 1$.
Si $t_k - t_{k-1} \leq \alpha$, alors
$t_{k+1} - t_k = M(L(t_k - t_{k-1})^p + M_2)(t_k - t_{k-1}) \leq M(L\alpha^p + M_2)(t_k - t_{k-1}) = h(t_k - t_{k-1}) < h\alpha < \alpha$ puisque $h < 1$.
Ainsi, $t_k - t_{k-1} \leq \alpha$ pour tout $k \geq 1$(car $t_1 = \alpha$).
On a donc montré par récurrence $t_k - t_{k-1} \leq \alpha$, pour tout $k \geq 1$.

De plus, si $t_{j+1} - t_j \leq h(t_j - t_{j-1}), \quad j = 1, 2, ..., k$, on a alors

$$t_{k+2} - t_{k+1} = M(L(t_{k+1} - t_k)^p + M_2)(t_{k+1} - t_k) \leq M(L\alpha^p + M_2)(t_{k+1} - t_k) = h(t_{k+1} - t_k).$$

Ce qui prouve que l'inégalité

$$t_{k+1} - t_k \leq h(t_k - t_{k-1})$$

est vraie pour tout $k = 0, 1, 2, ...$ On en déduit que

$$t_{k+1} - t_k \leq h(t_k - t_{k-1}) \leq h^2(t_{k-1} - t_{k-2}) \leq \cdots \leq h^k(t_1 - t_0) = h^k\alpha.$$

En sommant sur $j = 0, 1, \cdots, k-1$, on obtient

$$t_k \leq \alpha \sum_{j=0}^{k} h^j.$$

Le membre de droite étant la suite des sommes partielles de la série géométrique de raison $h < 1$ (donc convergente), la suite (t_k) converge vers un élément t^*, et

$$\lim t_k = t^* \leq \frac{\alpha}{(1-h)} < +\infty \text{ existe } .$$

Comme la suite $(t_k)_k$ est majorante de $(x_k)_k$, et par conséquent (x_k) est une suite de Cauchy de X, donc cette dernière converge vers un élément $x^* \in X$.

Pour tout $m > 0$,

$$\|x_m - x_0\| \leq t_m - t_0 \leq t^* \leq \frac{\alpha}{(1-h)} \leq \frac{r}{2} \text{ (car } 1 - h \geq \tfrac{2\alpha}{r}) \ .$$

La suite (x_k) est donc incluse dans les boules $I\!B_{t*}(x_0)$ et $I\!B_{\frac{r}{2}}(x_0)$. □

Soulignons que le théorème se différencie seulement du résultat de S. Bernard et al. [15] au niveau des hypothèses (c) et (d) où figure la borne uniforme d'une N-P-dérivée $A(x)$ de la multiapplication F.

5.3.1 Estimation de l'erreur et vitesse de convergence

Une estimation de l'erreur et la vitesse de convergence de notre méthode sont donc prouvées dans la proposition qui suit.

Proposition 5.3.3 *Soit une suite $(x_n)_n$ générée par l'itération* (5.3).

a) La méthode est au moins d'ordre 1 ;

b) Une estimation de l'erreur à l'ordre k est donnée par

$$\|x^* - x_k\| \leq \alpha \frac{h^k}{1-h}.$$

Preuve
De la relation $t_{k+1} - t_{t_k} \leq h^k \alpha$, on déduit que, pour $m \geq 1$, $k \geq 0$

$$\begin{aligned}
\|x_{k+m} - x_k\| &\leq \|(x_{k+1} - x_k) + (x_{k+2} - x_k) + ... + (x_{k+m-1} - x_{k+m-2}) + (x_{k+m} - x_{k+m-1})\| \\
&\leq \|x_{k+1} - x_k\| + \|x_{k+2} - x_{k-1}\| + ... + \|x_{k+m-1} - x_{k+m-2}\| + \|x_{k+m} - x_{k+m-1}\| \\
&\leq (t_{k+1} - t_k) + (t_{k+2} - t_{k+1}) + ... + (t_{k+m-1} - t_{k+m-2}) + (t_{k+m} - t_{k+m-1}) \\
&\leq \alpha h^k + \alpha h^{k+1} + ... + \alpha h^{k+m-2} + \alpha h^{k+m-1} \\
&= \alpha h^k (h^0 + ... + h^{m-2} + h^{m-1}) \\
&\leq \alpha h^k \sum_{j=0}^{m-1} h^j = \alpha h^k (1 - h^m)/(1-h).
\end{aligned}$$

La dernière inégalité tend vers $\alpha\frac{h^k}{1-h}$ quand $m \to +\infty$, indépendamment de k (car $h < 1$ implique que h^m tend vers 0).
D'où une estimation de l'erreur à l'ordre k

$$\|x^* - x_k\| \leq \alpha\frac{h^k}{1-h}.$$

Ceci prouve que la méthode converge super-linéairement. □

Remarque 5.3.4 *Si $G(x) = \{0\}$ et $K = \{0\}$, la méthode est identique à celle proposée dans [75].*

Remarque 5.3.5 *Si $G(x) = \{0\}$ et f est lisse, la méthode est analogue à celle introduite dans [83].*

Remarque 5.3.6 *Si $G(x) = \{0\}$, l'itération est exactement celle introduite dans [15].*

5.4 Cas où la fonction f est sous-analytique

Lorsque la partie univoque de (2.2) est sous-analytique, notre théorème sera encore valable si on remplace le Lemme 5.3.1 par la Proposition 1.7.12. Auquel cas, la suite majorante $(t_k)_k$ s'obtiendrait en remplaçant le nombre p par le nombre rationnel γ ; ce qui donnerait

$$t_{k+1} - t_k = M(L(t_k - t_{k-1})^{1+\gamma} + M_2(t_k - t_{k-1})).$$

La convergence super-linéaire de notre méthode s'en déduirait aisément.

Il est naturel de se demander que se passe-t-il si on perturbe la partie univoque f par une fonction g Lipschitz ? Nous allons répondre à travers ce qui suit.

5.5 Cas où la fonction f est perturbée par une fonction g lipschitzienne

Nous sommes donc ammenés à résoudre le problème

$$0 \in f(x) + g(x) + G(x) - K, \tag{5.8}$$

où g est l-lipschitzienne sur X_0.

-La suite $(x_k)_k$ que nous aurons à construire se déduira aisément de celle qu'on vient de proposer.

L' itéré x_{k+1} s'obtient, à partir de x_k, en ajoutant au membre de gauche de l'ensemble des contraintes du problème de minimisation convexe (5.3) l'image $g(x_k)$, c'est-à-dire x_{k+1} est solution de

$$\min\{\|x - x_k\| \ / \ f(x_k) + g(x_k) + (\Delta f(x_k) + A(x_k))(x - x_k) \in K - G(x_k)\}. \quad (5.9)$$

-L'existence et la convergence de la suite $(x_k)_k$ générée par notre méthode seront basées sur certaines propriétés que nous résumons dans la proposition suivante :

Théorème 5.5.1
Soit X_0 un sous-ensemble ouvert de $\mathbb{R}^n$. Considérons $f : \mathbb{R}^n \to \mathbb{R}^m$ et $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, $g : \mathbb{R}^n \to \mathbb{R}^m$ respectivement semi-lisse d'ordre p, N-P-différentiable et $l-$Lipschitz sur X_0. On suppose que F est à graphe fermé. De plus, supposons l'existence d'un point x_0 tel que le processus convexe $T(x_0)x = \Delta f(x_0)x - K$ soit défini sur $\mathbb{R}^n$ et surjectif, et des réels positifs B, M_1, M_2 et α vérifiant les propriétés suivantes :

(a) $\|g(x) - g(y)\| \leq l\|x - y\|$;

(b) $2BM_1 < 1$ *et* $\|T^{-1}(x_0)\| \leq B$;

(c) $\|\Delta f(y)\| \leq M_1$ *et* $\|A(y)\| \leq M_2$ *pour tous* $\Delta f(y) \in \partial f(y)$ *et* $y \in X_0$, $A(y)$ *est une N-P-dérivée de l'application multivoque G au point y ;*

(d) $\alpha < \frac{r}{2}, 0 < h = M(L\alpha^p + M_2) < 1 - \frac{2\alpha}{r}$, *où L et r sont donnés par le lemme précédent, avec* $M = \frac{B}{1-2BM_1}$;

(e) $\|x_1 - x_0\| \leq \alpha$ *avec x_1 est obtenu à partir de x_0 via l'algorithme précédente.*

Il existe t^* tel que si $\mathbb{B}_{t^*}(x_0) \subset X_0$, alors l'algorithme précédent génère une suite $(x_k)_k$ telle que pour tout $k \geq 0$, $x_k \in \mathbb{B}_{\frac{r}{2}}(x_0)$ et converge vers x^* tel que $f(x^*) + g(x^*) \in K - G(x^*)$. De plus, on a une estimation de l'erreur

$$\|x^* - x_k\| \leq \alpha \frac{h^k}{1-h}.$$

Preuve
En raisonnant comme dans la première partie de la preuve du Théorème 5.3.2, on prouve l'existence de $x_0, x_1, ..., x_k$ générée par l'algorithme (5.9) et des nombres réels strictement positifs $t_0, t_1, ..., t_k$ tels que

$$\|x_j - x_{j-1}\| \leq t_j - t_{j-1}, \quad j = 1, 2, ..., k$$

avec $t_j \geq t_{j-1}$.
D'autre part, on obtient également l'existence d'un x_{k+1} solution de (5.9).

Nous allons continuer à utiliser les arguments de la preuve du Théorème 5.3.2 pour montrer l'existence de x_{k+1} et t_{k+1} qui vérifient l'inégalité précédente.
Remarquons que le problème

$$(*) \quad f(x_k)+g(x_k)+(\Delta f(x_k))+A(x_k))(x-x_k) \in f(x_{k-1})+g(x_{k-1}) \\ +(\Delta f(x_{k-1})+A(x_{k-1})(x_k-x_{k-1})+K,$$

d'inconnue $x \in X$, peut s'écrire sous la forme

$$x-x_k \in T^{-1}(x_k)\Big(-f(x_k))-g(x_k)+f(x_{k-1})+g(x_{k-1})+\Delta f(x_{k-1})+A(x_{k-1})(x_k-x_{k-1})\Big).$$

Le membre de droite de ce dernier étant un ensemble convexe, fermé et non vide de l'espace réflexif X, il en existe une solution $\tilde{x}$ qui soit également solution du problème $(*)$.

Or, le membre de droite du problème $(*)$ étant contenu dans $K-G(x_{k-1}) \subset K-G(x_k)$, il s'en suit que tout x solution du problème $(*)$ est forcément un point de faisabilité de (5.9) pour $j = k-1$.
Il en résulte que $\tilde{x}$ est un point de faisabilité du problème (5.9).
Par suite

$$\|x_{k+1}-x_k\| \leq \|\tilde{x}-x_k\| \leq \|T^{-1}(x_k)\| \Big(\| -f(x_k)+f(x_k)+\Delta f(x_{k-1})(x_k-x_{k-1}) \\ +\|g(x_k)-g(x_{k-1})\|+A(x_{k-1})(x_k-x_{k-1})\Big).$$

Combinant le lemme précédent et des hypothèses de la proposition, on obtient

$$\|x_{k+1}-x_k\| \leq M\Big(L\|x_k-x_{k-1}\|^{1+p}+l\|x_k-x_{k-1}\|+M_2\|x_k-x_{k-1}\|\Big)$$

soit

$$\|x_{k+1}-x_k\| \leq M\Big(L\|x_k-x_{k-1}\|^{1+p}+(M_2+l)\|x_k-x_{k-1}\|\Big).$$

Ce qui conduit à la suite majorante $(t_k)_k$ définie par

$$t_{k+1}-t_k = M\Big(L(t_k-t_{k-1})^{1+p}+(M_2+l)\|x_k-x_{k-1}\|\Big)$$

avec $t_0 = 0$ et $t_1 = 1$.
La suite de la preuve est donc identique à celle de la preuve du Théorème 5.3.2 il suffit

de remplacer la borne uniforme M_2 des N-P-dérivées de F par $M_2 + l$. $\square$

Ainsi, le résultat de convergence s'obtient en ajoutant uniquement l'hypothèse $l-$Lipschitz de g à celles du Théorème 5.3.2.

Nous clôturons ce chapitre avec une remarque ne faisant pas partie du cadre dans lequel on est puisqu'elle consiste à prouver dans quelles mesures que le Théorème 5.3.2 de S. Bernard et al. [15] peut être adapté si la fonction univoque f était différentiable et à dérivée hölderienne. Cela sera légitime puisque, d'une part, cette condition impliquera l'obtention d'un résultat similaire au Lemme 5.3.1 où γ et λ substitueront à L et p. D'autre part, la suite majorante sera similaire à celle utilisée dans la convergence du théorème.

La remarque qui suit montrera que l'adaptation aura lieu en remplaçant le gradient généralisé de Clarke par la dérivée de Fréchet, puis appliquer la condition de Hölder au membre de droite du lemme utilisé par S. Bernard et al.[15].

5.6 Remarque : f Fréchet différentiable avec dérivée hölderienne

Dans cette section, nous allons montrer que, sous certaines hypothèses, notre méthode peut être adaptée pour résoudre un problème plus général que celui du problème (3.1) lorsque la dérivée de Fréchet satisfait une condition de Hölder

$$\|\nabla f(x) - \nabla f(x^{'})\| \leq \gamma \|x - x^{'}\|^{\lambda} \tag{5.10}$$

avec $0 < \lambda \leq 1$ pour tous $x, y \in X_0$.
Nous allons donc proposer une méthode itérative pour résoudre le problème

$$0 \in f(x) + G(x) - K, \tag{5.11}$$

où la dérivée de Fréchet de f satisfait une condition de Hölder.

Juste avant d'énoncer notre nouveau résultat de convergence, nous énonçons une proposition fondamentale

Proposition 5.6.1 *Considérons la suite* (t_k) *définie par*

$$\begin{cases} t_0 = 0,\ t_1 = \eta. \\ t_{k+1} - t_k = \beta\gamma(t_k - t_{k-1})^{1+\lambda} + \beta\alpha(t_k - t_{k-1}),\ \ \textit{pour}\ k = 0, 1, 2, \cdots, \end{cases}$$

avec α, β, η des réels strictement positifs.
Alors, la suite $(t_k)_k$ converge vers un nombre t^* si le nombre positif $h = \beta\eta^\lambda + \beta\alpha < 1$.
De plus, une majoration de l'erreur

$$t^* - t_k \leq \eta \frac{h^k}{1-h}, \quad k > 0.$$

Preuve
Nous allons procéder par récurrence afin de majorer $t_{k+1} - t_k$ par un nombre strictement positif qui est fonction de
Pour $k = 1$, on a

$$t_2 - t_1 = \Big(\beta\gamma(t_1 - t_0)^\lambda + \beta\alpha\Big)(t_1 - t_0) = \Big(\beta\gamma\eta^\lambda + \beta\alpha\Big)\eta = \eta h \leq \eta.$$

Observons d'abord que si $t_k - t_{k-1} \leq \eta$, alors

$$t_{k+1} - t_k = \Big(\beta\gamma(t_k - t_{k-1})^\lambda + \beta\alpha\Big)(t_k - t_{k-1}) = \Big(\beta\eta^\lambda + \beta\alpha\Big)\eta \leq \eta h \leq \eta.$$

Ainsi, on a montré que

$$t_{k+1} - t_k \leq \eta, \quad \text{pour tout } k \geq 1.$$

Supposons que

$$t_{j+1} - t_j \leq h(t_j - t_{j-1}), \quad j = 1, 2, ..., k.$$

Donc,

$$t_{k+2} - t_{k+1} = \Big(\beta\gamma(t_{k+1} - t_k)^\lambda + \beta\alpha\Big)(t_{k+1} - t_k) = \Big(\beta\gamma\eta^\lambda + \beta\alpha\Big)(t_{k+1} - t_k) = h(t_{k+1} - t_k).$$

Ceci prouve que

$$t_{k+1} - t_k = h(t_{k+1} - t_k), \quad k = 0, 1, ...$$

Par suite

$$t_{k+1} - t_k \leq (t_k - t_{k-1}) \leq h^2(t_{k-1} - t_{k-2}) \leq ... \leq h^k(t_1 - t_0) = h^k\eta.$$

En sommant sur $j = 0, 1, ..., k-1$,

$$t_k \leq \eta \sum_{j=0}^{k} h^j.$$

Puisque $h < 1$, en passant à la limite sur k, le membre de droite converge vers $\frac{\eta}{1-h}$.
Il en résulte que la suite (t_k) converge vers un nombre $t^* \leq \frac{\eta}{1-h}$.

D'autre part, pour tous $m \geq 1,\ k \geq 0$,

$$t_{k+m} - t_k \leq \sum_{j=0}^{m-1} t_{k+j+1} - t_{k+j} \leq \alpha \sum_{j=0}^{m-1} h^{k+j} = \eta h^k \sum_{j=0}^{m-1} h^j.$$

En faisant tendre $m \to +\infty$,

$$t^* - t_k \leq \eta \frac{h^k}{1-h}. \quad \square$$

Ainsi, la suite obtenue est bien similaire à la suite majorante des itérés de la méthode précédente.

Nous enchaînons maintenant avec notre résultat principal.

Théorème 5.6.2
Soit X_0 un sous-ensemble ouvert de $\mathbb{R}^n$. Considérons $f : \mathbb{R}^n \to \mathbb{R}^m$ et $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ respectivement semi-lisse d'ordre p et N-P-différentiable sur X_0. On suppose que G est à graphe fermé, puis l'existence d'un point x_0 tel que $T(x_0)x := \nabla f(x_0)x - K$ soit défini sur $\mathbb{R}^n$ et surjectif, et des réels positifs B, M_1, M_2 et α vérifiant les propriétés suivantes :

(a) $2BM_1 < 1$ *et* $\|T^{-1}(x_0)\| \leq B$,

(b) $\|\Delta f(y)\| \leq M_1$ *et* $\|A(y)\| \leq M_2$ *pour tous* $\Delta f(y) \in \partial f(y)$ *et* $y \in X_0$, $A(y)$ *est une N-P-dérivée de l'application multivoque G au point y ;*

(c) $0 < h = M(\gamma\alpha^\lambda + M_2) < 1$, *avec* $M = \frac{B}{1-2BM_1}$;

(d) $\|\nabla f(x) - \nabla f(x')\| \leq \gamma\|x - x'\|^\lambda$, *pour tous* $x,\ x' \in X_0$, $0 < \lambda \leq 1$;

(e) $\|x_1 - x_0\| \leq \alpha$ *avec x_1 est obtenu à partir de x_0 via l'algorithme précédent.*

Il existe t^* tel que si $\mathbb{B}_{t^*}(x_0) \subset X_0$, alors l'algorithme précédent génère une suite (x_k) qui converge vers x^* tel que $f(x^*) \in K - G(x^*)$.
De plus, on a une estimation de l'erreur

$$\|x^* - x_k\| \leq \alpha \frac{h^k}{1-h}, \quad k > 0.$$

Preuve
Des procédés similaires à ceux de la preuve de notre premier résultat de convergence nous permet de construire $x_0, x_1, ..., x_j$ solutions du problème (5.11) et une suite croissante de réels positifs $(t_k)_k$ tels que

$$x_k \in I\!B_{t^*}(x_0) \subset X_0 \text{ et } \|x_{k+1} - x_k\| \leq t_{k+1} - t_k, \quad k = 1, 2, ..., j.$$

Pour montrer l'existence d'un x_{j+1} qui vérifie l'inclusion et l'inégalité précédente, nous allons encore une fois utiliser une technique de la preuve de notre résultat de convergence.
Par analogie à l'inégalité (3.4) de la preuve, nous avons l'existence d'un $\tilde{x}$ tel que

$$\|\tilde{x} - x_j\| \leq \|\tilde{T}_{x_j}^{-1}\|\Big(\|f(x_{j-1}) - f(x_j) + \nabla f(x_{j-1})(x_j - x_{j-1})\| + \|A(x_{j-1})\|\|x_j - x_{j-1}\|\Big).$$

Compte tenu de la condition de Hölder que vérifie la dérivée de Fréchet de f, on a

$$\|f(x_{j-1}) - f(x_j) + \nabla f(x_{j-1})(x_j - x_{j-1})\|$$
$$= \Big\| \int_0^1 [\nabla f(\theta x_j + (1-\theta)x_{j-1}) - \nabla f(x_{j-1})](x_j - x_{j-1})\mathrm{d}\theta \Big\|$$
$$\leq \int_0^1 \gamma \|x_j - x_{j-1}\|^{\lambda}.\|x_j - x_{j-1}\|\,\theta^{\lambda}\mathrm{d}\theta = \frac{1}{1+\lambda}\gamma\|x_j - x_{j-1}\|^{1+\lambda} \leq \gamma\|x_j - x_{j-1}\|^{1+\lambda},$$

et par un raisonnement similaire utilisé dans la preuve précédente,

$$\|T^{-1}(x_k)\| \leq \frac{B}{1 - 2BM_1}.$$

Il vient

$$\|\tilde{x} - x_j\| \leq \frac{B}{1 - 2BM_1}\Big(\gamma\|x_j - x_{j-1}\|^{1+\lambda} + M_2\|x_j - x_{j-1}\|\Big).$$

Par suite, la définition de x_{j+1} nous conduit aux inégalités

$$\|x_{j+1} - x_j\| \leq \|\tilde{x} - x_j\| \leq M\Big(\gamma\|x_j - x_{j-1}\|^{1+\lambda} + M_2\|x_j - x_{j-1}\|\Big).$$

Nous considérons donc la suite $(t_k)_k$ définie par

$$\begin{cases} t_0 = 0,\ t_1 = \alpha. \\ t_{k+1} - t_k = M\Big[\gamma(t_k - t_{k-1})^{1+\lambda} + M_2(t_k - t_{k-1})\Big], \text{ pour } k = 0, 1, 2, \cdots \end{cases}$$

La convergence de la suite $(t_k)_k$ résulte de la Proposition 5.6.1. Ainsi, la suite des itérés $(x_k)_k$ générée par la méthode converge vers x^* solution du problème (5.11). □

Conclusion et perspectives

Ce travail porte sur l'utilisation du concept de processus convexe pour la résolution d'inclusions variationnelles. Notre champ de travail prend en compte des multiapplications qui sont différentiables et par conséquent leurs dérivées.
Ainsi le travail présenté dans cette thèse généralise les travaux antérieures dans lesquels la multiapplication était fixe et plus précisément un cône convexe fermé.
Les multiapplications de cette thèse sont à graphe fermé, à valeurs compactes et convexes, différentiables au sens de Nachi-Penot et on demande aux dérivées d'être uniformément bornées.
Nous obtenons différents résultats selon que la partie univoque soit lisse, lipschitzienne ou semi-lisse. Pour les différentes méthodes introduites, les vitesses de convergence semi-locales associées sont données.

Ce travail ouvre d'autres perspectives que ce soit sur le plan théorique ou numérique. Devant se terminer dans les délais, nous n'avons pas eu le temps d'implémenter les différentes méthodes mais ce sera fait dans les mois qui viennent. Un autre projet est de reprendre la théorie en utilisant d'autres types de dérivées de multiapplications. Enfin, il serait bon de voir l'incidence de ces résultats sur d'autres sujets en relation avec le contrôle optimal.

Bibliographie

[1] R. J. Aumann, *Integrals of set-valued functions,* J. Math. Anal. 12 (1965), 1-12.

[2] J.-P. Aubin, *Contingent derivatives of set valued maps and existence of solutions to non linear inclusions and differential inclusions,* Part A, in : Adv. in Math. Suppl. Stud., vol 7a, Academic Press, New York-London, 1981, 159-229.

[3] J.- P. Aubin, *Lipschitz behaviour of solutions to convex minimization problems,* Math. Oper. Res., 9 (1) (1984), 87-111.

[4] J.-P. Aubin and H. Frankowska, Set-valued analysis. Systems & Control : Foundations & Applications, 2. Birkhäuser Boston, Inc., Boston, MA, 1990.

[5] I. K. Argyros, Computational theory of iterative methods, Studies in Computational Mathematics, 2007.

[6] P. Brunovsky, *On the necessity of a certain convexity condition for lower closure of control problems.* SIAM J. Control **6** (1968), 174-185.

[7] T. F. Bridgland, *Trajectory integrals of set valued functions,* Pacific journal of mathematics, Vol. 33, No. 1, 1970.

[8] H. T. Banks and M. Q. Jacobs, *A différential calculus for multifunctions,* J. Math. Anal. Appl., 29 (1970), 246-272.

[9] F. S. de Blasi, *On the differentiability of multifunctions,* Pacific J. Math., 66 (1) (1976), 67-81.

[10] E. Bierstone and P. D. Milan, *Semianalytic and subanalytic sets,* IHES. Publ. Math., 67 (1998), 5-42.

[11] J. F. Bonnans, *Local convergence of Newton-type methods for variationnal inequalities an non linear programming,* Appl. Math. Optim., 29 (2) (1994), 161-186.

[12] R. L. Burden and J. D. Faires, *Numerical Analysis,* Seventh Edition, 2001.

[13] J. Bolte, A. Danilis and A. S. Lewis, *Tame mapping are semismooth,* Math. Program., Ser. B, 117 (2009), 5-19.

[14] S. Burnet, Méthode de résolution d'inclusions variationnelles sous hypothèse de stabilité, Ph D thesis, Université des Antilles et de la Guyane, 2012.

[15] S. Bernard, C. Cabuzel, S. P. Nuiro and A. Piétrus, *Extended Semismooth Newton Method for Functions with Values in a Cone*, Acta Appl. Math., (2016) 155 : 85-98.

[16] C. Castaing,"*Sur les multi-applications mesurables."* Doctoral Dissertation, Université de Caen, France, 1967.

[17] L. Cesary, *Existence theorems for optimal controls of the Mayer type.* SIAM J. Control 6 (1968), 517-552.

[18] F. H. Clarke, *Generalized gradients and applications,* Trans. Amer. Math. Soc. 205 (1975), 247-262.

[19] F. H. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983.

[20] E. Cătinas, *On some iterative methods for solving nonlinear equations,* Revue d'Analyse Numérique et de Théorie de l'Approximation, 23 (1) (1994), 47-53.

[21] C. Cabuzel and A. Piétrus, *Local convergence of Newton's method for subanalytic variationnal inclusions,* Positivity, 12 (3) (2008), 523-533.

[22] G. Debreu, *Integration of correspondences,* In "Proc. Fifth Berkeley Symposium on Mathematical Statistics and Probability," 351-372. University California Press, 1966.

[23] J. P. Dedieu, *Penalty function in subanalytic optimization,* Optimization, 26 (1992), 27-32.

[24] A. L. Dontchev and W. W. Hager, *An inverse mapping thorem for set valued maps,* Proc. Amer. Math. Soc., 121 (1994), 481-48.

[25] A. L. Dontchev, *Local convergence of the Newton method for generalized equation,* C.R.A.S., 322, Serie 1, (1996), 385-398.

[26] A. L. Dontchev, *Local of a Newton-type method based on partial linearization,* The Mathematics of numerical Analysis (Parc City, UT, 1995), Lecture in Appl. Math., 32, Amer. Math. Soc., Prividence, RI, 1996.

[27] A. L. Dontchev, *Uniform convergence of the Newton method for Aubin Continous maps,* Serdica Math. J., 22, Serie 1, (1996), 385-398.

[28] A. L. Dontchev, M. Quincampoix and N. Zlateva, *Aubin criterion for metric regularity,* J. of Convex Analysis, 13 (2) (2006), 281-297.

[29] A. L. Dontchev and R. T. Rockafellar, *Implicit Functions and Solution Mapping,* Springer Monographs in Mathematics, 2009.

[30] A. F. Filipov, *Classical solutions of differential equations with multivalued righhand side.* SIAM J. Control 5 (1967), 609-621.

[31] M. H. Geoffroy, S. Hilout and A. Piétrus, *Acceleration of convergence in Dontchev's iterative method for solving variational inclusions,* Serdica Math. J. 29 (2003), no 1, 45-54.

[32] M.-H. Geoffroy and A. Piétrus, An iterative method for perturbed generalized equations, C. R. Acad. Bulgare Sci., 57 (2004), no. 11, 7-12.

[33] M. H. Geoffroy and A. Piétrus, *Local convergence of some iterative methods for generalized equations,* J. Math. Anal. appl, 290 (2004), 497-505.

[34] M. Huruhara, *Intégration des applications mesurables dont la valeur est un compact convexe,* Funkcialaj Ekvacioj 10 (1967), 205-223.

[35] H. Hermes, *Calculus of set valued functions and control,* J. Math. Mech., 18 (1968), 47-60.

[36] H. Halkin and E. C. Hendricks, *Subintegrals of set-valued functions with semi-analytic graphs.* Proc. Nat. Sci., U.S.A. 59 (1968), 365-367.

[37] H. Halkin, *Mathematical programming without differentiability, in : Calculus of Variations and Control Theory,* Proc. Sympos., vol 36, Math. Res. Center, Univ. Wisconsin, New York, 1976, 279-287.

[38] J. B. Hiriart-Urruty, *New concepts in nondifferentiable programming,* Bull. Soc. Math. France 60 (1979), 57-85.

[39] A. H. Hamel, F. Heyde and B. Rudloff, *Set-valued risk measures for conical market models,* Math. Financ. Econ., 5 (1) (2011), 1-28.

[40] T. Hoheisel, C. Kanzow, B.S. Mordukhovic and H. Phan, *Generalized Newtons method based on graphical derivates,* Nonlinear Anal. 75 (2012), 1324-1340.

[41] T. X. D. Ha, *A Hausdorff-type distance, a directional derivative of a se-valued map and applications in set optimization,* Vietnam Academy of Science and Technology. (2017)

[42] A. Ioffe, *Nonsmooth Analysis : differential calculus of nondifferentiable mappings,* Trans. Amer. Math. Soc. 266 (1) (1981), 1-56.

[43] M. Q. Jacobs, *Measurable multivalued mapping and Lusin's theorem,* Trans. Am. Math. Soc., 134 (1968), 471-481.

[44] M. Q. Jacobs, *On the approximation of integrals of multivalued functions,* SIAM J. Control 7 (1969), 158-177.

[45] B. S. Jeffreys and H. Jeffreys, Methods of Mathematical Physics, 3rd ed., Cambrige University Press, England, 1988.

[46] C. Jean-Alexis, *A cubic method without second order derivative for solving variational inclusions,* C. R. Acad. Bulgare Sci., 59 (12) (2006), 1213-1218.

[47] J. Jahn and T. X. D. Ha, *New order relations in set optimization,* J. Optim. Theory Appl., 148 (2011), 209-236.

[48] C. H. Jeffrey, *Characterizing generalized derivatives of set-valued maps : Extending the tangential and normal approaches,* SIAM J. Control Optim. No 1(2013), 145-171.

[49] J. Jahn, *Directional derivative in set optimization with the less order relation,* Taiwanese J. Math. 19, no. 3 (2015), 737-757.

[50] L. V. Kantorovich, *On Newton's method,* Trudy Mat. Inst. Steklov, Vol. 28, (1949), 105-144.

[51] N. Kikuchi, *On Some fundamental theorems of contingent equations in connection with the control problems,* Publ. RIMS, Kyoto Univ., Ser. A, 3 (1967), 361-371.

[52] N. Kikuchi, *On contingent equations satisfying the Caratheodory type conditions,* Publ. RIMS, Kyoto Univ., Ser. A, 3 (1968), 361-371.

[53] D. Kuroiwa, Some notions of convexity for set-valued maps and their relations, Surikaisekikenkyusho Kokyuroku, 939 (1996), 168-172. Research on nonlinear analusis and convex analysis (Japanese) (Kyoto, 1995).

[54] C. Kanzow, *Some noninterior continuation methods for linear complementary problems,* SIAM Journal on Matrix Analysis and Applications, 17 (1996) 851-868.

[55] D. Kuroiwa, Natural Criteria of Set-valued Optimization, manuscrit, Shimane University, Japan, 1998.

[56] D. Kuroiwa, *On derivatives and convexity of set-valued maps and optimality conditions in set optimization,* J. Nonlinear Convex Anal. 10 (2009), 41-50.

[57] A. Lasota and Z. Opial, *An application of the Kakutani-Ky Fan theorem in the theory of ordinary differential equations,* Bull. Acad. Pol. Sc., Ser. Sci., Math., Astr. Phys., 13 (1965), 781-786.

[58] A. Lasota and C. Olech, *On the Closedness of the set of trajectories a control system,* Bull. Acad. Pol. Sci., Math., Astr. Phys., 14 (1966), 615-621.

[59] A. S. Levitin and B. T. Polyak, *Contrained minimization methods,* Mat. Fiz. 6 (5) (1966),787-823.

[60] D. T. Luc, *Theory of Vector optimization,* Lecture Notes in Economic and Mathematical systems, 319. Springer-Verlag, Berlin, 1989.

[61] I. P. Mysovskih, *On the convergence of L. V. Kantorovich's method of solution of functional equations and its applications,* Dokl. Akad. Nauk SSSR (N.S.), 70 (1950), 565 - 568.

[62] J. J. Moreau, Fonctionelles convexes, "Sem. sur les Équations aux Dérivées Partielles, II, 1966-1967." Collège de France, Paris.

[63] M. Martelli and A. Vignoli, *On differentiability of multi-valued maps,* Boll. Ital.,(4) 10 (1974), 701-712.

[64] B. S. Mordukhovich, *Maximum principle in the problem of time optimal response with nonsmooth contraints,* J. Appl. Math. Mech. 40 (6) (1976), 960-969 (1977) ; translated from Prikl. Mat. Mekh. (40) (1976), 1014-1023 (In Russian).

[65] R. Mifflin, *An Algorithm for contrained optimization with semismooth functions,* Mathematics of Operations Research, 2 (1977), 191-207.

[66] M. M. Mäkelä and Neittaanmäki, *Nonsmooth optimization : Analysis and Algorithm with Applications to Optimal Control,* University of Jyväskylä, Finland, 1992.

[67] B. S. Mordukhovic, *Complete characterization of openess properties of openess metric regularity and lipschitzian properties of multifunctions,* Trans. Amer. Math. Soc., 340 (1993), 1-36.

[68] B. S. Mordukhovich, Variational Analysis and Generalized differentiation, I : Basic Theory. Grundlehren Series (Fundamental Principles of Mathematical Sciences) 330 (2006), Springer, Berlin. *Local convergence of some iterative methods for generalized equations,* J. Math. Anal. appl, 290 (2004), 497-505.

[69] K. Nachi and J.P. Penot, *Inversion of multifunctions and differential inclusions,* Control Cybernet, 34 (2005), no. 3, 871-901.

[70] C. Olech, *A note concerning set-valued measurable functions,* Bull. Acad. Polon. Sci., Math., Astr. Phys. 13 (1965), 317-321.

[71] C. Olech, *Lexicographical order range of integral and "bang-bang" principle,* In "Mathematical Theory of Control." (A. V. Balakrishnan and L. W. Neustadt, Eds.) 35-45. Academic Press, New York, 1967.

[72] J. M. Ortega and W. C. Rheinboldt, *Iterative solution of nonlinear equations in several variables,* New York, Academic Press, 1970.

[73] A. Piétrus, *Generalized equations under mild differentiability condition,* Rev. R. Acad. Cienc. Exact. Fís. Nat., 94, 1 (2000), 15-18.

[74] G. Pascaline, *Convergence de Fisher et H-différentiabilité des applications multivoques,* Ph D thesis, Université des Antilles et de la Guyane, 2010.

[75] L. Qi and J. Sun, *A nonsmooth version of Newton's method,* Mathematical Programming, 58 (1993), 353-367.

[76] L. Qi and J.-S. Pang, *Nonsmooth equation : Motivation and Algorithms,* Journal of Mathematical Analysis and Application, 146 (1993), 271-279.

[77] C. H. J. Pang, *Generalized differentiation with positivity homogeneous maps : Applications in set-valued analysis and metric regularity,* Mathematics of operation research, Vol. 36, no. 3, 2011.

[78] A. Piétrus, *Non differentiable perturbed Newton's method for functions with values in a cone,* Investig. Oper., 35(1) (2012), 58-67.

[79] A. Piétrus and C. Jean-Alexis, *Newton-Secant method for functions with values in a cone,* Investig. Oper., (2013), 1-17.

[80] H. Rådström, *An embedding theorem for spaces of convex sets,* Proc. Amer. Math. Soc. 3 (1952), 165-169.

[81] W. C. Rheinboldt, *A unified convergence theory for a class of itérative processes,* SIAM J. Numer. Anal., 5(1) (1968), 42-63.

[82] S. M. Robinson, *Extension of Newton's Method to Nonlinear Functions with Values in a Cone,* Numer. Math. 19 (1972), 341-347.

[83] S. M. Robinson, *Normed convex processes,* Trans. American. Math. Soc., 174 (1972), 127-140.

[84] S. M. Robinson, *Generalized equation and their solution, part I : basic theory,* Math. Programming Study, 10 (1979), 128-141.

[85] S. M. Robinson, *Generalized equation and their solutions, part II : application to non linear programming,* Math. Programming Study, 19 (1982), 200-221.

[86] R. T. Rockafellar and R. J. B. Wets, Variational analysis, A series of Comprehensives Studies in Mathematics, 317 (1998), Springer.

[87] R. T. Rockafellar, Convex Analysis, Princeton University Press, 1970.

[88] M. Shatzman, Analyse Numérique : Une approche mathématique, Dunod, Paris, 2001.

[89] E. T. Whittaker and G. Robinson, The calculus of observations : A treative on numerical mathematics, 4th ed., Dover, New York, 1967.

[90] J. Warga, *Derivative containers, inverse functions,and controllability in : Calculus of Variations and Control Theory,* Proc. Sympos., vol. 36, Math. Res. Center, Univ. Wisconsin, Madisson, WI, 1975, Academic Press, New York, 1976, pp. 13-45 ; p.46 (Erratum).

[91] L. Qi., S.-Y. Wu and G. Zhou, *Semismooth Newton Method for Solving Semi-Infinite Programming Problems*, 27 (2003), 215-232.

[92] A. I. Zinchenko, *Some approximate method of solving equations with non-differentiable operators,* Dopovidi Akad. Nauk. Ukrain RSR,(1963), 156-161 (In Russian).