

HABILITATION À DIRIGER DES RECHERCHES
DE L'UNIVERSITÉ D'ANGERS

(SCIENCES DE L'INFORMATION ET DE LA COMMUNICATION)

MÉTHODES NUMÉRIQUES GARANTIES
POUR LA CLASSIFICATION DE FONCTIONS
ET LE CONTRÔLE OPTIMAL

Nicolas Delanoue
nicolas.delanoue@univ-angers.fr

Soutenue le mercredi 10 octobre 2018 devant le jury :

Président : Luc JAULIN (PR Lab-STICC - ENSTA Bretagne)

Rapporteurs : Didier HENRION (DR LAAS - Toulouse)
Nacim RAMDANI (PR PRISME - Université d'Orléans)
Stef GRAILLAT (PR LIP6 - Sorbonne Université)

Examineurs : Marie-Françoise ROY (PR IRMAR - Université de Rennes)
Luc JAULIN (PR Lab-STICC - ENSTA Bretagne)
Laurent HARDOUIN (PR LARIS - Université d'Angers)

LARIS EA 7315

Table des matières

| | | |
|----------|---|-----------|
| 1 | Présentation générale | 5 |
| 1.1 | Curriculum vitae | 5 |
| 1.1.1 | Etat civil | 5 |
| 1.1.2 | Titres universitaires | 5 |
| 1.1.3 | Activités professionnelles | 6 |
| 1.2 | Activités pédagogiques | 6 |
| 1.2.1 | Enseignements | 6 |
| 1.2.2 | Activités administratives | 7 |
| 1.3 | Activités de recherches | 8 |
| 1.3.1 | Thèse de doctorat | 8 |
| 1.3.2 | Travaux de recherches engagés depuis la thèse | 9 |
| 1.3.3 | Encadrement doctoral et scientifique | 13 |
| 1.3.4 | Animation scientifique | 14 |
| 1.3.5 | Invitation d'universités étrangères | 14 |
| 1.4 | Publications et communications | 15 |
| 1.4.1 | Publications dans des journaux internationaux | 15 |
| 1.4.2 | Chapitres d'ouvrage collectif | 16 |
| 1.4.3 | Conférences internationales | 16 |
| 1.4.4 | Séminaires et groupe de travail | 17 |
| 2 | Arithmétique des intervalles | 19 |
| 2.1 | Encadrement numérique pour des fonctions | 21 |
| 2.2 | Encadrement d'ensemble solution d'équations | 23 |
| 2.2.1 | SIVIA | 23 |
| 2.2.2 | Newton par intervalles | 25 |
| 2.3 | Encadrement d'intégrales | 28 |
| 3 | Classification des fonctions lisses | 33 |
| 3.1 | Introduction | 33 |
| 3.2 | Invariants locaux - Encadrement des singularités | 37 |
| 3.2.1 | Cas des fonctions de la droite sur la droite | 37 |
| 3.2.2 | Cas des fonctions du plan sur le plan | 38 |
| 3.2.3 | Cas des fonctions de l'espace dans l'espace | 42 |
| 3.3 | Invariants globaux | 44 |
| 3.3.1 | Exemple introductif en dimension 1 | 44 |
| 3.3.2 | Calcul numérique du contour apparent en dimension 2 | 48 |

| | | |
|----------|---|------------|
| 3.3.3 | Portrait de fonctions | 61 |
| 3.4 | Conclusion | 67 |
| 4 | Calcul optimal et théorie de la mesure | 69 |
| 4.1 | Introduction | 69 |
| 4.2 | Transport optimal | 71 |
| 4.2.1 | Introduction | 71 |
| 4.2.2 | Bornes inférieures à la valeur optimale | 72 |
| 4.2.3 | Bornes supérieures à la valeur optimale | 78 |
| 4.2.4 | Conclusion | 81 |
| 4.3 | Contrôle optimal | 82 |
| 4.3.1 | Introduction - État de l'art | 83 |
| 4.3.2 | Mesures d'occupation | 85 |
| 4.3.3 | Formulation linéaire d'un problème de contrôle optimal non linéaire | 90 |
| 4.3.4 | Relaxation en dimension finie via le calcul par intervalles | 93 |
| 4.3.5 | Exemples illustratifs | 96 |
| 4.3.6 | Conclusion | 100 |
| 4.4 | Conclusion | 101 |
| 5 | Conclusions | 103 |

Résumé

Les travaux présentés lors de cet HDR s'appuient sur un outil de calcul, l'arithmétique des intervalles, et proposent deux thématiques : le calcul des singularités d'une fonction pour la robotique et la discrétisation de problèmes de programmation linéaire en dimension infinie pour le contrôle. En robotique, la modélisation cinématique d'un robot est donnée par une fonction souvent lisse entre les variables de commande et la position du robot. L'étude de cette fonction est cruciale afin de comprendre le comportement de ce dernier. Dans ces travaux, des méthodes numériques garanties capables de calculer les singularités de telles fonctions sont proposées. D'un point de vue plus global, une méthodologie permettant de faire le portrait d'une fonction donnée a été développée. Par ailleurs, l'arithmétique des intervalles a également permis de proposer un schéma de discrétisation de problèmes de programmation linéaire en dimension infinie où l'optimum recherché est une mesure. Cette technique génère une suite croissante de bornes inférieures à la valeur de la solution du programme initial. Deux applications sont présentées : une autour du transport optimal, l'autre en contrôle optimal en utilisant la notion de mesures d'occupation.

Mots-clés :

Arithmétique des intervalles, Singularités, Classification, Contrôle optimal non linéaire, Programmation linéaire en dimension infinie, Optimisation

Introduction

Ce mémoire présente les travaux de recherche que j'ai effectués depuis ma thèse et a pour vocation de prouver aux membres de jury ma capacité à diriger des recherches. Dans ce court chapitre d'introduction, je commence par donner l'organisation générale de ce document d'habilitation à diriger des recherches (HDR) et j'adresse quelques remerciements.

Organisation du document

Dans le chapitre 1, mon CV est détaillé. Il décrit de façon générale les thèmes de recherche abordés, les étudiants encadrés dans leur formation à la recherche, ainsi que les fonctions administratives que j'ai occupées depuis mon recrutement. Enfin, il contient les références bibliographiques des communications dont je suis co-auteur. Le chapitre 2 rappelle les définitions autour de l'arithmétique des intervalles. Il contient aussi quelques résultats non encore publiés qui seront mis à profit dans les chapitres 3 et 4. Le chapitre 3 propose des méthodes numériques de classification des fonctions lisses. Ces méthodes trouvent leur intérêt en particulier en robotique. Dans le chapitre 4, on donne des schémas numériques de résolutions de programmes linéaires en dimension infinie. La démarche scientifique et chronologique de ces travaux de recherche est détaillée dans la section 1.3 nommée *Activités de recherches* du chapitre 1. Enfin, la dernière partie de ce manuscrit est constituée de perspectives générales de travaux envisagés.

Remerciements

Je voudrais tout d'abord remercier Didier Henrion, Nacim Ramdani, Stef Graillat, Marie-Françoise Roy et Luc Jaulin pour avoir accepté de participer au jury de cette habilitation.

Je tiens aussi à remercier tout spécialement Laurent Hardouin pour avoir été mon directeur de recherche lors de cette habilitation et m'avoir donné l'opportunité d'enseigner le contrôle optimal en master.

J'en profite aussi pour remercier tous les collègues de l'université avec lesquels j'ai déjà partagé tant de bons moments et de belles discussions (pas que scientifiques). Merci Mehdi Lhommeau, Sébastien Lagrange, Jean-Baptiste Fasquel, Sébastien Lahaye, Bertrand Cotentenceau, Jean-Louis Boimond, François Chapeau-Blondeau, Marie-Françoise Gérard, Philippe Lucidarme, Frédéric Guégnard, Anne Humeau et Christelle Jussien.

Enfin, un immense merci à toi, Julie, qui en plus de m'avoir soutenu à présenter ce travail, réussit toujours à créer toutes les conditions permettant de m'épanouir professionnellement et personnellement. Un grand merci à toi.

Chapitre 1

Présentation générale

Dans ce chapitre, sont rassemblés quelques renseignements administratifs concernant ma formation, les postes que j'ai occupés, mes activités de recherche et mes publications. La description scientifique de mes recherches, dont une présentation résumée est donnée au paragraphe 1.3, ne débute qu'au chapitre 2.

1.1 Curriculum vitae

1.1.1 Etat civil

| | |
|----------------------|--|
| Nom | Nicolas Delanoue |
| Naissance | 30 octobre 1980 à Angers |
| Nationalité | Française |
| Situation de famille | marié, 2 enfants. |
| Enseignement | Institut Universitaire de Technologie - Site de Belle-Beille 4, Bd Lavoisier, BP 42018, 49016 Angers Cedex |
| Recherche | Laboratoire Angevin de Recherche en Ingénierie des Sys- tèmes (LARIS) 62, Avenue Notre Dame du Lac, 49 100 Angers, |
| Téléphone | +33 2 44 68 75 69 |
| Portable | +33 6 64 63 00 33 |
| Email | nicolas.delanoue@univ-angers.fr |
| Page web | http://perso-laris.univ-angers.fr/~delanoue/ |

1.1.2 Titres universitaires

1998-2001 *Licence en Mathématiques pures*

Lieu UCO, Angers

2001-2002 *Master 1 : Mathématiques*

Lieu Université d'Angers

Mémoire Convergences fortes et faibles dans les espaces L^p .

2002-2003 *Master 2 : Mathématiques et applications*

| | |
|-------------------|--|
| Lieu | UFR sciences et techniques, département mathématiques, Université de Nantes, France |
| Option | Equations aux dérivées partielles |
| Mémoire de Master | Une méthode numérique pour la résolution d'EDP basée sur les volumes finis. |

2005 *CAPES de mathématiques.*

2003-2006 *Doctorat de l'université d'Angers*

suiivi en 2008 de la qualification aux fonctions de maître de conférences par les sections 27 et 61 du CNU.

Intitulé Algorithmes numériques pour l'analyse topologique.
Analyse par intervalles et théorie des graphes.

Soutenance le 14 décembre 2006.

Directeur Luc Jaulin

Membres du jury :

- Rapporteur : Michel PETITOT, LIFL, Laboratoire d'informatique, Lille, France
- Rapporteur : Frédéric BENHAMOU, LINA, Laboratoire d'informatique, Nantes, France.
- Examineur : Adam PARUSINSKI, LAREMA, Mathématiques, Angers, France.
- Examineur : Luc JAULIN, Directeur de thèse, ENSTA, Brest, France.
- Examineur : Bertrand COTTENCEAU, Co-directeur, LARIS, Angers, France.

1.1.3 Activités professionnelles

2000-2001 *Moniteur en informatique, UCO Angers*

Enseignement auprès des étudiants de Licence 1 sur la notion de complexité, les boucles et la récursivité ...

Mai 2003 à Septembre 2003 *Stage de Master 2 - ENITIAA Nantes, Mathématiques appliquées.* Une méthode numérique pour la résolution d'EDP basée sur les volumes finis. L'équation aux dérivées partielles résolue numériquement couple la mécanique des fluides et l'électromagnétisme. Des routines Fortran ont été développées pour compléter un logiciel de CFD (Computational Fluid Dynamics) le rendant capable d'intégrer ce couplage.

Depuis 2008 Maître de conférences à l'Institut Universitaire de Technologie d'Angers (IUT).

1.2 Activités pédagogiques

1.2.1 Enseignements

Depuis ma titularisation en tant que maître de conférences à l'IUT, j'ai donné environ 240 heures équivalent TD d'enseignements par an au sein de l'institut. J'ai par ailleurs dispensé des enseignements à la faculté des sciences ou bien à l'ISTIA (école d'ingénieur de l'université) en cycle préparatoire et en cycle ingénieur. En ce qui concerne les notions enseignées, on peut citer de façon non exhaustive :

- les mathématiques au département GEII avec la transformée de Laplace, les séries de Fourier, les intégrales impropres ...

- le calcul scientifique dans le cycle préparatoire de l’ISTIA avec la représentation des nombres sur machine, le conditionnement de matrices, les méthodes itératives de résolution de systèmes linéaires, l’optimisation via la méthode du gradient conjugué, ou encore l’analyse en composantes principales, l’interpolation de Lagrange ...
- le calcul par intervalles et ses applications auprès des étudiants du master SDS de l’université d’Angers,
- l’informatique avec un cours introductif à l’architecture matérielle et logicielle des ordinateurs,
- les systèmes d’exploitation auprès des étudiants du département GEII avec les processus, l’ordonnanceur, les systèmes de fichiers ...
- la modélisation et simulation de systèmes mécaniques articulées auprès des étudiants du cycle ingénieur de l’ISTIA avec les multiplicateurs de Lagrange ...
- les réseaux informatiques et télécommunication avec les modèle OSI, l’encapsulation, les adresses, les protocoles TCP, IP, IPv6, les technologies opérateurs, les protocoles de routage dynamique ...
- la programmation en perl, c, C++, scilab, matlab ...
- la contrôle optimal auprès des étudiants du master SDS avec l’équation d’Hamilton-Jacobi-Bellman, le principe du maximum de Pontryagin ...
- l’électronique numérique avec les câblages, les portes logiques, l’algèbre de Boole, le VHDL ...

J’ai aussi coorganisé et donné des enseignements proposés à des personnes en formation continue. On peut citer par exemple une formation de 5 jours auprès des informaticiens du CHU d’Angers ou bien la formation auprès du CUME¹. Cette formation avait pour objectif de consolider les connaissances réseaux des techniciens du ministères de l’éducation nationale.

En 2015, j’ai été sollicité par le centre Terre des sciences afin de donner une conférence grand public sur le fonctionnement d’Internet. Celle-ci s’est concrétisée par une intervention devant plus de cent personnes à l’institut municipal d’Angers avec pour titre : “Comment fonctionne le réseau des réseaux?”.

A partir de cette conférence grand public et en partenariat avec un collègue de l’ISTIA, Rémy Guyonneau, j’ai créé une capsule pédagogique dont le thème est le fonctionnement du réseau Internet. Il s’agit d’un ensemble de supports numériques comportant des vidéos, des documents et des questionnaires où l’apprenant peut avancer à son rythme en autonomie.

Je participe aussi au renouvellement des enseignements et de leurs supports.

1.2.2 Activités administratives

En plus des enseignements, j’ai eu l’occasion de participer à la vie du département, de la composante et plus généralement de l’université :

- De 2009 à 2012, Directeur des études du Département GEII (Génie électrique et informatique industrielle).
- De 2012 à 2015, Responsable du recrutement au département GEII,

1. Club des Utilisateurs de Micro-ordinateurs dans l’Éducation.

- Depuis 2015, Responsable de l’animation scientifique et culturelle du département GEII. Dans le cadre de cette responsabilité, j’ai coordonné la mise en place de plusieurs évènements :
 - Organisation annuelle de la nuit de l’électricité au cinéma puis à l’espace culturel de l’université,
 - Co-organisation de la conférence de Richard Stallman en partenariat avec la faculté des sciences et l’ISTIA (18 avril 2016),
 - Organisation biannuelle de la journée Post-trophées de la robotique. Cet évènement a pour objectif de faire découvrir l’université à plus de 200 enfants de primaire intéressés par la robotique. Les élèves participent à des ateliers en électronique, programmation, traitement du son, . . .
- Depuis 2013, Chargé de mission numérique pour l’institut universitaire de technologie. Le chargé de mission numérique est à l’interface des usagers (étudiants et personnels de l’IUT) et de la Direction Du Numérique de l’université (DDN). Il a un rôle de communication bidirectionnelle. En effet, il doit connaître les problématiques de chaque profil (personnels administratifs, étudiants, enseignants et enseignants-chercheurs) afin de synthétiser des besoins pour la DDN. D’un autre côté, il doit expliquer les raisons des décisions prises par l’université.

1.3 Activités de recherches

Mes travaux de recherche consistent en la création de méthodes numériques basées sur le calcul par intervalles. Cet outil développé dans les années 1960 par R.E. Moore permet de créer des algorithmes dits garantis. Les données manipulées par ce genre d’algorithmes sont des intervalles dont les bornes sont codables en machine. S’appuyant sur les propriétés de monotonie de fonctions, il est possible de tirer des conclusions pour des ensembles non dénombrables de points. Par exemple, il existe des algorithmes permettant de prouver qu’une fonction réelle est positive sur un intervalle donné. Cet outil m’a permis pendant la thèse de développer des algorithmes qui étudient la topologie d’un ensemble et d’autres qui montrent la stabilité d’un système dynamique. Ces travaux sont décrits dans la section 1.3.1. Depuis la thèse, j’ai continué à m’appuyer sur cet outil afin de développer deux thèmes : l’étude de fonctions et le contrôle optimal. Les travaux autour de l’étude de fonctions sont en quelque sorte dans la continuité de l’étude de la topologie des ensembles alors que les travaux en théorie du contrôle sont le résultat de mon intérêt pour les systèmes dynamiques. Ces travaux seront plus précisément introduits dans la section 1.3.2 avant d’être respectivement détaillés dans les chapitres 3 et 4.

1.3.1 Thèse de doctorat

2003-2006 THÈSE DE DOCTORAT

Algorithmes numériques pour l’analyse topologique. Analyse par intervalles et théorie des graphes.

Mots clé : analyse topologique, analyse par intervalle, algorithmes garantis, domaine d’attraction

Résumé : Le travail présenté dans cette thèse concerne d’une part, l’étude qualitative d’ensembles et d’autre part, celui de l’étude de la stabilité d’un système dynamique. Les méthodes numériques proposées combinent le calcul par intervalles et la théorie des graphes.

De nombreux problèmes, comme l’étude de l’espace des configurations d’un robot, se ramènent à une étude qualitative d’ensembles. Ici, la “taille” de l’ensemble importe peu, ce qui compte, c’est sa “topologie”. Les méthodes proposées calculent des invariants topologiques d’ensembles. Les ensembles considérés sont décrits à l’aide d’inégalités \mathcal{C}^∞ . L’idée maîtresse est de décomposer un ensemble donné en parties contractiles et d’utiliser l’homologie de Čech. Ces méthodes trouvent leurs applications par exemple en robotique et tout particulièrement en planification de trajectoires.

La seconde partie de la thèse concerne l’étude de points asymptotiquement stables des systèmes dynamiques (linéaires ou non). Plus largement, on propose une méthode pour approcher le bassin d’attraction d’un équilibre asymptotiquement stable. Dans un premier temps, on utilise la théorie de Lyapunov et le calcul par intervalles pour trouver effectivement un voisinage inclus dans le bassin d’attraction d’un point prouvé asymptotiquement stable. Puis, on combine, une fois de plus, la théorie des graphes et les méthodes d’intégration d’équations différentielles ordinaires pour améliorer ce voisinage et ainsi construire un ensemble inclus dans le bassin d’attraction de ce point.

1.3.2 Travaux de recherches engagés depuis la thèse

A la suite de la thèse, deux thèmes de recherche ont été menés en parallèle. L’un autour de la robotique et l’autre autour du contrôle des systèmes dynamiques. Ces contributions seront détaillées respectivement dans les chapitres 3 et 4.

En robotique, la modélisation de robots diffère selon les besoins d’étude et de simulation. En supposant toutes les masses nulles et toutes les pièces rigides, on peut modéliser un mécanisme articulé par une relation $g(x, y) = 0$ où x et y sont respectivement les variables de configuration et les variables de l’espace de travail. Dans le cas des robots sériels, la relation précédente prend la forme particulière suivante : $y = f(x)$. Autrement dit, il est possible de connaître la posture y du robot étant donnée sa configuration x . La figure 1.1 donne un exemple de robot série. A droite sur la figure est représenté un robot avec deux degrés de liberté. Les deux pièces mobiles sont reliées à l’aide de liaisons pivot. Les angles au niveau de ces liaisons sont notés x_1 et x_2 . L’espace des configurations de ce robot est donc le produit cartésien de deux cercles, c’est à dire un tore $\mathbb{R}^2/\mathbb{Z}^2$ (ici noté X). Pour ce mécanisme, la position y de l’organe terminal vit dans un plan et est fonction du couple (x_1, x_2) . L’expression de cette fonction est donnée par l’équation (1.1).

$$f : \begin{matrix} X \\ \left(\begin{matrix} x_1 \\ x_2 \end{matrix} \right) \end{matrix} \rightarrow \begin{matrix} \mathbb{R}^2 \\ \mapsto \left(\begin{matrix} 2 \cos(x_1) + \cos(x_1 + x_2) \\ 2 \sin(x_2) + \sin(x_1 + x_2) \end{matrix} \right) \end{matrix}. \quad (1.1)$$

En supposant les angles x_1 et x_2 bornés, la figure 1.2 donne une représentation graphique de cette fonction. Grace à cette représentation, on peut en déduire graphiquement

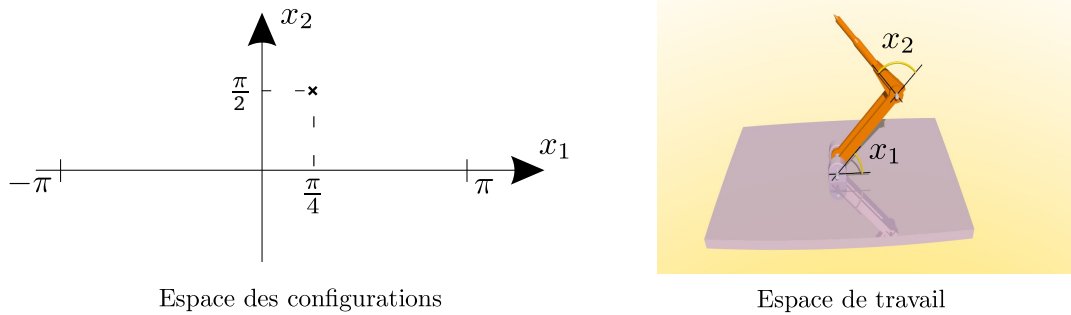


FIGURE 1.1 – Modèle géométrique direct.

que certains points de l'espace de travail admettent deux antécédents alors que d'autres n'en ont qu'un (voire pas du tout).

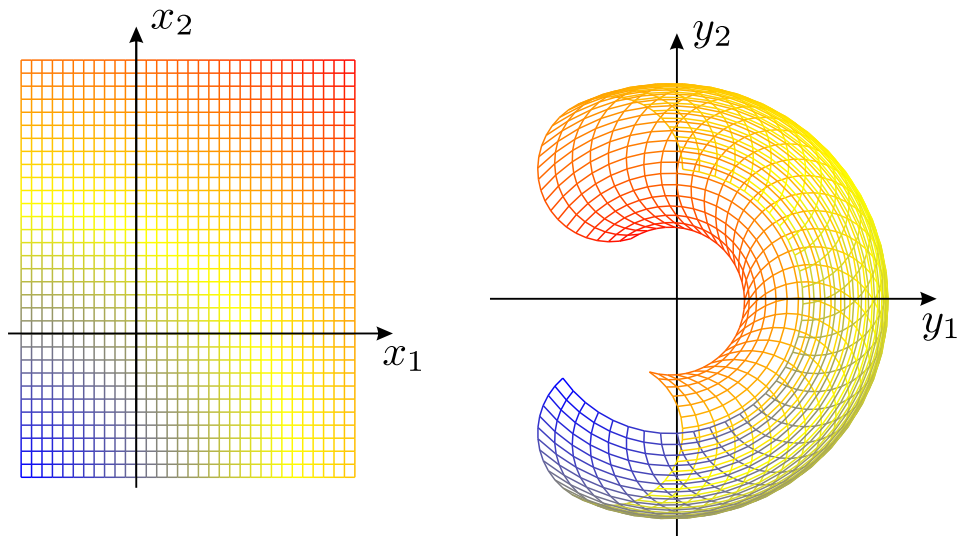


FIGURE 1.2 – Représentation graphique de la fonction (1.1).

Presque toutes les informations pertinentes d'un robot série sont contenues dans la fonction f qui le modélise. Les roboticiens nomment cette fonction le *modèle géométrique direct*. Plus exactement et pour que l'étude soit pertinente, il faut quotienter par la relation d'équivalence suivante : changement de variables lisses sur l'espace des configurations ou sur l'espace de travail. En effet, lors de la modélisation, le roboticien est libre de choisir n'importe quel système de coordonnées. Mathématiquement, le problème consiste donc en l'étude d'une fonction f donnée modulo une relation d'équivalence. Une autre façon de voir les choses consiste à se donner l'espace des fonctions lisses de X dans Y et de considérer l'action des difféomorphismes² de X et de Y sur cet espace de fonctions.

2. On rappelle ici que l'ensemble des difféomorphismes sur un ensemble X forme un groupe avec comme loi de composition interne la composition.

Ce problème est un sujet d'étude profond en mathématiques et a motivé de nombreux chercheurs [1, 2, 3, 4, 5]. Les premières propriétés qui survivent à cette action sont les singularités de f . En effet, si f admet une singularité d'une certaine nature, après une action, la nouvelle fonction aura une singularité de même nature. De plus, la topologie globale des espaces singuliers reste la même. Par conséquent, avec Sébastien Lagrange (LARIS), nous avons développé un algorithme capable de calculer la topologie de ces espaces singuliers pour des fonctions lisses entre des espaces de dimension 2. Ces travaux ont été publiés dans [6]. Seuls des exemples académiques avaient été traités lors de cette publication. D'un point de vue théorie de la calculabilité, il est indécidable de tester si deux fonctions sont dans la même classe. En effet, toute partie fermée de \mathbb{R}^n peut être le lieu singulier d'une fonction lisse. Un tel algorithme devrait être capable de montrer l'égalité de deux tels sous-ensembles, ce qui est impossible. Heureusement, les travaux de Arthur Sard, de Hassler Whitney, puis de René Thom, avec le théorème de transversalité, nous permettent d'espérer un nombre fini de types de singularités pour certaines dimensions. En ces dimensions dites "gentilles", l'ensemble de ces fonctions forme une partie résiduelle. Techniquement on parle alors de fonctions génériques.

À la suite de cette publication, j'ai coencadré avec Philippe Wenger (CNRS-LS2N) et Sébastien Lagrange la thèse de doctorat de Romain Benoit. Le sujet de thèse est la classification des robots. Ces travaux ont tout d'abord validé l'approche initiée dans [6] sur des cas réels issus de la robotique. En particulier, l'étude des singularités du robot, classiquement nommé 3R, a pu être menée avec ce genre de techniques numériques et ces résultats ont été publiés dans Mechanical Sciences [7]. Après la détermination de lieux singuliers d'une fonction modèle géométrique direct, nous avons proposé, toujours avec Romain Benoit, une représentation combinatoire d'une fonction donnée. L'idée consiste à générer pour une fonction donnée f une représentation discrète de cette fonction modulo les changements de variables sur la source ou le but. Ces travaux peuvent être vus comme une généralisation de la notion de tableau de variations enseignée au lycée. Nous avons baptisé ce concept de représentation discrète de fonctions sous le nom de *portrait* de fonction (par analogie avec la notion de *portrait de phase* qui existe pour les systèmes dynamiques).

Ce genre de description combinatoire s'avère très utile en robotique. En effet, dans le contexte de planification de trajectoires, connaître l'agencement topologique des singularités des robots est primordial puisqu'au voisinage de ces singularités, il est souvent possible de faire un choix dans l'espace de configuration afin de suivre une trajectoire dans l'espace de travail. Un autre exemple est celui des singularités nommées *fronce*.

D'un point de vue robotique, au voisinage d'un point fronce, il est possible pour un robot manipulateur série de changer de posture sans traverser de singularité [8]. Ces travaux de description globale d'une fonction ont fait l'objet d'une publication dans Robotica [9].

Les différents aspects de ces méthodes ont été présentés lors de plusieurs séminaires. L'approche générale de la classification de fonctions lisses a été présentée lors des Rencontres Arithmétiques de l'Informatique Mathématiques de 2013 à l'institut Henri Poincaré. Suite à cet exposé, j'ai participé au séminaire organisé par Guillaume Moroz (INRIA) dans le cadre du projet ANR SingCAST où j'ai proposé un algorithme capable de calculer la topologie du contour apparent d'une fonction générique de \mathbb{R}^2 dans \mathbb{R}^2 .

Par ailleurs, une autre partie de mes travaux porte sur les systèmes dynamiques et en

particulier sur le contrôle optimal. Un cours donné par Jean-Bernard Lasserre et Didier Henrion (LAAS Toulouse) m'a donné goût aux approches basées sur la théorie de la mesure. En effet, cette théorie permet de modéliser de nombreux problèmes (transports optimal, contrôle optimal) sous forme de programmes linéaires en dimension infinie. D'un point de vue effectif, l'approche de Jean-Bernard Lasserre propose de "discrétiser" le problème de programmation linéaire à l'aide des moments de cette mesure inconnue. Les contraintes du programme linéaire étant transformées en contraintes sur les moments. Combinés à des résultats relativement récents de géométrie semi-algébrique réelle (SOS, Putinar . . .), ces problèmes peuvent être numériquement résolus.

Étant donné que les outils de l'arithmétique par intervalles peuvent montrer qu'une fonction est positive sur un intervalle, j'ai tenté de proposer une alternative aux méthodes développées par Jean-Bernard Lasserre et Didier Henrion. Par soucis de simplicité, Mehdi Lhommeau et moi avons commencé par nous attaquer au problème de transport optimal. En effet, contrairement au problème de contrôle optimal, aucune contrainte différentielle n'apparaît. Ces travaux ont donné lieu à la publication [10] dans *Computational Optimization and Applications* où il est proposé une discrétisation du primal et du dual afin de calculer respectivement une borne inférieure et une borne supérieure de l'optimum. Ces travaux ont aussi fait l'objet d'un exposé lors de SWIM 2014.

Toujours dans ce contexte de théorie de la mesure, nous avons développé une méthode capable de calculer une borne inférieure au problème de contrôle optimal. L'approche est la suivante : grâce à la démarche de Vinter, le problème de contrôle optimal est formulé sous forme d'un problème de programmation linéaire en dimension infinie à l'aide de la notion de mesure d'occupation. Dans un second temps, il est discrétisé, plus ou moins finement, grâce à des bornes obtenues via l'arithmétique des intervalles. La résolution de ce dernier donne une borne inférieure à l'optimum du problème de contrôle optimal. Ces travaux ont été soumis pour publication [11] auprès de *Computational Optimization and Applications*. Les premières versions de cette approche ont aussi été diffusées lors du congrès international : *SIAM Conference on Control and Its Applications* en 2015.

Plus récemment, nous tentons d'appliquer ce genre de techniques au dual du problème de Vinter. Une telle résolution nous donnerait ce qui est appelé une sous solution de la fonction valeur solution de l'équation d'Hamilton-Jacobi-Bellman. L'équation aux dérivées partielles d'Hamilton-Jacobi-Bellman peut être vue comme une généralisation de l'algorithme de Bellman pour le calcul du plus court chemin dans un graphe au cas où le temps, l'état et les commandes sont continus. J'ai eu la chance d'enseigner la théorie du contrôle optimal auprès des étudiants du Master SDS à Angers en 2016.

Finalement, les différentes rencontres lors de déplacements, en particulier à Manchester, ont été l'occasion d'échanges scientifiques autour du problème SLAM qui apparaît en robotique mobile. L'acronyme anglais SLAM signifie localisation et cartographie simultanées. Un robot autonome doit au cours de sa mission, simultanément, construire une carte de l'environnement et l'utiliser afin de se localiser. Nous avons récemment publié, avec Alexandru Stancu and co, l'article [12] dans *Robotics and Autonomous Systems*. Dans ce papier, l'arithmétique par intervalles est utilisée afin de manipuler le support des distributions des erreurs de mesures.

Au sein de mon laboratoire, j'ai aussi eu l'occasion d'échanger sur des méthodes traitement d'images avec Jean-Baptiste Fasquel. Ces discussions ont donné lieu à la publication de l'article [13] sur la segmentation d'images dans *IEEE Transactions on Pattern Analysis*

and Machine Intelligence. L'originalité de l'approche consiste à imposer des contraintes qualitatives sur les régions à segmenter. Ces contraintes sont modélisées par un graphe G_0 qui code des relations d'inclusion et de photométrie entre les segments attendus. A partir d'une image, un graphe G_1 est généré pour lequel les nœuds sont les segments obtenus lors d'une sursegmentation. Puis, certains segments de cette sursegmentation sont fusionnés après avoir maximisé une fonction coût définie sur l'ensemble des sous-isomorphismes de graphes entre G_0 et G_1 . Toujours basé sur ces informations topologiques à priori et codées dans un graphe, nous avons aussi proposé une méthode guidant l'algorithme des k-means [14].

Dans la suite du document, je ne mettrai en avant que les travaux dont je suis à l'initiative.

1.3.3 Encadrement doctoral et scientifique

Thèse de Romain Benoit

| | |
|-------------------------|--|
| Doctorant | Romain BENOIT (octobre 2013 - novembre 2017) |
| Titre de la thèse | Analyse qualitative de robots. |
| Encadrement | 50% Philippe Wenger, 30% Nicolas Delanoue, 20% Sébastien Lagrange. |
| Financement | Allocation de recherche du ministère et monitorat. |
| Publications | 3 articles dans des revues avec comité de lecture, dont 2 en phase de soumission 1 congrès international avec actes et comités de lecture. |

Constitution du jury :

- Rapporteur : Luc JAULIN, Professeur des Universités,
- Rapporteur : Fabrice ROULLIER, Directeur de recherches INRIA,
- Examineur : Guillaume MOROZ, Chargé de recherche INRIA,
- Examineur : Michel COSTE, Professeur des Universités émérite,
- Examineur : Marie-Françoise ROY, Professeure des Universités émérite,
- Invité : Damien CHABLAT, Directeur de recherche CNRS.

Thèse de Luz Adriana Guzman

| | |
|------------------------------------|---|
| Doctorante | Luz Adriana Guzman (février 2018 - janvier 2021) |
| Titre provisoire de la thèse | Advanced Max+ Based Techniques in Optimization of Complex Dynamic Systems. |
| Encadrement | 40% Nicolas Delanoue, 60% Sébastien Lahaye, Vadim Azhmyakov. |
| Financement | Cotutelle avec l'université de Medellin, Colombie. |
| Publications | 1 congrès international avec actes et comités de lecture soumis. |

Stage de Master d'Auguste Sallen

| | |
|--------------------------|--|
| Etudiant de master | Auguste Sallen (Mai à Juillet 2017) |
| Titre du stage | Modélisation dynamique en robotique humanoïde - La méthode de Newton-Euler |
| Encadrement | 50% Nicolas Delanoue, 50% Philippe Lucidarme. |
| Financement | École normale supérieure Rennes. |

1.3.4 Animation scientifique

- Depuis 2016, je suis responsable de l'équipe robotique du LARIS. Cette responsabilité a pour objectif de coordonner les actions scientifiques menées par les 4 maîtres de conférences rattachés à cette équipe. On peut citer par exemple la co-écriture du dossier déposé auprès de l'Agence Nationale de la Recherche en 2018 autour d'un projet de robotique humanoïde nommé B4D2. Par ailleurs, j'ai eu l'occasion de m'investir, à divers degrés, dans le dépôt de deux dossiers retenus pour financement lors de l'Appels à Projets Recherche de l'Atlantic 2020 :

- Poodle, porté par Rémi Guyonneau, Université d'Angers - LARIS/SDO - en partenariat avec LARIS/ISISV, Moltech/Spectroscopie et IRHS/EMERSYS,
- Robibio, porté par Philippe Lucidarme, Université d'Angers - LARIS/SDO - en partenariat avec le LS2N/Rev. Le projet Robibio a été accepté dans la catégorie Challenge et est complémentaire au projet déposé auprès de l'ANR en robotique humanoïde.

De plus, notre équipe concrétise des partenariats industriels avec des entreprises comme OCTOPUS ROBOTS SA ou bien Balyo sous forme de thèses CIFRE ou bien de stages de master.

- Depuis 2017, je suis co-animateur (avec Abdallah Assi du département de mathématiques de la faculté des sciences) de l'axe *Fondements des mathématiques et du numérique* de la SFR MathSTIC. Cette SFR fédère 3 laboratoires de l'Université d'Angers : le laboratoire de mathématiques LAREMA, le laboratoire d'informatique LERIA et le LARIS.

1.3.5 Invitation d'universités étrangères

Mes différentes communications écrites ou orales m'ont permis d'être repéré, puis invité, par les deux chercheurs suivants qui sont renommés dans notre communauté :

- Warwick TUCKER : en 2008 lorsqu'il était responsable de l'équipe CAPA à l'Université de Bergen en Norvège, puis en 2012 lorsqu'il était à l'Université d'Uppsala en Suède,
- Alexandru STANCU à l'université de Manchester (Royaume-Uni) en 2014 avec qui j'ai par la suite coécrit l'article [12].

1.4 Publications et communications

1.4.1 Publications dans des journaux internationaux

- (1) Nicolas Delanoue, Luc Jaulin, and Bertrand Cottenceau. Using interval arithmetic to prove that a set is path-connected. *Theoretical Computer Science*, 351(1) :119 - 128, 2006. Real Numbers and Computers. [15]
- (2) Nicolas Delanoue, Luc Jaulin, and Bertrand Cottenceau. Guaranteeing the homotopy type of a set defined by non-linear inequalities. *Reliable Computing*, 13(5) :381-398, Oct 2007 [16]
- (3) Sébastien Lagrange, Nicolas Delanoue, and Luc Jaulin. On sufficient conditions of the injectivity : Development of a numerical test algorithm via interval analysis. *Reliable Computing*, 13(5) :409-421, Oct 2007. [17]
- (4) Sébastien Lagrange, Nicolas Delanoue, and Luc Jaulin. Injectivity analysis using interval analysis : Application to structural identifiability. *Automatica*, 44(11) :2959 - 2962, 2008. [18]
- (5) Nicolas Delanoue, Luc Jaulin, and Bertrand Cottenceau. An algorithm for computing a neighborhood included in the attraction domain of an asymptotically stable point. *Communications in Nonlinear Science and Numerical Simulation*, 21(1) :181 - 189, 2015. Numerical Computations : Theory and Algorithms (NUMTA 2013), International Conference and Summer School. [19]
- (6) Nicolas Delanoue and Sébastien Lagrange. A numerical approach to compute the topology of the apparent contour of a smooth mapping from \mathbb{R}^2 to \mathbb{R}^2 . *Journal of Computational and Applied Mathematics*, 271(Supplement C) :267 - 284, 2014. [6]
- (7) Nicolas Delanoue, Mehdi Lhommeau, and Philippe Lucidarme. Numerical enclosures of the optimal cost of the kantorovitch's mass transportation problem. *Computational Optimization and Applications*, 63(3) :855-873, Apr 2016. [10]
- (8) R. Benoit, N. Delanoue, S. Lagrange, and P. Wenger. Guaranteed detection of the singularities of 3r robotic manipulators. *Mechanical Sciences*, 7(1) :31-38, 2016. [7]
- (9) Mohamed Mustafa, Alexandru Stancu, Nicolas Delanoue, and Eduard Codres. Guaranteed slam - an interval approach. *Robotics and Autonomous Systems*, 2018 [12]
- (10) Jean-Baptiste Fasquel and Nicolas Delanoue. A graph based image interpretation method using a priori qualitative inclusion and photometric relationships. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017. [13]
- (11) Jean-Baptiste Fasquel and Nicolas Delanoue. An approach for sequential image interpretation using a priori binary perceptual topological and photometric knowledge and k-means based segmentation. *Journal of the Optical Society of America* 2018. [14]
- (12) Romain Benoit, Nicolas Delanoue, Sébastien Lagrange and Philippe Wenger, P. Combinatorial description of 2-dimensional kinematic functions : application to 3R robotic manipulators *Soumis à Robotica*, 2018. [9]
- (13) Nicolas Delanoue, Mehdi Lhommeau, and Sébastien Lagrange. Nonlinear optimal control : A numerical scheme based on occupation measures and interval analysis. *Soumis à Computational Optimization and Applications*, 2017. [11]

1.4.2 Chapitres d'ouvrage collectif

- (1) Nicolas Delanoue, Luc Jaulin, and Bertrand Cottenceau. *Counting the Number of Connected Components of a Set and Its Application to Robotics*, pages 93-101. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006 [20]
- (2) Nicolas Delanoue, Luc Jaulin, Laurent Hardouin, and Mehdi Lhommeau. *Guaranteed characterization of capture basins of nonlinear state-space systems* Informatics in Control, Automation and Robotics Lecture Notes in Electrical Engineering, 2009, Volume 24, Part III, 265-272. [21]

1.4.3 Conférences internationales

- (1) *Proving that a set is connected via interval analysis.*
N. DELANOUE, L. JAULIN, B. COTTENCEAU.
Workshop on state-of-the-art in scientific computing, Minisymposium, Interval methods, Para' 04 - Denmark - June 23, 2004. PARA'04. Copenhagen Denmark.
- (2) *Guaranteed numerical injectivity test via interval analysis.*
S. LAGRANGE, N. DELANOUE, L. JAULIN.
International Conference on Principles and Practice of Constraint Programming (CP'06)
- (3) *Garantir le type d'homotopie d'un ensemble défini par des inégalités.*
Exposé le 29 janvier 2007 lors des Journées Nationales de Calcul Formel (JNCF) 2007 CIRM, Luminy, 29 janvier - 2 février 2007.
- (4) *Attraction domain and nonlinear dynamical system using interval analysis.*
SWIM08 : a Small Workshop on Interval Methods, Montpellier 2008
- (5) *A new method for integrating ODE based on monotonicity (2010)*
N. DELANOUE, L. JAULIN.
SCAN 2010 14th GAMM-IMACS International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics ENS de Lyon, France, September 27-30, 2010
- (6) *Attraction domain of a nonlinear system using interval analysis and Lyapunov Theory (2011)*
N. DELANOUE, L. JAULIN.
School and Conference on Computational Methods in Dynamics - Trieste - Italy - 20 June 2011 - 8 July 2011
Invité par Warwick Tucker.
- (7) *Classification of mappings from \mathbb{R}^2 to \mathbb{R}^2 (2012)*
N. DELANOUE, S. LAGRANGE
Intervals Pavings and Applications - Uppsala University, Uppsala, Sweden - 15-17 October 2012
- (8) *Classification of stable maps from a simply connected subset of \mathbb{R}^2 to \mathbb{R}^2*
Exposé en juillet 2013 lors de SWIM 2013 - Small Workshop on Intervals Methods à Brest
- (9) *Interval analysis and Optimal Transport (2014)*
SWIM 2014, 7th Small Workshop on Interval Methods, Uppsala, Sweden

- (10) *A guaranteed numerical method to classify smooth mappings from \mathbb{R}^2 to \mathbb{R}^2*
 Invité à exposer par Alexandru Stancu lors SMART 2014 - 1st Small Symposium on Set-Membership : Applications, Reliability and Theory
 The University of Manchester, Aerospace Research Institute, Manchester, UK
- (11) *Optimal Control Via Occupation Measures and Interval Analysis (2015)*
 N. DELANOUE, M. LHOMMEAU, S. LAGRANGE.
 SIAM CT 15, 2015 SIAM Conference on Control and Its Applications (CT15)
- (12) *Optimal Control Via Occupation Measures and Interval Analysis*
 Exposé le 15 juin 2017 à Manchester lors du symposium SWIM-SMART 2017
 Summer Workshop on Interval Methods, and 3rd International Symposium on Set Membership - Applications, Reliability and Theory. The University of Manchester, Aerospace Research Institute, Manchester, UK
- (13) *Preliminary survey of backdrivable linear actuators for humanoid robots*
 PHILIPPE LUCIDARME, NICOLAS DELANOUE, FRANCK MERCIER, YANNICK AOUS-
 TIN, CHRISTINE CHEVALLEREAU, PHILIPPE WENGER
 ROMANSY 2018, 22nd CISM IFToMM Symposium on Robot Design, Dynamics and Control June 25-28, 2018 Rennes, France
- (14) *Robust Optimal Control of Linear-Type Dynamic Systems with Random Delays*
 VADIM AZHMYAKOV, ERIK I. VERRIEST, LUZ A. GUZMAN TRUJILLO, SEBASTIEN LAHAYE, NICOLAS DELANOUE
 Soumis 9th IFAC Symposium on Robust Control Design (ROCOND'18), Florianopolis, Brazil.

1.4.4 Séminaires et groupe de travail

- (1) *Prouver qu'un ensemble est connexe en utilisant le calcul par intervalles.*
 Exposé le 5 février 2004 à E.N.S. Cachan auprès du Groupe de Travail Ensembliste.
- (2) *Introduction au calcul par intervalles, exemple d'utilisation pour le π_0 d'un ensemble.*
 Exposé le 5 mai 2004 à Piriac/mer à l'occasion des Rencontres doctorales Mathématiques 2004.
- (3) *Garantir le type d'homotopie d'un ensemble défini par des inégalités.*
 Exposé le 5 novembre 2004 au Département et laboratoire de mathématiques de l'université d'Angers.
- (4) *Introduction à la topologie algébrique. Garantir le type d'homotopie d'un ensemble via le calcul par intervalles.*
 Exposé le 3 février 2005 lors du Groupe de travail sur les Méthodes Ensemblistes (Modélisation, Analyse et Conduite des Systèmes dynamiques).
- (5) *Prouver qu'un ensemble est connexe via l'analyse par intervalle.*
 Invité à exposer le 5 septembre 2005 lors des JN-MACS Lyon.
- (6) *Analyse par intervalles et théorie de Lyapunov*
 Exposé le 29 novembre 2005 lors du séminaire des doctorants du laboratoire de mathématiques d'Angers.
- (7) *Stability analysis of a nonlinear system using interval analysis.*
 Exposé le 16 mars 2006 lors du Groupe de travail sur les Méthodes Ensemblistes.

- (8) *Calcul par intervalles pour l'étude de propriétés topologiques d'un ensemble.*
Invité par Nathalie Revol à exposer le 6 avril 2006 à l'ENS de Lyon lors d'un séminaire.
- (9) *Calcul par intervalles et stabilité de systèmes dynamiques.*
Exposé le 11 mai 2006 lors des Rencontres Doctorales de Mathématiques Rennes 2006.
- (10) *La méthode de Newton (d'après le livre de J.P. Dedieu : Points fixes, Zéros et la Méthode de Newton).*
Exposé le 20 juin 2007 lors du séminaires des doctorants d'Angers (Laboratoire d'automatique).
- (11) *Set convexity analysis with Interval Analysis.*
Exposé le 13 mars 2008 lors Groupe de travail du GDR MACS du CNRS.
- (12) *Guaranteeing the homotopy type of a set defined by non-linear inequalities - Guaranteed numerical estimation of the attraction domain with interval analysis*
Invité par Warwick Tucker à donner deux exposés - le 20 décembre 2008 au Département de Mathématiques - Université de Bergen, Norvège.
- (13) *A new method for integrating ODE based on monotonicity*
Exposé le 10 juin 2009 lors de SWIM09 : Workshop on Interval Methods à EPFL Lausanne, Suisse
- (14) *Sur le calcul de la topologie d'objets en utilisant seulement de l'arithmétique d'intervalles.*
Invité à exposer le 27 octobre 2009 lors de RAIM'09 : 3es Rencontres Arithmétique de l'Informatique Mathématique - ENS Lyon.
- (15) *Computational Homology* Exposé des principaux résultats du livre *Computational Homology* de Tomasz Kaczynski, Konstantin Mischaikow, Marian Mrozek.
Séminaire du LARIS - Angers - février 2009
- (16) *Guaranteeing the homotopy type of a set defined by non-linear inequalities.*
Invité à exposer en février 2010 au CIRM Luminy Marseille lors de Topological Methods For The Study Of Discrete Structures
- (17) *Classification des applications lisses de \mathbb{R}^2 to \mathbb{R}^2*
Invité à exposer le 20 novembre 2013 lors de RAIM 2013 - Rencontres Arithmétiques de l'Informatique Mathématique
Institut Henri Poincaré, Paris
- (18) *Interval analysis and Optimal Transport*
Invité par Gilles Chabert le 23 juin 2014 à exposer lors du séminaire Contraintes et Géométrie aux Mine de Nantes
- (19) *Contour apparent d'applications lisses de \mathbb{R}^2 dans \mathbb{R}^2 via le calcul par intervalles.*
Invité à exposer par Guillaume Moroz lors du séminaire : Singular Curves and Surfaces Topology, Méthodes de subdivisions pour les systèmes singuliers en relation avec le projet <https://project.inria.fr/singcast/>
- (20) *Interval analysis and Optimal Transport*
Exposé le 6 juillet 2016 au LIP 6 (laboratoire d'informatique de Paris 6)

Chapitre 2

Arithmétique des intervalles

L'arithmétique des intervalles est une approche qui permet de formaliser des raisonnements manipulant des inégalités sur les nombres réels. Elle pourra être mise à profit pour générer des bornes tout d'abord pour les fonctions, puis pour les ensembles solutions d'équations, ou bien même pour des intégrales. A titre d'exemple, considérons le raisonnement suivant : si les deux nombres réels x et y sont supposés vérifier les inégalités $1 \leq x \leq 2$ et $4 \leq y \leq 6$, on peut en déduire que la quantité $x + y$ vérifie $5 \leq x + y \leq 8$. Nommons X l'intervalle $[1, 2]$ et Y l'intervalle $[4, 6]$, le raisonnement précédent s'écrit maintenant

$$x \in X \text{ et } y \in Y \Rightarrow x + y \in [5, 8].$$

Il est alors très tentant d'écrire l'addition des deux intervalles $X + Y$ et de poser $X + Y = [1, 2] + [4, 6] = [5, 8]$. C'est ce que propose le formalisme de l'arithmétique des intervalles. Avec ce formalisme, on pourra donc écrire :

$$x \in X \text{ et } y \in Y \Rightarrow x + y \in X + Y.$$

Dans ce chapitre, un intervalle $[\underline{x}, \bar{x}]$ est un ensemble qui peut être écrit $\{x \in \mathbb{R}^n, \underline{x} \leq x \leq \bar{x}\}$ avec \underline{x} et \bar{x} deux éléments de \mathbb{R}^n . Ici, la relation \leq doit être comprise composante par composante. Notons que cette définition implique que les intervalles ici considérés sont compacts. L'ensemble des intervalles est classiquement noté \mathbb{IR}^n . On peut étendre le raisonnement précédent, appliqué à l'addition, aux fonctions définies sur les réels. Le fait de disposer d'opérateurs ou de fonctions définies sur des intervalles est concrétisé par le concept de *fonction d'inclusion*.

Définition 1 (Fonction d'inclusion). Une fonction $[\varphi] : \mathbb{IR}^n \rightarrow \mathbb{IR}^m$ est une fonction d'inclusion pour $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ si $\forall [x] \in \mathbb{IR}^n, \varphi([x]) \subset [\varphi]([x])$ (où $\varphi([x])$ désigne l'image directe de $[x]$ par φ , i.e. $\varphi([x]) = \{\varphi(x) | x \in [x]\}$).

Cette définition est illustrée par la figure 2.1.

L'arithmétique des intervalles [22] donne une méthode effective pour construire une fonction d'inclusion. En effet, dès qu'une fonction φ est définie par une expression arithmétique, Arnold Neumaier prouve dans [23] qu'il est toujours possible de construire une fonction d'inclusion $[\varphi]$ pour φ . Cette possibilité d'encadrer l'image directe d'un intervalle par une fonction peut s'avérer très utile. On peut par exemple tenir ce genre de

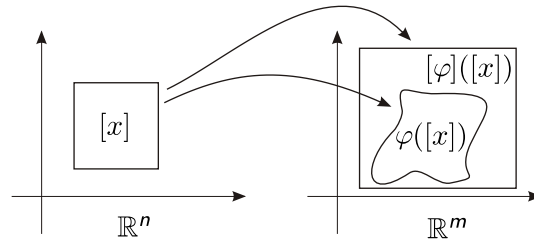


FIGURE 2.1 – Illustration d’une fonction d’inclusion.

proposition : si $0 \notin [\varphi]([x])$, alors on en conclut que $\forall x \in [x], \varphi(x) \neq 0$. Depuis les travaux initiaux de Ramon Moore [22] [24] qui a introduit cette nouvelle arithmétique, de nombreux algorithmes ont été développés dans différentes directions : on peut citer les algorithmes d’optimisation globale [25], des méthodes de résolution numériques pour les systèmes dynamiques non-linéaires [26], ou bien encore des techniques d’estimations dans le cadre des systèmes dynamiques [27] . . .

Comme les méthodes basées sur l’arithmétique des intervalles produisent des résultats garantis, certains ont élaboré des démonstrations de conjectures mathématiques. On peut citer par exemple les travaux de Hales qui en 2003 a démarré un projet nommé “Flyspeak project” (“Formal Proof of Kepler”). Plus exactement, son objectif est de vérifier automatiquement chaque étape de sa preuve de la conjecture de Kepler dont une partie est basée sur l’arithmétique des intervalles. On parle alors de preuve assistée par ordinateur.

Un autre exemple de preuve assisté par ordinateur est celui produit par Warwick Tucker dans le contexte des systèmes chaotiques. En effet, il a démontré l’existence de l’attracteur étrange pour le système différentiel ordinaire décrit par les équations de Lorentz. Sa méthode est partiellement basée sur l’utilisation de l’arithmétique des intervalles [28]. Un des points clés de sa démonstration a nécessité d’étendre les méthodes d’intégration différentielle garanties de [26] au cas des points d’équilibre hyperbolique.

Un autre exemple important est la généralisation de la méthode de Newton, cette généralisation est classiquement nommée : méthode de Newton par intervalles. Elle peut être utilisée pour trouver tous les zéros d’une fonction différentiable donnée $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n$. La méthode de Newton par intervalles crée une suite d’intervalles contenant les zéros de φ et jouie d’intéressantes propriétés : combinée avec le théorème de point fixe de Brouwer [29], elle peut démontrer l’existence et l’unicité d’un zéro de φ , on pourra consulter [30], [31] et la section 2.2.2.

Dans ce chapitre, on explique brièvement comment l’arithmétique des intervalles peut générer des bornes à une fonction φ donnée sur un intervalle donné. On rappellera dans la section 2.2 des techniques de résolutions d’équations. En d’autres termes, on cherchera à encadrer l’ensemble des solutions d’une équation. Finalement, avec cette possibilité d’encadrer une fonction φ , on verra dans la section 2.3 comment on peut encadrer la valeur d’une intégrale de la forme

$$\int_X \varphi(x) d\mu(x)$$

avec μ une mesure sur X .

2.1 Encadrement numérique pour des fonctions

Définition 2 (Arithmétique des intervalles). Considérons deux intervalles réels $[x] = [\underline{x}, \bar{x}]$ et $[y] = [\underline{y}, \bar{y}]$, i.e. $[x], [y] \in \mathbb{IR}$. Les 4 opérations $+$, $-$, \times et \div sont définies, comme dans [32], par :

$$\begin{aligned} [x] + [y] &= [\underline{x} + \underline{y}, \bar{x} + \bar{y}], \\ [x] - [y] &= [\underline{x} - \bar{y}, \bar{x} - \underline{y}], \\ [x] \times [y] &= [\min\{\underline{x}\underline{y}, \underline{x}\bar{y}, \bar{x}\underline{y}, \bar{x}\bar{y}\}, \max\{\underline{x}\underline{y}, \underline{x}\bar{y}, \bar{x}\underline{y}, \bar{x}\bar{y}\}], \\ [x] \div [y] &= [x] \times \left[\frac{1}{\underline{y}}, \frac{1}{\bar{y}} \right], \text{ si } \underline{y}\bar{y} > 0. \end{aligned} \quad (2.1)$$

Proposition 1. Les quatre opérations de l'arithmétique des intervalles (2.1) sont des fonctions d'inclusion pour les fonctions d'addition, soustraction, multiplication et de division définies sur les couples de nombres réels.

On peut aussi définir, comme faisant partie de l'arithmétique des intervalles, les opérations de négation et de puissance :

$$-[x] = [-\bar{x}, -\underline{x}], \quad (2.2)$$

$$[x]^n = \begin{cases} [1, 1] & \text{si } n = 0, \\ [\underline{x}^n, \bar{x}^n] & \text{si } n \text{ est pair ou } \underline{x} \geq 0, \\ [\bar{x}^n, \underline{x}^n] & \text{si } n \text{ est impair et } \bar{x} \leq 0, \\ [0, \max\{\underline{x}^n, \bar{x}^n\}] & \text{sinon.} \end{cases} \quad (2.3)$$

Il est important de noter que la propriété de fonction d'inclusion est compatible avec la composition de fonctions. Par conséquent, dès qu'une fonction φ est définie par son expression et qu'on dispose de fonctions d'inclusion pour les fonctions apparaissant dans son expression, on peut automatiquement générer une fonction d'inclusion $[\varphi]$ pour φ . Cette remarque peut être résumée par la proposition suivante :

Proposition 2. Si φ et ϕ sont des fonctions avec $[\varphi]$ et $[\phi]$ comme fonction d'inclusion, alors $[\varphi] \circ [\phi]$ est une fonction d'inclusion pour $\varphi \circ \phi$.

Démonstration. Soit $[x] \in \mathbb{IR}^n$, alors $\phi([x]) \subset [\phi]([x])$. Par conséquent $\varphi \circ \phi([x]) \subset [\varphi] \circ [\phi]([x])$. \square

Exemple 1. On se donne la fonction $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ définie par l'expression suivante :

$$\varphi(x) = \frac{3}{2}(1 - 2x + x^2).$$

La fonction $[\varphi] : \mathbb{IR} \rightarrow \mathbb{IR}$ définie par

$$[\varphi]([\underline{x}, \bar{x}]) = \left[\frac{3}{2}, \frac{3}{2} \right] \times \left([1, 1] - [2, 2] \times [\underline{x}, \bar{x}] + [\underline{x}, \bar{x}]^2 \right)$$

est une fonction d'inclusion pour φ .

En particulier, si \underline{x} est positif, alors on a

$$[\varphi]([\underline{x}, \bar{x}]) = \left[\frac{3}{2} - 3\bar{x} + \frac{3}{2}\bar{x}^2, \frac{3}{2} - 3\underline{x} + \frac{3}{2}\underline{x}^2 \right].$$

Remarque. On notera que les nombres réels $\frac{3}{2}$, 1 et 2 ont été *convertis* en intervalle et que les opérations initialement définies sur les nombres réels ont été remplacées par leur équivalent défini sur les intervalles. En pratique, on utilise aussi les nombres à virgule flottante et si un nombre réel α n'est pas représentable en machine, alors il est converti en un intervalle $[\underline{\alpha}, \bar{\alpha}]$ dont les bornes sont représentables et contenant α .

De cette manière, on se protège des erreurs d'arrondis de l'arithmétique des nombres à virgule flottante, et on obtient numériquement des bornes garanties.

Remarque. La méthode présentée précédemment génère ce qui est classiquement appelée la *fonction d'inclusion naturelle*. On notera la dangerosité pédagogique de parler de la fonction d'inclusion naturelle dans le sens où une même fonction φ écrite différemment donne naissance à deux fonctions d'inclusion naturelles. L'exemple suivant suffit à résumer cette remarque. Soit $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ définie par $\varphi(x) = (x+1)^2$, d'un point de vue fonctionnelle, φ est aussi égale à $\psi(x) = x^2 + 2x + 1$ alors que les deux fonctions d'inclusion

$$[\varphi]([\underline{x}, \bar{x}]) = ([\underline{x}, \bar{x}] + [1, 1])^2,$$

et

$$[\psi]([\underline{x}, \bar{x}]) = [\underline{x}, \bar{x}]^2 + 2[\underline{x}, \bar{x}] + [1, 1]$$

sont différentes. En effet, par exemple, avec $[\underline{x}, \bar{x}] = [-1, 1]$, on obtient d'un coté

$$[\varphi]([-1, 1]) = [1, 2]^2 = [1, 4]$$

alors que de l'autre

$$[\psi]([-1, 1]) = [0, 1] + [-2, 2] + [1, 1] = [-1, 4].$$

La remarque précédente illustre le fait que la fonction d'inclusion naturelle dépend de l'expression choisie de la fonction initiale. En changeant de point de vue, on peut dire qu'il existe plusieurs fonctions d'inclusion pour une même fonction. Notons qu'il est possible d'ordonner partiellement les fonctions d'inclusions, voire même de mettre une structure de treillis sur cet ensemble [33]. Dans le cas des fonctions réelles continues, comme les intervalles considérés sont des compacts de \mathbb{R} , la fonction ensembliste classiquement appelée image directe est la plus petite, c'est la meilleure des fonctions d'inclusion. Malheureusement, des résultats de calcul formel nous indiquent qu'elle est impossible de la générer algorithmiquement même pour la classe des fonctions polynomiales [34].

En général, on se contentera de la fonction d'inclusion naturelle. Notons qu'il existe d'autres techniques de génération de fonctions d'inclusion. Le lecteur intéressé par ces autres méthodes pourra consulter [23] ou bien [35]. La qualité de la fonction d'inclusion choisie influence le temps de calcul. Théoriquement parlant, il sera suffisant de savoir que la fonction d'inclusion est convergente pour montrer la convergence du schéma numérique associé.

Définition 3. Une fonction d'inclusion $[\varphi]$ de $\varphi : E \rightarrow \mathbb{R}$ est *convergente* si pour toute suite $\{X_k\}_{k \geq 0}$ d'ensembles de E :

$$\Delta[\varphi](X_k) \rightarrow 0 \text{ quand } \Delta X_k \rightarrow 0.$$

Dans cette dernière définition, Δ est la fonction réelle qui renvoie le diamètre d'un ensemble. Dans la suite du manuscrit, nous aurons besoin aussi d'une autre propriété que peut satisfaire une fonction d'inclusion. C'est la notion de monotonie.

Définition 4. Une fonction d'inclusion $[\varphi]$ of $\varphi : E \rightarrow \mathbb{R}$ est dite *monotone* si

$$\forall X, Y \subset E, X \subset Y \Rightarrow [\varphi](X) \subset [\varphi](Y).$$

La monotonie n'est, en général, pas satisfaite par les fonctions d'inclusion. Par contre, étant donnée une suite de décroissantes d'intervalles $\{X_n\}_{n \geq 0}$, on peut à partir d'une fonction d'inclusion $[\varphi]$ créer une fonction d'inclusion monotone $[\varphi]_{\downarrow}$ en posant

$$[\varphi]_{\downarrow}(X_n) = \bigcap_{j \leq n} [\varphi](X_j),$$

ou bien, de façon équivalente, itérativement :

$$[\varphi]_{\downarrow}(X_{n+1}) = [\varphi]_{\downarrow}(X_n) \cap [\varphi](X_{n+1}) \text{ avec } [\varphi]_{\downarrow}(X_0) = [\varphi](X_0).$$

La fonction $[\varphi]_{\downarrow}$ est d'inclusion pour φ et monotone sur l'ensemble $\{X_n\}_{n \geq 0}$. Cette technique pourra être étendue à une suite de partitions et sera mise en pratique lors du chapitre 4.

2.2 Encadrement d'ensemble solution d'équations

2.2.1 SIVIA

L'arithmétique des intervalles peut être employée pour résoudre des équations non linéaires de la forme $\varphi(x) = 0$ où φ est une fonction de $E \subset \mathbb{R}^n$ dans \mathbb{R}^m . L'idée est proche de la méthode informatique nommée brute force. Cette méthode construit un ensemble \bar{S} contenant les solutions de l'équation $\varphi(x) = 0$. Les grandes étapes sont les suivantes :

- on crée une partition finie \mathcal{X} de E ,
- on initialise \bar{S} à l'ensemble vide via l'instruction $\bar{S} \leftarrow \emptyset$,
- pour chaque élément X de la partition \mathcal{X} , on encadre $\varphi(X)$ par $[\varphi](X)$, deux situations incompatibles peuvent alors se produire :
 - soit $0 \in [\varphi](X)$, auquel cas, on ajoute X à \bar{S} en posant $\bar{S} \leftarrow \bar{S} \cup X$,
 - soit $0 \notin [\varphi](X)$, auquel cas, on décide d'oublier X car il est certain qu'aucune solution de $\varphi(x) = 0$ n'est contenue dans X .

Finalement, l'ensemble ainsi construit \bar{S} contient nécessairement les solutions de l'équation $\varphi(x) = 0$. Ce processus est illustré par l'exemple suivant.

Exemple 2. Considérons la fonction φ dont le graphe est donné par la figure 2.2.

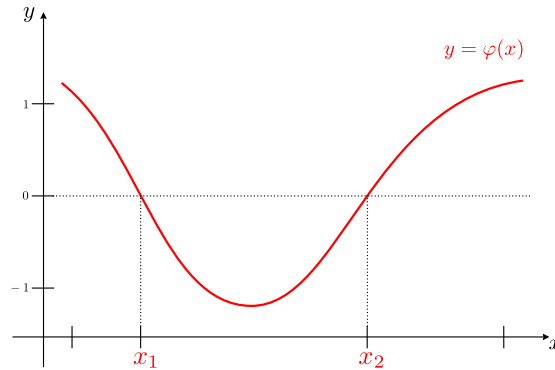


FIGURE 2.2 – Graphe de la fonction φ .

On cherche à résoudre l'équation $\varphi(x) = 0$, i.e. trouver les points nommés x_1 et x_2 . Via cette approche ensembliste, on cherche à encadrer de l'ensemble $S = \{x \mid \varphi(x) = 0\}$. Comme indiqué dans l'algorithme précédent, la méthode consiste tout d'abord à partitionner l'ensemble E avec des intervalles ici nommés X_i , $i = 1, \dots, 6$. L'évaluation des différents $[\varphi](X_i)$ est illustrée par la figure 2.3. En effet, on retrouve en gris les ensembles produits $X_i \times [\varphi](X_i)$, pour $i = 1, \dots, 6$. L'interprétation géométrique du fait que la fonction $[\varphi]$ est d'inclusion pour φ est la suivante : le graphe de φ est inclus dans la réunion des rectangles de couleur grise. Finalement, seuls les intervalles X_2 et X_5 ont une image par

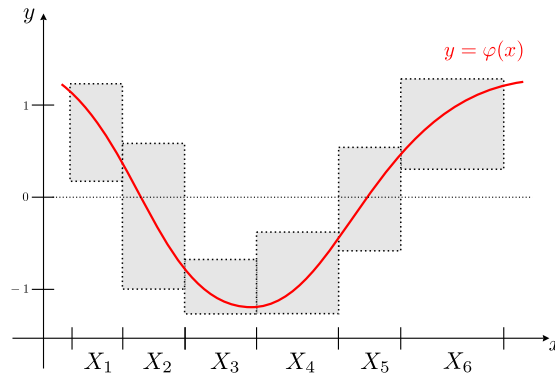


FIGURE 2.3 – Graphe de la fonction $[\varphi]$ définie sur $\{X_i\}_{i=1,\dots,6}$.

$[\varphi]$ qui contient 0. On en déduit que nécessairement les solutions de l'équation $\varphi(x) = 0$ sont dans l'ensemble $\bar{S} = X_2 \cup X_5$. Autrement dit, l'ensemble S des solutions de l'équation $\varphi(x) = 0$ est bien inclus dans \bar{S} , formellement $S \subset \bar{S}$.

Ce procédé peut être rappelé récursivement avec comme entrée non plus E mais l'ensemble \bar{S} récemment calculé jusqu'à obtenir une précision souhaitée. Le résultat sera une approximation de l'ensemble des solutions de $\varphi(x) = 0$. Ce genre d'approche a été formalisé et étendu par Luc Jaulin dans sa thèse pour des ensembles du type

$$\bigcup_i \bigcap_j \{x \in E \mid \varphi_{ij}(x) \leq 0\},$$

où la relation \leq est la relation d'ordre sur \mathbb{R} . Cette méthode se nomme *SIVIA* pour (Set Inversion Via Interval Analysis). On pourra consulter [33] pour plus de détails.

Exemple 3. La figure 2.4 illustre la sortie pour la résolution de l'équation $x^2 + y^2 + xy + 10 \leq 0$ avec (x, y) un point du pavé $E = [-10, 10] \times [-10, 10]$. Autrement dit, on encadre l'ensemble $S = \{(x, y) \in [-10; 10]^2 \mid \varphi(x, y) = x^2 + y^2 + xy + 10 \leq 0\}$ avec deux ensembles \underline{S} et \overline{S} tels que

$$\underline{S} \subset S \subset \overline{S}$$

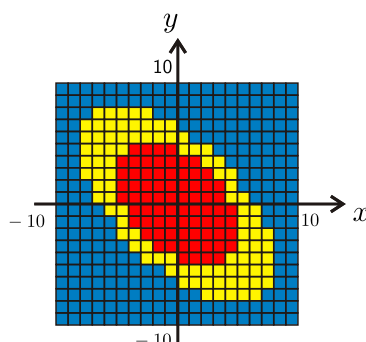


FIGURE 2.4 – Sur cette figure, les pavés inclus dans S sont représentés en rouge et leur réunion forme \underline{S} . Ceux qui n'intersectent pas S sont bleus. Les pavés jaunes sont ceux pour lesquels la méthode n'a pas permis de statuer et donc la réunion des jaunes et des rouges forment \overline{S} .

La raison pour laquelle il existe des pavés jaunes dans l'exemple précédent est double. Premièrement l'ensemble à encadrer est une ellipse et n'est pas la réunion de pavés, on se retrouve donc avec des pavés qui ne contiennent S qu'en partie. La seconde raison vient du fait que les fonctions d'inclusion peuvent être de mauvaise qualité dans le sens où elles contiennent plus, voire beaucoup plus que l'image directe. On nomme ce défaut récurrent des méthodes ensemblistes le *pessimisme*. Ce terme n'est sans doute pas le plus approprié. En effet, ce défaut n'enlève rien quant aux conclusions formelles que l'on peut tirer d'un raisonnement s'appuyant sur l'arithmétique des intervalles.

2.2.2 Newton par intervalles

La méthode présentée dans la section précédente pour résoudre une équation du type $\varphi(x) = 0$ ne génère qu'une sur-approximation de l'ensemble des solutions. Dans le cas d'un système carré, c'est à dire avec autant d'équations que d'inconnues, il est souvent possible de démontrer l'existence et l'unicité des solutions dans un intervalle donné. Cette méthode s'appuie sur un opérateur nommé l'*opérateur de Newton*.

Définition 5 (Opérateur de Newton par intervalles). Soient $[A] \in \mathbb{IR}^{n \times n}$ et $[b] \in \mathbb{IR}^n$, on note par $\Sigma([A], [b])$ un intervalle de \mathbb{IR}^n qui contient l'ensemble $\{x \in \mathbb{R}^n \mid Ax = b, A \in [A], b \in [b]\}$. Soit φ une fonction continûment différentiable de \mathbb{R}^n vers \mathbb{R}^n , $[x]$ un intervalle

de \mathbb{R}^n , et x_1 un élément de $[x]$. Avec $[d\varphi] : \mathbb{I}\mathbb{R}^n \rightarrow \mathbb{I}\mathbb{R}^{n \times n}$ une fonction d'inclusion pour la différentielle $d\varphi$, la fonction

$$\begin{aligned} N : \mathbb{I}\mathbb{R}^n &\rightarrow \mathbb{I}\mathbb{R}^n \\ [x] &\mapsto x_1 - \Sigma([d\varphi]([x]), \varphi(x_1)) \end{aligned} \quad (2.4)$$

est appelée *opérateur de Newton par intervalles*.

Combiné au théorème de point fixe de Brouwer, il est possible de démontrer les propositions suivantes.

Proposition 3. *Si $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ est une fonction continûment différentiable, $[x]$ un intervalle, x_1 et x^* deux points de $[x]$, alors*

$$\varphi(x^*) = 0 \Rightarrow x^* \in N([x]) \quad (2.5)$$

Démonstration. La preuve se trouve dans la thèse [35]. □

La proposition 3 indique que les solutions de l'équation $\varphi(x) = 0$ vivant dans $[x]$ sont nécessairement des éléments de $N([x])$. Elle ne donne aucune information sur l'unicité, ni même sur l'existence des solutions.

Exemple 4. On considère dans cet exemple la fonction réelle φ dont le graphe est représenté sur la figure 2.5.

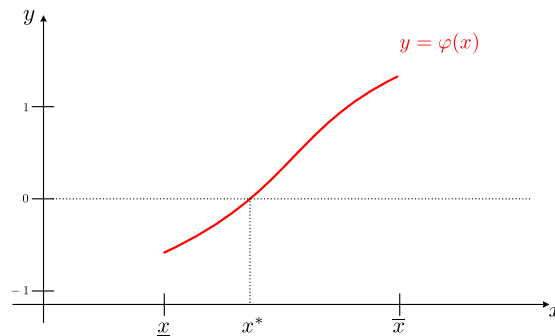


FIGURE 2.5 – Graphe de la fonction φ . Le nombre réel x^* est la seule solution de l'équation $\varphi(x) = 0$.

Notre objectif est de résoudre l'équation $\varphi(x) = 0$ sur l'intervalle $[\underline{x}, \bar{x}]$. On cherche donc à encadrer le nombre réel noté x^* dans la figure précédente. Géométriquement, la relation (2.4) consiste à tracer le cône de toutes les droites passant par le point de coordonnées $(x_1, f(x_1))$ et dont la pente est une pente des tangentes à φ sur l'intervalle $[\underline{x}, \bar{x}]$. Ceci est illustré par la figure 2.6.

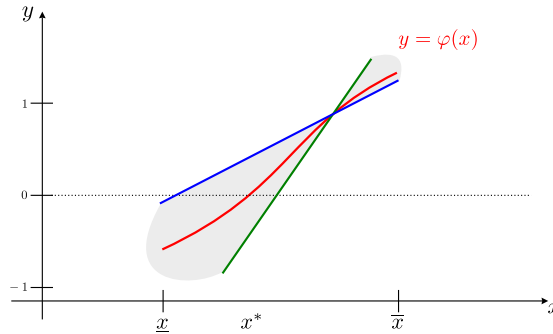


FIGURE 2.6 – Illustration graphique de la méthode de Newton par intervalles.

Finalement, l'intersection du cône avec l'axe des abscisses est donnée par $N([\underline{x}, \bar{x}])$.

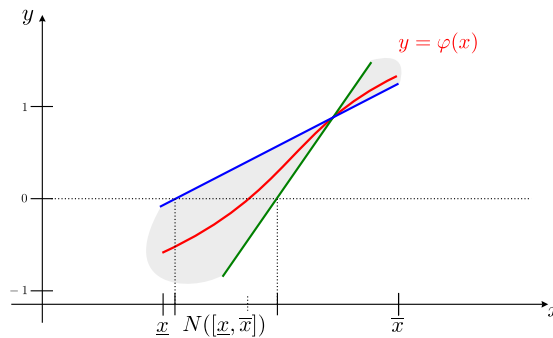


FIGURE 2.7 – Illustration graphique de la méthode de Newton par intervalles.

En accord avec la proposition 3, on vérifie géométriquement que la solution x^* est bien incluse dans $N([\underline{x}, \bar{x}])$. On pourra rappeler récursivement l'opérateur de Newton afin d'améliorer la précision de l'encadrement de x^* . En pratique, si l'ensemble $[d\varphi]([x])$ n'est pas inclus dans $GL(\mathbb{R}^n)$, on ne peut pas encadrer $\Sigma([d\varphi]([x]), \varphi(x_1))$ par un intervalle. Dans ce cas, l'intervalle $[x]$ est partitionné et la méthode de Newton est rappelée récursivement sur tous les intervalles issus de ce partage.

Proposition 4. Soient φ une fonction continûment différentiable sur $[x]$ et $x_1 \in [x]$. Si $[d\varphi]([x]) \subset GL(\mathbb{R}^n)$ et $N([x]) \subset [x]$ alors il existe un unique $x^* \in [x]$ tel que $\varphi(x^*) = 0$.

Démonstration. La preuve se trouve dans [35]. □

En pratique, lorsque les conditions de la proposition 4 sont satisfaites, on peut conclure quant à l'existence et l'unicité de la solution du problème $\varphi(x) = 0$. La proposition suivante anticipe sur le chapitre suivant avec une hypothèse de généricité qui sera rappelée ultérieurement.

Proposition 5. Soit φ une fonction $C^\infty(\mathbb{R}^n, \mathbb{R}^n)$ générique. Alors la méthode de Newton par intervalles termine. (Elle fournit un ensemble fini d'intervalles $\{[x]_i\}_i$ de sorte que chaque intervalle contienne un unique zéro de φ).

Démonstration. La preuve se trouve [35]. □

2.3 Encadrement d'intégrales

Étant données une mesure μ et une fonction mesurable φ , en s'appuyant sur les bornes obtenues via l'arithmétique des intervalles, on propose dans cette section une méthode permettant d'encadrer l'intégrale suivante :

$$\int_E \varphi(x) d\mu.$$

L'approche proposée est similaire à une méthode des rectangles. On donnera aussi des résultats de convergence sous certaines conditions portant sur μ et φ . Dans toute la suite du manuscrit, chaque ensemble sera nécessairement mesurable vis à vis de la mesure considérée. De même, chaque fonction sera supposée implicitement intégrable sur l'ensemble mesurable vis à vis de la mesure par rapport à laquelle on cherche à intégrer. On utilisera dans tout le manuscrit la notation bra-ket de Dirac pour écrire les intégrales, c'est à dire :

$$\langle \mu, \varphi \rangle = \int_E \varphi(x) d\mu.$$

Théorème 1. Soient μ une mesure positive sur E et φ une fonction réelle définie sur E . Supposons qu'il existe des nombres réels $\underline{\varphi}$ et $\overline{\varphi}$ tels que $\forall x \in E, \underline{\varphi} \leq \varphi(x) \leq \overline{\varphi}$, alors

$$\mu(E) \cdot \underline{\varphi} \leq \langle \mu, \varphi \rangle \leq \mu(E) \cdot \overline{\varphi}.$$

Démonstration. On pourra consulter le manuel de Walter Rudin [36], en particulier 1.24. □

Proposition 6. Soit μ une mesure positive sur E , φ une fonction réelle définie sur E et \mathcal{X} une partition finie de E . Pour tout élément X de \mathcal{X} , on se donne les nombres réels $\underline{\varphi}(X)$ et $\overline{\varphi}(X)$ tels que $\forall x \in X, \underline{\varphi}(X) \leq \varphi(x) \leq \overline{\varphi}(X)$. Alors

$$\sum_{X \in \mathcal{X}} \mu(X) \underline{\varphi}(X) \leq \langle \mu, \varphi \rangle \leq \sum_{X \in \mathcal{X}} \mu(X) \overline{\varphi}(X). \quad (2.6)$$

Démonstration. On écrit

$$\langle \mu, \varphi \rangle = \int_E \varphi(x) d\mu = \sum_{X \in \mathcal{X}} \int_X \varphi(x) d\mu.$$

Comme $\forall x \in X, \underline{\varphi}(X) \leq \varphi(x) \leq \overline{\varphi}(X)$, par le théorème 1, on en déduit $\forall X \in \mathcal{X}$:

$$\mu(X) \underline{\varphi}(X) \leq \int_X \varphi(x) d\mu \leq \mu(X) \overline{\varphi}(X).$$

En additionnant, on obtient la double inégalité voulue. □

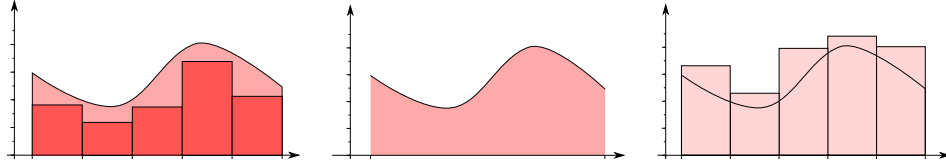


FIGURE 2.8 – Illustration de la proposition 6. La valeur de l'intégrale $\int_a^b f(x)dx$ (au centre) est bien comprise entre $\sum_{X \in \mathcal{X}} \underline{\varphi}(X) \lambda(X)$ (à gauche) et $\sum_{X \in \mathcal{X}} \overline{\varphi}(X) \lambda(X)$ (à droite).

La figure 2.8 illustre la proposition 6. Elle donne un exemple de fonction réelle définie sur \mathbb{R} que l'on cherche à intégrer sur un intervalle E vis-à-vis de la mesure de Lebesgue. Son intégrale, correspondant à l'aire algébrique sous la courbe, est bien comprise entre la somme des aires des rectangles de la figure de gauche et de la figure de droite.

Les bornes supposées connues dans les hypothèses de la proposition 6 pourront être obtenues par une méthode s'appuyant sur l'arithmétique des intervalles. Comme il n'est pas toujours possible de faire une partition \mathcal{X} de E de sorte que tous ses éléments soient des intervalles, on définit ici une extension de l'arithmétique des intervalles à l'ensemble des parties de E .

Définition 6. Une fonction $[\varphi] : 2^E \rightarrow \mathbb{I}\mathbb{R}$ est une *fonction d'inclusion* pour une fonction réelle $\varphi : E \rightarrow \mathbb{R}$ si $\forall X \subset E, \varphi(X) \subset [\varphi](X)$ (où 2^E désigne l'ensemble des parties de E).

Avec ces fonctions d'inclusion, on peut reformuler la proposition 6 de la façon suivante :

Proposition 7. Soit μ une mesure positive sur E , φ une fonction réelle E , \mathcal{X} une partition finie de E , et $[\varphi] = [\underline{\varphi}, \overline{\varphi}]$ une fonction d'inclusion pour φ alors

$$\sum_{X \in \mathcal{X}} \mu(X) \underline{\varphi}(X) \leq \langle \mu, \varphi \rangle \leq \sum_{X \in \mathcal{X}} \mu(X) \overline{\varphi}(X). \quad (2.7)$$

Les exemples suivants illustrent la proposition 7. Dans l'exemple 5, on a une partition fixe de E alors que dans l'exemple 6, le cardinal de la partition est paramétré par un nombre entier n . En augmentant n , on affinera la partition tout en espérant améliorer l'encadrement (2.7).

Exemple 5. Soit μ la mesure, absolument continue par rapport à la mesure de Lebesgue λ , définie pour chaque ensemble mesurable F par

$$\mu(F) = \int_F 3x^2 d\lambda.$$

Soit E l'ensemble $[1, 4]$ et $\mathcal{X} = \{X_i\}_{i \in \{1, 2, 3\}}$ la partition définie de sorte que X_i soit l'intervalle ouvert à droite $[i, i + 1[= \{x \mid i \leq x < i + 1\}$ pour $i = 1$ ou $i = 2$ et $X_3 = [3, 4]$. Supposons maintenant qu'on souhaite encadrer $\langle \mu, \varphi \rangle$ avec $\varphi(x) = 4x$. D'après la proposition 7, on a

$$\sum_{i=1}^3 \underline{\varphi}(X_i) \mu(X_i) \leq \langle \mu, \varphi \rangle \leq \sum_{i=1}^3 \overline{\varphi}(X_i) \mu(X_i), \quad (2.8)$$

où $\underline{\varphi}, \overline{\varphi}$ sont respectivement définies via l'arithmétique des intervalles par $\underline{\varphi}([\underline{x}, \overline{x}]) = 4\underline{x}$ et $\overline{\varphi}([\underline{x}, \overline{x}]) = 4\overline{x}$. Par conséquent, l'équation (2.8) devient

$$\sum_{i=1}^3 4i\mu(X_i) \leq \langle \mu, \varphi \rangle \leq \sum_{i=1}^3 4(i+1)\mu(X_i).$$

Comme $\mu(X_i) = \int_{X_i} 3x^2 d\lambda = (i+1)^3 - i^3 = 3i^2 + 3i + 1$, l'équation (2.8) est équivalente à

$$\sum_{i=1}^3 4i(3i^2 + 3i + 1) \leq \langle \mu, \varphi \rangle \leq \sum_{i=1}^3 4(i+1)(3i^2 + 3i + 1).$$

On en déduit :

$$624 \leq \langle \mu, \varphi \rangle \leq 876.$$

Exemple 6. Considérons la fonction f déjà rencontrée lors de l'exemple 1 et définissons la partition $\mathcal{X}_n = \{X_i\}_{1 \leq i \leq n}$ de sorte que $X_i = [\frac{i-1}{n}, \frac{i}{n}]$. On en déduit

$$\sum_{i=1}^n \lambda(X_i) \underline{\varphi}(X_i) \leq \int_0^1 f(x) d\lambda(x) \leq \sum_{i=1}^n \lambda(X_i) \overline{\varphi}(X_i),$$

c'est à dire :

$$\frac{1}{n} \sum_{i=1}^n \frac{3}{2} - 3\frac{i}{n} + \frac{3}{2} \left(\frac{i-1}{n} \right)^2 \leq \int_0^1 f(x) d\lambda(x) \leq \frac{1}{n} \sum_{i=1}^n \frac{3}{2} - 3\frac{i-1}{n} + \frac{3}{2} \left(\frac{i}{n} \right)^2.$$

Une question naturelle est de savoir si la méthode déduite de la proposition 7 est convergente. Autrement dit, on aimerait savoir si en raffinant la partition \mathcal{X}_n , on obtiendrait un encadrement de plus en plus précis de l'intégrale $\langle \mu, \varphi \rangle$. Afin de faire la preuve de la convergence d'une méthode d'intégration basée sur l'équation 2.7, il nous faut topologiser autant que faire se peut l'espace 2^E .

Supposons que E soit un sous-ensemble compact de l'espace métrique (\mathbb{R}^n, δ) . Avec $X \in 2^E$, on note par \overline{X} sa fermeture. La fonction réelle d définie sur les paires non vides de 2^E par $d(X_1, X_2) = d_H(\overline{X}_1, \overline{X}_2)$ est une pseudo-métrique sur 2^E avec d_H la distance de Hausdorff définie entre les sous-ensembles compacts. On rappelle que la métrique d_H est classiquement donnée par

$$d_H(X_1, X_2) = \max \left\{ \sup_{x_1 \in X_1} \inf_{x_2 \in X_2} \delta(x_1, x_2), \sup_{x_2 \in X_2} \inf_{x_1 \in X_1} \delta(x_1, x_2) \right\}.$$

On notera que cette distance d_H peut aussi être utilisée pour créer une topologie sur l'ensemble des intervalles. Finalement, avec X un élément de 2^E , on définit le diamètre de X par $\Delta(X) = \sup \{ \delta(x, y) \mid x, y \in X \}$.

Les définitions de monotonie et de convergence peuvent être étendues à ces nouvelles fonctions d'inclusion. Ces définitions seront nécessaires lors des preuves de convergence des schémas numériques d'intégration.

Définition 7. Une fonction d'inclusion $[\varphi]$ de $\varphi : E \rightarrow \mathbb{R}$ est dite *monotone* si

$$\forall X, Y \in 2^E, X \subset Y \Rightarrow [\varphi](X) \subset [\varphi](Y).$$

Définition 8. Une fonction d'inclusion $[\varphi]$ de $\varphi : E \rightarrow \mathbb{R}$ est dite *convergente* si pour toute suite $\{X_k\}_{k \geq 0}$ de sous-ensembles de E :

$$\Delta[\varphi](X_k) \rightarrow 0 \text{ quand } \Delta X_k \rightarrow 0.$$

Sous des hypothèses de pseudo-continuités de la fonction d'inclusion $[\varphi]$, la méthode basée sur la relation (2.7) peut être montrée convergente. Cette propriété est formalisée dans la proposition 8. Avant de prouver ce résultat de convergence, on définit la taille d'une partition comme étant le plus grand des diamètres de ses éléments. Formellement avec \mathcal{X} est une partition de E , on note par $\Delta(\mathcal{X})$ le nombre positif suivant :

$$\Delta(\mathcal{X}) = \max_{X \in \mathcal{X}} \Delta(X).$$

Proposition 8 (Convergence). Soient μ une mesure sur un compact E , φ une fonction réelle définie sur E et $\{\mathcal{X}_k\}_{k \geq 0}$ une suite de partition de E . Si $[\varphi] = [\underline{\varphi}, \overline{\varphi}]$ est une fonction d'inclusion continue pour φ vérifiant $\forall x \in E, [\varphi](\{x\}) = \{\varphi(x)\}$ alors

$$\sum_{X \in \mathcal{X}_k} \mu(X) \underline{\varphi}(X) \rightarrow \langle \mu, \varphi \rangle \text{ quand } \Delta(\mathcal{X}_k) \downarrow 0,$$

et

$$\sum_{X \in \mathcal{X}_k} \mu(X) \overline{\varphi}(X) \rightarrow \langle \mu, \varphi \rangle \text{ quand } \Delta(\mathcal{X}_k) \downarrow 0.$$

Démonstration. Comme E est compact, l'ensemble $K(E)$ défini par $K(E) = \{X \subset E \mid X \text{ est compact}\}$ équipé de la topologie induite par la distance de Hausdorff est compact. Avec $[\varphi]$ continue, par le théorème de Heine-Cantor, $[\varphi]$ est uniformément continue. C'est à dire :

$$\forall \varepsilon > 0 \exists \delta > 0 \forall X, Y \subset E, d(X, Y) < \delta \Rightarrow d([\varphi](X), [\varphi](Y)) < \varepsilon.$$

En particulier, avec $Y = \{x\} \subset A$ on a

$$\forall \varepsilon > 0 \exists \delta > 0 \forall X \subset E, d(X, \{x\}) < \delta \Rightarrow d([\varphi](X), \varphi(\{x\})) < \varepsilon.$$

Par conséquent,

$$\forall \varepsilon > 0 \exists \delta > 0 \forall X \subset E, \Delta(X) < \delta \Rightarrow \Delta([\varphi](X)) < \varepsilon.$$

Comme $\Delta \mathcal{X}_k \downarrow 0$ quand $k \rightarrow \infty$, on en déduit

$$\forall \varepsilon > 0 \exists k_0 > 0 \forall k \geq k_0 \forall X \in \mathcal{X}_k, (\overline{\varphi}(X) - \underline{\varphi}(X)) < \varepsilon.$$

En additionnant, on obtient :

$$\forall \varepsilon > 0 \exists k_0 > 0 \forall k \geq k_0 \sum_{X \in \mathcal{X}_k} (\overline{\varphi}(X) - \underline{\varphi}(X)) < \varepsilon \mu(E).$$

Finalement, les deux bornes de l'encadrement donné dans la proposition 6 convergent vers la même valeur. \square

Avant de clore cette section, on donne ici quelques corollaires des propositions précédentes qui nous seront utiles dans le cas du transport optimal. L'idée est de remplacer, dans les hypothèses de la proposition 8, la partition par un recouvrement dont les intersections sont de mesures nulles.

Définition 9. Soient μ une mesure et E un sous-ensemble de \mathbb{R}^n . Une famille finie $\{X_i\}_i$ d'intervalles de \mathbb{R}^n est qualifiée de μ -pavage si les conditions suivantes sont réunies :

$$E = \bigcup_i X_i \text{ et } i \neq j \Rightarrow \mu(X_i \cap X_j) = 0$$

Corollaire 1. Soit $\varphi : E \rightarrow \mathbb{R}$ une fonction intégrable via à vis de la mesure de Lebesgue λ et $\{X_i\}_{i \in I}$ un λ -pavage de E , alors

$$\sum_{i \in I} \underline{\varphi}(X_i) \lambda(X_i) \leq \int_X \varphi(x) d\lambda(x) \leq \sum_{i \in I} \overline{\varphi}(X_i) \lambda(X_i). \quad (2.9)$$

Démonstration. Soit X_i un élément du pavage. Comme $[\varphi]$ est une fonction d'inclusion pour φ , on peut écrire

$$\forall x \in X_i, \underline{\varphi}(X_i) \leq \varphi(x) \leq \overline{\varphi}(X_i)$$

avec $[\underline{\varphi}(X_i), \overline{\varphi}(X_i)] = [\varphi](X_i)$. Par conséquent, on en déduit

$$\int_{X_i} \underline{\varphi}(X_i) d\lambda(x) \leq \int_{X_i} \varphi(x) d\lambda(x) \leq \int_{X_i} \overline{\varphi}(X_i) d\lambda(x).$$

De façon équivalente, on a

$$\underline{\varphi}(X_i) \int_{X_i} d\lambda(x) \leq \int_{X_i} \varphi(x) d\lambda(x) \leq \overline{\varphi}(X_i) \int_{X_i} d\lambda(x).$$

Comme $\{X_i\}_i$ est un λ -pavage, on conclut

$$\sum_{i \in I} \underline{\varphi}(X_i) \lambda(X_i) \leq \int_X \varphi(x) d\lambda(x) \leq \sum_{i \in I} \overline{\varphi}(X_i) \lambda(X_i).$$

□

Corollaire 2 (Convergence). Soient X un élément de $\mathbb{I}\mathbb{R}^n$, $[\varphi]$ une fonction d'inclusion continue pour φ vérifiant $\forall x \in X, [\varphi](\{x\}) = \{\varphi(x)\}$ et $\{\mathcal{X}_n\}_{n \in \mathbb{N}}$ une suite de λ -pavage de X . Si $\Delta(\mathcal{X}_n)$ tend vers 0 quand n tend vers ∞ alors les deux bornes de l'équation (2.9) convergent vers la même valeur.

Chapitre 3

Classification des fonctions lisses

3.1 Introduction

Dans ce chapitre, on développe des méthodes numériques capables de distinguer des fonctions lisses. La motivation et les applications en robotique de cette classification ont été rappelées dans la section 1.3.2. Ces méthodes numériques s'appuient en grande partie sur l'opérateur de Newton donné dans le chapitre 2. On entend par fonction lisse toute fonction infiniment différentiable et notera par $\mathcal{C}^\infty(X, Y)$ l'ensemble de ces fonctions lorsque l'espace source est X et l'espace but Y . Comme dans toute classification, il est indispensable de se donner une relation d'équivalence afin de tester l'égalité de deux objets. Dans le cas des fonctions lisses, plusieurs équivalences ont été proposées dont la suivante que nous retiendrons dans un premier temps.

Définition 10 (Equivalence de fonctions lisses). Soient $f_1, f_2 : X \rightarrow Y$ deux fonctions lisses. La fonction f_1 est *équivalente* à f_2 s'il existe deux difféomorphismes $g : X \rightarrow X$ et $h : Y \rightarrow Y$ telle que $h \circ f_1 = f_2 \circ g$, i.e. telle que le diagramme suivant commute :

$$\begin{array}{ccc} X & \xrightarrow{f_1} & Y \\ \downarrow g & & \downarrow h \\ X & \xrightarrow{f_2} & Y \end{array}$$

On écrira $f_1 \sim f_2$ pour noter que f_1 et f_2 sont équivalentes.

On démontre facilement que la relation \sim est d'équivalence.

Exemple 7. Considérons les deux fonctions réelles suivantes : $f_1 : x \mapsto x^2$ et $f_2 : x \mapsto ax^2 + bx + c$ avec $a, b, c \in \mathbb{R}, a \neq 0$. On a $f_1 \sim f_2$, en effet, $h^{-1} \circ f_1 \circ g = f_2$ avec g et h^{-1} les deux difféomorphismes suivants $g : x \mapsto x + \frac{b}{2a}$ et $h^{-1} : y \mapsto ay - \frac{b^2}{4a} + c$.

Le théorème suivant, bien qu'évident, est l'essence même de la théorie des singularités.

Théorème 2. Soient f_1 et f_2 deux fonctions équivalentes avec h et g deux difféomorphismes vérifiant $f_2 \circ g = h \circ f_1$. Si x_1, x_2 sont deux éléments de X vérifiant $g(x_1) = x_2$ alors $\text{rank } df_1(x_1) = \text{rank } df_2(x_2)$.

Autrement dit, chaque singularité survit à la suite d'un changement de variables lisses sur l'espace source X ou sur l'espace but Y . Ces singularités vont donc intervenir dans la plupart des invariants cherchant à classifier les fonctions. On peut même étendre cette propriété locale à l'ensemble des singularités et leur relation de voisinage. On en déduit le corollaire suivant :

Corollaire 3. *Soient f_1 et f_2 deux fonctions lisses, si $f_1 \sim f_2$ alors leurs lieux singuliers S_{f_1} et S_{f_2} sont homéomorphes. De plus, leurs valeurs singulières, $f_1(S_{f_1})$ et $f_2(S_{f_2})$, sont aussi homéomorphes.*

Ce corollaire nous indique que les topologies des lieux singuliers et des valeurs singulières forment des invariants. L'exemple suivant nous donne une application directe de ce corollaire.

Exemple 8. Soient f_1 et f_2 étant définies par

$$\begin{aligned} f_1 : [-3, 3] &\rightarrow \mathbb{R} \\ x &\mapsto x + 1, \end{aligned}$$

et

$$\begin{aligned} f_2 : [-3, 3] &\rightarrow \mathbb{R} \\ x &\mapsto x^2 + 1. \end{aligned}$$

En accord avec le corollaire 3, les singularités sont préservées sous l'action des couples de difféomorphismes sur X et Y définie par

$$\begin{aligned} (g, h) : \mathcal{C}^\infty(X, Y) &\rightarrow \mathcal{C}^\infty(X, Y) \\ f &\mapsto h^{-1} \circ f \circ g \end{aligned} \quad (3.1)$$

Par conséquent, les fonctions f_1 et f_2 ne peuvent être équivalentes puisque f_2 admet une unique singularité en $x = 0$, i.e. $S_{f_2} = \{x = 0\}$, alors que $S_{f_1} = \emptyset$.

L'étude *topologique* des espaces singuliers est alors primordiale lors de la classification de fonctions lisses. Comme dit précédemment, les singularités survivent lors de changements de variables lisses. On peut aussi être tenté de distinguer les singularités les unes des autres avec le rang comme indiqué par le théorème 2. C'est ce que fait la classification de Thom-Boardman.

Définition 11 (Classes de Thom-Boardman). Soit i un entier, on définit l'ensemble $\Sigma^i(f)$ par

$$\Sigma^i(f) = \{x \in X \mid \ker df_x = i\}.$$

On dira que le point $x \in X$ est de classe Σ^i dès que x appartient à $\Sigma^i(f)$.

Remarque. On notera que, dans la définition 11, on utilise la dimension du noyau de la différentielle alors que dans le théorème 2, c'était le rang. Le théorème du rang [37] nous indique que cette différence n'en est pas vraiment une.

Par conséquent, pour une fonction lisse f donnée, son lieu singulier S_f se partitionne avec les différents sous-ensembles $\Sigma^i(f)$. Formellement, on a donc

$$S_f = \bigcup_{i \in \mathbb{N}} \Sigma^i \text{ et } i \neq j \Rightarrow \Sigma^i \cap \Sigma^j = \emptyset. \quad (3.2)$$

La réunion présente dans l'équation (3.2) est une réunion finie du fait que $\dim X$ et $\dim Y$ sont supposées finies.

Exemple 9. Dans le cas où X et Y sont de dimension 1, l'ensemble $\Sigma^2(f)$ est vide simplement parce que $2 > 1$ (de même que tous les autres $\Sigma^k(f) = \emptyset$ dès que $k \geq 2$). L'ensemble source X se décompose en $X = \Sigma^0(f) \cup \Sigma^1(f)$ avec $\Sigma^0(f) \cap \Sigma^1(f) = \emptyset$, puisque $\Sigma^1(f)$ est l'ensemble des points singuliers alors que $\Sigma^0(f)$ sont les points réguliers.

Ces types de singularités nous renseignent un peu plus sur une fonction f que simplement l'étude du lieu singulier. La topologie de chaque Σ^i sera donc cruciale dans une classification. On peut continuer à chercher plus d'informations sur f en regardant comment se comporte f quand elle est restreinte à Σ^i . Il se peut très bien que $f|_{\Sigma^i}$ ait encore des singularités. De façon moins formelle, on cherche en quelque sorte à savoir si f a des *singularités de singularités*. On définit donc *récurivement* les sous ensembles Σ^I de X où I est une suite finie de nombres entiers, via :

Définition 12 (Classes étendues de Thom-Boardman). Soit $\Sigma^I(f) = \Sigma^{i_1, i_2, \dots, i_k}(f)$ inclus dans X , alors l'ensemble $\Sigma^{i_1, i_2, \dots, i_{k+1}}(f)$ est défini par

$$\Sigma^{i_1, i_2, \dots, i_{k+1}}(f) = \Sigma^{i_{k+1}}(f|_{\Sigma^I(f)}).$$

La topologie des $\Sigma^I(f)$, avec I qui parcourt l'ensemble des mots finis, est une source d'information encore plus riche sur la fonction f que pouvait l'être la topologie du lieu singulier ou bien celle des valeurs singulières. De plus, les relations de voisinage (voire de recollement) entre les différents $\Sigma^I(f)$ sont très importantes car elles aussi survivent lors de l'action de changements sur le source et sur le but. La topologie des Σ^I et leur relation de voisinage peuvent être compilées dans un objet mathématique appelé *stratification* [3]. Cette notion de stratification sera rappelée dans la section 3.3 qui s'intéresse aux invariants globaux.

On pourra consulter la section 3.2.2 où on détaille le cas particulier où $\dim X = \dim Y = 2$ pour lequel Σ^1 et $\Sigma^{1,1}$ jouent des rôles particuliers. Le lecteur pourra aussi consulter le livre de V.I. Arnold [1] pour les propriétés de ces classes de Thom-Boardman. Par exemple, on prouve facilement que la suite finie d'entier I doit être nécessairement décroissante pour que $\Sigma^I(f)$ soit non vide. De plus, si I est un préfixe¹ de J , alors $\Sigma^I(f) \supset \Sigma^J(f)$.

Malheureusement, ces espaces de singularités Σ^i (ou plus généralement Σ^I) peuvent avoir des topologies extrêmement complexes. Il existe néanmoins des résultats positifs du côté de la théorie de la mesure comme par exemple le théorème de Sard [39] prouvé en 1942.

Théorème 3 (Sard). Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ une application lisse alors l'ensemble des valeurs critiques de f est de mesure de Lebesgue nulle, i.e. $\lambda(f(S_f)) = 0$.

Plus récemment et suite aux travaux de Hassler Whitney, John Mather et René Thom, des listes finies de singularités ont pu être dégagées pour des applications dites stables. Une application lisse est qualifiée de stable si toutes les applications dans un voisinage² suffisamment petit lui sont équivalentes. Il se trouve que l'ensemble des applications stables

1. Au sens de la théorie des langages [38].

2. Ce voisinage est correctement défini en mettant par exemple la topologie de Whitney sur l'espace $C^\infty(X, Y)$ [40].

forme une partie résiduelle de l'espace des fonctions lisses pour certains couples de dimension de l'espace but et de l'espace source. Par conséquent et pour ces dimensions, les singularités de ces applications stables sont très appréciées.

Une partie de ces résultats de classification se retrouve grâce au théorème de transversalité de René Thom :

Théorème 4 (Théorème de transversalité). *Soient X et Y des variétés lisses. Si W une sous-variété de l'espace des jets $J^k(X, Y)$, alors l'ensemble des applications lisses vérifiant $j^k f \bar{\cap} W$ forme une partie résiduelle de $C^\infty(X, Y)$ (au sens de la topologie de Whitney sur C^∞).*

Remarque. Le symbole $\bar{\cap}$ concrétise la notion de transversalité entre une fonction et une variété. Avec f une fonction, on note par $j^k f$ son jet dans l'espace $J^k(X, Y)$. C'est en quelque sorte une façon de géométriser les coefficients de Taylor jusqu'à l'ordre k en les réunissant dans un même objet. La notion d'espace résiduel peut être comparée au concept probabiliste "presque par partout". Malheureusement, la taille de l'espace des fonctions $C^\infty(X, Y)$ et sa topologie nous obligent à introduire cette définition d'ensemble résiduel. Techniquement, on retiendra qu'une partie d'un ensemble est résiduelle si elle contient une intersection dénombrable d'ouverts denses. Toutes ces définitions relatives à la théorie des catastrophes se trouvent dans le livre de Michel Demazure [40].

A première vue, le théorème de Thom paraît déconnecté de la notion de singularité. Il trouvera de nombreuses applications en théorie des singularités et donc en classification. Pour ce faire, il suffira par exemple de poser W , dans le théorème 4, comme étant la variété incluse dans J^k définie par df_x n'est pas de rang plein. Ce genre de technique sera illustré lors de la preuve de théorèmes de ce chapitre.

Faisons maintenant une dernière remarque d'ordre terminologique. Une façon de présenter les résultats s'est installée dans cette communauté. Tout d'abord, lorsqu'une fonction lisse f fait partie d'un ensemble résiduel, on dit que f est générique. Il est d'habitude³ de présenter une assertion comme "*L'ensemble des fonctions f qui vérifient la propriété P forme une partie résiduelle*" sous la forme "*Soit f une fonction générique, alors f vérifie la propriété P* ".

Organisation du chapitre

Etant donné un point x de X et une fonction lisse f , le fait que x appartienne, oui ou non, à une classe Σ^I est une propriété locale. En effet, il suffit de connaître f sur n'importe quel voisinage de x pour être en mesure de le savoir. La section 3.2, nommée *invariants locaux*, décrit des méthodes numériques pour encadrer les ensembles Σ^I pour différentes paires de dimensions d'espace source et but.

La topologie des différents Σ^I est quant à elle une information plus globale. En effet, il faut connaître f sur tout X pour pouvoir calculer les topologies des ensembles Σ^I . On verra qu'elle n'est pas toujours suffisante pour distinguer deux fonctions non équivalentes. La section 3.3, nommée *invariants globaux*, propose des méthodes pour construire des invariants qui ne sont pas uniquement basées de propriétés locales de la fonction f .

3. A mon goût, sans doute mauvaise.

3.2 Invariants locaux - Encadrement des singularités

Lorsque la dimension de l'espace but est 1, i.e. lorsque l'on s'intéresse aux fonctions réelles lisses, on retrouve des résultats proches de la théorie de Morse. Ce cas particulier sera rappelé dans la section 3.2.1 et servira d'introduction aux méthodes développées par la suite. Le couple $(\dim X, \dim Y) = (2, 2)$ a été très étudié par Hassler Whitney [4] et est traité dans la section 3.2.2. Finalement, la section 3.2.3 s'intéresse au couple $(\dim X, \dim Y) = (3, 3)$. Dans chacune de ces sous-sections, on rappellera les résultats généraux théoriques en faisant références aux ouvrages classiques comme [2], puis on donnera des systèmes d'équations caractérisant les ensembles Σ^I qui apparaissent génériquement. Les solutions de ces équations pourront finalement être encadrées via les techniques données dans la section 2.2.

3.2.1 Cas des fonctions de la droite sur la droite

Nous commençons dans cette section avec le cas des applications de \mathbb{R} dans \mathbb{R} . Ce cas simple et souvent familier nous servira de support aux situations plus compliquées en dimensions supérieures. Pour ces fonctions, la chute de rang de l'application différentielle signifie simplement que la différentielle s'annule. L'ensemble but étant de dimension 1, les singularités de f sont aussi classiquement appelées *points critiques*. Avec un système de coordonnées, ces points critiques forment les solutions du système carré⁴ suivant :

$$df_x = 0. \tag{3.3}$$

Dans ce cas, les applications stables vérifient la propriété suivante : tous leurs points critiques sont non dégénérés. Un point critique est dit non dégénéré si la hessienne de f ce point critique x est non dégénérée. Dans le cas des fonctions des fonctions de la droite sur la droite, cela signifie simplement que la dérivée seconde ne s'annule pas au endroit où la dérivée première est nulle.

Exemple 10. Considérons la fonction $x \mapsto x^2 - x$ avec comme espace source l'intervalle réel $[-2; 2]$ et comme espace but \mathbb{R} . L'ensemble $S_f = \{x \in [-2; 2] \mid f'(x) = 2x - 1 = 0\} = \{\frac{1}{2}\}$. Dans ce cas, on a $\Sigma^1 = S_f$ et $\Sigma^0 = [-2, 2] - \{\frac{1}{2}\}$. Cet exemple est tiré du manuscrit de thèse de Romain Benoit [41] où il est traité en détail.

Dans le cas des fonctions définies entre espaces de dimension 1, le théorème de transversalité est très proche d'un autre résultat de la théorie de Morse. Génériquement parlant, au voisinage d'une singularité, une fonction f est équivalente à la fonction carré $x \mapsto x^2$.

Par conséquent, dans le cas générique, le système (3.3) pourra être résolu par une méthode de Newton par intervalles puisque que la fonction $x \mapsto df_x$ est inversible au voisinage de ses zéros. On pourra consulter le livre de Michel Demazure [40] qui contient, en partie, une très bonne introduction à la notion de fonction de Morse.

4. Un système d'équation est dit carré s'il a autant d'équations que d'inconnues.

3.2.2 Cas des fonctions du plan sur le plan

Introduction

Pour les applications lisses du plan X et vers le plan Y , les points singuliers d'une application f sont les points x de X pour lesquels le rang de la différentielle df_x est inférieur à 2.

Localement, il n'existe que deux types génériques de singularités : les points plis et les fronces. En accord avec la classification de Thom-Boardman, les points plis forment l'ensemble $\Sigma^{1,0}$ alors que les fronces sont les éléments de $\Sigma^{1,1}$. En d'autres termes, si (x_0, y_0) est un point singulier d'une application générique f , alors il existe un voisinage de (x_0, y_0) telle que f soit localement équivalente à l'une des applications suivantes : $(x, y) \mapsto (x, y^2)$ ou bien $(x, y) \mapsto (x, xy - y^3)$. Les figures 3.1 et 3.2 illustrent chacune de ces formes normales.

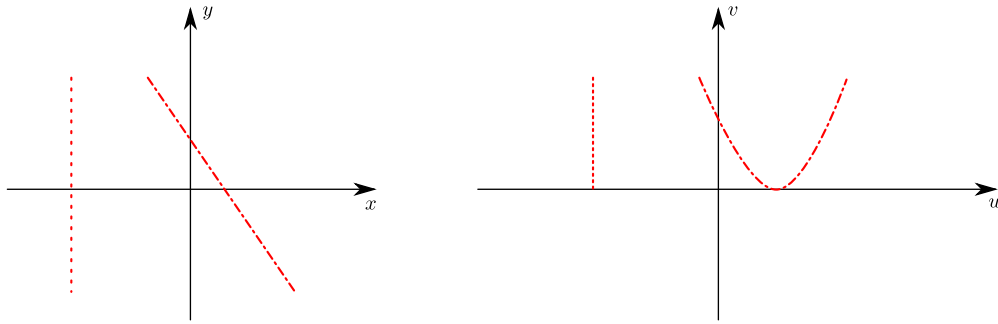


FIGURE 3.1 – Forme normale d'un point pli : $(x, y) \mapsto (u, v) = (x, y^2)$.

Sur la figure 3.1 est représentée une fonction de \mathbb{R}^2 dans \mathbb{R}^2 . Comme il n'est pas si courant de dessiner des fonctions du plan sur le plan, on donne ici une interprétation géométrique. Sur la gauche de la figure 3.1, l'espace source et deux droites avec des motifs différents comme s'il s'agissait de trajectoires $t \mapsto \gamma(t) = (x(t), y(t))$ sont représentés. A droite, on a l'image de ces trajectoires par f , c'est à dire que l'on dessine $t \mapsto (u(t), v(t)) = f(\gamma(t))$. Comme la fonction considérée est $(u, v) = (x, y^2)$, la droite d'équation $x = -1$, est envoyée par f sur la demi-droite d'équation $u = -1$ et $v \geq 0$. La seconde droite, oblique sur la gauche de la figure 3.1, est envoyée par f sur la parabole dans l'espace des (u, v) . On peut, très grossièrement, imaginer le comportement de cette fonction comme si elle pliait une feuille (x, y) le long de l'axe $y = 0$.

La figure 3.2 quant à elle présente une fronce. Sur la gauche, la parabole rouge en pointillés décrit l'ensemble singulier S_f de f . Tous les points de cette parabole sont des points de Σ^1 . Cet ensemble Σ^1 se partitionne en deux sous-ensembles $\Sigma^{1,1}$ et $\Sigma^{1,0}$. L'ensemble $\Sigma^{1,1}$ est un singleton qui ne contient que l'origine, ce point est qualifié de fronce. C'est la seule singularité de la fonction $f|_{\Sigma^1}$. Tous les autres points de la parabole forment $\Sigma^{1,0}$ et ont un comportement local semblable à un point pli.

Les valeurs singulières sont représentées à droite en rouge.

D'un point de vue moins rigide, on peut dire que localement, une fonction générique avec singularité ressemble à l'une des figures 3.3 ou 3.4.

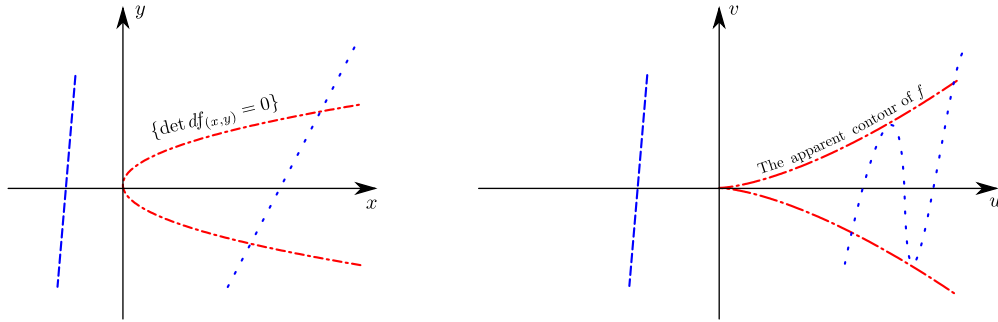


FIGURE 3.2 – Forme normale d’une fronce, $f : (x, y) \mapsto (u, v) = (x, xy - y^3)$.

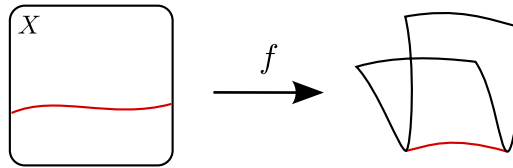


FIGURE 3.3 – Une courbe de point pli.

Ces deux formes normales ont été identifiées par Hassler Whitney et se retrouvent comme cas particulier du théorème de transversalité de René Thom. Concrètement, on peut énoncer :

Théorème 5 (Propriétés des applications génériques). *Une application lisse générique f de \mathbb{R}^2 vers \mathbb{R}^2 vérifie la propriété suivante : l’ensemble des points singuliers, noté $S_f = \{x \in X \mid \det df_x = 0\}$, est une courbe régulière de \mathbb{R}^2 . De plus, la courbe S_f est composée de points pli et de fronces ; le sous-ensemble des fronces est discret.*

Démonstration. La preuve se trouve dans [40]. La preuve donnée dans [6] se veut plus proche des de la méthode effective capable d’encadrer S_f et les fronces. \square

Mise en équation

Comme présenté précédemment, de façon générique, il n’existe qu’un nombre fini de cas. L’espace singulier $S_f (= \Sigma^1)$ se décompose en $\Sigma^{1,0}$ et $\Sigma^{1,1}$. Comme l’équation qui caractérise S_f est simplement $\det df_x = 0$, on pourra, dans le cas générique, calculer un encadrement de Σ^1 avec une méthode comme SIVIA en résolvant l’équation scalaire $\det df_x = 0$. La plupart des points seront des points de $\Sigma^{1,0}$ sauf un nombre discret d’entre eux qui seront des éléments de $\Sigma^{1,1}$, les fronces. Dans cette section, nous nous focalisons sur les fronces. On propose une méthode afin de les encadrer rigoureusement et de les dénombrer. Une formulation algébrique des fronces équivalente à la définition classique est donnée. Soit f une application générique et p un point singulier, i.e. $p \in S_f = \{x \in X \mid \det df_x = 0\}$. A la vue de la discussion précédente, soit le noyau de df_p est inclus dans l’espace tangent $T_p S_f$, soit non. Algébriquement, on se retrouve dans l’une des deux situations suivantes :

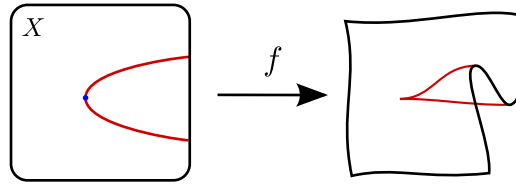


FIGURE 3.4 – Une courbe de point pli avec une fronce.

$$\text{i) } T_p S_f \oplus \ker df_p = T_p X,$$

$$\text{ii) } T_p S_f = \ker df_p.$$

où $T_p Z$ dénote l'espace tangent à x de Z .

On remarque que si p est une singularité vérifiant la condition i, alors p est appelé un point pli, c'est à dire un élément de $\Sigma^{1,0}$. Le théorème de Whitney [42] énonce qu'il existe un voisinage V de p tel que $f|V$ est équivalent à l'application $(x, y) \mapsto (x, y^2)$.

Le cas ii correspond à une fronce, i.e. à un élément de $\Sigma^{1,1}$, et est plus compliqué. Choisissons un champ de vecteurs ξ ne s'annulant pas, tel que $\xi_x \in \ker df_x$. Par hypothèse, ξ_p est tangent à S_f lorsque p est une fronce. La nature de la singularité en p dépend de l'ordre de contact que ξ a avec S_f en p . Soit k une application réelle lisse sur X telle que $k|S_f = 0$ et $dk_p \neq 0$ et considérons la fonction $w : s \mapsto dk_s \cdot \xi_s$. Dans le cas ii, par hypothèse, on a donc $w(p) = 0$. Notons que l'ordre de ce zéro ne dépend ni du choix de ξ ni de k . Finalement, un point p est appelé fronce simple si ce zéro est simple. Le second théorème de Whitney énonce que si p est une fronce simple alors il existe un voisinage V de p telle que $f|V$ est équivalente à l'application

$$(x, y) \mapsto (x, xy + y^3).$$

La contribution principale de cette section est la proposition 9. Cette proposition donne un système carré d'équations qui caractérise les points de $\Sigma^{1,1}$. L'intérêt principal de cette condition sous forme de systèmes d'équations carré est de nous permettre de le résoudre numériquement avec une méthode de Newton par intervalles. Finalement, pour une application lisse générique donnée, on pourra encadrer les fronces et les dénombrer.

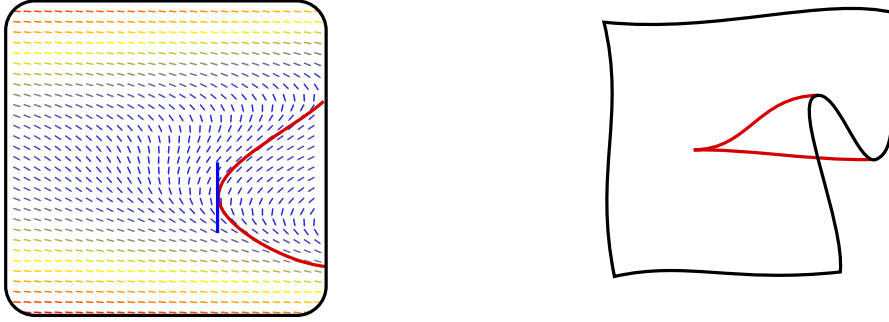
Proposition 9. Soit f une fonction lisse générique de \mathbb{R}^2 dans \mathbb{R}^2 , on dénote par c l'application suivante :

$$\begin{aligned} c & : X \rightarrow \mathbb{R}^2 \\ p & \mapsto df_p \xi_p, \end{aligned} \tag{3.4}$$

où ξ est le champ de vecteurs défini par $\xi_p = \begin{pmatrix} \partial_2 \det df_p \\ -\partial_1 \det df_p \end{pmatrix}$.

Si $c(p) = 0$ et dc_p est inversible alors p est une fronce simple. De plus cette condition est localement nécessaire.

Démonstration. Prouvons d'abord la suffisance : $\xi_p \neq 0$ et $c(p) = 0$ signifie que $\ker df_p \neq \{0\}$, c'est à dire que $p \in S$. Comme $\mathbb{R}\xi_p = T_p S_f$, l'équation $c(p) = 0$ est équivalente à $T_p S = \ker df_p$.

FIGURE 3.5 – Le champs de vecteurs ξ est tangent à S_f .

La locale nécessité vient de la généricité de f . Soit $p \in S_f$ une fronce, alors génériquement l'application lisse $p \mapsto d \det df_p$ ne s'annule pas en p . Par conséquent, il existe un voisinage V de p telle que $\forall p' \in V, \xi_{p'} \neq 0$. Restreint à ce voisinage, $df_p \xi_p = 0$ implique que $\ker df_p = T_p S_f (= \mathbb{R} \xi_p)$. Comme le point fronce p est supposé simple, on peut en déduire que dc est inversible en p . \square

Exemple 11. Considérons dans cet exemple l'application lisse $f : (x, y) \mapsto (u, v) = (x, xy - y^3)$. L'ensemble singulier de la fonction f est définie par $\{(x, y) \in \mathbb{R}^2 \mid \det df_{x,y} = 0\}$, i.e. comme $df_{x,y} = \begin{pmatrix} 1 & 0 \\ y & x - 3y^2 \end{pmatrix}$, on obtient $S_f = \{(x, y) \in \mathbb{R}^2 \mid x - 3y^2 = 0\}$. Ce lieu singulier est la courbe régulière représentée en pointillé rouge et à gauche sur la figure 3.2. Le champ de vecteurs ξ est défini par $\xi_{x,y} = -6y\partial_x - \partial_y$, ou encore :

$$\xi_{x,y} = \begin{pmatrix} -6y \\ -1 \end{pmatrix}.$$

Dans ce cas, l'application c de la proposition 9 s'écrit

$$c : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} -6y \\ -3y^2 - x \end{pmatrix}.$$

On en déduit que la fonction f admet une seule fronce en $(0, 0)$.

3.2.3 Cas des fonctions de l'espace dans l'espace

Dans cette section, on s'intéresse aux fonctions de \mathbb{R}^3 dans \mathbb{R}^3 . Génériquement, les singularités sont des types suivants Σ^1 , $\Sigma^{1,1}$, $\Sigma^{1,0}$, $\Sigma^{1,1,1}$ et $\Sigma^{1,1,0}$. Les autres Σ^I étant vides pour des raisons de dimension des espaces de matrices des rangs correspondants.

Ces espaces singuliers non vides sont inclus les uns dans les autres selon l'arbre représenté sur la figure 3.6. Sur cette figure, la relation d'inclusion entre deux ensembles A et B est représentée par $B \rightarrow A$ pour indiquer que $A \subset B$.

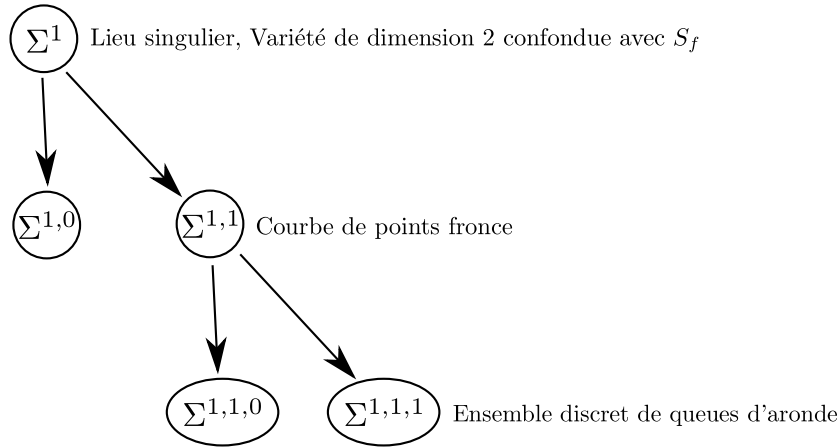


FIGURE 3.6 – Relation d'inclusion entre les différents Σ^I dans le cas générique.

En effet, génériquement, l'ensemble Σ^1 est l'ensemble des points singuliers et forme une surface incluse dans l'espace tridimensionnel source. Les points de $\Sigma^{1,0}$ forment aussi une surface. Ses points sont des sortes de points plis autour desquels deux volumes séparés par $\Sigma^{1,0}$ vont être envoyés sur le même volume comme le ferait l'application $(x, y, z) \mapsto (x, y, z^2)$. Le reste de Σ^1 , c'est à dire $\Sigma^{1,1}$, est courbe. Cette courbe est composée de froinces formant $\Sigma^{1,1,0}$ accompagné de points, encore plus spéciaux, appelés *queues d'aronde*. L'ensemble des queues d'aronde $\Sigma^{1,1,1}$ est génériquement discret.

Mise en équation

Dans cette section, nous proposons trois systèmes d'équations capable de caractériser respectivement Σ^1 , $\Sigma^{1,1}$ et $\Sigma^{1,1,1}$. Tout d'abord, génériquement l'ensemble Σ^1 est confondu avec S_f , par conséquent, on en déduit :

Théorème 6. *Génériquement, l'ensemble Σ^1 est composé de l'ensemble des solutions de l'équation*

$$\det df_x = 0 \quad (3.5)$$

L'équation 3.5 est une équation scalaire définie sur une espace de dimension 3. Cette surface, solution de 3.5, pourra être encadrée avec une méthode type SIVIA.

Concernant l'ensemble $\Sigma^{1,1}$, il s'agit de trouver les points de Σ^1 pour lesquels la fonction $f|_{\Sigma^1}$ voit son rang chuter de 1. D'un point de vue géométrique, les points $\Sigma^{1,1}$

sont les points x pour lesquels il existe un vecteur de l'espace tangent $T_x\Sigma^1$ aussi dans le noyau de la différentielle de f . Un peu à la manière de la proposition 9, pour trouver les points x de $\Sigma^{1,1}$, on caractérise à la fois le champ de droites noyau de df_x et le champ de plans tangents à Σ^1 .

En supposant que l'espace source soit une partie ouverte de \mathbb{R}^3 , on peut exploiter la structure euclidienne et remarquer qu'un vecteur du noyau de df_x est simultanément orthogonal aux trois vecteurs lignes de df_x . Comme f est supposée de rang 2 en $x \in \Sigma^1$, une de ces trois lignes est une combinaison linéaire des deux autres. Encore à cause de la dimension 3, le noyau est donc forcément engendré par le produit vectoriel de ces deux dernières lignes. Nommons ce vecteur qui engendre $\ker df_x$ par w et regardons, maintenant, l'espace tangent à Σ^1 . A cause de la structure euclidienne, le covecteur $\nabla \det df_x$ est normal au plan tangent $T_x\Sigma^1$. On peut donc écrire :

$$T_x\Sigma^1 = \{\xi \in \mathbb{R}^3 \mid (\nabla \det df_x) \cdot (\xi - x) = 0\}$$

On en déduit le théorème suivant

Théorème 7. *L'ensemble $\Sigma^{1,1}$ est composé de l'ensemble des solutions de*

$$\begin{cases} \det df_x = 0, \\ (\nabla \det df_x) \cdot w_x = 0, \end{cases} \quad (3.6)$$

avec w_x l'un des vecteurs non nuls de la famille

$$\{w^{ij} \in \mathbb{R}^3 \mid w^{ij} = l_x^i \wedge l_x^j \text{ et } \{i, j\} \subset \{1, 2, 3\}\}$$

où l_x^i désigne la i ème ligne de df_x .

En ce qui concerne les points de $\Sigma^{1,1,1}$, la mise en équation de ces points nommés *queue d'aronde* est donnée par le théorème suivant :

Théorème 8. *L'ensemble $\Sigma^{1,1,1}$ est composé de l'ensemble des solutions de*

$$\begin{cases} \det df_x = 0, \\ (\nabla \det df_x) \cdot w_x = 0, \\ \nabla((\nabla \det df_x) \cdot w_x) \cdot w_x = 0, \end{cases} \quad (3.7)$$

avec w_x est l'un des vecteurs non nuls de la famille

$$\{w \in \mathbb{R}^3 \mid w = l_i \wedge l_j \text{ et } \{i, j\} \subset \{1, 2, 3\}\}$$

où l_i désigne la i ème ligne de df_x .

La preuve se trouve dans le manuscrit de thèse [41]. Les idées sont globalement les mêmes que pour les autres cas. On exploite la structure euclidienne pour caractériser l'espace tangent à $\Sigma^{1,1}$ et on cherche w engendrant le noyau de df_x et élément de $T_x\Sigma^{1,1}$. Le système d'équations (3.7) d'inconnue x est carré et inversible au voisinage de ses solutions. La méthode de Newton par intervalles peut donc idéalement montrer l'existence et l'unicité d'une queue d'aronde dans un pavé tridimensionnel donné. Un exemple est traité dans le manuscrit de thèse [41]. On y trouvera aussi les temps de calcul de la méthode de Newton pour trouver les zéros de la fonction définie par le système carré (3.7) sur quelques exemples.

3.3 Invariants globaux

Etant donné un point x de X , il est possible de décider si x est élément de Σ^f en étudiant f sur n'importe lequel des voisinages de x . Tous les résultats effectifs de la section précédente 3.2, qui cherchaient à encadrer les ensembles de Thom-Boardman Σ^f ne s'intéressent qu'aux propriétés locales de f . Dans cette section, l'adjectif *global* est utilisé en opposition à l'adjectif *local*. La topologie des ensembles Σ^f est quant à elle une information globale. Si on y ajoute, les relations de voisinages entre les différents Σ^f , on obtient une stratification de l'espace source [3]. Malheureusement, ce n'est pas toujours suffisant pour distinguer deux fonctions. En effet, si deux points singuliers ont la même image par f , alors toute conjugaison de f aura nécessairement deux points singuliers ayant la même image. Cette information n'est pas contenue dans la topologie des ensembles de Thom-Boardman et ce genre de propriété découle de la proposition suivante :

Proposition 10. *Soient f_1 et f_2 deux fonctions équivalentes de telle sorte que le diagramme suivant commute :*

$$\begin{array}{ccc} X & \xrightarrow{f_1} & Y \\ \downarrow g & & \downarrow h \\ X & \xrightarrow{f_2} & Y \end{array}$$

Si y_1 et y_2 sont deux éléments de Y vérifiant $y_2 = h(y_1)$ alors les ensembles $f_1^{-1}(y_1)$ et $f_2^{-1}(y_2)$ sont homéomorphes.

Dans cette section, on présente des invariants globaux. Tout d'abord, on notera l'existence de la notion de l'espace des multi-jets pour prendre en compte les paires de points ayant la même image. Dans ce contexte, le théorème de transversalité a été généralisé pour prendre en compte simultanément plusieurs points dans l'espace source. Il se nomme le théorème de transversalité multijet.

Théorème 9 (Théorème de transversalité multi-jet). *Si X et Y sont des variétés lisses et W une sous variétés de l'espace des multi-jets $J_{(n)}^k(X, Y)$ alors l'ensemble des applications lisses $T_W = \{f \in C^\infty(X, Y) \mid j_{(n)}^k f \pitchfork W\}$ est un sous-ensemble résiduel de $C^\infty(X, Y)$.*

Nous le mettrons à profit dans les différentes sections qui suivent. Afin de préparer le lecteur aux outils déployés dans ces approches, la section 3.3.1 traite du cas, sans doute plus commun, des fonctions de la droite dans la droite.

La section 3.3.2 présente un algorithme pour calculer la topologie du lieu singulier d'une application du plan sur le plan (aussi appelée le contour apparent). La section 3.3.3 donne une méthodologie pour construire ce que nous avons appelé le *portrait d'une fonction*, sorte de photographie combinatoire de la fonction étudiée.

3.3.1 Exemple introductif en dimension 1

Ce cas est souvent familier car enseigné au lycée. En effet, la plupart des élèves apprennent à construire ce qui est appelé le tableau de variation d'une fonction. Cette construction s'appuie principalement sur le lemme suivant :

Lemme 1. Soit f une fonction lisse de $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ et X un intervalle de \mathbb{R} . Si df_x est non nulle sur X alors f est injective sur X .

Ce résultat énonce qu'il est possible de décider d'une propriété globale de la fonction f (son injectivité) uniquement en s'appuyant sur un continuum de propriétés locales. Il est intéressant de remarquer que cette propriété n'est plus vraie par exemple dans le cas des fonctions de $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^2)$. Il suffira de regarder la courbe $\gamma : t \mapsto (\cos(t), \sin(t))$ pour s'en convaincre. Autour de ces questions, on peut aussi faire mention de la conjecture jacobienne [43] même si elle s'intéresse plus au cas des corps algébriquement clos.

Comme indiqué dans l'introduction, si une fonction f admet des singularités ayant la même image, alors il en sera de même pour toute fonction équivalente à f . De plus, la position relative des singularités les unes par rapport aux autres ne changera pas. Cette dernière remarque découle du fait que tout difféomorphisme h ou g est forcément monotone (en dimension 1). En particulier, on remarque que les deux fonctions f_1 et f_2 de la figure 3.7, ont des lieux singuliers homéomorphes, i.e. $S_{f_1} \simeq S_{f_2}$ alors qu'elles ne sont pas équivalentes. En effet, la fonction f_1 admet deux optimums globaux qui sont situés

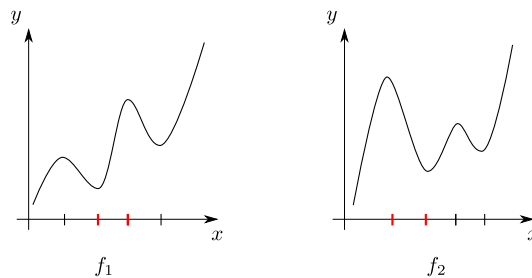


FIGURE 3.7 – Deux fonctions non équivalentes ayant le même lieu singulier.

entre deux autres de ses singularités alors que pour la fonction f_2 , ses optimums globaux sont situés à une extrémité de son espace source. Par conséquent, pour connaître (et décrire) une fonction lisse f en espace de dimension 1, on propose la procédure suivante.

- On commence par trouver ses singularités, i.e. S_f .
- Puis, on partitionne l'espace source le long de $f^{-1}(f(S_f))$ afin d'extraire les positions relatives des singularités dans le source et leur position relative dans l'espace but.
- L'étape suivante consiste à coder, dans un graphe⁵, les relations de voisinage entre les singularités,
- finalement, via la fonction f , on stocke dans une fonction discrète définie sur l'ensemble des points créés à partir de X et à valeur dans ceux créés à partir de Y .

Afin d'illustrer cette procédure, étudions attentivement la fonction donnée dans l'exemple 12.

5. Graphes de la théorie des graphes à ne pas confondre avec la notion de graphe de fonctions.

Exemple 12. Considérons la fonction f représentée par la figure 3.8. La première étape

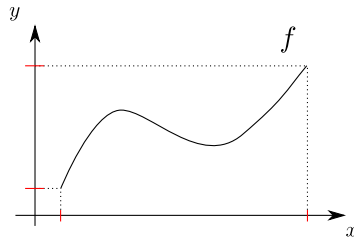


FIGURE 3.8 – Graphe de la fonction f .

consiste à trouver les points critiques de f . Géométriquement, ce sont les points où la tangente au graphe est horizontale. Ces points sont nommés ici a^2 et a^3 .

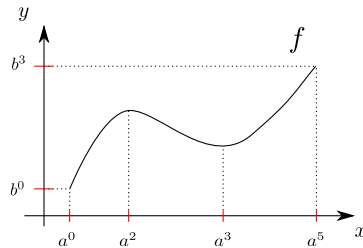


FIGURE 3.9 – Singularités S_f de f et leur image $f(S_f)$.

L'étape suivante consiste à calculer les images de ces points critiques, puis leurs réciproques. Cette étape est illustrée par la figure 3.10. La géométrie de la fonction f crée donc le point supplémentaire a_1 comme il est dans la réciproque de l'image de la singularité a_3 . On se retrouve donc avec l'ensemble $f^{-1}(f(S_f))$ composé des points a^1, a^2 et a^3 . Remarquons aussi que l'on a aussi ajouté les points a^0, a^5 extrêmes de l'espace source pour coder sa frontière, ainsi que leur image (b_0 et b_3). Dans l'étape suivante, on crée les deux graphes illustrés par les arcs présents sur la figure 3.11. Ces deux graphes codent respectivement les relations de voisinage entre les points a^i et b^j . Par construction et à cause de la dimension 1, si deux points a^i et a^j sont directement connectés, il ne peut y avoir d'autre point a^k intermédiaire entre a^i et a^j .

Finalement, les relations induites par f entre d'un côté les points de l'ensemble $A = \{a^0, a^1, a^2, a^3, a^4, a^5\}$ et de l'autre $B = \{b^0, b^1, b^2, b^3\}$ sont codées par une fonction discrète

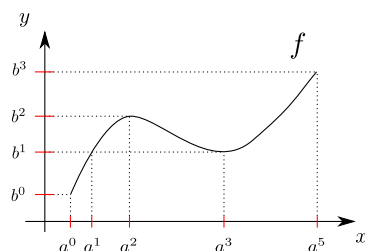
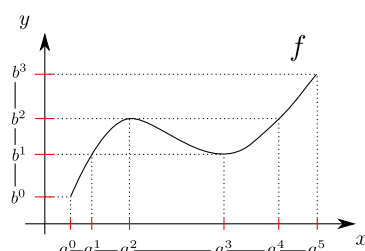
FIGURE 3.10 – Calcul de l'ensemble $f^{-1}(f(S_f))$.

FIGURE 3.11 – Graphes de voisinage entre les points de la contre image de l'image des points critiques.

$F : A \rightarrow B$. Dans le cas de l'exemple ici traité, cette fonction F est donnée par

$$\begin{array}{rcl}
 F & : & A \rightarrow B \\
 & & a^0 \mapsto b^0 \\
 & & a^1 \mapsto b^1 \\
 & & a^2 \mapsto b^2 \\
 & & a^3 \mapsto b^1 \\
 & & a^4 \mapsto b^2 \\
 & & a^5 \mapsto b^3
 \end{array}$$

La fonction F ainsi que les graphes de voisinage créés sont des informations très riches sur la fonction f étudiée. On dira que la fonction discrète F est le *portrait*⁶ de f . Toutes ces notions seront clairement définies et généralisées dans la section 3.3.3.

Remarque. Dans le cas des fonctions lisses du cercle dans le cercle, on pourra noter l'existence de la notion de *mots de Gauss*. Le lecteur intéressé par cette construction pourra consulter l'article de J.A. Moya-Pérez [44] ainsi que les références citées.

Concernant le calcul effectif de ce portrait, il pourra se faire grâce aux méthodes présentées dans la section 3.2.1. En effet, pour une fonction générique, ses points critiques

6. Ce choix de terminologie a été fait pour faire écho à la notion de portrait de phase présent dans la théorie des équations différentielles ordinaires.

sont non dégénérés. C'est à dire que l'équation scalaire d'inconnue $x \in \mathbb{R}$

$$df_x = 0. \quad (3.8)$$

forme un système carré localement inversible au voisinage de ses solutions. De plus, le théorème de transversalité, étendu au cas multi-jet, nous permet de conclure que pour une fonction générique f , deux points critiques différents n'ont pas la même image. Formellement, on écrit $\forall x_1, x_2 \in S_f, x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$. En résolvant le système carré $c = 0$, on trouvera les points de $f^{-1}(f(S_f))$ avec c la fonction suivante :

$$c : X \times X - \Delta(X) \rightarrow \mathbb{R}^2 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} df(x_1) \\ f(x_1) - f(x_2) \end{pmatrix}. \quad (3.9)$$

Dans l'équation (3.9), $\Delta(X)$ désigne le sous-ensemble de $X \times X$ défini par $\{(x_1, x_2) \in X \times X \mid x_1 = x_2\}$ et est souvent appelé la diagonale de $X \times X$.

L'hypothèse de non dégénérescence des points critiques de f ainsi que le fait que deux singularités aient des images différentes forment les conditions pour qu'une fonction f soit qualifiée de *fonction de Morse*. En conclusion, pour les fonctions de Morse (qui forment un ensemble résiduel), on pourra encadrer aussi finement que nécessaire leurs points critiques ainsi que leur position relative via la méthode de Newton par intervalles en résolvant (3.8) et (3.9). On pourra donc numériquement calculer leur portrait.

Remarque. En ce qui concerne la recherche des zéros de (3.9), une méthode de Newton par intervalles basée sur une subdivision naïve de $X \times X$ ne suffira pas. En effet, pour un élément (α, α) de la diagonale ΔX , on a

$$dc_{(\alpha, \alpha)} = \begin{pmatrix} f''(\alpha) & 0 \\ f'(\alpha) & f'(\alpha) \end{pmatrix}.$$

Si, en plus, le nombre réel α est un élément de Σ^1 , alors $f'(\alpha) = 0$, et donc

$$dc_{(\alpha, \alpha)} = \begin{pmatrix} f''(\alpha) & 0 \\ 0 & 0 \end{pmatrix}.$$

Et dans ce cas, la méthode de Newton par intervalles ne pourra jamais conclure. Par conséquent, dans le processus de subdivision de $X \times X$, pour tout pavé contenant la diagonale ΔX , on testera si la condition du lemme 1 est satisfaite avec comme fonction f' . L'injectivité de f' ainsi démontrée nous suffira à savoir qu'aucune paire de points (x_1, x_2) ne vérifie

$$x_1 \neq x_2 \text{ et } df_{x_1} = df_{x_2}.$$

Ce genre de technique sera étendue dans la section suivante.

3.3.2 Calcul numérique du contour apparent en dimension 2

Dans cette section, nous essayons de caractériser plus globalement des fonctions entre variétés de dimension 2. L'objectif est de calculer la topologie de l'image du lieu singulier $f(S_f)$. Dans le cas des fonctions plan sur plan, cet ensemble est appelé le contour apparent.

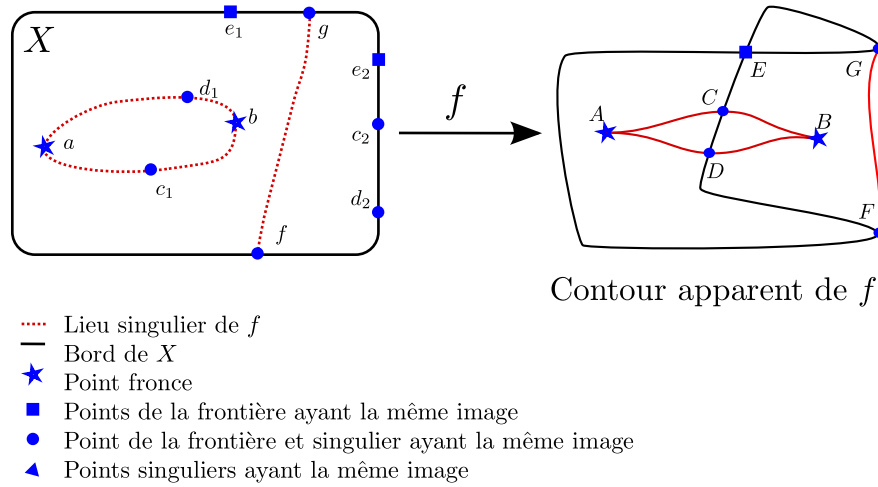


FIGURE 3.12 – Le contour apparent de f .

Plus exactement, on considère le cas où l'espace source X est homéomorphe à la boule unité de \mathbb{R}^2 . A titre d'introduction, on considère l'exemple illustré par la figure 3.12.

L'espace source X a un bord ∂X régulier. Les points en pointillé rouge forment une courbe qui est la réunion des points singuliers de f . Certains points sont en plus particuliers, en effet, cette fonction a deux fronces, notées a et b . La présence de ces points pourra être détectée par le jeu d'équations présenté dans la proposition 9. L'existence des points c_1 et d_1 va aussi survivre aux changements de variables sur l'espace source ou sur l'espace but. En effet, les points c_1 et c_2 ont la même image tout en vérifiant les propriétés suivantes :

- le point c_1 est singulier,
- le point c_2 appartient au bord de X .

Notre objectif est de référencer toutes ces situations globales puis de proposer des méthodes numériques capables de démontrer l'existence de ces points (tout en les encadrant). Le théorème 10 liste tous les cas :

Théorème 10 (Propriétés d'une application générique). *Soit X un compact simplement connexe de \mathbb{R}^2 dont le bord ∂X est régulier. Une application générique lisse de f de X dans \mathbb{R}^2 vérifie les propriétés suivantes :*

- i) l'ensemble des points singuliers, noté $S = \{x \in X \mid \det df_x = 0\}$, est une courbe régulière de X , de plus les éléments de S sont des points pli ou bien des fronces ; l'ensemble des fronces est dit discret,*
- ii) trois points singuliers différents n'ont pas la même image ; deux points singuliers ayant la même image sont des points pli et s'intersectent transversalement,*
- iii) trois points de la frontière différents n'ont pas la même image ; deux points de la frontière ayant la même image s'intersectent transversalement,*
- iv) trois points différents de la réunion de la courbe singulière et de la frontière n'ont pas la même image. Si un point de la courbe singulière et du bord ont la même image, alors le point singulier est un point pli et ils s'intersectent transversalement.*

v) si la courbe singulière et le bord s'intersectent, alors ce point est un point pli et de plus les tangentes à la courbe singulière et à la frontière sont différentes.

Démonstration. Pour prouver ce théorème, il suffit de montrer que l'ensemble des applications lisses vérifiant les conditions i), ii), iii), iv) et v) forme un sous-ensemble résiduel de $\mathcal{C}^\infty(X, \mathbb{R}^2)$. Comme être résiduel est stable par intersection finie, il suffit de montrer que chacun des ensembles correspondant à chaque cas est résiduel dans $\mathcal{C}^\infty(X, \mathbb{R}^2)$ pour conclure.

- Le cas i) a déjà été démontré lors du théorème 10.
- Considérons maintenant le cas ii). On doit démontrer que l'ensemble des applications lisses vérifiant la condition ii) est résiduel dans $\mathcal{C}^\infty(X, \mathbb{R}^2)$. La preuve s'appuie sur le théorème 9. Dans notre situation, il suffit de considérer le triple 1-jet de f défini par

$$\begin{aligned} j_{(3)}^1 f & : \quad \Delta_{(3)}X & \rightarrow & \quad J_{(3)}^1(X, \mathbb{R}^2) \\ & (x_1, x_2, x_3) & \mapsto & (x_1, f(x_1), df_{x_1}, x_2, f(x_2), df_{x_2}, x_3, f(x_3), df_{x_3}) \end{aligned}$$

où $\Delta_{(3)}X$ est le sous-ensemble X^3 avec des éléments deux à deux distincts :

$$\Delta_{(3)}X = \{(x_1, x_2, x_3) \in X^3 \mid i \neq j \Rightarrow x_i \neq x_j\}$$

et

$$J_{(3)}^1(X, \mathbb{R}^2) = \{(x_i, y_i, u_i)_{1 \leq i \leq 3} \in (J^1(X, \mathbb{R}^2))^3 \mid (x_1, x_2, x_3) \in \Delta_{(3)}\}$$

Considérons la sous variété W_2 de $J_{(3)}^1(X, \mathbb{R}^2)$ définie par

$$W_2 = \{(x_1, y_1, u_1, x_2, y_2, u_2, x_3, y_3, u_3) \mid \det u_1 = 0 \wedge \det u_2 = 0 \wedge \det u_3 = 0 \wedge y_1 = y_2 \wedge y_2 = y_3\}.$$

En accord avec le théorème de transversalité multijet, T_{W_2} est résiduel dans $\mathcal{C}^\infty(X, \mathbb{R}^2)$. De plus, W_2 est de codimension 7, et $\Delta_{(3)}$ est de dimension 6. Par conséquent, avec f dans T_{W_2} , $\{(x_i)_{1 \leq i \leq 3} \in \Delta_{(3)}X \mid j_{(3)}^1 f \bar{\cap} W_2\}$ est vide. En d'autres termes, trois points singuliers de f n'ont pas la même image.

Considérons maintenant la sous variété W'_2 de $J_{(2)}^1$ définie par

$$W'_2 = \{(x_1, y_1, u_1, x_2, y_2, u_2) \mid \det u_1 = 0 \wedge \det u_2 = 0 \wedge y_1 = y_2\}.$$

W'_2 est de codimension 4, par conséquent f est dans $T_{W'_2}$ alors l'ensemble des paires de points différents de S est de dimension 0 (*i.e.* codimension 4 dans $\Delta_{(2)}X$). De plus, $W'_2 \cap \{\text{im } u_1 = \text{im } u_2\}$ est de codimension supérieure à 4, alors deux points singuliers ayant la même image réalisent un croisement normal.

- Les autres cas peuvent être facilement prouvés à partir de la version multi-jet du théorème de transversalité en suivant la même approche.

□

Le cas i du théorème 10 est illustré par les figures 3.13 et 3.14.

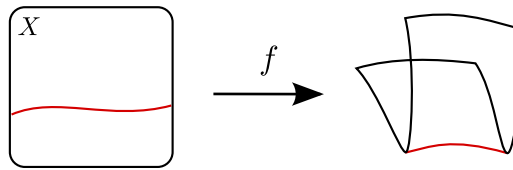


FIGURE 3.13 – Une courbe de point pli.

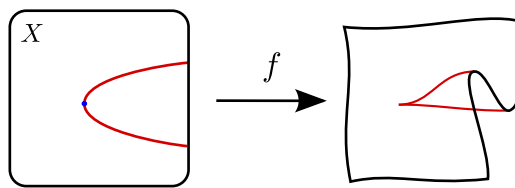


FIGURE 3.14 – Une courbe de point pli avec une fronce.

En accord avec le théorème 10, notre analyse générique doit considérer les 2 multi-jets d'une application lisse donnée. Les cas ii, iii, iv et v sont respectivement illustrés par les figures 3.15, 3.16, 3.17, 3.18.

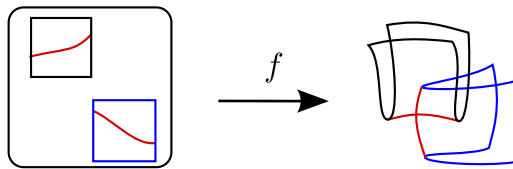


FIGURE 3.15 – Génériquement, deux courbes de points pli se croisent transversalement, cas ii du théorème 10.

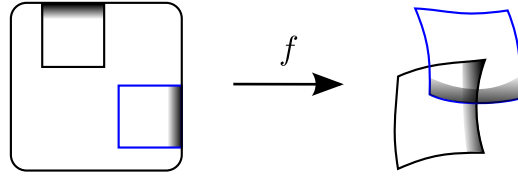


FIGURE 3.16 – Génériquement, deux courbes frontières se croisent transversalement, cas iii du théorème 10.

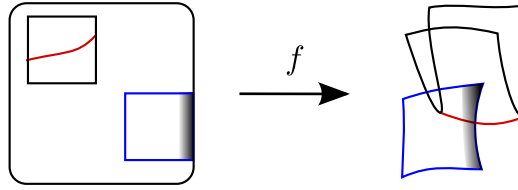


FIGURE 3.17 – Génériquement, une courbe frontière et une courbe de point pli se croisent transversalement, cas iv du théorème 10.

Dans cette sous-section, on donne aussi une condition suffisante pour qu'un voisinage donné d'une singularité ne contienne pas deux points singuliers différents ayant la même image. Autrement dit, la condition suivante peut suffire pour montrer qu'une fonction restreinte à une courbe est injective. L'idée est de trouver une direction que la courbe image suit. Ces trois prochains lemmes sont en quelque sorte des généralisations du lemme 1.

Lemme 2 (Injective). *Soit δ une courbe lisse $\delta : [0, 1] \ni t \mapsto \delta(t) \in \mathbb{R}^n$, et notons par w la fonction définie par $w : t \mapsto \alpha \cdot \delta(t)$ où α est une forme linéaire sur \mathbb{R}^n . Si $\dot{w}(t) \geq 0$ et si l'ensemble $\{t \mid \dot{w}(t) = 0\}$ est fini alors δ est injective.*

Démonstration. Les lignes de niveau $D_\lambda = \{x \in \mathbb{R}^n \mid \alpha \cdot x = \lambda\}_{\lambda \in \mathbb{R}}$ forment une foliation de l'espace but avec des hyperplans. La condition $\dot{w}(t) \geq 0$ et le fait que l'ensemble $\{t \mid \dot{w}(t) = 0\}$ soit fini impliquent que $t \neq t' \Rightarrow \delta(t) \in D_\lambda$ et $\delta(t') \in D_{\lambda'}$ avec $\lambda \neq \lambda'$. Comme $\lambda \neq \lambda' \Rightarrow D_\lambda \cap D_{\lambda'} = \emptyset$, on en conclut que δ est injective. \square

Dans notre cas, le lemme précédent ne pourra pas être directement appliqué à $f|_{S_f}$ car, en général, nous ne disposons pas de paramétrisation explicite de S_f . On donne donc ici le lemme 3.

Lemme 3. *Soit $f : X \rightarrow \mathbb{R}^n$ une application lisse et X une espace compact de \mathbb{R}^2 . Soit $\Gamma : X \rightarrow \mathbb{R}$ une submersion telle que l'ensemble $S = \{x \in X \mid \Gamma(x) = 0\}$ soit contractile [45]. Si*

$$\forall J \in \tilde{d}f_X \cdot \begin{pmatrix} \partial_2 \Gamma(X) \\ -\partial_1 \Gamma(X) \end{pmatrix}, J \text{ est un morphisme injectif}$$

alors $f|_S$ est un plongement.

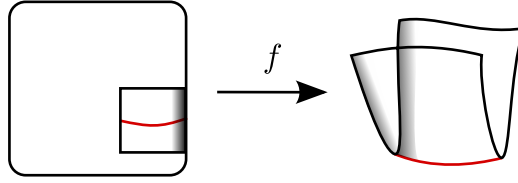


FIGURE 3.18 – Une courbe frontière et une courbe de point pli se croisent transversalement dans l’espace source, cas v du théorème 10.

Démonstration. La preuve se trouve dans [6]. \square

Remarque. Pour les notions de plongement, submersion, le lecteur pourra consulter [40].

Malheureusement, cette condition suffisante ne peut être vérifiée si l’ensemble X contient une fronce. Dans ce cas, on pourra employer le lemme 4 qui donne une condition encore plus technique.

Lemme 4. *Soit $f : X \rightarrow \mathbb{R}^2$ une application générique lisse avec X un compact simplement connexe de \mathbb{R}^2 . Supposons qu’il existe une unique fronce simple p_0 à l’intérieur de X , on note par α un covecteur normal à $\text{im } df_{p_0}$ non nul, et ξ un champ de vecteur ne s’annulant pas tel que $\forall p \in S, \xi_p \in T_p S$ (S contractile).*

Si la fonction $g = \sum \alpha_i \xi^3 f_i : X \rightarrow \mathbb{R}$ ne s’annule pas $f|_S$ est injective. Cette condition est localement nécessaire.

Ici, le champ de vecteurs ξ est vu comme la dérivation de $\mathcal{C}^\infty(X)$ définie par

$$\xi = \sum \xi_i \partial_{x_i}.$$

Démonstration. Considérons l’application $\delta : t \mapsto f \circ \gamma$ avec γ une paramétrisation de S telle que $\dot{\gamma} = \xi$ et $\gamma(0) = p_0$. On note par w la fonction $w : t \mapsto \alpha \cdot \delta(t)$. Par définition, $f \circ \gamma$ admet une fronce simple pour $t = 0$, en d’autres termes $\dot{\delta}(0) = 0$ (et $\dot{w}(0) = 0$). On sait aussi que $\ddot{w}(0) = \alpha \cdot \ddot{\delta}(0) = 0$ car $\ddot{\delta}(0)$ appartient à $\text{im } df_{p_0}$. Supposons maintenant que la fonction g ne s’annule pas. On remarque alors que $g \circ \gamma = \ddot{w}$.

Finalement, les assertions suivantes sont vraies :

1. $\dot{w}(0) = 0$,
2. $\ddot{w}(0) = 0$,
3. $t \mapsto \ddot{w}(t)$ ne s’annule pas.

A partir de 1 et 2, on peut conclure que 0 est un extremum de \dot{w} . Supposons que $g > 0$, alors 3 implique que 0 est l’unique minimum de \dot{w} . Par conséquent, grâce à la proposition 2, on en déduit que $f|_S$ est injective. Pour l’autre cas, *i.e.* $g < 0$, il suffit de considérer $-w$ pour conclure. \square

On peut remarquer que d’un point de vue temps de calcul, calculer algébriquement l’expression de g est souvent trop coûteuse en temps et en espace. L’exemple suivant illustre l’approche et peut être résolu à la main. En pratique, pour montrer que la restriction à S d’une application f avec des fronces est injective, on utilisera la proposition précédente en la combinant avec des techniques de Différentiation automatique [46].

Exemple 13. On considère l'application $(x, y) \mapsto (f_1, f_2) = (x, y^3 - xy)$. Le point de coordonnées $(0, 0)$ est une fronce simple. Dans ce cas, on a : $df = \begin{pmatrix} 1 & 0 \\ -y & 3y^2 - x \end{pmatrix}$, $d \det df = 3y^2 - x$ et $\xi = 6y\partial_x + 1\partial_y$. Pour montrer que $f|_S$ est injective, on doit seulement montrer que la fonction g ne s'annule pas.

$$g = \alpha_1 \xi^3 f_1 + \alpha_2 \xi^3 f_2$$

avec $\alpha = (\alpha_1, \alpha_2) = (0, 1)$ qui est orthogonal à $\text{im } df_{(0,0)}$. Finalement, comme $\xi^3 f_2 = -12$, on en déduit que $f|_S$ est injective. La figure 3.19 illustre cet exemple.

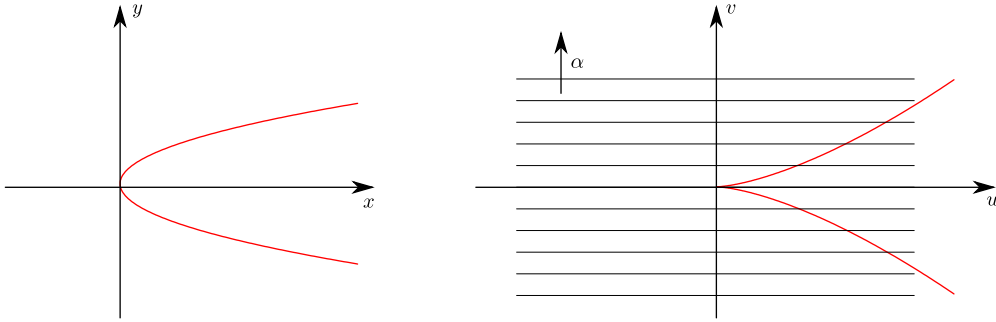


FIGURE 3.19 – $f|_S$ est une application injective. L'ensemble des $\{y = \lambda\}_{\lambda \in \mathbb{R}}$ est une foliation de l'ensemble but.

Mise en équation - Détection des paires de points singuliers ayant la même image - cas ii

L'objectif de ce paragraphe est de proposer un système d'équations pour le cas ii du théorème 10. On cherche donc une méthode pour encadrer les couples de points plus différents ayant la même image. C'est à dire la situation illustrée par la figure 3.15. En d'autres termes, on cherche à encadrer l'ensemble

$$S^{\Delta 2} = \{(\alpha, \beta) \in S_f \times S_f - \Delta S_f \mid f(\alpha) = f(\beta)\} / \simeq$$

où ΔX est la diagonale de X et \simeq est la relation d'équivalence définie par $(\alpha_1, \beta_1) \simeq (\alpha_2, \beta_2) \Leftrightarrow (\alpha_1, \beta_1) = (\beta_2, \alpha_2)$ ou $(\alpha_1, \beta_1) = (\alpha_2, \beta_2)$. L'approche proposée est proche de celle donnée pour le cas unidimensionnel et s'appuie sur l'utilisation de la proposition 3. L'idée principale est de partitionner $X \times X$ tout en considérant la relation \simeq . On définit tout d'abord l'application nommée *folds* par

$$\begin{aligned} \text{folds} : \quad X \times X &\rightarrow \mathbb{R}^4 \\ \left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \right) &\mapsto \begin{pmatrix} \det df(x_1, y_1) \\ \det df(x_2, y_2) \\ f_1(x_1, y_1) - f_1(x_2, y_2) \\ f_2(x_1, y_1) - f_2(x_2, y_2) \end{pmatrix} \end{aligned} \quad (3.10)$$

La relation entre l'ensemble cherché $S^{\Delta 2}$ et l'application $folds$ est donnée par :

$$S^{\Delta 2} = folds^{-1}(\{0\}) - \Delta S_f / \simeq .$$

En effet, si les deux premières composantes de la fonction $folds$ s'annulent alors (x_1, y_1) et (x_2, y_2) sont singuliers. De plus, si les deux dernières composantes s'annulent alors (x_1, y_1) et (x_2, y_2) ont la même image.

Comme la relation d'équivalence \simeq sera prise en compte dans le processus de subdivision, il nous reste plus qu'à encadrer $folds^{-1}(\{0\}) - \Delta S_f$. L'approche proposée s'appuie sur une méthode de Newton pour trouver les zéros de la fonction $folds$ et permettra de montrer l'existence et l'unicité d'un couple de point pli ayant la même image.

En utilisant naïvement la méthode de Newton pour enfermer les zéros de la fonction $folds$, elle échouera. Il faut nécessairement qu'au voisinage des zéros de $folds$, cette fonction soit inversible. Cette condition n'est pas remplie pour les points appartenant à ΔS_f . En effet, pour toute paire (α, α) de ΔS_f , la différentielle de $folds$ est conjuguée à

$$\begin{pmatrix} a & b & 0 & 0 \\ 0 & 0 & a & b \\ a_{11} & a_{12} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{21} & a_{22} \end{pmatrix}$$

qui est singulière car $\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \det df(\alpha) = 0$.

En d'autres termes, comme tout pavé de la forme $[x] \times [x]$ contient ΔS , la méthode de Newton échouera. Pour contourner cette difficulté, on remarque que l'assertion $\{(\alpha, \beta) \in S \times S - \Delta S \mid f(\alpha) = f(\beta)\} \cap [x] \times [x] = \emptyset$ est équivalente à $f|_S \cap [x]$ est un plongement. Notre algorithme testera si la condition suffisante du lemme 3 est satisfaite en posant $\Gamma = \det df$, et démontrera que $f|_S \cap [x]$ est un plongement.

Mise en équation - Points de la frontière dont les images s'intersectent - cas iii

Cette section concerne le cas iii du théorème 10 : deux points appartenant à la frontière peuvent avoir la même image. Cette situation est illustrée par la figure 3.16. On suppose ici que la frontière est déterminée comme étant le lieu des zéros d'une fonction Γ , i.e. :

$$\partial X = \{x \in X \mid \Gamma(x) = 0\}$$

La fonction réelle Γ est supposée être une submersion au moins sur ∂X . Plus précisément, on cherche à encadrer l'ensemble

$$\partial X^{\Delta 2} = \{(\alpha, \beta) \in \partial X \times \partial X - \Delta \partial X \mid f(\alpha) = f(\beta)\} / \simeq ,$$

où \simeq est la même relation d'équivalence que celle définie dans le paragraphe précédent. Ici encore, on utilisera l'opérateur de Newton par intervalles pour trouver les zéros d'une fonction. Dans ce cas, la fonction suivante caractérise génériquement $\partial X^{\Delta 2}$.

$$\begin{array}{ccc}
X \times X & \rightarrow & \mathbb{R}^4 \\
\left(\begin{array}{c} x_1 \\ y_1 \end{array} \right), \left(\begin{array}{c} x_2 \\ y_2 \end{array} \right) & \mapsto & \left(\begin{array}{c} \Gamma(x_1, y_1) \\ \Gamma(x_2, y_2) \\ f_1(x_1, y_1) - f_1(x_2, y_2) \\ f_2(x_1, y_1) - f_2(x_2, y_2) \end{array} \right)
\end{array} \quad (3.11)$$

Les étapes de l'algorithme pour encadrer $\partial X^{\Delta 2}$ sont presque les mêmes que pour $S^{\Delta 2}$. Génériquement, si α et β sont deux éléments de ∂X ayant la même image, les droites $df_\alpha T_\alpha \Gamma$ et $df_\beta T_\beta \Gamma$ ne sont pas confondues, par conséquent, la différentielle de la fonction (3.11) est inversible au voisinage de ses zéros. Comme dans le cas précédent, on traite la diagonale avec attention. On élimine les pavés de la forme $[\alpha] \times [\alpha]$ en testant la condition suffisante du lemme 3.

Mise en équation - La courbe pli et le bord s'intersectent - cas iv et v

Cette section traite des deux derniers cas pouvant se produire génériquement, i.e. les iv et v du théorème 10. Comme précédemment, on donne une méthode pour encadrer l'ensemble suivant

$$A = \{(\alpha, \beta) \in S \times \partial X \mid f(\alpha) = f(\beta)\}.$$

On considère deux sous-cas. Le premier correspond à la situation où le couple (α, β) sont deux points différents alors que dans le second, $\alpha = \beta$, i.e. un point pli est sur la frontière.

Pour la présentation, on décompose l'ensemble A en deux sous-ensembles

$$A - \Delta A = \{(\alpha, \beta) \in A \mid \alpha \neq \beta\},$$

et

$$\Delta A = \{(\alpha, \beta) \in A \mid \alpha = \beta\}.$$

L'ensemble $A - \Delta A$ peut être encadré, via l'opérateur de Newton, en cherchant les zéros de la fonction suivante :

$$\begin{array}{ccc}
X \times X & \rightarrow & \mathbb{R}^4 \\
\left(\begin{array}{c} x_1 \\ y_1 \end{array} \right), \left(\begin{array}{c} x_2 \\ y_2 \end{array} \right) & \mapsto & \left(\begin{array}{c} \det df(x_1, y_1) \\ \Gamma(x_2, y_2) \\ f_1(x_1, y_1) - f_1(x_2, y_2) \\ f_2(x_1, y_1) - f_2(x_2, y_2) \end{array} \right).
\end{array} \quad (3.12)$$

Cette méthode ne peut pas être utilisée pour les points de ΔA comme la différentielle n'est pas inversible. En effet, les deux dernières lignes sont identiques. Cependant, les zéros de la fonction suivante

$$\begin{array}{ccc}
X & \rightarrow & \mathbb{R}^2 \\
\left(\begin{array}{c} x_1 \\ y_1 \end{array} \right) & \mapsto & \left(\begin{array}{c} \det df(x_1, y_1) \\ \Gamma(x_1, y_1) \end{array} \right)
\end{array} \quad (3.13)$$

sont les intersections recherchées, i.e. entre la courbe de points plis $\Sigma^{1,0}$ et le bord ∂X dans l'ensemble source. Génériquement, les tangentes en ces points ne sont pas confondues, ce qui assure que la différentielle de (3.13) est inversible au voisinage de ces points.

Supposons que la méthode de Newton réussisse à montrer l'existence et l'unicité d'un zéro de (3.13) pour un domaine $[x]$. Autrement dit, on aura uniquement prouvé que $\Delta A \cap [x]$ est un point. Ce n'est pas suffisant, en effet, on a un problème similaire à ce qui se passait pour la courbe de point pli. Par conséquent, on doit aussi démontrer que ce point est l'unique dans A . Pour ce faire, on emploie une condition suffisante donnée par un lemme technique dans [6].

Synthèse : connecter les éléments

Cette section donne un algorithme capable de construire un graphe homéomorphe au contour apparent d'une fonction générique lisse donnée. L'idée principale est tout d'abord d'appliquer les algorithmes précédents pour rigoureusement encadrer les points importants : fronces, nœuds, i.e. les paires de points plis ayant la même image . . . Ces points seront les nœuds d'un graphe dont les arcs seront les relations de voisinage données par le fait qu'une paire de points est directement connectée par une courbe singulière ou bien par la frontière de X .

Soit f une fonction lisse générique définie sur un sous-ensemble compact simplement connecté X de \mathbb{R}^2 dans \mathbb{R}^2 . Définissons par X_0 la réunion des fronces et des points doubles singuliers et frontières. Formellement, on a donc :

$$X_0 = \left\{ x \in X \mid \begin{array}{l} \text{ou } x \text{ est une fronce} \\ \text{ou } \exists \tilde{x}, (x, \tilde{x}) \in S^{\Delta 2} \\ \text{ou } \exists \tilde{x}, (x, \tilde{x}) \in \partial X^{\Delta 2} \\ \text{ou } \exists \tilde{x}, (x, \tilde{x}) \in A \text{ ou } (\tilde{x}, x) \in A \end{array} \right\}.$$

On pose aussi Y_0 comme étant $f(X_0)$. Soit X_1 l'ensemble $S_f \cup \partial X - X_0$ et $Y_1 = f(X_1)$. Génériquement, les ensembles X_0 et Y_0 sont discrets et vérifient $\#Y_0 \leq \#X_0$. De plus (Y_0, Y_1) et (X_0, X_1) forment, respectivement, une stratification du contour apparent et de $S \cup \partial X$. Par construction, $f|_{X_1}$ est injective, i.e. X_1 et Y_1 sont homéomorphes. L'application f induit une relation sur X définie par $\alpha f \beta \Leftrightarrow f(\alpha) = f(\beta)$. Cette relation est une relation d'équivalence et le quotient X/f est donc bien défini. Génériquement, les deux assertions suivantes sont vraies :

- X_0/f est homéomorphe à Y_0 ,
- $(X_0 \cup X_1)/f$ est homéomorphe au contour apparent $Y_0 \cup Y_1$.

En pratique, pour calculer la topologie du contour apparent, on divise tout d'abord l'ensemble source X avec un recouvrement $\{p_i\}$ puis on colle les pièces ensemble. La sortie de cette première étape est un graphe simple⁷ homéomorphe à l'ensemble singulier $S \cup \partial X = X_0 \cup X_1$. A partir de cette sortie et grâce à la relation d'équivalence induite par f , un graphe homéomorphe au contour apparent est calculé. Le théorème suivant résume le principe de la méthode :

Théorème 11. *Soit $P = \{p_i\}_{1 \leq i \leq n}$ un recouvrement tel que*

- i) $S \cup \partial X \subset \cup_i p_i$,*
- ii) $\forall (p, q), p \cap q \neq \emptyset \Rightarrow (S \cup \partial X) \cap p \cap q$ est simplement connexe,*
- iii) $\forall p, X \cap p$ contient au plus un élément de X_0 ,*

7. Un graphe simple est un graphe non orienté sans arêtes multiples et sans boucle.

Soit \mathcal{X} la relation sur $\{p_i\}_{1 \leq i \leq n}$ définie par

$$p\mathcal{X}q \Leftrightarrow (S \cup \partial X) \cap p \cap q \text{ est simplement connexe.}$$

Définissons la relation d'équivalence f sur $\{p_i\}$ par

$$pfq \Leftrightarrow f(X_0 \cap p) = f(X_0 \cap q) \text{ et } X_0 \cap p \neq \emptyset,$$

alors \mathcal{X}/f est homéomorphe au contour apparent.

Démonstration. La preuve se trouve dans [6]. □

En pratique, la méthode subdivise récursivement X jusqu'à ce que le recouvrement créé satisfasse les conditions du théorème 11. Les nœuds du graphe sont obtenus en testant les conditions présentées dans les sections précédentes. La condition suffisante présentée dans [15] est employée pour garantir qu'une courbe est simplement connexe. On obtient alors numériquement la topologie du contour apparent \mathcal{X}/f .

Exemples

Cette section donne deux exemples pour illustrer l'approche. L'exemple 14 est académique alors que l'exemple 15 est un exemple numérique. Ce dernier a été traité automatiquement par le programme développé pour démontrer la faisabilité de l'approche.

Exemple 14. Considérons l'application suivante :

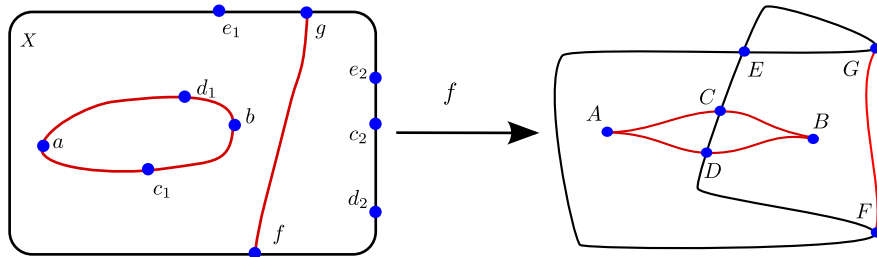


FIGURE 3.20 – Une stratification de l'ensemble critique et le contour apparent de f .

Pour cette fonction générique, X_0 est l'ensemble fini $\{a, b, c_1, c_2, d_1, d_2, e_1, e_2, f, g\}$ et X_1 a 11 composantes connexes. Plus précisément, a et b sont des fronces, S^{Δ^2} est vide, (c_1, c_2) et (d_1, d_2) sont dans A , (e_1, e_2) appartient à ∂X^{Δ^2} , et (f, f) et (g, g) sont dans ΔA . Les éléments de l'ensemble image $Y_1 = f(X_0)$ sont notés $\{A, B, C, D, E, F, G\}$ de sorte que chaque lettre majuscule soit l'image de sa lettre minuscule par f (i.e. $A = f(a)$, $B = f(b)$, ...). Les seuls éléments de X_0 qui sont en relation via la relation d'équivalence f sont donc : $c_1 f c_2, d_1 f d_2$ et $e_1 f e_2$. Par conséquent, X_0/f peut être identifié avec $\{a, b, c_1, d_1, e_1, f, g\}$ qui est homéomorphe à $\{A, B, C, D, E, F, G\}$. L'ensemble Y_1 a aussi 11 composantes connexes car $f|_{X_1}$ est injective.

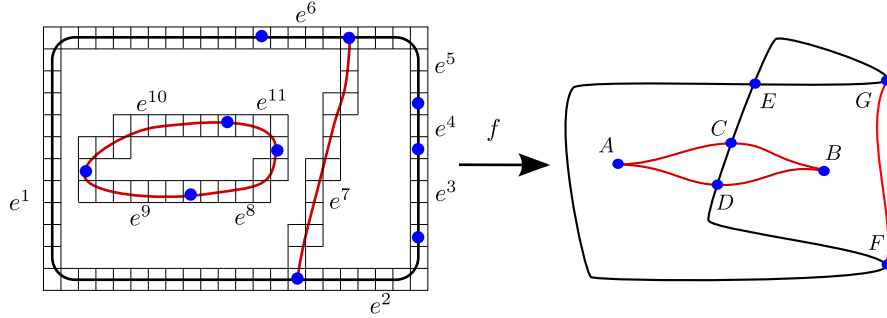


FIGURE 3.21 – Recouvrement de l'ensemble critique S_f .

La figure 3.21 propose un recouvrement $\{p_i\}$ de $S \cup \partial X$ vérifiant les conditions du théorème 11.

A partir de ce recouvrement, et de la relation f , on peut formellement calculer la matrice d'incidence du graphe \mathcal{X} :

$$\mathcal{X} = \begin{matrix} & e^1 & e^2 & e^3 & e^4 & e^5 & e^6 & e^7 & e^8 & e^9 & e^{10} & e^{11} \\ \begin{matrix} a \\ b \\ c_1 \\ c_2 \\ d_1 \\ d_2 \\ e_1 \\ e_2 \\ f \\ g \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix} \quad (3.14)$$

Ce graphe a été calculé en oubliant les nœuds du graphe de la figure 3.21 ne contenant pas de points critiques. Cette opération réduit le nombre de nœuds sans changer la topologie tout en étant compatible avec la relation f . Finalement, il peut être quotienté par la relation d'équivalence f :

$$\mathcal{X}/f = \begin{matrix} & e^1 & e^2 & e^3 & e^4 & e^5 & e^6 & e^7 & e^8 & e^9 & e^{10} & e^{11} \\ \begin{matrix} A \\ B \\ C \\ D \\ E \\ F \\ G \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix} \quad (3.15)$$

Formellement, ce quotient est obtenu ligne par ligne. Chaque ligne Y de \mathcal{X}/f est simplement l'addition des lignes y_i de \mathcal{X} telles que $f(y_i) = Y$, e.g. $C = c_1 + c_2$. D'après le théorème 11, le graphe \mathcal{X}/f est homéomorphe au contour apparent de f .

Exemple 15. Contrairement à l'exemple précédent où la fonction était décrite par une figure, ici f est donnée par une expression. De plus, cet exemple a été traité par le solveur implémenté pour montrer la faisabilité de la méthode. Soit f la fonction lisse suivante :

$$f : [-1, 4] \times [-0.4, 0.6] \rightarrow \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 0.5x + (1 - y^2) \cos(x + 0.2y) \\ (1 - y^2) \sin(x) \end{pmatrix}$$

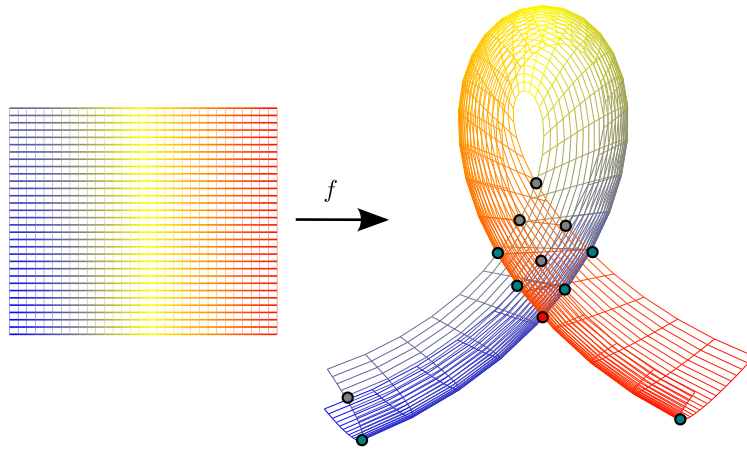


FIGURE 3.22 – Illustration de la fonction de l'exemple 15.

Les calculs ont été menés sur un ordinateur 64 bit avec 3887 MB de RAM et un processeur Intel Core i7. Le programme et son code source sont disponibles sur le web via l'url <http://perso-laris.univ-angers.fr/~delanoue/>. La bibliothèque d'arithmétique par intervalles utilisée est celle développée par W. Hofschuster, W. Krämer nommé filib++ Interval Library version 3.0.2 (Voir [47]). Le programme a démontré que X_0 a 22 éléments. Plus précisément, la sortie du solveur est la suivante :

1. le nombre de fronces est 0,
2. le nombre d'intersections de la courbe point pli dans l'espace but est 1,
3. le nombre d'intersections de la courbe frontière dans l'espace but est 5,
4. le nombre d'intersections de point pli avec la courbe frontière est 6.

La topologie du contour apparent est donnée par :

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Le temps de calcul total est 1 min 33 s. Pour calculer la topologie de $S \cup \partial X$, l'espace source a été divisé avec un pavage régulier de 27^2 pavés. Le programme a été baptisé Thom en l'honneur du mathématicien René Thom.

3.3.3 Portrait de fonctions

L'invariant calculé dans la section précédente, bien que relativement riche, n'est pas suffisant pour distinguer toute paire de fonctions non équivalentes. En effet, il existe des couples de fonctions non équivalentes ayant le même contour apparent. Un exemple de tel couple est présenté dans le manuscrit de thèse de Romain Benoit [41].

Lors de la thèse de Romain Benoit, il a donc décidé de développer une méthode capable d'encore mieux discerner les fonctions. Ceci a été formalisé par la notion de *portrait de fonction*. Cette notion déjà introduite dans l'introduction de la section 3.3 s'appuie sur l'idée, classiquement employée en topologie algébrique, de *triangulation*⁸. De façon informelle, le portrait d'une fonction sera une application entre triangulations. D'une certaine manière, faire le portrait d'une fonction consiste à la trianguler.

Définition 13. (Triangulation ou complexe simplicial abstrait) Soit \mathcal{N} un ensemble fini de symboles $\{a^0, a^1, \dots, a^n\}$. Une triangulation \mathcal{K} un sous-ensemble des parties de \mathcal{N} vérifiant

$$\sigma \in \mathcal{K} \Rightarrow \forall \sigma_0 \subset \sigma, \sigma_0 \in \mathcal{K} \quad (3.16)$$

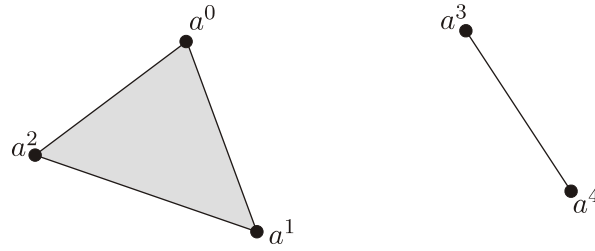
Exemple 16. L'ensemble

$$\mathcal{K} = \left\{ \emptyset, \{a^0\}, \{a^1\}, \{a^2\}, \{a^3\}, \{a^4\}, \{a^0, a^1\}, \{a^1, a^2\}, \{a^0, a^2\}, \{a^3, a^4\}, \{a^0, a^1, a^2\} \right\}.$$

est une triangulation.

La figure 3.23 donne une réalisation géométrique de cette triangulation.

8. Une triangulation peut être vu comme une généralisation de la notion de graphe. En effet, dans un graphe, on a des nœuds (des singletons) et des arcs (des paires) qui les relie. Dans une triangulation, on a des nœuds, des arcs, des triangles (des triplets), des tétraèdres (des quadruplets) ...

FIGURE 3.23 – Une réalisation géométrique de \mathcal{K} .

L'énumération complète des éléments d'une triangulation \mathcal{K} n'est pas nécessaire. Il suffit d'énumérer ses faces (*i.e.* les simplexes de dimension maximale). En effet, supposons que $\{a^0, a^1, a^2\} \in \mathcal{K}$, la relation (3.16) implique que les ensembles $\{a^0\}$, $\{a^1\}$, $\{a^2\}$, $\{a^0, a^1\}$, $\{a^1, a^2\}$, $\{a^0, a^2\}$ sont aussi des éléments \mathcal{K} . Cette remarque suggère la notation suivante.

Notation 1. Avec $\mathcal{V} = \{a^0, a^1, \dots, a^n\}$ une collection finie d'éléments (les nœuds abstraits) et $2^{\mathcal{V}}$ l'ensemble des parties de \mathcal{V} , une triangulation est une partie de $2^{\mathcal{V}}$ vérifiant (3.16). Soit $\{\sigma_1, \dots, \sigma_m\}$ une partie de $2^{\mathcal{V}}$ (pas nécessairement une triangulation), on note par $\sigma_1 + \dots + \sigma_m$ la triangulation suivante :

$$\sigma_1 + \dots + \sigma_m := \bigcup_{i=1}^{i=m} 2^{\sigma_i}$$

Avec cette notation, la triangulation de l'exemple 16 peut être écrite $a^0 a^1 a^2 + a^3 a^4$. Étant donnée une triangulation \mathcal{K} , le sous-ensemble composé des singletons est appelé le 0-squelette et noté \mathcal{K}^0 . Une application entre les 0-squelette de deux triangulations \mathcal{K} et \mathcal{L} est dite *simpliciale* si elle vérifie la propriété suivante :

$$\{a^0, a^1, \dots, a^n\} \in \mathcal{K} \Rightarrow \{F(a^0), F(a^1), \dots, F(a^n)\} \in \mathcal{L}. \quad (3.17)$$

Exemple 17. Soient \mathcal{K} et \mathcal{L} deux triangulations définies par $\mathcal{K} = a_0 a_1 + a_1 a_2 + a_2 a_3$ et $\mathcal{L} = b_0 b_1 + b_1 b_2$. L'application F définie par

$$F : \begin{array}{l} a^0 \mapsto b^0 \\ a^1 \mapsto b^1 \\ a^2 \mapsto b^2 \\ a^3 \mapsto b^1 \end{array}$$

est une application simpliciale.



Une application simpliciale F de \mathcal{K}^0 vers \mathcal{L}^0 peut être naturellement étendue en une application de \mathcal{K} vers \mathcal{L} avec :

$$F(\sigma) := \{F(a^i) \mid a^i \in \sigma\}.$$

Dans ce chapitre, on ne distinguera pas une application simpliciale de son extension. Comparer une application simpliciale et une application lisse ne semble pas possible a priori. En effet, une application simpliciale est une fonction envoyant les éléments d'un ensemble fini dans un autre ensemble fini. La régularité de ces fonctions n'a pas de sens immédiat puisqu'aucune topologie n'est donnée sur ces ensembles. Une bonne façon de pouvoir comparer une fonction lisse f et une application simpliciale consiste à construire une réalisation géométrique de F .

La notion de réalisation géométrique est une construction classique [45] qui topologise une triangulation. Cette construction s'appuie sur le théorème de plongement de Whitney [48]. Ce théorème montre que toute triangulation (abstraite) peut être plongée dans \mathbb{R}^n avec n suffisamment grand. L'image d'un plongement d'une triangulation \mathcal{K} donnée est appelée *réalisation géométrique* de cette triangulation. On notera par $|\mathcal{K}|$ la réalisation géométrique de la triangulation \mathcal{K} .

De la même manière, une application simpliciale $F : \mathcal{K} \rightarrow \mathcal{L}$ peut être réalisée avec des applications affines entre les simplexes $|\mathcal{K}|$ et $|\mathcal{L}|$. La condition (3.17) assure que toute réalisation géométrique F est continue. Une réalisation géométrique est donc une fonction lisse par morceaux. En effet, sa restriction à chaque simplexe est affine. Nous utiliserons cette approche pour coder une fonction lisse à équivalence près.

Afin d'y parvenir, on considère une relation d'équivalence entre fonctions continues, nommée conjugaison topologique, définie avec des homéomorphismes à la place des difféomorphismes dans la définition 10.

Définition 14. Soient f_1 et f_2 deux applications continues. Alors f_1 et f_2 sont topologiquement conjuguées s'il existe des homéomorphismes $g : X \rightarrow X'$ et $h : Y \rightarrow Y'$ de sorte que le diagramme

$$\begin{array}{ccc} X & \xrightarrow{f_1} & Y \\ \downarrow g & & \downarrow h \\ X' & \xrightarrow{f_2} & Y' \end{array}$$

commute. Dans ce cas, on notera par $f_1 \sim_0 f_2$ pour dire que f_1 et f_2 sont topologiquement conjuguées.

Cette relation de conjugaison topologique est plus faible que la relation d'équivalence dans le sens où $f_1 \sim f_2 \Rightarrow f_1 \sim_0 f_2$ pour toute paire (f_1, f_2) de fonctions lisses. Nous pouvons maintenant comparer une application simpliciale et une application lisse.

Définition 15 (Portrait). Soit f une application lisse et F une application simpliciale, on dit que F est le *portrait* de f si une réalisation géométrique de F est topologiquement conjuguée à f .

Exemple 18. L'application simpliciale donnée dans l'exemple 17 est un portrait de l'application lisse $f : [-4, 2] \ni x \mapsto x^2 - 1 \in \mathbb{R}$. En effet, l'application simpliciale

$$F : \begin{array}{l} a^0 \mapsto b^0 \\ a^1 \mapsto b^1 \\ a^2 \mapsto b^2 \\ a^3 \mapsto b^1 \end{array}$$

peut être réalisée par

$$|F| : [-2; 1] \rightarrow [0, 2]$$

$$x \mapsto \begin{cases} -x, & \text{si } x \in [-2, -1] \\ -x, & \text{si } x \in [-1, 0] \\ x, & \text{si } x \in [0, 1]. \end{cases}$$

Les applications g et h^{-1} sont des homéomorphismes et font commuter les diagrammes suivants :

$$\begin{array}{ccc} [-2, 1] & \xrightarrow{|F|} & [0, 2] \\ \downarrow g & & \downarrow h \\ [-4, 2] & \xrightarrow{f} & [-1, 15] \end{array}$$

avec

$$g : \begin{array}{l} [-2, 1] \rightarrow [-4, 2] \\ x \mapsto 2x, \end{array}$$

et

$$h^{-1} : \begin{array}{l} [-1, 15] \rightarrow [0, 2] \\ x \mapsto \frac{1}{2}\sqrt{y+1}. \end{array}$$

Exemple 19. Considérons les deux triangulations $\mathcal{K} = a_0(a_1a_2 + a_2a_3 + a_3a_4 + a_4a_5 + a_5a_6 + a_6a_1)$, $\mathcal{L} = b_0(b_1b_2 + b_2b_3 + b_3b_4 + b_4b_1)$. L'application $F : \mathcal{K}^0 \rightarrow \mathcal{L}^0$ définie par : $F(a_0) = b_0, F(a_1) = b_1, F(a_2) = b_2, F(a_3) = F(a_5) = b_3, F(a_4) = F(a_6) = b_4$ s'étend en une application simpliciale de \mathcal{K} dans \mathcal{L} . Soit p une fonction simple d'une application de \mathbb{R}^2 dans \mathbb{R}^2 , alors il existe un voisinage V de p tel que F soit le portrait de $f|_V$.

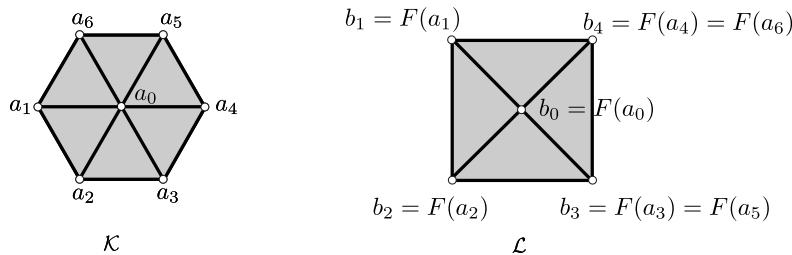


FIGURE 3.24 – Le portrait d'une fonction \mathbb{R}^2 to \mathbb{R}^2 .

Calculer le portrait d'une application lisse f donnée est un problème ambitieux. Le contour apparent calculé dans la section précédente vérifie le théorème suivant :

Théorème 12. *En reprenant, les notations du théorème 11, pour tout portrait F de f , le 1-squelette de $\text{im } F$ contient un sous-graphe homéomorphe à \mathcal{X}/f .*

Romain Benoit a proposé dans l'article [9] et son manuscrit de doctorat une méthodologie pour construire un portrait. Sa méthode consiste à tout d'abord partitionner l'espace source X et l'espace but Y en composantes connexes. Le domaine X est partitionné à l'aide des Σ^I , puis Y est lui aussi partitionné à l'aide des images $f(\Sigma^I)$. Dans une dernière étape, X est repartitionné à l'aide des contreimages de $f(\Sigma^I)$, i.e. avec $f^{-1}(f(\Sigma^I))$. Chaque élément de la partition \mathcal{X} est mis en correspondance avec un élément de la partition de \mathcal{Y} via la fonction f .

Afin de proposer le portrait d'une fonction f , les espaces topologiques éléments des différentes partitions sont triangulés. La difficulté principale consiste à trianguler les morceaux de sorte à satisfaire simultanément les deux propriétés suivantes :

- Soient A et B deux éléments de la partition de X . Si A est une partie du bord de B , la triangulation de A doit être incluse dans la triangulation de B .
- Soit Z est un élément de la partition de Y , les composantes connexes de la réciproque de Z par f doivent être conjointement triangulées de sorte qu'on puisse définir une application simpliciale sur chaque morceau s'envoyant sur la triangulation de Z .

Ces deux contraintes ont été levées en proposant des modèles de triangulations paramétrées par la topologie d'ensemble à trianguler. Afin d'illustrer la démarche, regardons le cas particulier donné par la figure 3.25. Sur cet exemple, f est une fonction du tore⁹ dans le plan.

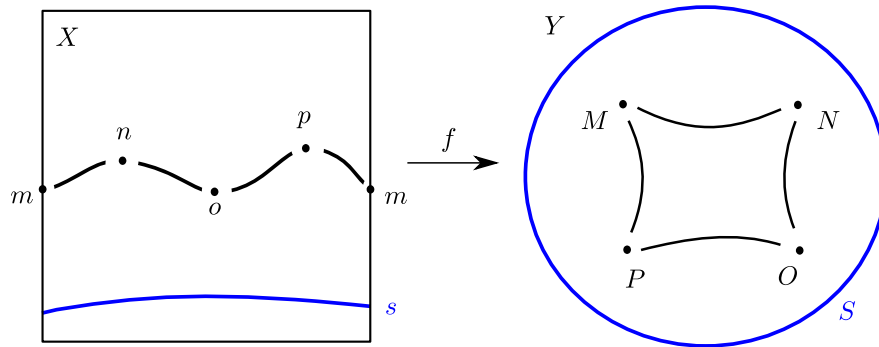


FIGURE 3.25 – Une application f du tore dans le plan.

Ici, la courbe singulière notée s est composée uniquement de point pli. Ce cercle (topologique) s , composante connexe de Σ^1 est envoyé bijectivement sur le cercle S par f . La seconde composante connexe de Σ^1 est aussi un cercle sur lequel se trouvent 4 fronces notées m, n, o et p . Ces 4 fronces sont envoyées bijectivement sur les points nommés M, N, O et P dans l'espace but. Chacun des 4 segments reliant ses fronces est une composante connexe de $\Sigma^{1,0}$ et est aussi envoyé bijectivement sur chacun des segments

9. Pour la représentation graphique, on a présenté le source en mettant le tore à plat, il reste donc à quotienter le plan par la relation \simeq avec \simeq la relation sur \mathbb{R}^2 définie par $(a_1, a_2) \simeq (b_1, b_2) \Leftrightarrow \forall i, a_i - b_i \in \mathbb{Z}$. Il est donc normal que le point m apparaissent deux fois.

dans le but. Il reste à observer deux composantes connexes bidimensionnels dans l'espace source qui font partie de Σ^0 . Chacune de ces deux régions est topologiquement un cylindre ($S^1 \times [0, 1]$).

Au voisinage des fronces représentées sur la figure 3.25, il se passe nécessairement quelque chose qui ressemble au portrait local donné par la figure 3.24. En effet, en regardant l'image réciproque par f de $f(\Sigma^{1,1})$ et $f(\Sigma^{1,0})$, on voit apparaître une structure plus riche. Ces points, images réciproques des fronces, sont nommés avec une apostrophe sur la figure 3.26 :

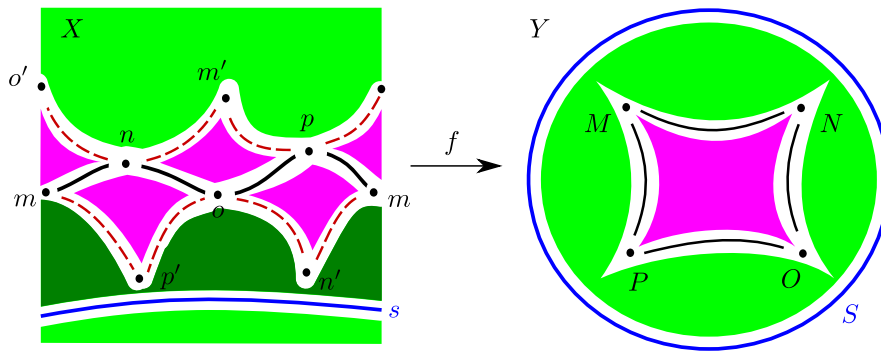


FIGURE 3.26 – Une application f du tore dans le plan.

Sur la figure 3.26, les 4 disques topologiques sont envoyés par f bijectivement sur le disque en violet à droite alors que les deux régions vertes sont envoyées sur le cylindre vert. Les segments pointillés sont eux aussi envoyés par f bijectivement sur les segments à droite. En particulier, le segment op' est envoyé sur le segment OP .

A partir de cet objet, on peut construire un portrait de f en triangulant astucieusement et conjointement X et Y . Le résultat de ce genre de méthode est donné par la réalisation du portrait illustré par la figure 3.27.

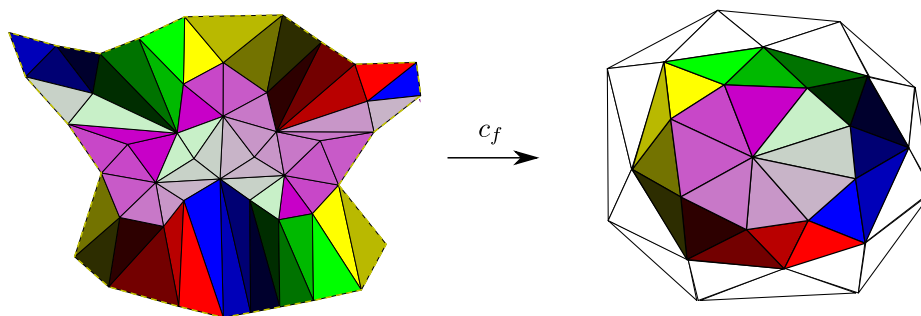


FIGURE 3.27 – Portrait de f .

Cette méthodologie de création de portraits a été appliquée pour réaliser tous les portraits génériques des fonctions décrivant le robot 3R. Ces résultats se trouvent compiler dans la publication [9] et complètent l'étude donnée par Philippe Wenger dans [49]. D'un point de vue robotique, le portrait est aussi un élément déterminant dans l'étude de ces

mécanismes [50].

A notre connaissance, aucun algorithme n'a jamais été proposé pour automatiquement créer le portrait d'une fonction lisse donnée. Ce genre d'algorithme serait central en théorie de la classification.

On remarquera que le portrait n'est pas, stricto sensu, un invariant. En effet, pour une même fonction, il est possible de donner plusieurs portraits. Évidemment, les réalisations de ces portraits seront nécessairement topologiquement conjuguées. En quotientant l'ensemble des portraits par la relation de conjugaison topologique, on obtient bien un invariant. Mais là n'était pas l'objectif. Même si théoriquement, une stratification semble être plus économe pour décrire une fonction, il ne faut pas perdre de vue que les strates peuvent avoir des topologies extrêmement compliquées. Nous ne souhaitons pas laisser cette complexité dans les strates. Le portrait d'une fonction se veut exhaustif comme peut l'être une triangulation homéomorphe à une variété.

Même si un portrait code de façon discrète une fonction modulo \sim_0 , il semble difficile d'avoir des algorithmes efficaces pour tester l'égalité de deux fonctions en testant l'équivalence topologique de leur portrait. En effet, les triangulations et les portraits sont des objets qui généralisent le concept de graphe, et comme décider de l'existence d'un isomorphisme entre deux graphes est de classe NP, il y a de fortes chances qu'on ne connaisse pas avant longtemps un algorithme efficace pour décider de l'équivalence topologique de deux portraits.

3.4 Conclusion

Dans ce chapitre, nous avons proposé, dans un premier temps, des méthodes numériques capables d'encadrer les espaces de Thom-Boardman. Les topologies de ces ensembles sont des invariants pour la classification des fonctions. Dans un second temps, un algorithme pour calculer le contour apparent d'une fonction du plan sur le plan a été donné. Finalement, comme le contour apparent n'était pas un invariant suffisamment fort, le concept de portrait de fonction a été introduit.

Nous envisageons aujourd'hui des méthodes capables de faire le portrait de surface incluse dans \mathbb{R}^4 décrite par une relation de la forme $g(x, y) = 0$ avec x et y deux éléments de \mathbb{R}^2 . D'un point de vue robotique, cela permettrait de faire le portrait de mécanismes comme le manipulateur plan parallèle 3RPR. Pour cette classe, il faudrait approfondir les recherches du côté mathématiques pour connaître les points qui joueraient le rôle des fronces et des nœuds. Puis, il faudra proposer une méthode de calcul d'une triangulation de la surface décrite par g tout en étant capable de la projeter conjointement sur une triangulation de X et sur une triangulation de Y .

Par ailleurs, la méthodologie développée, dans le cadre de la robotique, lors de la thèse de Romain Benoit pour la création de portrait de fonctions n'a pas été implémentée. Néanmoins, elle est déjà appréciée par les roboticiens qui y voient une sorte de carte d'identité du robot. Afin de vérifier rapidement l'importance de cette approche pour la communauté, le doctorant n'a réalisé que des portraits de robot se laissant modéliser par des fonctions de \mathbb{R}^2 dans \mathbb{R}^2 . En perspective, il serait souhaitable d'automatiser sous forme d'algorithme la démarche initiée et de la généraliser en dimension supérieure.

Chapitre 4

Calcul optimal et théorie de la mesure

4.1 Introduction

J'ai eu la chance de participer à l'un des cours sur l'optimisation polynomiale et le contrôle organisés par le GdR 3273 Mathématiques de l'Optimisation et Applications dont les intervenants étaient : Didier Henrion et Jean-Bernard Lasserre. Leur cours m'a redonné goût à la théorie de la mesure dans le sens où il m'a permis de comprendre que de nombreux problèmes pouvaient se modéliser via un programme linéaire en dimension infinie. En effet, chacun des problèmes de cette longue liste¹ en fait partie :

- trouver le minimum global d'une fonction sur un sous-ensemble de \mathbb{R}^n ,
- résoudre des équations,
- calculer la mesure de Lebesgue d'un sous-ensemble $S \subset \mathbb{R}^n$,
- calculer une borne supérieure à $\mu(S)$ parmi toutes les mesures μ vérifiant certaines conditions sur ses moments,
- calculer la valeur optimale d'un problème de contrôle optimal,
- évaluer la valeur du transport optimal de Kantorovitch,
- évaluer le critère ergodique associé à une chaîne de Markov,
- évaluer des intégrales multivariées,
- calculer des équilibres de Nash,
- évaluer la valeur de $\hat{f}(x)$ où x est un point de \mathbb{R}^n et \hat{f} l'enveloppe convexe de la fonction f
- ...

Tous ces problèmes peuvent se modéliser avec l'aide de la théorie de la mesure de façon plus ou moins directe. Par exemple, le problème de transport optimal de Kantorovitch est déjà classiquement modélisé via la théorie de la mesure [52]. Regardons, à titre d'exemple, comment un problème d'optimisation globale unidimensionnelle non linéaire se laisse modéliser avec la théorie de la mesure. On considère donc le problème suivant :

$$\begin{aligned} \min_{x \in \mathbb{R}} \quad & \varphi(x) \\ \text{tel que} \quad & x \in X \end{aligned} \tag{4.1}$$

1. Cette liste est tirée du livre de Jean-Bernard Lasserre [51].

où X est pour cet exemple un intervalle compact $[a, b]$. Par soucis de simplicité, on suppose que la fonction φ est continue. L'idée est de résoudre ce problème en l'écrivant sous une autre forme :

$$\begin{aligned} \min_{\mu \in \mathcal{M}(\mathbb{R})} \quad & \int_{\mathbb{R}} \varphi(x) d\mu \\ \text{tel que} \quad & \int_X d\mu = 1 \\ & \mu \in \mathcal{M}^+(X) \end{aligned} \tag{4.2}$$

Dans le problème (4.2), l'ensemble $\mathcal{M}(\mathbb{R})$ désigne l'espace vectoriel des mesures sur \mathbb{R} alors que l'ensemble $\mathcal{M}^+(X)$ désigne le cône des mesures positives dont le support est X . La contrainte $\int_X \mu = 1$ impose que le poids de X vis-à-vis de μ est 1. A priori, le problème 4.2 semble plus compliqué que le problème 4.1. En effet, on ne cherche plus un nombre réel x mais une mesure de probabilité μ à support dans X de sorte à minimiser $\int_{\mathbb{R}} \varphi(x) d\mu$.

La fonction coût $\mu \mapsto \int_{\mathbb{R}} \varphi(x) d\mu$ dans (4.2), aussi notée $\mu \mapsto \langle \varphi, \mu \rangle$, est linéaire et les contraintes présentes dans le problème (4.2) sont aussi linéaires. Finalement, le problème (4.2) est programme linéaire *en dimension infinie*.

Il n'est pas très difficile de se convaincre que les deux problèmes ont la même valeur optimale. En effet, étant donnée une densité de probabilité μ à support dans X , l'intégrale $\int_{\mathbb{R}} \varphi(x) d\mu$ n'est autre que la moyenne de φ pondérée par les poids de μ . Quel que soit μ , cette moyenne est nécessairement supérieure à la valeur minimale de φ . Si x^* désigne un optimum du problème (4.1), et μ une mesure de probabilité, on a donc l'inégalité

$$\int_{\mathbb{R}} \varphi(x) d\mu \geq \varphi(x^*).$$

Cette borne inférieure est atteinte en posant μ une masse de Dirac centrée en x^* . En effet, cette mesure, notée δ_{x^*} , vérifie $\int_{\mathbb{R}} d\delta_{x^*} = 1$, $\delta_{x^*} \in \mathcal{M}^+(X)$ et surtout :

$$\int_{\mathbb{R}} \varphi d\delta_{x^*} = \langle \varphi, \delta_{x^*} \rangle = \varphi(x^*).$$

Ce qui démontre que les deux problèmes (4.1) et (4.2) atteignent la même valeur optimale. Une fois modélisé sous la forme d'un programme linéaire en dimension infinie, l'approche de Jean-Bernard Lasserre consiste à discrétiser le programme linéaire en dimension infinie en s'intéressant aux moments de la mesure μ recherchée. Lorsque les données sont polynomiales, ce qui sera par exemple le cas dans le problème (4.1) si φ est un polynôme de la variable x , l'intégrale fonction coût sur les mesures de (4.2) se transforme en somme finie sur les moments (avec autant de termes dans la somme que le degré du polynôme). Grâce à des résultats relativement récents de géométrie semi-algébrique réelle, il est possible d'écrire des conditions nécessaires aux contraintes présentes dans le programme linéaire en dimension infinie. Ces conditions nécessaires portent sur les moments de μ et cette méthodologie donne naissance à des problèmes d'optimisation semi-définie positive. Il est même possible d'imbriquer ces programmes les uns dans les autres. Ceci est appelé une hiérarchie et produit généralement une suite croissante de valeurs optimales convergentes vers la valeur optimale du programme initial. Les fondements et les applications de cette

théorie, qui résout des problèmes des moments généralisés (*generalized moment problems*, GMP) sont décrits dans l'excellent livre [51].

Dans ce chapitre, une approche basée sur le calcul par intervalles pour approcher deux types de programmes linéaires en dimension infinie est proposée. Tout d'abord celui du problème de transport de masse de Kantorovitch dans la section 4.2 puis celui du contrôle optimal dans la section 4.3. Ce choix de présenter deux problèmes est celui qui a été adopté chronologiquement. En effet, dans le cadre du transport optimal, aucun effort supplémentaire de modélisation n'est nécessaire puisqu'il se formule, depuis les travaux de Leonid Kantorovitch [52], via la théorie de la mesure. Le problème de contrôle optimal est quant à lui plus délicat puisqu'il nécessite, un peu comme le problème 4.1 une réécriture via la théorie de la mesure. De plus, les aspects équations différentielles doivent être pris en compte. Les détails de cette modélisation seront présentés dans la section 4.3. Une dernière remarque avant de présenter ces deux problèmes. Avant de commencer ces travaux, nous avons tenté de résoudre des problèmes d'optimisation globale via une méthode basée sur la discrétisation du problème (4.2). Cette tentative s'est trouvée infructueuse comparée à l'énorme quantité de résultats disponibles en optimisation.

4.2 Transport optimal

Le problème de transport optimal a été formalisé par le mathématicien français Gaspard Monge en 1781. Depuis les travaux de Leonid Kantorovitch, celui-ci a été généralisé et est formulé avec la théorie de la mesure. En utilisant des bornes calculées via l'arithmétique des intervalles, on propose dans cette section, une discrétisation garantie du problème de transport de masses de Kantorovitch. Notre discrétisation est spatiale : les supports des mesures sont partitionnés. Le problème est relâché en un programme linéaire en dimension finie de sorte que son optimum donne une borne inférieure au problème initial. Finalement, en utilisant la dualité de Kantorovitch et l'arithmétique des intervalles, une méthode permettant d'obtenir une borne supérieure à l'optimum est aussi fournie.

4.2.1 Introduction

Le transport optimal est un sujet de recherche mathématiques qui a débuté avec ce qu'on appelle maintenant la théorie de Monge "des remblais et déblais" en 1781. Dans les années 1940, Kantorovitch [52] donna une formulation plus moderne à ce problème. De nos jours, le problème consiste à minimiser le coût du transport de deux distributions de masses μ et ν . Sans perte de généralité, la masse totale à déplacer peut être supposée unitaire. Par conséquent, le problème de transport optimal est formulé de la façon suivante :

$$\mathcal{T}(\mu, \nu) = \min_{\pi \in \Gamma(\mu, \nu)} \int_{X \times Y} c(x, y) d\pi(x, y), \quad (4.3)$$

où $\Gamma(\mu, \nu)$ est la collection des mesures de probabilité sur le produit cartésien $X \times Y$ avec comme marginales μ sur X et ν sur Y . Ce problème est un programme linéaire en dimension infinie avec des contraintes convexes. Depuis les découvertes de Yann Brenier dans les années 1990 [53, 54], le transport optimal a été remis sur le devant de la scène

par les mathématiciens et tout particulièrement par Cédric Villani [55]. On pourra aussi consulter Lawrence C. Evans [56] pour une introduction.

La contribution principale de cette section est de fournir des approximations de l'optimum \mathcal{T} . Cette approximation est garantie dans le sens où nous calculerons deux nombres réels $\underline{\mathcal{T}}$ et $\overline{\mathcal{T}}$ de sorte que

$$\underline{\mathcal{T}} \leq \mathcal{T} \leq \overline{\mathcal{T}}.$$

L'approche, détaillée plus tard, peut être résumée de la façon suivante. La borne inférieure $\underline{\mathcal{T}}$ est obtenue en discrétisant spatialement les mesures μ et ν , l'intégrale dans (4.3) est donc remplacée par une somme. Ce schéma numérique et ses propriétés sont présentés dans la sous-section 4.2.2. Les encadrements nécessaires de c , μ et ν sont calculés via l'arithmétique des intervalles (voir le chapitre 2).

Pour la borne supérieure, $\overline{\mathcal{T}}$, le schéma numérique s'appuie sur le dual du problème (4.3). Cette partie est détaillée dans la section 4.2.3. Dans chacune de ces sections, des exemples illustratifs sont donnés.

On pourra comparer cette approche à celle proposée par Dubuc dans [57]. Les deux approches sont voisines mais le schéma de Dubuc nécessite le calcul des optimums de la fonction c sur les sous-ensembles de $X \times Y$.

La relaxation en dimension finie ici présentée est aussi différente de celle proposée dans [58, 59]. Dans [58], les mesures sont approchées par des mesures discrètes (des sommes de Dirac). Comme dans [58], notre méthode génère une suite de programmes linéaires en dimension finie telle que l'optimum \mathcal{T} soit la limite des valeurs optimales de ces programmes. On notera aussi l'existence d'une méthode basée sur une descente du gradient [60].

Par ailleurs, cette contribution peut être comparée avec les résultats obtenus via la méthodologie développée par J.B. Lasserre [51] (Voir Proposition 7.7 p. 177). En effet, son approche basée sur les moments des mesures crée une suite de programmes semi-définis. La suite obtenue $\{\mathcal{T}_i\}_{i \in \mathbb{N}}$ des optimums des programmes semi-définis est croissante et converge vers \mathcal{T} . Le schéma basé sur les moments peut être vu comme une décomposition spectrale, *i.e.* en fréquence, alors que l'approche ici présentée peut être qualifiée de spatiale. L'approche de Jean-Bernard Lasserre suppose que la fonction c est polynomiale et que les supports des mesures μ et ν sont des ensembles semi-algébriques. On peut aussi citer les auteurs suivants qui résolvent numériquement le problème de transport optimal : Anderson [61], Benamou [62] et Mérigot [63].

4.2.2 Bornes inférieures à la valeur optimale

Cette section contient une des deux contributions principales de ce chapitre. Le théorème suivant permet de générer un programme linéaire en dimension finie de sorte que sa valeur optimale soit une borne inférieure à la valeur optimale du problème (4.3). Le théorème est énoncé et sa preuve donnée. Il est illustré par un exemple académique. Finalement, avec des hypothèses de continuité et de compacité, on prouve la convergence du schéma. Autrement dit, on génère une suite de bornes inférieures convergente vers $\mathcal{T}(\mu, \nu)$.

Relaxation en dimension finie

Théorème 13 (Relaxation). Soient μ et ν (avec comme support respectif X et Y) deux mesures absolument continues par rapport à la mesure de Lebesgue λ . Si $\{X_i\}_{1 \leq i \leq n}$ et $\{Y_j\}_{1 \leq j \leq m}$ sont des λ -pavages de X et Y . Supposons que $\mu(X_i) \in [\underline{\mu}_i, \bar{\mu}_i]$, $\nu(Y_j) \in [\underline{\nu}_j, \bar{\nu}_j]$, et qu'il existe des nombres réels $\underline{c}_{ij} \in \mathbb{R}$ tel que $\forall x, y \in X_i \times Y_j, \underline{c}_{ij} \leq c(x, y)$. Si on définit le réel \underline{T} par

$$\begin{aligned} \underline{T} &= \inf_{\pi_{ij} \in \mathbb{R}^n \otimes \mathbb{R}^m} \sum_{i,j} \underline{c}_{ij} \pi_{ij} \\ \text{tel que} \quad \forall i, \underline{\mu}_i &\leq \sum_j \pi_{ij} \leq \bar{\mu}_i, \\ \forall j, \underline{\nu}_j &\leq \sum_i \pi_{ij} \leq \bar{\nu}_j, \\ \forall i, \forall j, \pi_{ij} &\geq 0, \end{aligned} \tag{4.4}$$

alors

$$\underline{T} \leq \mathcal{T}(\mu, \nu).$$

Remarque. Le programme linéaire (4.4) est composé de $n \times m$ variables notées π_{ij} . En effet, le produit tensoriel $\mathbb{R}^n \otimes \mathbb{R}^m$ est isomorphe à $\mathbb{R}^{n \times m}$.

Démonstration. La première partie de la preuve consiste à démontrer que la fonction coût de (4.3) est plus grande que la fonction coût de (4.4). Soit π un plan de transfert de μ à ν alors, la collection $\{X_i \times Y_j\}_{\{(i,j) \in I \times J\}}$ est un π -pavage. En conséquence, on a :

$$\sum_{i,j} \int_{X_i \times Y_j} c(x, y) d\pi(x, y) = \int_{X \times Y} c(x, y) d\pi(x, y).$$

Comme $\forall x, y \in X_i \times Y_j, \underline{c}_{ij} \leq c(x, y)$, on en déduit que

$$\underline{c}_{ij} \pi(X_i \times Y_j) \leq \int_{X_i \times Y_j} c(x, y) d\pi(x, y).$$

Donc :

$$\sum_{i,j} \underline{c}_{ij} \pi_{ij} \leq \int_{X \times Y} c(x, y) d\pi(x, y)$$

avec $\pi_{ij} = \pi(X_i \times Y_j)$.

Dans la seconde partie de la preuve, on démontre que les contraintes de (4.3) impliquent celles de (4.4). Soit π une mesure de probabilité sur $X \times Y$ avec μ et ν comme marginales. Pour tout i , on a

$$\sum_j \pi(X_i \times Y_j) = \mu(X_i).$$

Comme $\mu(X_i)$ est supposé appartenir à $[\underline{\mu}_i, \bar{\mu}_i]$, on en déduit :

$$\forall i, \underline{\mu}_i \leq \sum_j \pi_{ij} \leq \bar{\mu}_i$$

avec $\pi_{ij} = \pi(X_i \times Y_j)$. Finalement, comme π est une mesure de probabilité sur $X \times Y$, alors $\pi_{ij} \geq 0$. \square

Un exemple académique

On considère les deux mesures de probabilité μ et ν , avec $[0, 1]$ comme support, définies par $\mu = dx$ et $\nu = \frac{3}{2}(1 - y^2)dy$. Ces deux mesures de probabilité sont représentées sur la figure 4.1. Pour ce problème de transport, on traite le cas où c est le carré de la distance euclidienne $c(x, y) = (x - y)^2$.

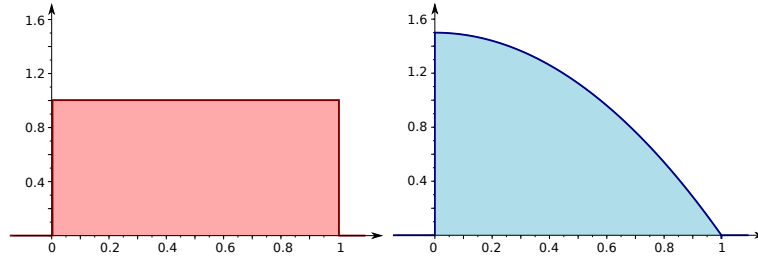


FIGURE 4.1 – Les mesures de probabilité μ et ν .

Divisons l'intervalle $[0, 1]$ avec un pavage régulier de sorte qu'il contienne $n = 6$ éléments

$$X_i = \left[\frac{i-1}{n}, \frac{i}{n} \right],$$

et créons un autre pavage nommé Y_j identique à X_i avec $i, j \in \{1, \dots, n\}$. En accord avec le théorème 13, on génère le programme linéaire suivant :

$$\begin{aligned} \mathcal{I} &= \min_{\pi_{ij} \in \mathbb{R}^n \otimes \mathbb{R}^m} \sum_{i,j} c_{ij} \pi_{ij} \\ \text{tel que} \quad \forall i, \underline{\mu}_i &\leq \sum_j \pi_{ij} \leq \bar{\mu}_i, \\ \forall j, \underline{\nu}_j &\leq \sum_i \pi_{ij} \leq \bar{\nu}_j, \\ \forall i, \forall j, \pi_{ij} &\geq 0, \end{aligned} \tag{4.5}$$

avec les données suivantes :

$$\left(c_{ij} \right)_{1 \leq i, j \leq n} = \begin{pmatrix} 0.000 & 0.000 & 0.027 & 0.111 & 0.250 & 0.444 \\ 0.000 & 0.000 & 0.000 & 0.027 & 0.111 & 0.250 \\ 0.027 & 0.000 & 0.000 & 0.000 & 0.027 & 0.111 \\ 0.111 & 0.027 & 0.000 & 0.000 & 0.000 & 0.027 \\ 0.250 & 0.111 & 0.027 & 0.000 & 0.000 & 0.000 \\ 0.444 & 0.250 & 0.111 & 0.027 & 0.000 & 0.000 \end{pmatrix}.$$

$$\begin{aligned} \underline{\mu} &= (0.166, 0.166, 0.166, 0.166, 0.166, 0.166), \\ \bar{\mu} &= (0.167, 0.167, 0.167, 0.167, 0.167, 0.167), \\ \underline{\nu} &= (0.247, 0.233, 0.205, 0.164, 0.108, 0.039), \\ \bar{\nu} &= (0.248, 0.234, 0.207, 0.165, 0.109, 0.040). \end{aligned}$$

Les bornes inférieures \underline{c}_{ij} ont été calculées en utilisant l'arithmétique par intervalles. Les données $\underline{\nu}, \underline{\mu}, \bar{\nu}$ et $\bar{\mu}$ ont été générées en appliquant la procédure d'intégration présentée dans le corollaire 1.

Ce programme linéaire a été résolu via l'algorithme du simplexe en utilisant le solveur GLPK [64]. Ce solveur, associé à la librairie gmp [65], implémente la méthode du simplexe en s'appuyant sur une arithmétique rationnelle exacte. D'après le théorème 13 et le fait que le solveur retourne comme valeur optimale 1.085×10^{-3} pour le problème (4.5), on en conclut que la valeur optimale \mathcal{T} du programme linéaire en dimension infinie vérifie :

$$1.085 \times 10^{-3} \leq \mathcal{T}.$$

En affinant la partition, cette procédure génère de meilleures bornes inférieures. La figure 4.2 donne les bornes inférieures obtenues pour différentes valeurs de n où n est le cardinal de la partition considérée.

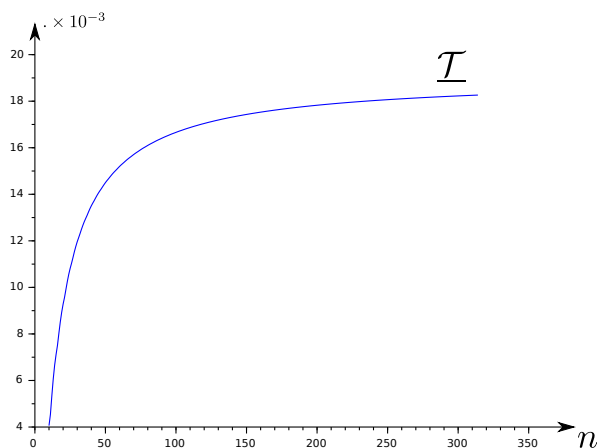


FIGURE 4.2 – Bornes inférieures de \mathcal{T} en fonction de n (avec $m = n$).

Pour cet exemple, des pavages réguliers de X et de Y ont été choisis avec la même cardinalité, i.e. $m = n$. A n fixé, le programme linéaire est de dimension n^2 et est composé de $2n$ contraintes linéaires.

Convergence

Proposition 11 (Convergence). *Soient X, Y deux éléments de $\mathbb{I}\mathbb{R}^k$, $\mu = fd\lambda$, $\nu = gd\lambda$ deux mesures avec comme support X et Y . On suppose que les fonctions $[f]$, $[g]$, $[c] = [\underline{c}, \bar{c}]$ sont des fonctions d'inclusion continues f, g, c vérifiant $\forall x \in X \forall y \in Y, [f](\{x\}) = \{f(x)\}$, $[g](\{y\}) = \{g(y)\}$ et $[c](\{(x, y)\}) = \{c(x, y)\}$. Soient $\{X_i\}_{i \in I}$ et $\{Y_j\}_{j \in J}$ des λ -pavages de X et Y . Soit $h = \max\{\Delta\{X_i\}, \Delta\{Y_j\}\}$. Si $\underline{\mathcal{T}}$ est la valeur optimale de (4.4) alors*

$$\lim_{h \rightarrow 0} \underline{\mathcal{T}} = \mathcal{T}$$

Démonstration. On note par $\tilde{\mathcal{T}}$ la valeur optimale suivante :

$$\begin{aligned} \tilde{\mathcal{T}} &= \inf_{\tilde{\pi}_{ij} \in \mathbb{R}^n \otimes \mathbb{R}^m} \sum_{i,j} \tilde{c}_{ij} \tilde{\pi}_{ij} \\ \text{tel que} \quad &\forall i, \sum_j \tilde{\pi}_{ij} = \mu(X_i), \\ &\forall j, \sum_i \tilde{\pi}_{ij} = \nu(Y_j), \\ &\forall i, \forall j, \tilde{\pi}_{ij} \geq 0, \end{aligned} \tag{4.6}$$

où $\tilde{c}_{ij} = \sup\{c(x, y) \mid (x, y) \in X_i \times Y_j\}$.

L'idée principale consiste à démontrer les deux assertions suivantes. Premièrement, les inégalités $\underline{\mathcal{T}} \leq \mathcal{T} \leq \tilde{\mathcal{T}}$ sont vraies et secondement, la différence $\tilde{\mathcal{T}} - \underline{\mathcal{T}}$ converge vers 0 quand h tend vers 0.

L'inégalité $\underline{\mathcal{T}} \leq \mathcal{T}$ a été démontrée dans le théorème 13. Montrons maintenant l'égalité $\mathcal{T} \leq \tilde{\mathcal{T}}$. Soit $\tilde{\pi}^* = (\tilde{\pi}_{ij}^*)$ une solution optimale du programme (4.6). A partir de $\tilde{\pi}^*$, on définit une mesure $\pi \in \Gamma(\mu, \nu)$ vérifiant $\forall i \forall j, \pi(X_i \times Y_j) = \tilde{\pi}_{i,j}^*$. On a

$$\int_{X \times Y} c(x, y) d\pi(x, y) = \sum_{i,j} \int_{X_i \times Y_j} c(x, y) d\pi(x, y).$$

Comme $\forall (x, y) \in X_i \times Y_j, c(x, y) \leq \tilde{c}_{ij}$, on en déduit :

$$\int_{X \times Y} c(x, y) d\pi(x, y) \leq \sum_{i,j} \int_{X_i \times Y_j} \tilde{c}_{ij} d\pi(x, y) = \sum_{i,j} \tilde{c}_{ij} \tilde{\pi}_{i,j}^* = \tilde{\mathcal{T}},$$

avec $\pi \in \Gamma(\mu, \nu)$. C'est à dire que pour chaque solution optimale de (4.6), il est possible de créer une solution admissible de (4.3). Cette solution admissible de (4.3) est plus petite que $\tilde{\mathcal{T}}$. En particulier, avec \mathcal{T} la valeur optimale de (4.3), on en conclut $\mathcal{T} \leq \tilde{\mathcal{T}}$.

Montrons maintenant que $\tilde{\mathcal{T}} - \underline{\mathcal{T}}$ converge vers 0 quand h tend vers 0. L'idée principale est de démontrer que les deux quantités $\tilde{\mathcal{T}} - \underline{\mathcal{T}}$ et $\mathcal{T} - \underline{\mathcal{T}}$ convergent vers 0 quand h tend vers 0 avec \mathcal{T} la valeur optimale du programme linéaire en dimension finie :

$$\begin{aligned} \mathcal{T} &= \inf_{\pi_{ij} \in \mathbb{R}^n \otimes \mathbb{R}^m} \sum_{i,j} c_{ij} \pi_{ij} \\ \text{tel que} \quad &\forall i, \sum_j \pi_{ij} = \mu(X_i), \\ &\forall j, \sum_i \pi_{ij} = \nu(Y_j), \\ &\forall i, \forall j, \pi_{ij} \geq 0. \end{aligned} \tag{4.7}$$

On a :

$$\underline{\mathcal{T}} \leq \underline{\mathcal{T}} \leq \mathcal{T} \leq \tilde{\mathcal{T}}.$$

Ici, on suit la démarche de la preuve de la proposition 2. Grâce au théorème de Heine-Cantor, comme $[c]$ est uniformément continue, on en déduit

$$\forall \varepsilon > 0 \exists \delta > 0, h < \delta \Rightarrow (\tilde{c}_{ij} - c_{ij}) < \varepsilon.$$

Par conséquent, $\forall(\pi_{ij}) \in \mathbb{R}^n \otimes \mathbb{R}^m$ vérifiant

$$\begin{cases} \forall i, \sum_j \pi_{ij} = \mu(X_i), \\ \forall j, \sum_i \pi_{ij} = \nu(Y_j), \end{cases}$$

on a

$$\forall \varepsilon > 0 \exists \delta > 0, h < \delta \Rightarrow \sum_{ij} (\tilde{c}_{ij} - c_{ij}) \pi_{ij} < \varepsilon \mu(X). \quad (4.8)$$

En particulier, avec (π_{ij}) une solution optimale de (4.7), on écrit

$$\forall \varepsilon > 0 \exists \delta > 0, h < \delta \Rightarrow \sum_{ij} \tilde{c}_{ij} \pi_{ij} - \sum_{ij} c_{ij} \pi_{ij} < \varepsilon \mu(X), \quad (4.9)$$

i.e.

$$\forall \varepsilon > 0 \exists \delta > 0, h < \delta \Rightarrow \sum_{ij} \tilde{c}_{ij} \pi_{ij} - \underline{\mathcal{T}} < \varepsilon \mu(X),$$

i.e.

$$\forall \varepsilon > 0 \exists \delta > 0, h < \delta \Rightarrow \sum_{ij} \tilde{c}_{ij} \pi_{ij} < \varepsilon \mu(X) + \underline{\mathcal{T}}.$$

Comme $\tilde{\mathcal{T}} \leq \sum_{ij} \tilde{c}_{ij} \pi_{ij}$, on a

$$\forall \varepsilon > 0 \exists \delta > 0, h < \delta \Rightarrow \tilde{\mathcal{T}} < \varepsilon \mu(X) + \underline{\mathcal{T}}.$$

Par conséquent,

$$\forall \varepsilon > 0 \exists \delta > 0, h < \delta \Rightarrow \tilde{\mathcal{T}} - \underline{\mathcal{T}} < \varepsilon \mu(X).$$

Montrons maintenant que $\underline{\mathcal{T}} - \underline{\mathcal{T}}$ converge vers 0 quand h tend vers 0. Les nombres réels $\underline{\mathcal{T}}$ et $\underline{\mathcal{T}}$ sont les valeurs optimales des programmes linéaires en dimension finie (4.4) et (4.7). En utilisant les notations matricielles, on a :

$$\begin{aligned} \underline{\mathcal{T}} &= \inf \{ c^T \pi \mid b = A\pi = \bar{b}, \quad \pi \geq 0 \}, \\ \underline{\mathcal{T}} &= \inf \{ c^T \pi \mid b - \underline{\varepsilon} \leq A\pi \leq b + \bar{\varepsilon}, \quad \pi \geq 0 \} \end{aligned}$$

avec

$$A = \underbrace{\begin{pmatrix} \underbrace{1 \cdots 1}_m & & & & \\ & \ddots & & & \\ & & & & \underbrace{1 \cdots 1}_m \\ \hline 1 & & & 1 & \\ & \ddots & \cdots & & \ddots \\ \underbrace{\quad \quad \quad}_m & & & \underbrace{\quad \quad \quad}_m & \\ \hline & & & & \end{pmatrix}}_{n \text{ fois}}, \pi = \begin{pmatrix} \pi_{11} \\ \vdots \\ \pi_{1m} \\ \pi_{21} \\ \vdots \\ \pi_{2m} \\ \vdots \\ \pi_{m1} \\ \vdots \\ \pi_{nm} \end{pmatrix}, b = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \\ \nu_1 \\ \vdots \\ \nu_m \end{pmatrix},$$

et

$$\underline{\epsilon} = \begin{pmatrix} \mu_1 - \underline{\mu}_1 \\ \vdots \\ \mu_n - \underline{\mu}_n \\ \nu_1 - \underline{\nu}_1 \\ \vdots \\ \nu_m - \underline{\nu}_m \end{pmatrix}, \bar{\epsilon} = \begin{pmatrix} \bar{\mu}_1 - \mu_1 \\ \vdots \\ \bar{\mu}_n - \mu_n \\ \bar{\nu}_1 - \nu_1 \\ \vdots \\ \bar{\nu}_m - \nu_m \end{pmatrix}.$$

Par conséquent, les deux programmes linéaires partagent la même fonction coût c , et la matrice de contraintes A . Comme les fonctions d'inclusion $[f]$ et $[g]$ sont supposées vérifier, $\forall x \in X, [f](\{x\}) = \{f(x)\}$ et $\forall y \in Y, [g](\{y\}) = \{g(y)\}$, la proposition 2 permet d'affirmer que les nombres réels $\underline{\epsilon}$ et $\bar{\epsilon}$ peuvent aussi petits que l'on souhaite. Le nombre réel $\underline{\mathcal{T}}$ peut être vu comme la valeur optimale d'un programme linéaire voisin de (4.7).

En utilisant les bornes proposées par [66], on montre que $\underline{\mathcal{T}}$ converge vers \mathcal{T} dès que $\underline{\epsilon}$ et $\bar{\epsilon}$ convergent vers 0. □

4.2.3 Bornes supérieures à la valeur optimale

Dans cette section, on utilise la dualité de Kantorovitch pour proposer une borne supérieure à la valeur optimale. On rappelle tout d'abord que le théorème de Kantorovitch énonce le fait suivant :

Théorème 14 (Dualité de Kantorovitch). *Soient μ et ν deux mesures de probabilité et $c : X \times Y \rightarrow \mathbb{R}_+ \cup \{+\infty\}$ une fonction semi-continue inférieurement. On définit*

$$\begin{aligned} \mathcal{T}'(\mu, \nu) &= \sup_{\varphi \in \mathcal{C}(X), \psi \in \mathcal{C}(Y)} \int_X \varphi d\mu + \int_Y \psi d\nu, \\ \text{tel que} & \quad \varphi(x) + \psi(y) \leq c(x, y), \end{aligned}$$

alors

$$\mathcal{T}(\mu, \nu) = \mathcal{T}'(\mu, \nu).$$

Avec $\mathcal{C}(X)$ l'ensemble des fonctions réelles continues et bornées sur X .

La preuve de ce théorème peut être trouvée dans [67].

Relaxation en dimension finie

Théorème 15 (Relaxation). *Soient μ et ν (à supports compacts X et Y) deux mesures absolument continues par rapport à la mesure de Lebesgue. Soit $\{X_i\}_i$ et $\{Y_j\}_j$ deux λ -pavages de X et de Y . On suppose que $\mu(X_i) \leq \bar{\mu}_i$, $\nu(Y_j) \leq \bar{\nu}_j$, et $\forall (x, y) \in X_i \times Y_j, c(x, y) \leq \bar{c}_{ij}$,*

$$\begin{aligned} \bar{\mathcal{T}} &= \sup_{(\varphi_i) \in \mathbb{R}^n, (\psi_j) \in \mathbb{R}^m} \sum_i \varphi_i \bar{\mu}_i + \sum_j \psi_j \bar{\nu}_j \\ \text{tel que} & \quad \varphi_i + \psi_j \leq \bar{c}_{ij}, \end{aligned} \tag{4.10}$$

alors

$$\mathcal{T} \leq \bar{\mathcal{T}}.$$

Démonstration. Soit (φ, ψ) un élément de $\mathcal{C}(X) \times \mathcal{C}(Y)$. Comme chaque X_i et Y_j est compact, alors il existe des nombres réels $\varphi_i = \sup_{x \in X_i} \varphi(x)$, $\psi_j = \sup_{y \in Y_j} \psi(y)$ tel que

$$\int_{X_i} \varphi(x) d\mu \leq \varphi_i \mu(X_i) \text{ et } \int_{Y_j} \psi(y) d\nu \leq \psi_j \nu(Y_j).$$

Comme $\mu(X_i)$ est supposé inférieur à $\bar{\mu}_i$, on en déduit :

$$\int_X \varphi d\mu + \int_Y \psi d\nu \leq \sum_i \varphi_i \bar{\mu}_i + \sum_j \psi_j \bar{\nu}_j.$$

La seconde partie de la preuve consiste à démontrer que la condition $\forall(x, y), \varphi(x) + \psi(y) \leq c(x, y)$ implique $\forall(i, j), \varphi_i + \psi_j \leq \bar{c}_{ij}$. Supposons que $\forall(x, y), \varphi(x) + \psi(y) \leq c(x, y)$, alors

$$\forall(x, y) \in X_i \times Y_j, \varphi(x) + \psi(y) \leq c(x, y) \leq \bar{c}_{ij}.$$

Comme les ensembles X_i et Y_j sont supposés compacts, alors il existe $(\xi, \zeta) \in X_i \times Y_j$ tel que

$$\varphi(\xi) = \sup_{x \in X_i} \varphi(x) = \varphi_i, \text{ et } \psi(\zeta) = \sup_{y \in Y_j} \psi(y) = \psi_j.$$

On conclut avec

$$\varphi_i + \psi_j \leq \bar{c}_{ij}.$$

□

Un exemple académique

Ici, on traite le même exemple que dans la section précédente. Considérons le pavage régulier de l'intervalle unitaire $[0, 1]$. Le théorème 15 permet de créer, avec $m = n = 6$, le programme linéaire en dimension finie suivant :

$$\begin{aligned} \bar{T} = & \sup_{(\varphi_i) \in \mathbb{R}^n, (\psi_j) \in \mathbb{R}^m} \sum_i \varphi_i \bar{\mu}_i + \sum_j \psi_j \bar{\nu}_j \\ & \text{tel que } \varphi_i + \psi_j \leq \bar{c}_{ij} \end{aligned} \quad (4.11)$$

avec les données suivantes :

$$\bar{\mu} = (0.167, 0.167, 0.167, 0.167, 0.167, 0.168),$$

$$\bar{\nu} = (0.248, 0.234, 0.207, 0.165, 0.109, 0.040),$$

et

$$(\bar{c}_{ij})_{1 \leq i, j \leq n} = \begin{pmatrix} 0.028 & 0.112 & 0.250 & 0.445 & 0.695 & 1.000 \\ 0.112 & 0.028 & 0.112 & 0.250 & 0.445 & 0.695 \\ 0.250 & 0.112 & 0.028 & 0.112 & 0.250 & 0.445 \\ 0.445 & 0.250 & 0.112 & 0.028 & 0.112 & 0.250 \\ 0.695 & 0.445 & 0.250 & 0.112 & 0.028 & 0.112 \\ 1.000 & 0.695 & 0.445 & 0.250 & 0.112 & 0.028 \end{pmatrix}.$$

Comme pour l'exemple de la section 4.2.2, les nombres réels \bar{c}_{ij} ont été calculés avec l'arithmétique des intervalles. Ce programme linéaire admet $\bar{\mathcal{T}} = 0.091$ pour valeur optimale. On déduit, grâce au théorème 15, que

$$\mathcal{T} \leq 91 \times 10^{-3}.$$

Cette dernière borne supérieure peut être améliorée en augmentant m et n . La figure 4.3 illustre le fait que ces bornes inférieures sont de plus en plus précises avec n (ici on a $m = n$).

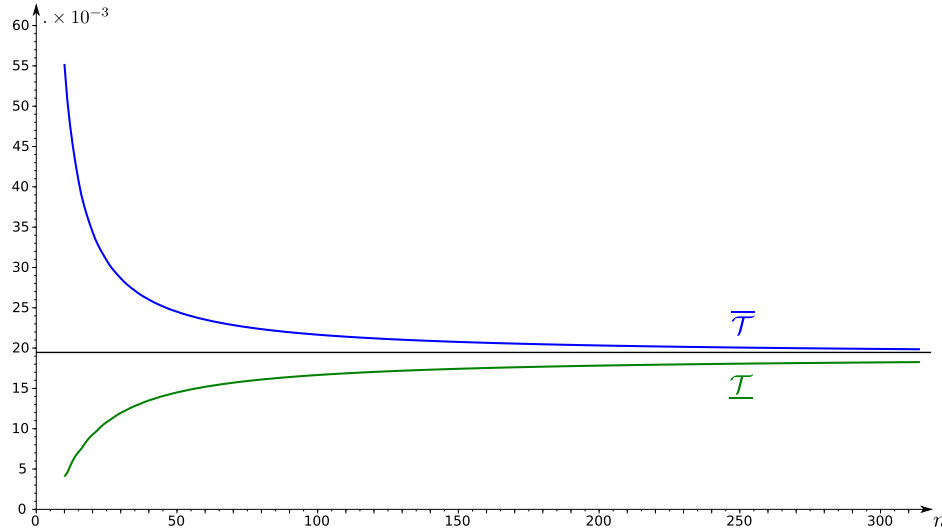


FIGURE 4.3 – Encadrement de \mathcal{T} en fonction de n (avec $m = n$).

En général, le programme linéaire a $m+n$ variables et est composé de $m \times n$ contraintes.

Remarque. Les vecteurs $\bar{\mu}$ et $\bar{\nu}$ doivent être générés avec attention. En effet, si $\sum_i \bar{\mu}_i \neq \sum_j \bar{\nu}_j$, alors les solutions de (4.11) peuvent ne pas être bornées. On remarque que le nombre réel $\bar{\mu}_6$ a été choisi égal à 0.168 à la place de 0.167 dans le but d'imposer l'égalité $\sum_i \bar{\mu}_i = \sum_j \bar{\nu}_j$.

Remarque. Comme la fonction c est convexe et les mesures μ et ν sont à support sur la droite réelle, on prouve que

$$\mathcal{T} = \int_0^1 |G^{-1}(t) - F^{-1}(t)|^2 dt,$$

où F^{-1} et G^{-1} sont les réciproques des fonctions de répartition de f et g , i.e.

$$F(x) = \int_{-\infty}^x d\mu, \text{ et } G(x) = \int_{-\infty}^x d\nu.$$

La preuve de ce résultat peut être trouvée dans [67]. Dans notre cas, on a

$$\begin{aligned} F(x) &= x, & \forall x \in [0, 1], \\ G(x) &= \frac{3}{2}x - \frac{1}{2}x^3, & \forall x \in [0, 1], \end{aligned}$$

et

$$\begin{aligned} F^{-1}(t) &= t, & \forall t \in [0, 1], \\ G^{-1}(t) &= 2 \cos \left(\frac{\arccos(-t)}{3} + \frac{4\pi}{3} \right), & \forall t \in [0, 1]. \end{aligned}$$

Finalement, avec une technique classique d'intégration, on obtient :

$$\mathcal{T} = \int_0^1 |G^{-1}(t) - F^{-1}(t)|^2 dt \simeq 19,0476 \times 10^{-3}.$$

Cette valeur est représentée par la ligne horizontale sur la figure 4.3 et est évidemment comprise entre les deux bornes calculées.

Applications numériques

Dans cette section, on s'intéresse au problème suivant

$$\mathcal{T}(\mu, \nu) = \min_{\pi \in \Gamma(\mu, \nu)} \int_{X \times Y} |x - y| d\pi(x, y), \quad (4.12)$$

avec deux mesures de probabilité μ et ν définies par $\mu = dx$ et $\nu = \frac{3}{2}(1 - y^2)dy$, dont le support est $[0, 1] = X = Y$. On considère une subdivision régulière de l'intervalle unité avec n intervalles. La figure 4.4 montre les bornes obtenues avec la méthode proposée.

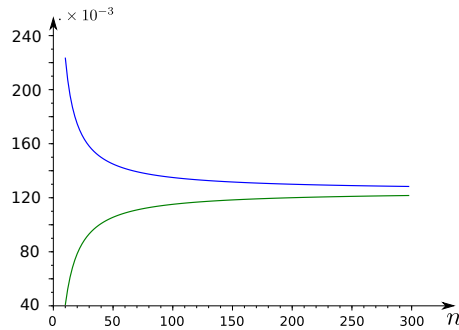


FIGURE 4.4 – Bornes inférieures et supérieures de \mathcal{T} en fonction n (avec $m = n$).

Finalement, comme indiqué dans l'introduction, cette approche n'est pas limitée aux données polynomiales. Le dernier exemple traité est le suivant :

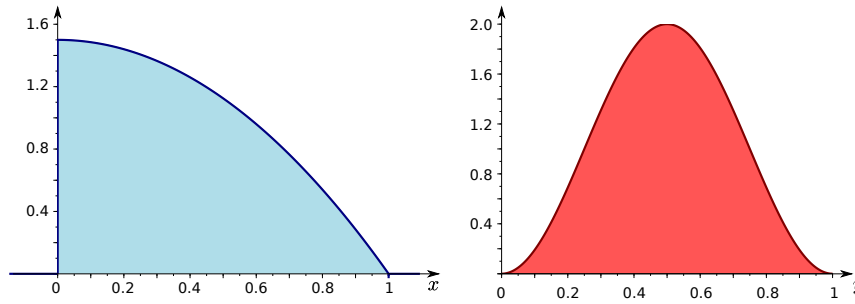
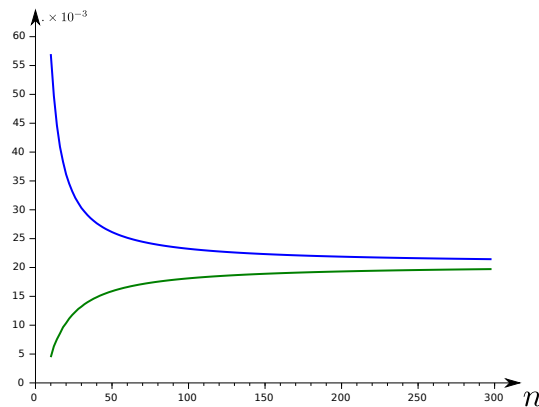
$$\mathcal{T}(\mu, \nu) = \min_{\pi \in \Gamma(\mu, \nu)} \int_{X \times Y} (x - y)^2 d\pi(x, y), \quad (4.13)$$

avec μ et ν à support $[0, 1] = X = Y$ définies par $\mu = \frac{3}{2}(1 - x^2)dx$ et $\nu = 2 \sin^2(\pi y)dy$. Les deux mesures sont représentés par la figure 4.5.

Les bornes inférieures et supérieures de \mathcal{T} sont données sur la figure 4.6.

4.2.4 Conclusion

Dans cette section, deux schémas de discrétisation ont été proposés pour encadrer la valeur optimale du problème de transport de Kantorovitch. Notre approche est relativement générale dans le sens où seule l'existence d'une fonction d'inclusion pour la fonction c et la possibilité d'encadrer numériquement μ et ν sont nécessaires (voir section 2.3).

FIGURE 4.5 – Mesures de probabilité μ et ν .FIGURE 4.6 – Bornes inférieures et supérieures de \mathcal{T} en fonction n pour $\mu = \frac{3}{2}(1-x^2)dx$, $\nu = 2\sin^2(\pi y)dy$ et $c(x, y) = (x-y)^2$.

Le coté effectif de ces algorithmes a été confirmé par une implémentation en C++. Ce code a permis de construire les figures comme la figure 4.6. Il est disponible sur le web <http://perso-laris.univ-angers.fr/~delanoue/>.

4.3 Contrôle optimal

Cette section présente un schéma d'approximation pour les problèmes de contrôle optimal en utilisant des programmes linéaires de dimension finie et l'arithmétique des intervalles. La présentation se décompose en deux étapes. Tout d'abord en suivant l'approche de Vinter [68] et en utilisant des mesures d'occupation, le problème de contrôle optimal est écrit sous la forme d'un problème de programmation linéaire de dimension infinie (formulation faible). Dans un second temps, grâce à l'arithmétique des intervalles, le programme linéaire en dimension infinie est relâché en un programme linéaire en dimension finie.

Le fait que la valeur optimale du programme linéaire en dimension finie fournisse une borne inférieure au problème de contrôle optimal est formalisé dans le théorème 17.

De plus, selon la qualité de la discrétisation et la taille de la famille de fonctions test choisie, les valeurs optimales de la suite de programmes linéaires forment une suite qui converge vers la valeur optimale du problème de contrôle optimal initial. La démonstration

sera donnée dans la preuve du théorème 19, de plus des exemples illustreront le principe de la méthodologie.

D'une certaine manière, la méthode ici proposée, peut être vue comme complémentaire aux méthodes directes qui donnent des bornes supérieures. Les bornes inférieures calculées par notre méthode peuvent être utilisées comme critère de qualité d'une commande obtenue par une méthode directe.

4.3.1 Introduction - État de l'art

Dans toute la suite de cette section, on s'intéresse au problème de contrôle optimal suivant :

$$\begin{aligned}
 J^*(t, x_t) &= \min_{u: [t, T] \rightarrow U} \int_t^T h(\tau, x(\tau), u(\tau)) d\tau + H(x(T)) \\
 \text{tel que} \quad &x(t) = x_t \\
 &\dot{x}(\tau) = f(\tau, x(\tau), u(\tau)), \forall \tau \in [t, T], \\
 &x(\tau) \in X, \forall \tau \in [t, T], \\
 &x(T) \in K. \\
 &u(\tau) \in U, \forall \tau \in [t, T].
 \end{aligned} \tag{4.14}$$

Dans cette formulation, l'ensemble $U \subset \mathbb{R}^m$ est compact et permet d'imposer une zone où la commande doit être choisie, le nombre réel t est l'instant initial, $x : (t, T) \rightarrow \mathbb{R}^n$ est classiquement nommée trajectoire, $u : (t, T) \rightarrow \mathbb{R}^m$ est l'action (ou bien la commande). La fonction $h : (0, T) \times \mathbb{R}^n \times U \rightarrow \mathbb{R}$ est appelée le coût courant (ou lagrangien), alors que la fonction $H : \mathbb{R}^n \rightarrow \mathbb{R}$ est nommée le coût final. Le champ de vecteurs $f : (t, T) \times \mathbb{R}^n \times U \rightarrow \mathbb{R}^n$ est supposé Lipschitz et borné. L'élément x_t de X est la condition initiale. L'ensemble X où l'état peut vivre est supposé compact. Par simplicité, on suppose ici que l'instant final $T \in \mathbb{R}$ est fixe. L'ensemble K est appelé la cible et correspond à une région où doit se trouver l'état à l'instant final.

En s'appuyant sur le principe d'optimalité de Bellman [69], on peut écrire l'équation aux dérivées partielles suivantes classiquement appelée l'équation de Hamilton-Jacobi-Bellman (HJB) :

$$\frac{\partial J^*}{\partial t} = - \min_{u \in U} \left\{ h(t, x, u) + \frac{\partial J^*}{\partial x} \cdot f(t, x, u) \right\}. \tag{4.15}$$

L'équation (4.15), avec la condition finale $J^*(T, x) = H(x)$ forme un système dynamique en dimension infinie. Sa solution J^* est classiquement appelée la fonction valeur car $J^*(t, x_t)$ est la valeur optimale obtenue en résolvant l'équation (4.14) partant de l'état x_t à l'instant t . L'équation (4.15) a été étudiée depuis de nombreuses années. Du point de vue théorique, l'existence, l'unicité et la régularité des solutions de (4.15) a été le point de départ de la théorie des solutions de viscosité. En effet, la fonction valeur n'est pas nécessairement différentiable, même dans certains cas simples. Les équations HJB n'ont en général pas de solution classique [70, 71, 72].

La théorie des solutions de viscosité reste un outil puissant pour les problèmes de contrôle optimal. Néanmoins, pour la plupart des exemples, il n'y a pas de solution explicite – comprendre expression algébrique – à cette équation et donc des méthodes numériques sont nécessaires.

De nombreux types d'algorithmes numériques pour résoudre le problème de contrôle optimal (4.14) ont été proposés. Le premier type est basé sur la discrétisation du problème de contrôle optimal (4.14) en un programme non linéaire en dimension finie (et grande). Dans cette approche, l'équation différentielle ordinaire est d'abord remplacée par un schéma numérique issu de la théorie classique des méthodes de résolution numérique d'EDO [73, 74]. Le problème de programmation non linéaire est ensuite résolu, le plus souvent, par un algorithme de programmation quadratique séquentiel ou une méthode de point intérieur [75]. La discrétisation peut aussi être choisie afin d'obtenir un problème de programmation dynamique de dimension finie qu'il reste à résoudre (voir par exemple [76]).

Le deuxième type d'approche consiste à résoudre l'équation HJB (4.15) par un schéma de discrétisation d'équations aux dérivées partielles [77, 78, 79]. L'approximation semi-lagrangienne est aussi une technique numérique pour résoudre l'équation (4.15) : des méthodes d'interpolation sont combinées avec des schémas numériques d'intégration, comme Runge-Kutta, de l'équation différentielle ordinaire afin de générer une approximation de la fonction valeur [80].

Une autre approche classique pour traiter le problème (4.14) est s'appuyer sur le principe du maximum de Pontryagin. Cet outil théorique est un cas particulier de l'équation d'Euler-Lagrange du calcul de variations. Le principe du maximum de Pontryagin fournit une condition *nécessaire* pour qu'un contrôle soit optimal. D'un point de vue technique, cette approche crée une équation différentielle ordinaire de dimension $2n$ avec des valeurs aux limites. Ce problème peut être résolu par des méthodes de tir afin de faire coïncider les deux extrémités. Ce type d'approche utilise classiquement la méthode de Newton pour trouver des racines. En général, il faut une bonne approximation du contrôle optimal pour trouver la solution optimale en raison de la faible taille du bassin d'attraction de l'itération de la méthode Newton.

Le lecteur est invité à consulter [81] afin de connaître les détails mathématiques de la construction de l'équation (4.15) ou bien l'équation différentielle ordinaire obtenue via le principe du maximum de Pontryagin.

Par ailleurs, une méthode ensembliste a été développée par Oliver Junge et Hinke Osinga dans [82]. Leur méthode se décompose en deux phases. Dans une première étape, l'espace des états est partitionné en sous-ensemble noté X_i avec $i \in I$. Puis un graphe orienté valué est construit. Ses noeuds sont les ensembles X_i et les poids sur ses arcs sont des nombres réels inférieurs au coût permettant de relier X_i et X_j . Finalement, un algorithme standard de plus court chemin de la théorie des graphes permet de construire une approximation de la fonction valeur et des trajectoires optimales approchées.

Dans la littérature, on peut aussi citer l'approche récente s'appuyant sur la linéarité tropicale de l'équation de HJB. On trouvera, dans les travaux de Akian et al., un analogue max-plus de la méthode des éléments finis de Petrov-Galerkin. On pourra consulter [83] ainsi que les références citées.

La méthode ici présentée ne cherche pas à directement résoudre (4.15) ou bien à exploiter directement le principe du maximum. D'un point de vue théorique, le problème de contrôle optimal (4.14) *non linéaire* peut être formulé sous la forme d'un programme *linéaire* en dimension infinie.

L'espace de recherche de ce programme est l'espace des mesures positives. L'équation différentielle apparaissant dans (4.14) s'écrit sous la forme d'une contrainte linéaire

classiquement nommée équation de Liouville [67]. Cette approche par programmation linéaire pour le contrôle optimal est décrite dans [84] et une généralisation aux inclusions différentielles a été proposée par Vinter. On pourra aussi consulter [68, 85, 86].

Dans [87], un schéma d'approximation pour les programmes linéaires en dimension infinie est donné. De plus, sous certaines hypothèses, la valeur du programme linéaire en dimension infinie peut être approchée, autant que souhaité, par les valeurs optimales obtenues des programmes en dimension finie. Pour des aspects plus effectifs, on pourra aussi consulter [88, 89].

Dans le contexte polynomial, *i.e.* si le champ de vecteurs f , les fonctions h et H sont polynomiaux et les ensembles X, U et K sont des semi-algébriques, une approche algébriquement élégante et efficace a été proposée par J.B. Lasserre [51] (voir le chapitre 10). L'idée est de relâcher le programme linéaire en dimension infinie en cherchant les moments de la mesure optimale. Les fondements et les applications de cette théorie qui résout des *generalized moment problems* sont décrits dans [51]. La résolution de ces hiérarchies peut aider à la synthèse de contrôleur dans le contexte des systèmes dynamiques [90, 91].

Organisation des sections suivantes La section 4.3.2 rappelle les définitions de mesures d'occupation et comment à partir d'une équation différentielle ordinaire donnée, on peut formuler une équation linéaire sur les mesures d'occupation associées. A partir de ces résultats, la section 4.3.3 donne la formulation faible du problème (4.14). Finalement, les méthodes d'intégration d'une fonction vis-à-vis d'une mesure énoncées au chapitre 2 nous permettent, dans la section 4.3.4, de relâcher ce programme en un programme en dimension finie. Des exemples illustrent la méthode et sont donnés dans la section 4.3.5. Comme notre approche n'est pas initialement limitée au cas polynomial, on proposera pour le second exemple une famille de fonctions tests seulement différentiables. L'avantage principal de cette classe de fonctions est sa capacité à générer des matrices de contraintes creuses.

4.3.2 Mesures d'occupation

Dans cette section, on explique comment des mesures d'occupation peuvent être obtenues à partir d'une courbe. En particulier, on donne la définition des mesures d'occupation quand la courbe est paramétrée, *i.e.* quand la courbe est définie par une fonction continue d'un intervalle réel dans un espace topologique.

Dans le cas où la courbe est solution d'une équation différentielle ordinaire, une contrainte linéaire peut être construite sur les mesures d'occupation associées à la courbe. Le détail de cette méthode pour obtenir cette équation linéaire est décrite dans la proposition 13.

Comme classiquement en théorie de la mesure, étant donnée une mesure, on peut définir l'intégrale d'une fonction vis-à-vis de cette mesure. Dans le cas particulier d'une mesure d'occupation μ associée à la courbe γ , l'intégration d'une fonction φ vis-à-vis de μ est égale l'intégration temporelle de $\varphi \circ \gamma$. Cette relation entre intégration spatiale et intégration temporelle est formalisée par la proposition 12.

Mesures d'occupation associé à une courbe paramétrée

Étant donnée une courbe paramétrée $\gamma : [0, T] \ni t \rightarrow \gamma(t) \in E$, il est possible de définir deux mesures d'occupation. La première notée μ_γ compte, comme pourrait le faire un chronomètre, le temps qu'une trajectoire reste à l'intérieur d'un ensemble donné. La seconde mesure ν_γ renseigne sur l'endroit où termine cette trajectoire. Ces mesures μ_γ et ν_γ sont formellement définies par les définitions 16 et 17.

Définition 16 (Mesure d'occupation μ_γ sur E associée à une courbe paramétrée γ). Soit $\gamma : [0, T] \mapsto E$ une courbe paramétrée, on définit μ_γ la mesure sur E par

$$\mu_\gamma(X) = \int_0^T 1_X(\gamma(t)) dt \text{ avec } 1_X(x) = \begin{cases} 1 & \text{si } x \in X, \\ 0 & \text{autrement.} \end{cases}$$

La définition 16 est illustrée par la figure 4.7. L'interprétation géométrique est la suivante, la trajectoire rentre dans l'ensemble X_1 à l'instant t_1 et en sort à l'instant t_2 . Par conséquent, $\mu_\gamma(X_1) = t_2 - t_1$.

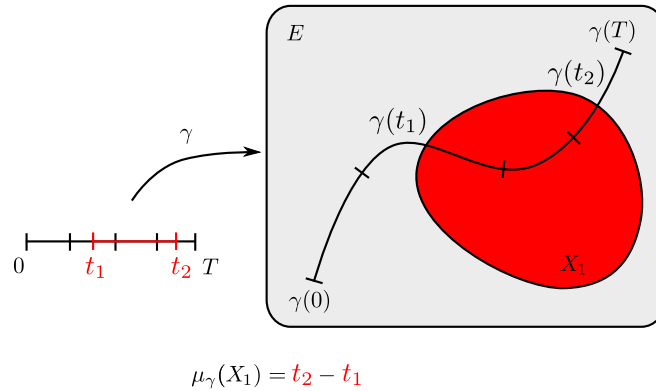


FIGURE 4.7 – Mesure d'occupation μ_γ associé à la courbe γ .

Exemple 20. Soit γ la courbe paramétrée suivante

$$\begin{aligned} \gamma : [0, \pi] &\rightarrow \mathbb{R}^2 \\ t &\mapsto \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}. \end{aligned} \quad (4.16)$$

Si B_O est la boule unitaire euclidienne \mathbb{R}^2 , alors on a $\mu_\gamma(B_O) = \pi$. On peut aussi écrire, par exemple,

$$\mu_\gamma(\{(x, y) \in \mathbb{R}^2 \mid x \geq 0\}) = \frac{\pi}{2}.$$

On remarque que le support de la mesure μ_γ associée à la courbe paramétrée γ est le sous-ensemble $\gamma([0, T])$ classiquement appelée courbe géométrique. Afin de bien comprendre cette notion de mesure d'occupation, on peut faire un parallèle avec la mesure de Dirac. En effet, une mesure de Dirac concentre une masse en un ensemble ponctuel alors qu'en général une mesure d'occupation répartit une masse donnée linéiquement, c'est à dire le long d'un ensemble unidimensionnel. Sur l'exemple précédent, la mesure d'occupation μ_γ est définie sur \mathbb{R}^2 , elle n'est ni absolument continue vis-à-vis de la mesure de Lebesgue sur \mathbb{R}^2 , ni une somme de Dirac.

Définition 17 (Mesure d'occupation à l'instant final ν_γ sur E associée à la courbe γ). Soit $\gamma : [0, T] \mapsto E$ une courbe paramétrée, on définit ν_γ par

$$\nu_\gamma(X) = \delta_{\gamma(T)}(X),$$

avec δ_x une masse de Dirac centrée sur le point x .

Cette notion est illustrée par la figure 4.8.

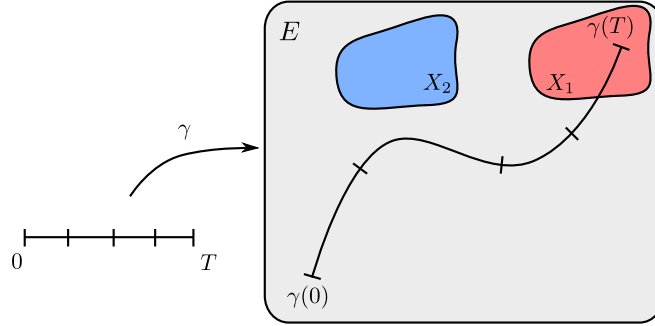


FIGURE 4.8 – Comme $\gamma(T) \notin X_2$, $\nu_\gamma(X_2) = 0$ alors que $\nu_\gamma(X_1) = 1$.

Exemple 21. Soit γ la courbe paramétrée suivante

$$\begin{aligned} \gamma : [0, 1] &\rightarrow \mathbb{R} \\ t &\mapsto t. \end{aligned} \tag{4.17}$$

Dans ce cas, la mesure d'occupation μ_γ n'est rien d'autre que la classique mesure de Lebesgue λ restreinte à l'intervalle $[0, 1]$ et la mesure d'occupation finale ν_γ est une masse de Dirac centrée sur $x = 1$ classiquement notée δ_1 .

Intégration vis-à-vis d'une mesure d'occupation

Cette section propose de rappeler la relation qui existe entre l'intégration d'une fonction donnée vis-à-vis d'une mesure d'occupation et l'intégration temporelle.

Proposition 12. Soit $\varphi : E \mapsto \mathbb{R}$ une fonction et $\gamma : [0, T] \mapsto E$ une courbe, alors

$$\langle \mu_\gamma, \varphi \rangle = \int_X \varphi(x) d\mu_\gamma = \int_0^T \varphi(\gamma(t)) dt. \tag{4.18}$$

La proposition précédente montre que l'intégration temporelle de $\varphi \circ \gamma$ peut être remplacée par une intégration spatiale de φ sur E vis-à-vis de la mesure μ_γ . On peut une fois de plus remarquer que le support de μ_γ est $\gamma([0, T])$ qui est souvent un sous-ensemble unidimensionnel de X . En particulier, chaque sous-ensemble de X qui n'intersecte pas $\gamma([0, T])$ ne compte pas dans l'intégration spatiale vis-à-vis de μ_γ . La proposition 12 est illustrée par la figure 4.9.

Démonstration. Voir [51] par exemple. □

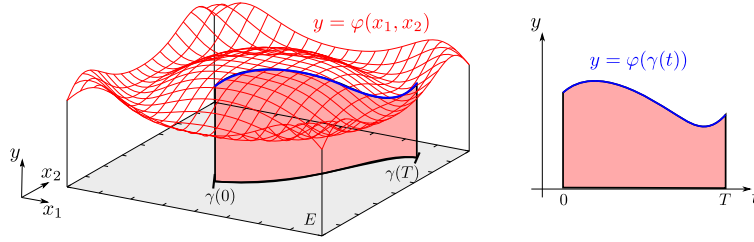


FIGURE 4.9 – L'intégration temporelle de $\varphi \circ \gamma$ peut être remplacée par une intégration spatiale de φ mais avec la mesure μ_γ .

Equation différentielle ordinaire et mesures d'occupation

Dans cette section, on décrit la procédure qui permet d'écrire une relation linéaire sur les mesures d'occupation μ et ν associées à une courbe solution d'un problème de Cauchy. La méthode est tout à fait générale puisqu'elle ne nécessite pas la résolution de l'équation différentielle ordinaire. Étant donnée une fonction test, on obtiendra une équation linéaire.

Proposition 13. *Soit*

$$\begin{cases} \dot{x}(t) &= f(x), \\ x(0) &= x_0. \end{cases} \quad (4.19)$$

le problème de Cauchy avec $f : X \rightarrow \mathbb{R}^n$ un champ de vecteurs Lipschitzien et $x_0 \in X$ une condition initiale. Soient μ et ν les mesures d'occupation associées à la courbe paramétrée $[0, T] \ni t \mapsto x(t) \in X$ solution de (4.19). Alors le couple (μ, ν) vérifie l'équation linéaire suivante :

$$\mathcal{L}'(\mu, \nu) = \delta_{x_0}. \quad (4.20)$$

L'opérateur \mathcal{L}' est défini par dualité :

$$\langle \mathcal{L}'(\mu, \nu), \varphi \rangle = \langle (\mu, \nu), \mathcal{L}\varphi \rangle,$$

et son dual $\mathcal{L} : \mathcal{C}^1(X) \rightarrow \mathcal{C}^0(X) \times \mathcal{C}^1(X)$ est

$$\mathcal{L} : \varphi \mapsto (x \mapsto -\nabla_x \varphi(x) \cdot f(x), x \mapsto \varphi(x)) \quad (4.21)$$

avec φ une fonction continûment différentiable définie sur l'espace des états X .

Démonstration. Soit φ une fonction continûment différentiable. En accord avec le théorème fondamental d'analyse, on peut écrire

$$\varphi(x(T)) - \varphi(x(0)) = \int_0^T \frac{d}{dt} \varphi(x(t)) dt.$$

Grâce à la règle de Leibniz, on écrit

$$\varphi(x(T)) - \varphi(x(0)) = \int_0^T \nabla_x \varphi(x(t)) \cdot \dot{x}(t) dt.$$

Par conséquent, avec une courbe paramétrée $t \mapsto x(t)$ solution du problème de Cauchy (4.19), on a :

$$\varphi(x(T)) - \varphi(x(0)) = \int_0^T \nabla_x \varphi(x(t)) \cdot f(x(t)) dt.$$

La proposition 12 permet d'écrire cette dernière équation avec les mesures d'occupation associées à la courbe $t \mapsto x(t)$ solution de (4.19) :

$$\int_X \varphi(x) d\nu - \varphi(x_0) = \int_X \nabla_x \varphi(x) \cdot f(x) d\mu.$$

On peut remarquer que la variable indépendante t n'apparaît plus dans cette dernière relation.

En utilisant le formalisme de Dirac, on peut écrire

$$\langle \nu, \varphi \rangle + \langle \mu, -\nabla_x \varphi(x) \cdot f(x) \rangle = \langle \delta_{x_0}, \varphi \rangle.$$

On en déduit :

$$\langle (\mu, \nu), \mathcal{L}\varphi \rangle = \langle \delta_{x_0}, \varphi \rangle.$$

Par dualité, il existe une application linéaire \mathcal{L}' telle que

$$\langle \mathcal{L}'(\mu, \nu), \varphi \rangle = \langle \delta_{x_0}, \varphi \rangle.$$

Comme cette relation est vraie pour tout $\varphi \in \mathcal{C}^1(x)$, alors le couple (μ, ν) vérifie la relation linéaire suivante

$$\mathcal{L}'(\mu, \nu) = \delta_{x_0}.$$

□

On peut noter que la relation (4.20) entre les mesures peut aussi s'écrire, via la théorie de la dualité et des fonctions tests sous la forme suivante

$$\frac{\partial \mu}{\partial t} + \operatorname{div} f\mu + \nu = \delta_0. \quad (4.22)$$

L'équation (4.22) est classiquement appelée équation de Liouville. Des détails supplémentaires à propos de cette équation peuvent être trouvés dans [67] (théorème 5.34).

Exemple 22. Considérons le problème de Cauchy suivant

$$\begin{cases} \dot{x}(t) &= 1, \\ x(0) &= 0. \end{cases} \quad (4.23)$$

Supposons $T = 1$, l'équation (4.20), avec la fonction test $\varphi(x) = 1$, implique simplement que (μ, ν) vérifie la contrainte linéaire suivante :

$$\int_{[0,1]} 1 d\nu - \int_{[0,1]} 0 d\mu = 1.$$

De plus, n'importe quelle fonction test lisse peut être utilisée pour générer une contrainte linéaire sur (μ, ν) .

$$\begin{aligned} \varphi(x) = 1 &\Rightarrow \int_{[0,1]} 1 d\nu - \int_{[0,1]} 0 d\mu = 1, \\ \varphi(x) = x &\Rightarrow \int_{[0,1]} x d\nu - \int_{[0,1]} 1 d\mu = 0, \\ \varphi(x) = x^2 &\Rightarrow \int_{[0,1]} x^2 d\nu - \int_{[0,1]} 2x d\mu = 0^2, \\ &\vdots \\ \varphi(x) = x^n &\Rightarrow \int_{[0,1]} x^n d\nu - \int_{[0,1]} nx^{n-1} d\mu = 0^n. \end{aligned}$$

La solution du problème (4.23) est évidemment la courbe paramétrée γ présentée dans l'exemple 21. Dans ce cas, les mesures d'occupation associées à x sont les suivantes :

- μ est λ , *i.e.* la mesure de Lebesgue restreinte à l'intervalle $[0, 1]$,
- ν est δ_1 , *i.e.* une mesure de Dirac centrée sur $x = 1$.

Le lecteur pourra vérifier que le couple (λ, δ_1) vérifie toutes les contraintes précédentes.

Dans cet exemple, comme le champ de vecteurs est polynomial, les contraintes présentées peuvent être vue comme des relations linéaires entres les moments de (μ, ν) . On peut noter que le proposition 13 ne suppose pas la polynomialité des fonctions test. On peut donc aussi écrire, avec $\varphi = \sin$, la contrainte suivante sur le couple (μ, ν) :

$$\int_{[0,1]} \sin(x) d\nu - \int_{[0,1]} \cos(x) d\mu = \sin(0).$$

Dans la section 4.3.5, la proposition 13 sera employée avec des fonctions test qui sont seulement \mathcal{C}^1 .

Exemple 23. Considérons le système dynamique 2-dimensionnel de Van der Pol :

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} y \\ (1-x^2)y - x \end{pmatrix} \quad (4.24)$$

Soit φ la fonction lisse suivante $\varphi : \mathbb{R}^2 \ni (x, y) \mapsto x^2 + y \in \mathbb{R}$, on suppose que la condition initiale est $(x_0, y_0) = (2, 1)$, et que l'instant final est $T = 1$. La proposition 13 permet d'écrire que le couple de mesures d'occupation (μ, ν) associé à la solution de l'équation (4.24) vérifie l'équation linéaire suivante :

$$\langle \nu, x^2 + y \rangle + \langle \mu, - \begin{pmatrix} 2x \\ 1 \end{pmatrix} \cdot \begin{pmatrix} y \\ (1-x^2)y - x \end{pmatrix} \rangle = x_0^2 + y_0,$$

i.e.

$$\langle \nu, x^2 + y \rangle + \langle \mu, -2xy - (1-x^2)y + x \rangle = 5.$$

4.3.3 Formulation linéaire d'un problème de contrôle optimal non linéaire

Dans cette section, on rappelle comment un problème de contrôle optimal peut être écrit sous la forme d'un programme linéaire en dimension infinie. L'idée est de contraindre le couple de mesures de sorte qu'il soit d'occupation, puis de chercher le couple qui est optimal. Il se trouve que l'intégrale temporelle apparaissant dans la fonction coût du problème de contrôle peut se mettre sous la forme d'une fonction réelle linéaire sur les mesures.

La démarche peut être vue comme une compilation des résultats présentés dans la

section précédente. Considérons le problème de contrôle optimal suivant :

$$\begin{aligned}
J^* &= \min_{u: [0, T] \rightarrow U} \int_0^T h(t, x(t), u(t)) dt + H(x(T)) \\
\text{tel que} \quad & x(0) = x_0, \\
& \dot{x}(t) = f(t, x(t), u(t)), \forall t \in [0, T], \\
& x(t) \in X, \forall t \in [0, T], \\
& x(T) \in K, \\
& u(t) \in U, \forall t \in [0, T].
\end{aligned} \tag{4.25}$$

On appelle *formulation faible* du problème (4.25) le programme linéaire suivant :

$$\begin{aligned}
J_\infty^* &= \min_{\mu \in \mathcal{M}, \nu \in \mathcal{N}} \langle \mu, h \rangle + \langle \nu, H \rangle \\
\text{tel que} \quad & \mathcal{L}'(\mu, \nu) = \delta_{(0, x_0)},
\end{aligned} \tag{4.26}$$

avec \mathcal{M} et \mathcal{N} l'ensemble des mesures positives avec comme support respectif $[0, T] \times X \times U$ et K . Tout d'abord, on peut remarquer que chaque contrôle admissible u génère une courbe paramétrée $t \mapsto x^u(t)$ solution du problème de Cauchy :

$$\begin{cases} x(0) = x_0, \\ \dot{x}(t) = f(t, x(t), u(t)), \forall t \in [0, T], \\ x(t) \in X, \forall t \in [0, T], \\ u(t) \in U, \forall t \in [0, T]. \end{cases} \tag{4.27}$$

Par conséquent, tout contrôle admissible u donne naissance à un couple de mesures d'occupation (μ, ν) . Plus précisément, comme dans la proposition 13, la mesure ν est une masse de Dirac centrée sur le point $x^u(T)$. La mesure μ , appelée mesure d'occupation sur l'espace d'état-commande, est la mesure d'occupation associée à la courbe paramétrée définie par :

$$\begin{aligned}
\gamma : [0, T] &\rightarrow [0, T] \times X \times U, \\
t &\mapsto (t, x^u(t), u(t)).
\end{aligned}$$

Les contraintes imposées sur l'état et sur la commande par l'intermédiaire des ensembles X , K et U sont écrites en supposant que les supports de μ et ν sont respectivement $[0, T] \times X \times U$ et K . Le point le plus difficile à comprendre est comment le système dynamique contrôlé (4.27) peut être réécrit sous la forme d'une contrainte linéaire sur l'ensemble des couples de mesures (μ, ν) :

$$\mathcal{L}'(\mu, \nu) = \delta_{(0, x_0)}. \tag{4.28}$$

La démarche est la même que celle donnée dans la preuve de la Proposition 13. Soit φ une fonction réelle différentiable sur $[0, T] \times X$, on a :

$$\varphi(T, x^u(T)) - \varphi(0, x^u(0)) = \int_0^T \frac{d}{dt} \varphi(t, x^u(t)) dt.$$

Via la règle de Liebniz, on obtient :

$$\varphi(T, x^u(T)) - \varphi(0, x^u(0)) = \int_0^T \frac{\partial \varphi}{\partial t}(t, x^u(t)) + \nabla_x \varphi(t, x^u(t)) \cdot \dot{x}^u(t) dt.$$

Comme x^u vérifie le problème de Cauchy (4.27), on peut écrire :

$$\begin{aligned}\varphi(T, x^u(T)) - \varphi(0, x_0) &= \int_0^T \frac{\partial \varphi}{\partial t}(t, x^u(t)) + \nabla_x \varphi(t, x^u(t)) \cdot f(t, x^u(t), u(t)) dt \\ \varphi(T, x^u(T)) - \int_0^T \frac{\partial \varphi}{\partial t}(t, x^u(t)) + \nabla_x \varphi(t, x^u(t)) \cdot f(t, x^u(t), u(t)) dt &= \langle \delta_{(0, x_0)}, \varphi \rangle\end{aligned}$$

Comme prouvé dans la proposition 12, l'intégration temporelle peut être remplacée par une intégration spatiale vis-à-vis des deux mesures d'occupation μ et ν :

$$\int_K \varphi(T, x^u(T)) d\nu - \int_{[0, T] \times X \times U} \frac{\partial \varphi}{\partial t}(t, x) + \nabla_x \varphi(t, x) \cdot f(t, x, u) d\mu = \langle \delta_{(0, x_0)}, \varphi \rangle. \quad (4.29)$$

Avec l'opérateur linéaire $\mathcal{L} : \mathcal{C}^1([0, T] \times X) \rightarrow \mathcal{C}^0([0, T] \times X \times U) \times \mathcal{C}^1(X)$ défini par

$$\mathcal{L}\varphi = \left((t, x, u) \mapsto -\frac{\partial \varphi}{\partial t}(t, x) - \nabla_x \varphi(t, x) \cdot f(t, x, u), x \mapsto \varphi(T, x) \right), \quad (4.30)$$

L'équation (4.29) peut être réécrite $\langle (\mu, \nu), \mathcal{L}\varphi \rangle = \langle \delta_{(0, x_0)}, \varphi \rangle$. Par dualité, il existe un opérateur linéaire \mathcal{L}' tel que $\langle (\mu, \nu), \mathcal{L}\varphi \rangle = \langle \mathcal{L}'(\mu, \nu), \varphi \rangle$. Finalement, $\forall \varphi \in \mathcal{C}^1([0, T] \times X)$, $\langle \mathcal{L}'(\mu, \nu), \varphi \rangle = \langle \delta_{(0, x_0)}, \varphi \rangle$.

En d'autres termes, pour chaque contrôle admissible u , il existe un couple de mesures d'occupation (μ, ν) vérifiant (4.28).

La fonction coût $(\mu, \nu) \mapsto \langle \mu, h \rangle + \langle \nu, H \rangle$ dans l'équation (4.26) est exactement la même que celle de l'équation (4.25). Encore une fois, l'intégration temporelle de (4.25) est remplacée par une intégration spatiale vis-à-vis des mesures μ sur l'ensemble $[0, T] \times X \times U$ dans l'équation (4.26). En suivant les mêmes idées, la valeur $H(x(T))$ dans (4.25) peut être réécrite comme $\langle \nu, H \rangle$ car ν est une mesure de Dirac centrée en $x(T)$.

En conclusion, le problème de contrôle optimal (4.25) peut être formulé comme un programme linéaire en dimension infinie (4.26). Notons qu'en général, sa valeur optimale J_∞^* est inférieure à la valeur optimale de J^* du problème de contrôle optimal (4.25). C'est pour cette raison que (4.26) est nommée la formulation faible de (4.25). Sous hypothèse de convexité, les deux valeurs coïncident. Dans [68], Vinter démontre le théorème suivant :

Théorème 16. *Supposons que le problème (4.26) est faisable. Si pour tout $(t, x) \in [0, T] \times X$, le sous-ensemble $V_{t,x} = \{f(t, x, u) \mid u \in U\}$ et la fonction à valeur réelle $g_{t,x}$ définie par $g_{t,x}(v) = \inf_{u \in U} \{h(t, x, u) \mid v = f(t, x, u)\}$ sont convexes alors*

$$J_\infty^* = J^*$$

et le problème de contrôle optimal (4.25) admet une solution optimale.

Notons que, par définition, le poids total $\nu(K)$ est nécessairement 1 et que celui de $[0, T] \times X \times U$ est $\mu([0, T] \times X \times U) = T$.

4.3.4 Relaxation en dimension finie via le calcul par intervalles

Dans cette section, la procédure d'encadrement présentée dans la proposition 8 est appliquée dans le but de relâcher le programme linéaire en dimension infinie comme (4.26). On donne aussi la preuve que le schéma numérique est une borne inférieure à la valeur optimale de (4.26).

Avec des hypothèses de continuités similaires à celles données dans la proposition 2, le schéma numérique proposé est montré convergent.

Relaxation en dimension finie

Théorème 17. *Considérons*

$$\begin{aligned} J^* &= \min_{\mu \in \mathcal{M}, \nu \in \mathcal{N}} \langle \mu, h \rangle + \langle \nu, H \rangle \\ \text{tel que} \quad & \mathcal{L}'(\mu, \nu) = \delta_{(0, x_0)}. \end{aligned} \quad (4.31)$$

Soit $\{X_i\}_{i \in I}$ une partition finie de $[0, T] \times X \times U$ et $\{Y_j\}_{j \in J}$ une partition finie de K . Supposons que \underline{h}_i et \underline{H}_j sont des bornes inférieures de h et H sur X_i et Y_j . De plus, on suppose que $\mathcal{P} = \{\varphi\}$ est une famille finie de fonctions différentiables avec t et x comme variables indépendantes, considérons le programme linéaire suivant :

$$\begin{aligned} \underline{J}^* &= \min_{\mu_i, \nu_j \in \mathbb{R}^+} \sum_{i \in I} \mu_i \underline{h}_i + \sum_{j \in J} \nu_j \underline{H}_j \\ \text{tel que} \quad & \forall \varphi \in \mathcal{P}, \sum_{i \in I} \mu_i \underline{\psi}_i + \sum_{j \in J} \nu_j \underline{\varphi}_j \leq \varphi(0, x_0) \leq \sum_{i \in I} \mu_i \bar{\psi}_i + \sum_{j \in J} \nu_j \bar{\varphi}_j, \\ & \text{où } \psi = -\frac{\partial \varphi}{\partial t} - \frac{\partial \varphi}{\partial x} f(t, x, u), \\ & \text{and } \forall i \in I, \psi(X_i) \subset [\underline{\psi}_i, \bar{\psi}_i], \forall j \in J, \varphi(Y_j) \subset [\underline{\varphi}_j, \bar{\varphi}_j], \end{aligned} \quad (4.32)$$

Alors

$$\underline{J}^* \leq J^*.$$

Démonstration. La preuve est proche de celle présentée dans le contexte du transport optimal. Les détails se trouvent dans la publication [11]. \square

Convergence

Dans cette section, on montre comment le théorème 17 peut donner naissance à un schéma numérique générant une suite croissante de bornes inférieures convergeant vers la valeur optimale du problème (4.31).

L'idée maîtresse consiste à améliorer séquentiellement la partition et l'ensemble des contraintes. Une partition \mathcal{X}_2 d'un ensemble E est un raffinement d'une partition \mathcal{X}_1 de E si chaque élément de \mathcal{X}_2 est un sous-ensemble de l'un des éléments de \mathcal{X}_1 . De façon informelle, cela signifie que \mathcal{X}_2 est une meilleure fragmentation que \mathcal{X}_1 , ou que \mathcal{X}_2 est plus fine que \mathcal{X}_1 . Dans ce cas, on écrit $\mathcal{X}_1 \geq \mathcal{X}_2$.

Soit \mathcal{P} une famille finie de fonctions tests φ pour lesquelles on dispose de fonctions d'inclusion notées $[\varphi] = [\underline{\varphi}, \bar{\varphi}]$. Notons \mathcal{Z} le couple $(\mathcal{X}, \mathcal{Y})$ de sorte que \mathcal{X} soit une partition

de $E = [0, T] \times X \times U$ et \mathcal{Y} une partition de K . On note alors par $\mathbb{P}(\mathcal{Z}, \mathcal{P})$ le programme linéaire en *dimension infinie* suivant et par $J^*(\mathcal{Z}, \mathcal{P})$ sa valeur optimale.

$$\begin{aligned}
& \min_{\mu \in \mathcal{M}, \nu \in \mathcal{N}} \quad \langle \mu, h \rangle + \langle \nu, H \rangle \\
& \text{tel que} \quad \forall \varphi \in \mathcal{P}, \quad \sum_{X \in \mathcal{X}} \mu(X) \underline{\psi}(X) + \sum_{Y \in \mathcal{Y}} \nu(Y) \underline{\varphi}(Y) \leq \varphi(0, x_0), \\
& \quad \text{et } \varphi(0, x_0) \leq \sum_{X \in \mathcal{X}} \mu(X) \bar{\psi}(X) + \sum_{Y \in \mathcal{Y}} \nu(Y) \bar{\varphi}(Y), \\
& \quad \text{où } \psi = -\frac{\partial \varphi}{\partial t} - \frac{\partial \varphi}{\partial x} f(t, x, u).
\end{aligned} \tag{4.33}$$

Théorème 18 (Croissance). *Les assertions suivantes sont vraies :*

i) $\mathcal{P}_1 \subset \mathcal{P}_2 \Rightarrow J^*(\mathcal{Z}, \mathcal{P}_1) \leq J^*(\mathcal{Z}, \mathcal{P}_2)$,

ii) Si $\forall \varphi \in \mathcal{P}$, la fonction d'inclusion $[\varphi]$ est monotone alors

$$\mathcal{Z}_1 \geq \mathcal{Z}_2 \Rightarrow J^*(\mathcal{Z}_1, \mathcal{P}) \leq J^*(\mathcal{Z}_2, \mathcal{P}).$$

Démonstration. i) Supposons que $\mathcal{P}_1 \subset \mathcal{P}_2$. Donc toute contrainte de $\mathbb{P}(\mathcal{Z}, \mathcal{P}_1)$ est aussi une contrainte de $\mathbb{P}(\mathcal{Z}, \mathcal{P}_2)$. C'est à dire que, l'ensemble faisable $\mathbb{P}(\mathcal{Z}, \mathcal{P}_2)$ est inclus dans celui de $\mathbb{P}(\mathcal{Z}, \mathcal{P}_1)$. Comme les deux fonctions coût sont identiques, on a nécessairement $J^*(\mathcal{Z}, \mathcal{P}_1) \leq J^*(\mathcal{Z}, \mathcal{P}_2)$.

Montrons maintenant ii). L'ensemble \mathcal{Z}_i est un couple de partitions $(\mathcal{X}_i, \mathcal{Y}_i)$. Soit φ un élément arbitraire de \mathcal{P} , et considérons la contrainte de $\mathbb{P}(\mathcal{Z}_2, \mathcal{P})$, i.e.

$$\sum_{X \in \mathcal{X}_2} \mu(X) \underline{\psi}(X) + \sum_{Y \in \mathcal{Y}_2} \nu(Y) \underline{\varphi}(Y) \leq \varphi(0, x_0) \leq \sum_{X \in \mathcal{X}_2} \mu(X) \bar{\psi}(X) + \sum_{Y \in \mathcal{Y}_2} \nu(Y) \bar{\varphi}(Y), \tag{4.34}$$

Comme \mathcal{X}_2 est un raffinement de \mathcal{X}_1 , alors il existe une application $\pi : \mathcal{X}_2 \rightarrow \mathcal{X}_1$ telle que pour tout $X_1 \in \mathcal{X}_1$,

$$X_1 = \bigcup_{X_2 \in \pi^{-1}(\{X_1\})} X_2$$

Comme la fonction d'inclusion $[\varphi]$ est supposée monotone, on en déduit que pour tout $X_2 \in \pi^{-1}(\{X_1\})$,

$$\underline{\varphi}(X_1) \leq \underline{\varphi}(X_2) \text{ et } \bar{\varphi}(X_2) \leq \bar{\varphi}(X_1).$$

La mesure μ est supposée être une mesure positive, on a donc l'inégalité suivante

$$\mu(X_1) \underline{\varphi}(X_1) \leq \sum_{X_2 \in \pi^{-1}(\{X_1\})} \mu(X_2) \underline{\varphi}(X_2).$$

Le même genre d'inégalité peut être obtenue pour \mathcal{Y}_1 et \mathcal{Y}_2 . En additionnant, on en déduit la relation suivante

$$\sum_{X \in \mathcal{X}_1} \mu(X) \underline{\psi}(X) + \sum_{Y \in \mathcal{Y}_1} \nu(Y) \underline{\varphi}(Y) \leq \varphi(0, x_0) \leq \sum_{X \in \mathcal{X}_1} \mu(X) \bar{\psi}(X) + \sum_{Y \in \mathcal{Y}_1} \nu(Y) \bar{\varphi}(Y). \tag{4.35}$$

Par conséquent, l'équation (4.34) implique l'équation (4.35). En d'autres termes, l'ensemble faisable de $\mathbb{P}(\mathcal{Z}_2, \mathcal{P})$ est inclus dans $\mathbb{P}(\mathcal{Z}_1, \mathcal{P})$. La fonction coût est identique dans les deux problèmes, on en déduit $J^*(\mathcal{Z}_1, \mathcal{P}) \leq J^*(\mathcal{Z}_2, \mathcal{P})$. \square

Théorème 19. Soit $\{\mathcal{P}_k\}_{k \geq 0}$ une suite croissante d'ensembles finis telle que $\cup_{k \geq 0} \text{vect } \mathcal{P}_k$ est faiblement dense dans l'ensemble des fonctions tests². Supposons que toutes les fonctions d'inclusion des fonctions tests sont monotones et convergentes. Si les suites de partition $\{\mathcal{X}_k\}_{k \geq 0}$ et $\{\mathcal{Y}_k\}_{k \geq 0}$ de E et K , sont croissantes et vérifient $\Delta \mathcal{X}_k \rightarrow 0$ et $\Delta \mathcal{Y}_k \rightarrow 0$ quand $k \rightarrow \infty$, alors

$$J^*(\mathcal{Z}_k, \mathcal{P}_k) \uparrow J^*.$$

Démonstration. La preuve relativement technique se trouvent dans la publication [11]. \square

Corollaire 4. Considérons le programme linéaire en dimension finie suivant et notons $\underline{J}^*(\mathcal{Z}, \mathcal{P})$ sa valeur optimale :

$$\begin{aligned} \min_{\mu_X \in \mathbb{R}^+, \nu_Y \in \mathbb{R}^+} \quad & \sum_{X \in \mathcal{X}} \mu_X \underline{h}(X) + \sum_{Y \in \mathcal{Y}} \nu_Y \underline{H}(Y) \\ \text{tel que} \quad & \forall \varphi \in \mathcal{P}, \sum_{X \in \mathcal{X}} \mu_X \underline{\psi}(X) + \sum_{Y \in \mathcal{Y}} \nu_Y \underline{\varphi}(Y) \leq \varphi(0, x_0), \\ & \text{et } \varphi(0, x_0) \leq \sum_{X \in \mathcal{X}} \mu_X \bar{\psi}(X) + \sum_{Y \in \mathcal{Y}} \nu_Y \bar{\varphi}(Y), \\ & \text{où } \psi = -\frac{\partial \varphi}{\partial t} - \frac{\partial \varphi}{\partial x} f(t, x, u). \end{aligned} \tag{4.36}$$

En supposant que les hypothèses du théorème 19 sont satisfaites. Si les fonctions d'inclusion $[h]$ et $[H]$ sont continues et convergentes alors

$$\underline{J}^*(\mathcal{Z}_k, \mathcal{P}_k) \uparrow J^* \text{ quand } k \rightarrow \infty.$$

Démonstration. Le raisonnement est le suivant. Premièrement, on définit le nombre réel $\bar{J}^*(\mathcal{Z}, \mathcal{P})$, puis on démontre que

$$\underline{J}^*(\mathcal{Z}, \mathcal{P}) \leq J^*(\mathcal{Z}, \mathcal{P}) \leq \bar{J}^*(\mathcal{Z}, \mathcal{P}). \tag{4.37}$$

Dans un second temps, on démontre que les deux bornes de l'équation (4.37), c'est à dire $\underline{J}^*(\mathcal{Z}_k, \mathcal{P}_k)$ et $\bar{J}^*(\mathcal{Z}_k, \mathcal{P}_k)$, convergent vers la même valeur quand k tend vers $+\infty$.

Le théorème 19 donne : $J^*(\mathcal{Z}_k, \mathcal{P}_k) \uparrow J^*$ quand $k \rightarrow \infty$. Avec $\mathcal{Z} = (\mathcal{X}, \mathcal{Y})$ telle que \mathcal{X} soit une partition finie de $E = [0, T] \times X \times U$, \mathcal{Y} une partition finie de K et \mathcal{P} une collection finie de fonctions test, on définit le nombre réel $\bar{J}^*(\mathcal{Z}, \mathcal{P})$ par

$$\begin{aligned} \min_{\mu_X \in \mathbb{R}^+, \nu_Y \in \mathbb{R}^+} \quad & \sum_{X \in \mathcal{X}} \mu_X \bar{h}(X) + \sum_{Y \in \mathcal{Y}} \nu_Y \bar{H}(Y) \\ \text{telle que} \quad & \forall \varphi \in \mathcal{P}, \sum_{X \in \mathcal{X}} \mu_X \bar{\psi}(X) + \sum_{Y \in \mathcal{Y}} \nu_Y \bar{\varphi}(Y) \leq \varphi(0, x_0), \\ & \text{et } \varphi(0, x_0) \leq \sum_{X \in \mathcal{X}} \mu_X \underline{\psi}(X) + \sum_{Y \in \mathcal{Y}} \nu_Y \underline{\varphi}(Y), \\ & \text{où } \psi = -\frac{\partial \varphi}{\partial t} - \frac{\partial \varphi}{\partial x} f(t, x, u). \end{aligned} \tag{4.38}$$

2. On note par $\text{vect } A$ l'ensemble des combinaisons linéaires de vecteurs de A

Les programmes linéaires en dimension finie (4.36) et (4.38) partagent le même ensemble faisable. Pour chaque μ, ν , via la proposition 6, on écrit :

$$\sum_{X \in \mathcal{X}} \mu(X) \underline{h}(X) + \sum_{Y \in \mathcal{Y}} \nu(Y) \underline{H}(Y) \leq \langle \mu, h \rangle \leq \sum_{X \in \mathcal{X}} \mu(X) \bar{h}(X) + \sum_{Y \in \mathcal{Y}} \nu(Y) \bar{H}(Y).$$

Par optimalité, on en déduit que pour chaque partition finie \mathcal{X} et pour chaque collection de fonctions test \mathcal{P} ,

$$\underline{J}^*(\mathcal{Z}, \mathcal{P}) \leq J^*(\mathcal{Z}, \mathcal{P}) \leq \bar{J}^*(\mathcal{Z}, \mathcal{P}).$$

La preuve de la croissance de la suite $\{J^*(\mathcal{Z}_k, \mathcal{P}_k)\}_{k \geq 0}$ est la même que celle donnée dans la preuve du théorème 19. Montrons maintenant que les deux bornes $\underline{J}^*(\mathcal{Z}_k, \mathcal{P}_k)$ et $\bar{J}^*(\mathcal{Z}_k, \mathcal{P}_k)$ de $J^*(\mathcal{Z}_k, \mathcal{P}_k)$ convergent vers la même valeur. Considérons la suite des solutions optimales $\{(\mu_k^*, \nu_k^*)\}_{k \geq 0}$ des programmes linéaires en dimension finie (4.36), i.e.

$$\underline{J}^*(\mathcal{Z}_k, \mathcal{P}_k) = \sum_{X \in \mathcal{X}_k} \mu_k^*(X) \underline{h}(X) + \sum_{Y \in \mathcal{Y}_k} \nu_k^*(Y) \underline{H}(Y). \quad (4.39)$$

Considérons la suite des nombres réels $\{L_k\}_k$, définie par

$$L_k = \sum_{X \in \mathcal{X}_k} \mu_k^* \bar{h}(X) + \sum_{Y \in \mathcal{Y}_k} \nu_k^* \bar{H}(Y). \quad (4.40)$$

La solution optimale (μ_k^*, ν_k^*) of (4.36) est une solution faisable de (4.38). Par optimalité, on peut en déduire que pour tout $k \in \mathbb{N}$,

$$\underline{J}^*(\mathcal{Z}_k, \mathcal{P}_k) \leq \bar{J}^*(\mathcal{Z}_k, \mathcal{P}_k) \leq L_k.$$

Comme les deux fonctions d'inclusion de h et H sont continues et convergentes, d'après la proposition 2, les deux suites définies par (4.39) et (4.40) convergent vers la même valeur. On conclut par $\underline{J}^*(\mathcal{Z}_k, \mathcal{P}_k) \uparrow J^*$ quand $k \rightarrow \infty$. \square

4.3.5 Exemples illustratifs

Exemple 1

On considère dans cette section un système intégrateur unidimensionnel, $\dot{x} = u$ avec x l'état et tel que le contrôle $t \mapsto u(t)$ vérifie la contrainte $|u(t)| \leq 1$, pour tout $t \in [0, T]$ et $T = 1$. De plus, l'état doit vérifier que $|x(t)| \leq 3$ pour tout $t \geq 0$. En utilisant les notations précédentes, on écrit $X = K = [-3, 3]$ et $U = [-1, 1]$.

Via l'approche décrite dans de ce chapitre, le problème de contrôle optimal suivant

$$\begin{aligned} J^* &= \min_{u: [0, T] \rightarrow [-1, 1]} \int_0^T x^2(\tau) d\tau + x^2(T) \\ \text{tel que} & \quad x(0) = 2.82, \dot{x} = u, \\ & \quad x(\tau) \in X, \\ & \quad u(\tau) \in U, \\ & \quad x(T) \in K, \end{aligned} \quad (4.41)$$

peut se reformuler sous la forme du programme linéaire en dimension infinie suivant :

$$\begin{aligned}
J^* &= \min_{\mu, \nu \in \mathcal{M} \times \mathcal{N}} \langle \mu, x^2 \rangle + \langle \nu, x^2 \rangle \\
\text{tel que} \quad &\mathcal{L}^*(\mu, \nu) = \delta_{(0, x_0)}, \\
&\text{supp}(\mu) \subset [0, 1] \times [-3, 3] \times [-1, 1], \\
&\text{supp}(\nu) \subset [-3, 3].
\end{aligned} \tag{4.42}$$

Puis, ce dernier programme est relâché par le programme suivant en dimension finie :

$$\begin{aligned}
\underline{J} &= \min_{\mu_i, \nu_j \in \mathbb{R}^+} \sum_{i \in I} \mu_i \underline{h}_i + \sum_{j \in J} \nu_j \underline{H}_j \\
\text{tel que } \forall \varphi \in \mathcal{P} \quad &\sum_{i \in I} \mu_i \underline{\psi}_i + \sum_{j \in J} \nu_j \underline{\varphi}_j \leq \varphi(0, x_0) \leq \sum_{i \in I} \mu_i \overline{\psi}_i + \sum_{j \in J} \nu_j \overline{\varphi}_j, \\
&\text{avec } \psi = -\frac{\partial \varphi}{\partial t} - \frac{\partial \varphi}{\partial x} f(t, x, u),
\end{aligned} \tag{4.43}$$

où

— l'ensemble \mathcal{P} est la collection suivante de fonction $\mathcal{C}^1([0, T] \times X)$:

$$\mathcal{P} = \{1, t, x, t^2, tx, x^2, t^3, t^2x, t^1x^2, x^3\},$$

— les nombres réels \underline{h}_i , \underline{H}_j , $\underline{\psi}_i$, $\overline{\psi}_i$ et $\underline{\varphi}_j$, $\overline{\varphi}_j$ sont obtenus via l'arithmétique des intervalles. On notera que l'expression de la fonction ψ a été calculée via une méthode formelle à partir de φ et f .

Pour cette implémentation, l'ensemble $[0, 1] \times X \times U$ a été partitionné avec un pavage régulier. Pour chaque dimension, l'espace a été subdivisé en n éléments. Le nombre total de variables μ_i est donc n^3 alors que le nombre de variables ν_j est n . En résumé, le programme linéaire (4.43) a $n^3 + n$ inconnues et 20 contraintes d'inégalités (comme $\text{card } \mathcal{P} = 10$). Les résultats des valeurs optimales obtenues sont illustrés sur la figure 4.10 en fonction de n .

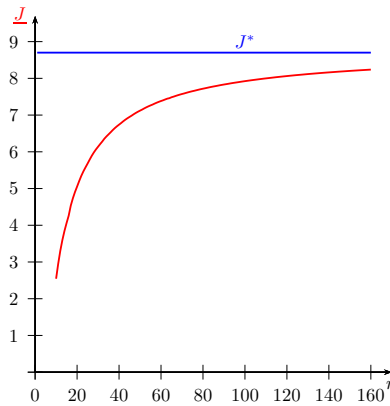


FIGURE 4.10 – Bornes inférieures garanties de J^* obtenue via la résolution du programme (4.42) avec $\text{card}(\{X_i\}) = n^3 + n$.

Exemple 2

Dans cette section, on considère le problème de contrôle optimal classiquement appelé “turnpike”. Ce problème est finement analysé dans la section 22.2 de [92] :

$$\begin{aligned}
 J^* &= \inf_{u:[0,2] \rightarrow U} \int_0^2 x(t) + u(t) dt \\
 \text{tel que} \quad &x(0) = 0 \\
 &\dot{x}(t) = 1 + x(t) - x(t)u(t), \forall t \in [0, 2], \\
 &u(t) \in U = [0, 3], \forall t \in [0, 2].
 \end{aligned} \tag{4.44}$$

Pour ce problème, l’optimum est atteint par une unique commande qui est constante par morceaux. En utilisant le principe du maximum, on peut montrer que la commande optimale u^* est définie par :

$$\begin{aligned}
 u^* : [0, 2] &\rightarrow [0, 3] \\
 t &\mapsto \begin{cases} 2 & \text{si } t \in [\ln(2), 2 - \ln(2)], \\ 0 & \text{autrement.} \end{cases}
 \end{aligned} \tag{4.45}$$

Par conséquent,

$$\begin{aligned}
 J^* &= \int_0^2 x(t) + u^*(t) dt, \\
 &= \int_0^{\ln(2)} e^t - 1 dt + \int_{\ln(2)}^{2-\ln(2)} 1 + 2 dt + \int_{2-\ln(2)}^2 4e^{-2+t} - 1 dt, \\
 &= 9 - 8 \ln(2),
 \end{aligned}$$

On en déduit la valeur approchée suivante :

$$J^* \simeq 3.45 \tag{4.46}$$

Afin de traiter le problème (4.44) avec le schéma numérique présenté précédemment, il nous faut un compact X où l’état doit rester. On peut faire l’analyse suivante : à partir de l’équation différentielle qui gouverne l’état, on peut déduire que $\dot{x} > 0$ quand x est proche de 0. Par conséquent, $x(t) > 0$ pour $t > 0$. Maintenant, on sait que $x(t) \geq 0$ pour tout $t \in [0, 2]$, comme u est supposée appartenir à $[0, 3]$, alors quel que soit l’action u , la vitesse \dot{x} vérifie $\dot{x}(t) \leq 1 + x(t)$. Donc, pour toute commande on aura $x(t) \leq e^t - 1$. Finalement, on peut supposer que $\forall t \in [0, 2], x(t) \in [0, e^2 - 1]$. Notons par X l’intervalle $[0, e^2 - 1]$, les hypothèses du théorème 16 sont satisfaites, alors la valeur optimale J^* du problème (4.44) peut être obtenue en résolvant le programme linéaire suivant :

$$\begin{aligned}
 J_\infty^* &= \min_{\mu \in \mathcal{M}, \nu \in \mathcal{N}} \langle \mu, h \rangle + \langle \nu, H \rangle \\
 \text{tel que} \quad &\mathcal{L}'(\mu, \nu) = \delta_{(0, x_0)}
 \end{aligned} \tag{4.47}$$

avec μ une mesure dont le support est inclus dans $[0, 2] \times [0, e^2 - 1] \times [0, 3]$. Le support de ν est quant à lui supposé être inclus dans $[0, e^2 - 1]$. On choisira donc X comme étant

l'intervalle $[0, e^2 - 1]$. Dans cet exemple où le coût terminal n'intervient pas, il suffit de poser H identiquement nulle.

Lors des différentes résolutions numériques, l'approche utilisée dans l'exemple précédent ne fut pas très efficace. En effet, pour une partition avec $m (= \text{card}\mathcal{X} + \text{card}\mathcal{Y})$ éléments et une collection \mathcal{P} de p fonctions incluses dans $\mathbb{R}[t, x]$, il faut calculer mp nombres réels pour définir les p contraintes. Cette étape de construction du programme linéaire était coûteuse en temps de calcul et en mémoire. Dans le but de réduire la dimension du programme, comme h et f ne dépendent pas explicitement de t , une formulation faible a été obtenue en partitionnant "seulement" l'ensemble $X \times U$. Cette simplification entraîne, dans le cas d'une partition équilibrée pour chaque axe, un programme avec $\text{card}\mathcal{X} + \text{card}\mathcal{Y} = n^2 + n$ variables.

De plus, et encore pour accélérer le temps de calcul et optimiser la mémoire, une collection \mathcal{P} de $\mathcal{C}^1(X)$ peut être choisie de sorte à réduire le nombre de coefficients non nuls apparaissant dans les contraintes. Autrement dit, cette nouvelle collection essaie de rendre la matrice des contraintes creuse.

En effet, on propose de s'intéresser à une famille de fonctions construites à partir de l'interpolation d'Hermite. Avec x_0 et x_1 deux nombres réels, on définit la fonction φ_{x_0, x_1} suivante :

$$\varphi_{x_0, x_1}(x) = \begin{cases} 0 & \text{si } x < x_0, \\ (x_0 - x)^2(-2x - x_0 + 3x_1)(x_1 - x_0)^{-3} & \text{si } x = x_0 \leq x \leq x_1, \\ 1 & \text{si } x_1 < x. \end{cases}$$

Cette fonction est différentiable et est illustrée sur la figure 4.11. Sa dérivée s'annule pour tout $x \notin [x_0, x_1]$. Lorsque les valeurs x_0 et x_1 sont bien choisies, de nombreux coefficients de la matrice des contraintes de (4.32) seront nuls.

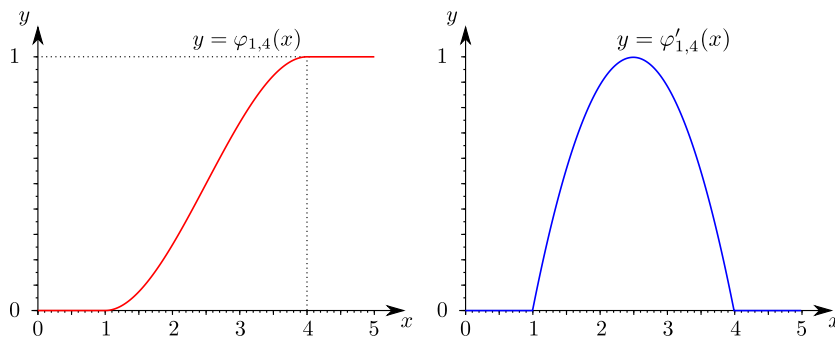


FIGURE 4.11 – Sur la gauche, la fonction différentiable $\varphi_{1,4}$ et sur la droite sa dérivée $\varphi'_{1,4}$.

A partir de cette fonction particulière φ_{x_0, x_1} , un peu comme dans la théorie des ondelettes, on la translate et la dilate pour générer de nouvelles fonctions linéairement indépendantes (sans forcément chercher à les rendre orthogonales). Les translations ont pour objectif de déplacer les coefficients non nuls alors que les dilatations engendrent des changements d'échelle.

D'un point de vue plus technique, on remarque que chaque intervalle $[x_0, x_1]$ définit une fonction différentiable φ_{x_0, x_1} . Pour générer des contraintes, on génère une collection

d'intervalles $\{p\}$ puis on crée la famille de fonctions test associées $\mathcal{H} = \{\varphi_p\}$. Pour être encore plus précis, avec N et T deux entiers, \mathcal{H} sera notée $\mathcal{H}(N, T)$ et sa construction se décompose en deux étapes. Premièrement un pavage régulier $P(N)$, avec N éléments, est généré en découpant l'intervalle X :

$$P(N) = \left\{ \left[\underline{X} + \frac{j}{N}, \underline{X} + \frac{j+1}{N} \right] \mid j \in \{-1, \dots, N-1\} \right\}$$

Continuant notre analogie avec la théorie des ondelettes, ceci peut être interprété comme une étape de changement d'échelle. Puis, dans une seconde étape, pour chaque élément p de $P(N)$, T nouveaux ensembles p_k sont générés en translatant p comme ceci :

$$p_k = p + \frac{k}{T} \Delta(p).$$

Encore une fois, ceci correspondrait à un décalage dans la théorie des ondelettes.

Finalement, à chaque ensemble p_k on associe la fonction différentiable φ_{p_k} dans l'ensemble $\mathcal{H}(N, T)$. La cardinalité de $\mathcal{H}(N, T)$ est donc NT . En pratique et pour cet exemple, l'ensemble $X \times U$ est d'abord divisé avec un pavage régulier de 170^2 pavés et K a été partitionné avec 250 intervalles de même taille. Le programme linéaire est donc de dimension $250^2 + 250 = 62\,750$. Concernant les contraintes, c'est la famille de fonctions suivantes qui a permis de créer les contraintes

$$\mathcal{P} = \mathcal{H}(1, 6) \cup \mathcal{H}(1, 30) \cup \mathcal{H}(6, 25) \cup \mathcal{H}(11, 20) \cup \{x \mapsto x^i \mid i \in \{0, \dots, 5\}\}$$

Tout ceci a donné une matrice de contraintes pour (4.32) avec 970 lignes, 62 750 colonnes et 10 270 993 coefficients non nuls (à comparer avec 970×62750). La valeur optimale obtenue après quelques secondes est : $\underline{J} = 3.312888487$ (ce qui est tout à fait cohérent avec le théorème 17 qui assure que $\underline{J} \leq J^*$, on pourra se rappeler du résultat numérique donnée par l'équation (4.46) où $J^* \simeq 3.45$).

Logiciel

Un solveur a été développé en C++. Comme dans le cas du transport optimal, ce logiciel s'appuie sur la librairie d'arithmétique des intervalles nommée `filib` [47], et génère un programme linéaire. Ce programme linéaire est ensuite résolu par le solveur `GLPK` [64] avec la librairie `gmp` [65] afin de limiter les erreurs qui pourraient apparaître dues à la représentation machine de nombres à virgules flottantes. Ce logiciel et son code source sont disponibles sur le web via l'url <http://perso-laris.univ-angers.fr/~delanoue/>

4.3.6 Conclusion

Dans cette section, une méthode basée sur l'arithmétique des intervalles et la notion de mesure d'occupations a été proposée pour résoudre le problème de contrôle optimal. Cette méthode construit une suite de programmes linéaires pour lesquels la suite des valeurs optimales forme une suite croissante de bornes inférieures à la valeur optimale cherchée. La convergence du schéma numérique a été démontrée. Pour l'instant, dans les méthodes développées le pavage est affiné d'un programme à l'autre en découpant l'ensemble des pavés du premier pavage. D'une partition à la suivante, il pourrait sans doute être plus

efficace de stocker les pavés où la fonction coût a le plus chuté et de continuer à creuser provisoirement dans cette direction. Cela permettrait probablement de diminuer les temps de calcul tout en limitant la quantité de mémoire.

Cette méthode pourrait facilement être étendue à la situation où on impose des contraintes conjointes sur l'état et l'action. En effet, une première étape basée sur l'algorithme SIVIA (voir la section 2.2.1) pourrait générer un pavage recouvrant l'ensemble défini par

$$\{(x, t, u) \in X \times [0, T] \times U \mid \bigcup_i \bigcap_j g_{ij}(x, t, u) \leq 0\} \text{ avec } g_{ij} : X \times [0, T] \times U \rightarrow \mathbb{R}.$$

Puis dans un second temps, la méthode illustrée ici pourrait être appliquée en supposant que les supports des mesures d'occupation μ et ν soient inclus dans le pavage retourné par SIVIA.

En perspective, il pourrait être pertinent de résoudre le problème dual du programme linéaire (4.26). Il est passionnant d'apprendre que ce programme dual est très proche de l'équation HJB. Avec des techniques proches de celles présentées pour les bornes supérieures du problème de transport optimal (section 4.2.3), il serait donc possible de générer des sous-solutions de l'équation d'HJB, i.e. des fonctions qui sont inférieures à la fonction valeur. Malheureusement la méthode donnée dans la section 4.2.3 ne peut se transposer directement puisqu'il faut être capable de simultanément borner une fonction et ses dérivées partielles sur un compact donné. Néanmoins et en suivant une approche classique en automatique, cette sous-solution pourrait servir à synthétiser une commande. Le lecteur pourra, encore une fois, consulter le livre [51] concernant le dual du problème (4.26) ou bien l'article [90] pour la synthèse de contrôle.

4.4 Conclusion

Dans ce chapitre, plusieurs schémas numériques de résolution de deux programmes linéaires en dimension infinie ont été proposés. Il serait important d'assembler ces résultats en une méthodologie générale qui permettrait de générer des schémas numériques d'approximation afin d'approcher n'importe quel programme linéaire dont la variable est une mesure.

L'idéal serait d'être capable, comme dans la proposition 11 et ou bien le théorème 19, de créer une suite croissante et convergente vers la valeur optimale du programme initial. Cela permettrait de résoudre des problèmes issus de la liste présentée dans l'introduction de ce chapitre.

Chapitre 5

Conclusions

Dans ce mémoire, j'ai essayé de démontrer ma capacité à diriger des recherches. Parmi les travaux effectués ces dernières années, j'ai développé deux thèmes : l'un autour de la classification de fonctions et l'autre autour des programmes linéaires en dimension infinie. Le dénominateur commun de ces approches relativement éloignées est l'arithmétique par intervalles. Chacun de ces thèmes a ses propres conclusions et perspectives qui sont présentées dans les chapitres précédents. Néanmoins, ces deux thèmes peuvent être rapprochés de la façon suivante. Etant donnée une fonction lisse de X dans Y et une mesure de probabilité ν sur Y , si f est injective alors la formule de changement de variables induit, via f , une unique mesure de probabilité μ sur X définie par la relation suivante :

$$\mu(A) = \nu(f(A)), \quad (5.1)$$

où A est un ensemble mesurable de X . En général, si la fonction f n'est pas injective, il existe plusieurs mesures μ sur X compatibles avec f et ν , c'est à dire vérifiant $\mu(A) = \nu(f(A))$. L'idée serait d'exploiter ce défaut pour générer une homologie sur les espaces de mesures. En effet, l'opération duale de (5.1), nommée *pushforward* de mesures, est une opération linéaire de $\mathcal{M}(X)$ dans $\mathcal{M}(Y)$ définie par

$$\begin{aligned} f_* : \mathcal{M}(X) &\rightarrow \mathcal{M}(Y) \\ \mu &\mapsto \nu = f_*\mu, \end{aligned}$$

avec

$$\nu(B) = (f_*\mu)(B) = \mu(f^{-1}(B)).$$

Les propriétés spectrales de cet opérateur nous renseigneront probablement sur les singularités de f .

Deux de mes travaux développés ces dernières années ont donné naissance à des schémas numériques pour le problème de contrôle optimal et pour le problème de transport optimal. Le dénominateur commun de ces approches est une modélisation via la théorie de la mesure sous forme de programmes linéaires en dimension infinie sur les mesures. Pour l'instant, ces deux méthodes n'ont pas été unifiées en une approche générale. Aujourd'hui, il me semble envisageable de formaliser des résultats théoriques afin de discrétiser tout programme linéaire en dimension infinie portant sur les mesures. Les applications pourraient être aussi variées que celles données dans le livre de J.B. Lasserre (Moments, Positive Polynomials and Their Applications) qui propose une approche algébrique.

De façon plus audacieuse, j'aimerais adapter les méthodes J. B. Lasserre afin de chercher des invariants des champs de vecteurs. En particulier, il me semble possible de trouver des mesures invariantes pour l'opérateur de Liouville. Ceci nous permettrait de détecter la présence de cycles limites, ou bien de trouver des fonctions de Darboux.

Le chapitre 3 traite de la classification de fonctions. Pour les fonctions stables (et génériques), nous avons proposé des encadrements des ensembles de Thom-Boardman pour les fonctions entre espace de dimensions i avec $i = 1, 2$ ou 3 . De plus, nos méthodes sont capables de prouver l'existence et l'unicité pour les espaces zéro dimensionnel. En particulier, on peut dénombrer les fronces $\Sigma^{1,1}$ pour les applications du plan sur le plan et pour les queues d'arondes $\Sigma^{1,1,1}$ pour les applications de l'espace dans l'espace. Cette capacité à démontrer ces propriétés semble pouvoir être généraliser en un théorème d'existence et de décidabilité. J'aimerais proposer un résultat théorique du type : toute propriété stable est démontrable au sens d'une machine de Turing. J'insiste sur la machine utilisée pour me mettre en opposition avec les résultats proposés dans le livre de Lenore Blum and al. [93]. En effet, on trouvera dans cet ouvrage des résultats de décidabilité et de complexité qui supposent l'existence de représentations exactes des nombres réels sur machine. Il me semble que le concept de généricité et de transversalité est celui qui permettra de démontrer que pour une propriété stable et générique donnée, il existe un algorithme capable de la démontrer.

Par ailleurs, je compte continuer à travailler en robotique et autour des systèmes dynamiques. En effet, depuis l'année passée, nous avons initié avec Philippe Lucidarme (LARIS) un travail en commun avec des collègues du LS2N de Nantes : Christine Chevallereau, Philippe Wenger et Yannick Aoustin. Notre objectif est l'étude et la création d'un robot humanoïde. L'originalité de ce robot repose sur l'utilisation de nouveaux moteurs linéaires qui ont la particularité d'être réversibles. Ces nouveaux moteurs réversibles permettent d'imaginer un robot marcheur capable de poser le pied avec une vitesse non nulle lors de l'impact. Nous envisageons la marche rapide, voire la course. Le projet ne fait que de débiter et de nombreux challenges scientifiques ont déjà été identifiés lors de la création de la preuve de concept : modélisation, simulation, dimensionnement, optimisation de la géométrie et contrôle. Sans être à l'initiative de ce projet, j'ai l'intention de fortement m'impliquer en particulier sur la partie commande. La marche étant un phénomène cyclique, la mise en place d'une commande nécessite de montrer que le système bouclé admet un cycle limite stable. J'ai déjà, par le passé, mis en œuvre des algorithmes de calcul de valeurs propres de l'application de premier retour de Poincaré. Un dossier de Projet de Recherche Collaborative a été déposé auprès de l'Agence Nationale de la Recherche.

Enfin, avec Sébastien Lahaye et Vadim Azhmyakov, nous avons commencé depuis début 2018 l'encadrement de la thèse de doctorat de Luz Adriana Guzman. Pour cette thèse, nous envisageons d'exploiter nos compétences en algèbre tropicale pour donner naissance à une nouvelle méthode de résolution de l'équation d'Hamilton-Jacobi-Bellman. En effet, comme précisé dans la section 4.3, l'équation HJB est linéaire dans ce genre d'algèbre.

Bibliographie

- [1] Vladimir I. Arnold, Sabir M. Gusein-Zade, and Alexander N. Varchenko. *Singularities of Differentiable Maps : Volume I : The Classification of Critical Points Caustics and Wave Fronts*. Monographs in Mathematics. Birkhäuser Boston, 1985.
- [2] Martin Golubitsky and Victor Guillemin. *Stable mappings and their singularities*. Graduate texts in mathematics. Springer-Verlag, 1974.
- [3] John N. Mather. How to stratify mappings and jet spaces. In Oscar Burlet and Felice Ronga, editors, *Singularités d'Applications Différentiables*, pages 128–176, Berlin, Heidelberg, 1976. Springer.
- [4] Hassler Whitney. On singularities of mappings of euclidean spaces. i. mappings of the plane into the plane. *Annals of Mathematics*, 62(3) :374–410, 1955.
- [5] René Thom. Ensembles et morphismes stratifiés. *Bull. Amer. Math. Soc.*, 75(2) :240–284, 1969.
- [6] Nicolas Delanoue and Sébastien Lagrange. A numerical approach to compute the topology of the apparent contour of a smooth mapping from \mathbb{R}^2 to \mathbb{R}^2 . *Journal of Computational and Applied Mathematics*, 271(Supplement C) :267–284, 2014.
- [7] Romain Benoit, Nicolas Delanoue, Sébastien Lagrange, and Philippe Wenger. Guaranteed detection of the singularities of 3R robotic manipulators. *Mechanical Sciences*, 7(1) :31–38, 2016.
- [8] Philippe Wenger. Cuspidal and noncuspidal robot manipulators. *Robotica*, 25(6) :677–689, 2007.
- [9] Romain Benoit, Nicolas Delanoue, Sébastien Lagrange, and Philippe Wenger. Combinatorial description of 2-dimensional kinematic functions : application to 3R robotic manipulators. *Soumis à Robotica*, 2018.
- [10] Nicolas Delanoue, Mehdi Lhommeau, and Philippe Lucidarme. Numerical enclosures of the optimal cost of the kantorovitch’s mass transportation problem. *Computational Optimization and Applications*, 63(3) :855–873, 2016.
- [11] Nicolas Delanoue, Mehdi Lhommeau, and Sébastien Lagrange. Nonlinear optimal control : A numerical scheme based on occupation measures and interval analysis. *Soumis à Computational Optimization and Applications*, 2018.
- [12] Mohamed Mustafa, Alexandru Stancu, Nicolas Delanoue, and Eduard Codres. Guaranteed slam - an interval approach. *Robotics and Autonomous Systems*, 100 :160–170, 2018.

- [13] Jean-Baptiste Fasquel and Nicolas Delanoue. A graph based image interpretation method using a priori qualitative inclusion and photometric relationships. *Accepté pour publication dans IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018.
- [14] Jean-Baptiste Fasquel and Nicolas Delanoue. An approach for sequential image interpretation using a priori binary perceptual topological and photometric knowledge and k-means based segmentation. *Accepté pour publication dans Journal of the Optical Society of America*, 2018.
- [15] Nicolas Delanoue, Luc Jaulin, and Bertrand Cotteceau. Using interval arithmetic to prove that a set is path-connected. *Theoretical Computer Science*, 351(1) :119–128, 2006.
- [16] Nicolas Delanoue, Luc Jaulin, and Bertrand Cotteceau. Guaranteeing the homotopy type of a set defined by non-linear inequalities. *Reliable Computing*, 13(5) :381–398, 2007.
- [17] Sébastien Lagrange, Nicolas Delanoue, and Luc Jaulin. On sufficient conditions of the injectivity : Development of a numerical test algorithm via interval analysis. *Reliable Computing*, 13(5) :409–421, 2007.
- [18] Sébastien Lagrange, Nicolas Delanoue, and Luc Jaulin. Injectivity analysis using interval analysis : Application to structural identifiability. *Automatica*, 44(11) :2959–2962, 2008.
- [19] Nicolas Delanoue, Luc Jaulin, and Bertrand Cotteceau. An algorithm for computing a neighborhood included in the attraction domain of an asymptotically stable point. *Communications in Nonlinear Science and Numerical Simulation*, 21(1) :181–189, 2015.
- [20] Nicolas Delanoue, Luc Jaulin, and Bertrand Cotteceau. *Counting the number of connected components of a set and its application to robotics*, pages 93–101. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.
- [21] Nicolas Delanoue, Luc Jaulin, Laurent Hardouin, and Mehdi Lhommeau. *Guaranteed Characterization of Capture Basins of Nonlinear State-Space Systems*, pages 265–272. Springer, Berlin, Heidelberg, 2009.
- [22] Ramon E. Moore. *Interval analysis*. Prentice-Hall series in automatic computation. Prentice-Hall, 1966.
- [23] Arnold Neumaier. *Interval Methods for Systems of Equations*. Cambridge Middle East Library. Cambridge University Press, 1990.
- [24] Ramon E. Moore. Practical aspects of interval computation. *Aplikace matematiky*, 13(1) :52–92, 1968.
- [25] Leocadio G Casado, José A Martínez, Inmaculada García, and Ya D. Sergeyev. New interval analysis support functions using gradient information in a global minimization algorithm. *Journal of Global Optimization*, 25(4) :345–362, 2003.
- [26] Rudolf J. Lohner. Enclosing the solutions of ordinary initial and boundary value problems. In E. W. Kaucher, U. W. Kulisch, and C. Ullrich, editors, *Computer Arithmetic : Scientific Computation and Programming Languages*, pages 255–286. Wiley-Teubner, Stuttgart, 1987.

- [27] Tarek Raïssi, Nacim Ramdani, and Yves Candau. Set membership state and parameter estimation for systems described by nonlinear differential equations. *Automatica*, 40(10) :1771–1777, 2004.
- [28] Warwick Tucker. The lorenz attractor exists. *Comptes Rendus de l'Académie des Sciences-Series I-Mathematics*, 328(12) :1197–1202, 1999.
- [29] Luitzen Egbertus Jan Brouwer. Über abbildung von mannigfaltigkeiten. *Mathematische Annalen*, 71(1) :97–115, 1911.
- [30] Gotz Alefeld. Inclusion methods for systems of nonlinear equations-the interval newton method and modifications. In J. Herzberger, editor, *Topics in Validated Computations*, pages 7–26. Elsevier Science BV, 1994.
- [31] Nathalie Revol. Interval newton iteration in multiple precision for the univariate case. *Numerical Algorithms*, 34(2) :417–426, 2003.
- [32] R. Baker Kearfott. Interval computations : Introduction, uses, and resources. *Euro-math Bulletin*, 2(1) :95–112, 1996.
- [33] Luc Jaulin, Michel Kieffer, Olivier Didrit, and Eric Walter. *Applied Interval Analysis : With Examples in Parameter and State Estimation, Robust Control and Robotics*. Springer London, 2001.
- [34] A.A. Gaganov. Computation complexity of the range of a polynomial in several variables. *Cybernetics*, 21(4) :418–421, 1985.
- [35] Nicolas Delanoue. *Algorithmes numériques pour l'analyse topologique. Calcul par intervalle et théorie des graphes*. PhD thesis, Ecole Doctorale d'Angers, Angers, 2006.
- [36] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.
- [37] Roger Godement. *Cours d'algèbre*. Collection Enseignement des sciences. Hermann, 1969.
- [38] John E. Hopcroft, Rajeev Motwani, and Jeffrey D. Ullman. *Introduction to Automata Theory, Languages, and Computation (3rd Edition)*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2006.
- [39] Arthur Sard. The measure of the critical values of differentiable maps. *Bulletin of the American Mathematical Society*, 48(12) :883–890, 1942.
- [40] Michel Demazure. *Bifurcations and Catastrophes : Geometry of Solutions to Nonlinear Problems*. Universitext. Springer Berlin Heidelberg, 2013.
- [41] Romain Benoit. *Analyse qualitative de robots*. PhD thesis, Ecole Doctorale d'Angers, Angers, 2017.
- [42] Whitney Hassler. On singularities of mappings of euclidean spaces. i. mappings of the plane into the plane. *Annals of Mathematics*, 62(3) :374–410, 1955.
- [43] Steve Smale. Mathematical problems for the next century. *Mathematical Intelligence*, 20 :7–15, 1998.
- [44] Juan A. Moya-Pérez and Juan J. Nuño Ballesteros. The link of a finitely determined map germ from \mathbb{R}^2 to \mathbb{R}^2 . *Journal of the Mathematical Society of Japan*, 62(4) :1069–1092, 2010.

- [45] C.R.F. Maunder. *Algebraic Topology*. Dover Books on Mathematics Series. Dover Publications, 1996.
- [46] Andreas Griewank and Andrea Walther. *Evaluating Derivatives : Principles and Techniques of Algorithmic Differentiation*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, second edition, 2008.
- [47] Filib++ interval library <http://www2.math.uni-wuppertal.de/~xsc/software/filib.html>.
- [48] Hassler Whitney. The self-intersections of a smooth n -manifold in $2n$ -space. *Annals of Mathematics*, 45(2) :220–246, 1944.
- [49] Philippe Wenger. Classification of 3R Positioning Manipulators. *ASME - Journal of Mechanical Design*, 120(2) :327–332, 1998.
- [50] Philippe Wenger. Uniqueness domains and regions of feasible paths for cuspidal manipulators. *IEEE Transactions on Robotics*, 20(4) :745–750, 2004.
- [51] Jean-Bernard Lasserre. *Moments, Positive Polynomials and Their Applications*. Imperial College Press optimization series. Imperial College Press, 2010.
- [52] Leonid V. Kantorovitch. On the translocation of masses. *Doklady. Akad. Nauk SSSR*, 37 :199–201, 1942.
- [53] Yann Brenier. Décomposition polaire et réarrangement monotone des champs de vecteurs. *C. R. Acad. Sci. Paris Sér. I Math.*, 305(19) :805–808, 1987.
- [54] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Communications on Pure and Applied Mathematics*, 44 :375–417, 1991.
- [55] Cédric Villani. *Optimal Transport : Old and New*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2008.
- [56] Lawrence C. Evans. Partial differential equations and monge-kantorovich mass transfer (survey paper). In *Current Developments in Mathematics, 1997, International Press*, 1999.
- [57] Serge Dubuc and Issa Kagabo. Numerical solutions of the mass transfer problem. *Rairo-operations Research*, 40(1) :1–17, 2006.
- [58] J. Rigoberto Gabriel and Gabriel Beatris. An approximation scheme for the mass transfer problem. *Morfismos*, 10 :1–13, 2006.
- [59] J. Rigoberto Gabriel, Juan González-Hernández, and Raquiel R. López-Martínez. Numerical approximations to the mass transfer problem on compact spaces. *IMA Journal of Numerical Analysis*, 30(4) :1121–1136, 2010.
- [60] Rick Chartrand, Brendt Wohlberg, Kevin R. Vixie, and Erik M. Bollt. A gradient descent solution to the monge-kantorovich problem. *Applied Mathematical Sciences*, 3(22) :1071–1080, 2009.
- [61] Edward J. Anderson and Andrew B. Philpott. Duality and an algorithm for a class of continuous transportation problems. *Mathematics of Operations Research*, 9(2) :222–231, 1984.
- [62] Jean-David Benamou. Numerical resolution of an “unbalanced” mass transport problem. *ESAIM : Mathematical Modelling and Numerical Analysis*, 37(5) :851–868, 2003.

- [63] Quentin Mérigot. A Multiscale Approach to Optimal Transport. *Computer Graphics Forum*, 30(5) :1583–1592, 2011.
- [64] Glpk : Gnu linear programming kit <http://www.gnu.org/software/glpk/>.
- [65] The gnu multiple precision arithmetic library <https://gmplib.org/>.
- [66] Stephen M. Robinson. Bounds for error in the solution set of a perturbed linear program. *Linear Algebra and its Applications*, 6 :69–81, 1973.
- [67] Cédric Villani. *Topics in Optimal Transportation*. Graduate studies in mathematics. American Mathematical Society, 2003.
- [68] Richard B. Vinter. Convex duality and nonlinear optimal control. *SIAM Journal on Control and Optimization*, 31(2) :518–538, 1993.
- [69] Richard E. Bellman. On the theory of dynamic programming. *Proceedings of the National Academy of Sciences of the United States of America.*, 38(8) :716–719, 1952.
- [70] Michael G. Crandall, Lawrence C. Evans, and Pierre-Louis Lions. Some properties of viscosity solutions of Hamilton-Jacobi equations. *Transactions of the American Mathematical Society*, 282(2) :487–502, 1984.
- [71] Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions. User’s guide to viscosity solutions of second order partial differential equations. *Bull. Am. Math. Soc., New Ser.*, 27(1) :1–67, 1992.
- [72] Michael G. Crandall and Pierre-Louis Lions. Viscosity solutions of Hamilton-Jacobi equations. *Transactions of the American Mathematical Society*, 277(1) :1–42, 1983.
- [73] Ernst Hairer, Syvert P. Nørsett, and Gerhard Wanner. *Solving Ordinary Differential Equations I (2Nd Revised. Ed.) : Nonstiff Problems*. Springer-Verlag New York, Inc., New York, NY, USA, 1993.
- [74] Ernst Hairer, Syvert P. Nørsett, and Gerhard Wanner. *Solving Ordinary Differential Equations II : Stiff and Differential-Algebraic Problems*. Lecture Notes in Economic and Mathematical Systems. Springer, 1993.
- [75] John T. Betts. *Practical Methods for Optimal Control and Estimation Using Nonlinear Programming*. Society for Industrial and Applied Mathematics, second edition, 2010.
- [76] Maurizio Falcone. Numerical solution of dynamic programming equations. In Appendix A in the volume M. Bardi and I. Capuzzo Dolcetta, editors, *Optimal control and viscosity solutions of Hamilton–Jacobi–Bellman equations*. Birkhäuser, Boston, 1997.
- [77] Chiehshen S. Huang, Song Wang, and Kok L. Teo. Solving Hamilton-Jacobi-Bellman equations by a modified method of characteristics. *Nonlinear Analysis : Theory, Methods & Applications*, 40(1-8) :279–293, 2000.
- [78] Shouyang Wang, F Gao, and Kok L. Teo. An upwind finite-difference method for the approximation of viscosity solutions to hamilton-jacobi-bellman equations. *IMA Journal of Mathematical Control and Information*, 17(2) :167–178, 2000.
- [79] Michael G. Crandall and Pierre-Louis Lions. Two approximations of solutions of Hamilton-Jacobi equations. *Mathematics of Computation*, 43(167) :1–19, 1984.

- [80] Maurizio Falcone and Roberto Ferretti. *Semi-Lagrangian Approximation Schemes for Linear and Hamilton-Jacobi Equations*. Society for Industrial and Applied Mathematics, 2013.
- [81] Lawrence C. Evans. An introduction to mathematical optimal control theory, version 0.2 <http://math.berkeley.edu/~evans/>.
- [82] Oliver Junge and Hinke M. Osinga. A set oriented approach to global optimal control. *ESAIM : Control, optimisation and calculus of variations*, 10(2) :259–270, 2004.
- [83] Marianne Akian, Stéphane Gaubert, and Asma Lakhoua. The max-plus finite element method for solving deterministic optimal control problems : Basic properties and convergence analysis. *SIAM Journal on Control and Optimization*, 47(2) :817–848, 2008.
- [84] Daniel Hernández-Hernández, Onésimo Hernández-Lerma, and Michael Taksar. The linear programming approach to deterministic optimal control problems. *Applcaciones Mathematicae*, 24(1) :17–33, 1996.
- [85] J.E. Rubio. *Control and optimization : the linear treatment of nonlinear problems*. Nonlinear science. Manchester University Press, 1986.
- [86] E. Balder. Control and optimisation : The linear treatment of nonlinear problems (j. e. rubio). *SIAM Review*, 30(4) :663–663, 1988.
- [87] Onésimo Hernández-Lerma and Jean-Bernard Lasserre. Approximation schemes for infinite linear programs. *SIAM Journal on Optimization*, 8(4) :973–988, 1998.
- [88] Vladimir Gaitsgory and Sergey Rossomakhine. Linear programming approach to deterministic long run average problems of optimal control. *SIAM Journal on Control and Optimization*, 44(6) :2006–2037, 2006.
- [89] Vladimir Gaitsgory and Sergey Rossomakhine. Averaging and linear programming in some singularly perturbed problems of optimal control. *Applied Mathematics & Optimization*, 71(2) :195–276, 2015.
- [90] Didier Henrion, Jean-Bernard Lasserre, and Carlo Savorgnan. Nonlinear optimal control synthesis via occupation measures. In *Proceedings of the IEEE Conference on Decision and Control*, pages 4749–4754, 2008.
- [91] Jean-Bernard Lasserre, Didier Henrion, Christophe Prieur, and Emmanuel Trélat. Nonlinear Optimal Control via Occupation Measures and LMI-Relaxations. *SIAM Journal on Control and Optimization*, 47(4) :1643–1666, 2008.
- [92] Francis Clarke. *Functional Analysis, Calculus of Variations and Optimal Control*. Graduate Texts in Mathematics. Springer London, 2013.
- [93] Lenore Blum, Felipe Cucker, Michael Shub, and Stephen Smale. *Complexity and Real Computation*. Springer New York, 1998.