

Rapport de Magistère

Juliette Chevallier

Septembre 2012 – Octobre 2016

La vérité est une ligne tracée entre les erreurs.
– **Franz Anton Mesmer** –

Remerciements

Avant de commencer, je voudrais remercier tous les professeurs qui m'ont encadrée durant ces quatre années passées à Orsay – en particulier Hervé Lemeur, Pierre-Guy Plamondon et Joan Glaunès – ainsi que Stéphanie Allasonnière pour avoir accepté de m'encadrer durant les trois prochaines années.

Je voudrais également remercier toutes les personnes qui ont compté pour moi durant ces quatre années et qui se sont toujours montrées présentes alors même que je ne l'étais pas toujours. Ceux qui, comme moi, "subissaient" : je pense particulièrement à Perinne pour ces soirées à boire du thé et l'Écosse qui ne nous a pas résisté, Céline pour ces cinq mois de stage et plus encore, Pierre pour me supporter quand je suis insupportable et ces "remontrances" quand je voulais renoncer, Léo pour avoir parfait ma culture cinématographique dilettante, Diane pour ces bières partagées, Enrica pour l'Italie, Samuel pour le reste, et ceux que j'oublie... Et, surtout, ceux qui ont cru bon de me croire sur parole : Xavier pour m'avoir poussée à venir étudier à Orsay, Micka et Nico pour leur patience quand je mettais jusqu'à plusieurs mois à leur répondre et ma famille qui m'a toujours soutenue malgré son incompréhension notable de tout cela.

Enfin, je voudrais remercier Christophe Giraud pour sa présence et son aide dans ma recherche de thèse et Antoine Chambert-Loir pour m'avoir fait découvrir un domaine des mathématiques que je me pensais inaccessible et répondu à mes (nombreuses !) questions.

Le présent document se décompose en sept parties, plus ou moins indépendantes.

La première relate mon cursus au sein du magistère d'Orsay ainsi que mon expérience d'apprentissage hors-mur, à savoir ma participation à l'iGEM Competition 2014. J'ai effectué quatre années à Orsay dont une année de césure pour passer l'agrégation de mathématiques que j'ai obtenue.

La deuxième présente les modèles élémentaires de la théorie des espaces de formes et de l'appariement difféomorphique de formes anatomiques. J'ai découvert ce domaine lors du cours d'Alain Trounev *Géométrie et Espaces de formes* du second trimestre du master Mathématiques pour les Sciences du Vivant. J'ai travaillé lors de mon stage de M2 sur la mise au point d'une nouvelle technique d'appariement de formes (tout ceci est relaté dans la septième et dernière partie de ce document)

Je poursuis en thèse dans cette voie mais en travaillant cette-fois sur des données longitudinales. Ce type de modèle étant très récent (2015) et n'ayant pas eu le temps d'en travailler tous les fondements mathématiques (ma thèse débutant le 1 octobre...) j'ai préféré me concentrer sur la description des modèles valables en un seul instant t , qui soutiennent de toute façon les seconds et dont la compréhension est primordiale.

Enfin, les parties suivantes correspondent à une sélection des travaux que j'ai rédigés lors de ces quatre années au magistère : à savoir mon TER de L3, celui de M1 et de M2 ; un compte-rendu de lecture d'article réalisé dans le cadre du cours *Optimisation et Simulation numérique* du premier semestre de M2 et, enfin, mon mémoire de stage concluant mes 5 mois de stage effectués au MAP 5, avec Joan Glaunès.

Sommaire

Cursus au sein du Magistère de Mathématiques d'Orsay : Licence Mathématiques Fondamentales et Appliquées ; Participation à l'iGEM Competition 2014 ; Master Mathématiques pour les Sciences du Vivant.	3
Présentation d'un domaine de recherche : Espaces de formes et recalage difféomorphique. Application à l'imagerie médicale et l'anatomie computationnelle.	19
TER de L3 : Modélisation mathématiques d'une infection par le virus du VIH. Encadré par Hervé Lemeur, Université Paris-Sud.	37
TER de M1 : Représentations linéaires du groupe symétrique. Encadré par Pierre-Guy Plamondon, Université Paris-Sud.	53
Lecture d'article : Autour du problème de sériation. Encadré par Filippo Santambrogio, Université Paris-Sud.	95
TER de M2 : Espaces de formes et courbure. Encadré par Joan Alexis Glaunès, Université Paris-Descartes.	137
Mémoire de M2 : Modèles de pseudo métamorphoses pour le recalage difféo- morphique de courbes et surfaces Encadré par Joan Alexis Glaunès, Université Paris-Descartes.	201
Références	289

– Première partie –

Cursus au sein du Magistère de Mathématiques d'Orsay

Apprentissage hors-mur

Licence Mathématiques Fondamentales et Appliquées

Participation à l'iGEM Competition

Master Mathématiques pour les Sciences du Vivant

Sommaire – Partie I

Magistère de Mathématiques d'Orsay	3
	6
I L3 Mathématiques Fondamentales et Appliquées	6
I.1 Premier semestre	6
I.2 Deuxième semestre	7
II M1 Mathématiques Fondamentales et Appliquées	7
II.1 Premier semestre	8
II.2 Deuxième semestre	8
III Apprentissage hors-mur et année de césure	9
III.1 Participation à l'iGEM Competition	9
III.1.a Le concours iGEM	9
III.1.b L'équipe iGEM Paris-Saclay	10
III.1.c Ma participation à l'iGEM	11
III.1.d Bilan de cette expérience	14
III.1.e Références	15
III.2 Agrégation de mathématiques	15
IV M2 Mathématiques pour les Sciences du Vivant	16
IV.1 Premier trimestre	16
IV.2 Deuxième trimestre	16
IV.3 Troisième trimestre	17
V Conclusion	17

Cursus au sein du Magistère

Avant toute chose et car cela éclaire mes choix de matières/filière au sein de ma scolarité à Orsay, il me semble important de préciser mes motivations à venir y étudier.

J'ai intégré le magistère de mathématiques d'Orsay à l'issue de deux années de classes préparatoires au lycée Pierre Corneille de Rouen. Ce choix a été pour moi la concrétisation de mon intérêt nouvellement acquis pour les mathématiques et leur possible application aux sciences du vivant.

En effet, j'ai eu la chance lors de ces deux années d'apprendre auprès de professeurs qui ont su me donner envie de comprendre les mathématiques et les sciences plus généralement et pas seulement les apprendre bêtement. Aussi, je ne pouvais me résoudre à intégrer une école d'ingénieur et me suis informée sur la possibilité de poursuivre un enseignement mathématique que j'espérais (sans avoir été déçue sur ce point !) de qualité.

J'avais à l'issue du lycée mis de côté mon attrait pour les sciences du vivant et plus particulièrement la médecine ; à l'issue de la prépa, celui-ci s'est refait sentir et j'ai cherché un "moyen" d'allier mathématiques et médical : l'ouverture du Master Mathématiques pour les Sciences du Vivant à Orsay, qui me semblait tout adapté, a achevé de me convaincre d'intégrer le magistère d'Orsay...

I. L3 Mathématiques Fondamentales et Appliquées

La première année est une année plutôt généraliste et la spécialisation assez ténue.

1. Premier semestre

Le premier semestre est un semestre de tronc commun permettant d'acquérir les bases (relativement solides si on a compris l'ensemble des cours !) nécessaires pour la suite et est composé des cours obligatoires suivants :

- Algèbre,
- Topologie et calcul différentiel,
- Intégration de Lebesgue,
- Programmation, algorithmique et théorie de la complexité.

A cela, s'ajoute le choix d'une option parmi :

- Mathématiques et biologie,
- Théorie des graphes.

Au vu de mon projet d'intégrer le Master MathSV, j'ai choisis le premier. En fait, j'aurais aimé pouvoir remplacer le cours *Programmation, algorithmique et théorie de la*

complexité par celui de *Théorie des graphes*, ayant déjà suivi l'option informatique en prépa ; mais, cela n'était pas permis par le magistère.

Ce premier semestre m'a confirmée dans mon choix d'avoir intégré le magistère ; j'ai notamment particulièrement apprécié le cours d'*Intégration de Lebesgue*.

2. Deuxième semestre

Le tronc-commun se poursuit au second semestre via les cours :

- Analyse complexe,
- Équations différentielles,
- Probabilité,
- Analyse de Fourier.

et le choix d'une option parmi :

- Algèbre effective & TER,
- Informatique théorique.

J'ai choisi la première option car je désirais effectuer un TER – Travaux Encadrés de Recherche – afin d'avoir un premier contact avec la recherche. J'ai effectué mon TER en binôme avec Perinne Rouxel, sous la direction d'Hervé Lemeur. Une copie de mon rapport se trouve à la page 37 du présent document. Ce TER portait sur la modélisation mathématique par des techniques d'équations différentielles ordinaires d'une infection par le virus de l'immunodéficience humaine. Même si nous n'avons pas eu le temps d'approfondir le sujet, j'ai beaucoup apprécié ce travail qui a été pour nous l'occasion de produire pour la première fois un travail d'original et en autonomie partielle.

Enfin, le cursus de magistère différait du cursus classique par la présence d'un cours, dit spécifique au magistère, d'*Introduction à la topologie générale*. Ce cours s'est révélé quasiment indispensable en M1 où la plupart des enseignants le considéraient comme acquis. Malgré son caractère un peu décousu, je pense qu'il m'a apporté beaucoup mathématiquement.

II. M1 Mathématiques Fondamentales et Appliquées

Le M1 est l'occasion de choisir parmi un nombre conséquent de matières celles qui nous intéressent et ainsi se préparer au mieux pour notre M2 et/ou l'agrégation.

1. Premier semestre

Comme je désirais passer l'agrégation l'année suivante, j'ai choisi de suivre la matière *Analyse et Mathématiques générales 1*. Ensuite, n'ayant que peu apprécié le cours de probabilité du deuxième semestre de L3, j'ai choisi de suivre le cours de *Théorie des distributions*. Je regrette de n'avoir pu suivre le cours d'*Algèbre* mais ce dernier avait lieu en même temps que le cours d'*Analyse et Mathématiques générales*. J'ai par ailleurs découvert la théorie des représentations linéaires via ce cours et grâce à Pierre-Guy Plamondon, domaine que j'ai particulièrement apprécié au point d'en faire mon TER de M1, et sur lequel j'ai eu la chance d'être interrogée à l'oral de l'agrégation ¹!

A cela, s'ajoute le second cours spécifique au magistère *Compléments de théorie spectrale et d'analyse harmonique* dont j'ai préféré la seconde partie.

2. Deuxième semestre

Au second semestre, j'ai poursuivi mon apprentissage des équations aux dérivées partielles entamé lors du cours de *Théorie des distributions* par le suivi du cours *Problème d'évolution*. Le contenu du cours *Analyse et Mathématiques générales 2* étant vraiment trop orienté agrégation et désireuse de poursuivre en master recherche après l'agrégation j'ai choisi de suivre le cours de *Géométrie* d'Antoine Chambert-Loir. Je ne regrette pas ce choix car ce cours, notamment de par son exigence en terme de travail, m'a fait progresser énormément mathématiquement et a ma préférence sur l'année. En parallèle et du fait de mes choix de matières, j'ai naturellement choisis le MAO *Calcul scientifique*, présentant les bases de la résolution numérique des équations aux dérivées partielles et permettant une ouverture vers d'autres sciences, telles que la physique.

Enfin, le cours spécifique au magistère de ce semestre était une *Introduction aux systèmes dynamiques*. J'ai trouvé ce cours très intéressant quoique manquant parfois de "rigueur" mathématique, ce qui a rendu son suivi par moment difficile.

J'ai effectué mon deuxième TER en binôme avec Gabriel Laumosne et sous la direction de Pierre-Guy Plamondon autour des *Représentations linéaires du groupe symétrique*, page 53. Cette fois-ci pas de travail innovant mais la réécriture d'une preuve très lacunaire en s'efforçant de ne pas "tricher", *i.e* de ne pas aller chercher de preuve détaillée dans une référence autre... Cette fois encore, j'ai vraiment apprécié le travail de TER, d'où ma volonté de débiter une thèse aujourd'hui.

1. Je profite donc de cette note en bas de page pour remercier Pierre-Guy Plamondon pour ses qualités pédagogiques, sans lesquelles je n'aurais probablement pas aussi bien réussi cette épreuve...

III. Apprentissage hors-mur et année de césure

En lieu et place de mon stage de magistère, j'ai décidé de participer à l'iGEM Competition 2014, concours qui s'est étalé de novembre 2013 à octobre 2014.

Mon stage officiel a eu lieu en laboratoire, à l'IGM, du 23 juin au 10 août ; pour autant, le travail mathématique, philosophique, artistique et informatique s'est étalé sur la totalité de la période sus-citée.

1. Participation à l'iGEM Competition

Le concours iGEM est ouvert à tout étudiant n'ayant pas dépassé l'obtention du Master, quelle que soit sa filière. Il consiste en la réalisation d'un projet autour de la biologie synthétique et de la génétique.

Ce projet doit s'inscrire dans une ou plusieurs grande(s) thématique(s) telle(s) que les questions environnementales, et plus particulièrement les problèmes que posent le renouveau énergétique aujourd'hui, ou bien la santé publique : médecine, nutrition, autosuffisance alimentaire... Les équipes peuvent également choisir de travailler sur le développement de logiciels en lien avec la biologie synthétique, de réfléchir à de nouvelles applications pratiques -ou non- de la biologie synthétique voire, comme c'est le cas pour notre équipe cette année, de faire la "promotion" de la biologie synthétique auprès du grand public notamment par le biais de la pratique artistique.

La biologie synthétique, discipline relativement récente quoique en plein essor, est une forme de biologie tournée vers l'ingénierie, dont le but est, d'une part, de comprendre le vivant en le construisant et, d'autre part, de faciliter les manipulations sur les organismes vivants en utilisant des procédés propres à l'ingénierie tels que la standardisation et l'automatisation.

Les participants de l'iGEM reçoivent du MIT un kit de composants génétiques élémentaires et standardisés, appelés Biobricks. Les Biobricks distribuées ont été réalisées par les équipes des années précédentes. Ce sont des segments d'ADN, sous la forme de plasmides, capables de s'intégrer facilement à une espèce particulière de micro-organismes, la plupart du temps des bactéries. Les nouvelles équipes sont alors notamment en charge de créer de nouvelles Biobricks ou bien d'améliorer celles des années précédentes.

À noter que l'iGEM est entièrement Open Source.

a. Le concours iGEM

Le travail à fournir se décompose en plusieurs tâches, plus ou moins autonomes, et chaque étudiant se spécialise dans une ou plusieurs d'entre elles. Certaines, comme le

Cursus au sein du Magistère

sponsoring, sont pratiquées par l'ensemble de l'équipe. Sans ordre particulier et sans se prétendre exhaustif, on peut citer :

- La recherche de financements, indispensables à l'achat de matériels de laboratoire, aux frais d'inscriptions, de communication et de déplacement lors de la soutenance devant le jury (à Boston pour notre année) et donc à la mise en oeuvre d'un projet iGEM ;
- La communication auprès d'un public diversifié sur l'équipe, notre projet et la biologie synthétique en général. Cela passe notamment par la réalisation d'événement(s) de sensibilisation, d'intervention(s) dans des lycées ou bien par la participation à des manifestations telles que la fête de la science ou bien le festival CURIOSITAS, qui a eu lieu du 4 ou 9 octobre 2014 sur le Campus Paris-Saclay ;
- La réalisation en laboratoire d'une ou plusieurs Biobricks, l'établissement de protocoles expérimentaux et l'étude scientifique d'une problématique ;
- La réflexion philosophique et éthique autour des questions que pose le projet présenté. Cette réflexion peut prendre la forme d'essais philosophiques, d'entretiens avec des personnes impliquées -ou non- dans la biologie synthétique, etc ;
- La mise en place d'un cahier de laboratoire et la mise en ligne de l'ensemble de nos résultats de recherche sur internet (wiki) et la création d'un site internet pour représenter l'équipe ;
- La gestion de l'équipe et de la communication interne, de la trésorerie de l'association et des différentes réservations pour le voyage inhérent à la soutenance devant un jury aux États-Unis ;
- La modélisation mathématique liée à notre travail, que ce soit pour prédire de la faisabilité de certaines idées ou pour illustrer certaines observations faites en laboratoire ;
- De plus, de par sa nature artistique, notre projet exigeait une réflexion artistique et l'invention d'un concept et d'une présentation originale du projet scientifique en lui-même.

b. L'équipe iGEM Paris-Saclay

L'année 2014 signait la participation de l'université pour la troisième année consécutive au concours et a été récompensée par une médaille d'or.

L'équipe comptait une vingtaine d'étudiants membres venant de filières variées : de biologie mais également de physique, chimie, mathématique, informatique, bio-informatique, cancérologie, pharmacie, mécanique... Nous étions par ailleurs encadrés par cinq chercheurs de l'IGM – Jean-Luc Pernodet, Sylvie Lautru, Claire Toffano-Nioche, Olivier

Namy et Philippe Bouloc – et deux doctorantes – Solene Ithurbide et Alice Devigne.

Cette année, l'équipe avait décidé de concourir dans la catégorie *Art&Design*. Nous utilisons donc les outils de la biologie synthétique pour un projet de nature artistique voué à poser des questions éthiques. Plus exactement, nous souhaitons réaliser un citron moulé avec du milieu nutritif pour bactéries, dans lequel seraient disposées des bactéries génétiquement modifiées. L'objectif était que ces bactéries soient vertes et passent naturellement au jaune, comme si le citron artificiel mûrissait. Elles devaient également sentir une odeur proche de celle d'un vrai citron.



FIGURE 1. – Logo de l'équipe iGEM Paris-Saclay

Nous avons par la présente l'ambition de soulever plusieurs problématiques telles que le devenir de l'agro-alimentaire de nos jours et de l'éventualité, avec l'ère des imprimantes 3D qui s'annonce, d'une imprimante à "imprimer des aliments" à partir de cartouches nutritives. Notre projet posait également la question de la représentation du vivant, au sens propre comme au sens figuré. Plus généralement, nous nous sommes focalisés sur les liens qui existent entre pratiques scientifiques et artistiques.

c. Ma participation à l'iGEM

J'ai travaillé sur l'aspect modélisation mathématique du projet, en collaboration avec Pierre Roux et Eugène Ndiaye, tous deux magistériens.

Je me suis plus particulièrement intéressée à l'établissement d'un modèle de croissance pour la population bactérienne habitant notre citron artificiel. Nous avons, à cet effet, étudié différents ouvrages et publications sur le sujet [5, 29, 39]; en particulier le livre d'Eric Renshaw, *Modelling Biological Populations in Space and Time* [32] et celui de Norman T.J Bailey, *The Elements of Stochastic Processes with Applications to the Natural Sciences* [4].

Plutôt que de reproduire ici l'ensemble des équations, je joins un lien vers le wiki que nous avons réalisé lors de ce concours et qui résume l'ensemble de nos travaux :

Cursus au sein du Magistère

http://2014.igem.org/Team:Paris_Saclay [22] et particulièrement la section modélisation du présent wiki : http://2014.igem.org/Team:Paris_Saclay/Modeling. L'intérêt est double : je me suis en effet beaucoup investie dans la réalisation de ce wiki comme expliqué ci-dessous.

Je me suis en effet investie dans divers domaines allant de la biologie aux mathématiques en passant par l'informatique, l'éthique, le management de l'équipe en général, la cohérence artistique du projet, la levée de fond et l'organisation d'événements sociaux en lien avec le concours iGEM avec, par exemple, la mise en oeuvre d'une rencontre des différentes équipes françaises, sur notre campus, le weekend du 2-3 août 2014.



FIGURE 2. – Univers graphique pour le projet de l'année 2014

Plus spécifiquement, je me suis assez considérablement investie dans la gestion de l'équipe, ce qui implique une connaissance globale du rôle des uns et des autres et de l'avancement des différents sous-projets pour permettre une meilleure coordination au sein de l'équipe. Cela s'est avéré extrêmement chronophage et bien plus complexe que je

ne le pensais auparavant ; mais riche d'apprentissage sur comment travailler en équipe et veiller au bon déroulement des tâches, dès lors qu'une vingtaine de personnes sont concernées et ne sont alors plus aussi à même qu'en petit comité de communiquer les unes avec les autres.

Ce stage a été aussi pour moi l'occasion de me former en programmation web car j'ai dû apprendre, pour les besoins de la cause, l'HTML et le CSS ainsi que la marche à suivre générale qui accompagne la création d'un nouveau site internet "ambitieux" de nos jours.

Le développement de notre wiki, hébergé par la fondation iGEM à l'adresse sus-citée, s'est déroulé en trois temps :

- i. La première étape consistait en la mise en place d'un design pour notre wiki : il fallait décider de la forme que l'on voulait lui donner, etc. Cette tâche était d'autant plus importante que nous avons décidé, pour l'année 2014, de concourir dans la catégorie *Art&Design* et que l'unité artistique de notre projet était donc capitale. Cette tâche a été majoritairement réalisée par trois autres étudiants – un étudiant suivant un cursus très accès programmation, une étudiante passionnée de dessin et un troisième maîtrisant l'art de la tablette graphique – et moi-même. Il ne s'agissait pas uniquement de réfléchir aux couleurs de la police mais à la forme globale que nous souhaitions donner à notre wiki. Outre la limitation graphique se posait également les questions de l'ergonomie du site internet et de la faisabilité, côté programmation, des différentes idées. En effet, la politique de l'iGEM étant l'open source, notre wiki ne devait comporter que des langages basiques tels que l'HTML ou le CSS pour permettre une modification par le plus grand nombre.
- ii. La seconde étape a consisté en la réalisation des différents éléments conceptualisés lors de la première étape sur des logiciels de dessin.
- iii. Enfin, la programmation à proprement parler du wiki et l'intégration des différents éléments graphiques dans ce dernier.

J'ai également joué un rôle important dans la partie éthique puisque j'ai participé à la réalisation de deux "questionnaires" éthiques :

- i. Le premier, dont la paternité impute en grande partie à Eugène Ndiaye, consiste en un court questionnaire, d'une dizaine de questions, portant sur le BioArt et la représentation du vivant en général. Nous avons décidé de le diffuser auprès des autres équipes iGEM afin de récolter leur impression sur le sujet. La principale contribution d'Eugène a consisté en l'explication aux biologistes de la nécessité de cibler une population type si l'on désire obtenir des résultats statistiques valides, ceux-ci n'ayant qu'une très vague notion des statistiques.
- ii. Le second, qui m'a personnellement plus intéressée, même si son intérêt mathématique est clairement plus modique, est plus accès sur la thématique du BioArt et du

lien que l'on peut faire entre pratique artistique et scientifique. Plus exactement, il s'agissait d'une trame pour nous permettre de mener à bien les entretiens que nous avons organisés auprès d'un public averti (scientifiques intéressés par l'art et inversement) que d'un réel questionnaire.

Enfin, au même titre que mon collègue Pierre Roux, j'ai fait partie des douze étudiants, sélectionnés par l'équipe pour la représenter devant le jury, à Boston, du 27 octobre au 3 novembre 2014. Cela a été pour moi l'occasion de me former à la réalisation d'un exposé mathématique – puisqu'il s'agissait d'exposer le fruit de nos travaux en modélisation – en anglais, devant un public pluridisciplinaire.

d. Bilan de cette expérience

En premier lieu, ma participation à l'iGEM m'a confirmée dans mon choix de poursuivre mes études en M2 MathSV car j'ai aimé l'aspect interdisciplinaire de mon travail mathématique, d'autant que j'éprouve un réel intérêt pour les sciences du vivant d'une manière générale.

En second lieu, ma participation à l'iGEM a été, pour moi, l'occasion de mettre des mots sur la notion de travail en équipe : je n'avais en effet jamais eu l'occasion de monter un projet aussi construit, chronophage et abouti avec autant de personnes. D'autant plus avec des étudiants non-mathématiciens qui ont souvent des manières d'aborder les choses différentes des étudiants que j'ai eu l'habitude de côtoyer dès lors qu'il s'agit de travail à plusieurs, depuis mon arrivée à l'université Paris-Sud.

D'un point de vu purement mathématique, cela a été pour moi l'occasion de m'intéresser aux principes mathématiques de la modélisation déterministe et stochastique, en l'occurrence de l'évolution d'une population, qui m'étaient inconnus, n'ayant suivi ni cours de probabilité, ni cours de statistiques en M1. Je cerne également mieux le travail du mathématicien qui souhaite travailler en étroite collaboration avec des chercheurs en biologie.

Cela a été également l'occasion pour moi de m'introduire à la microbiologie, une connaissance suffisante des mécanismes en jeux étant nécessaire à l'exercice de la modélisation rigoureuse. J'ai aussi profité de cette expérience pour, enfin, me lancer dans l'apprentissage du codage web et parfaire mon anglais.

En résumé, je dirais que bien que mon "stage" ait été plus chronophage et éprouvant qu'un stage plus conventionnel, je ne le regrette pas car l'expérience s'est avérée enrichissante.

e. Références

- [4] Norman T.J Bailey. *The Elements of Stochastic Processes with Applications to the Natural Sciences*. Wiley Classics Library. Wiley-Blackwell, 1990.
- [5] Nicolas Burq and Patrick Gérard. Contrôle optimal des équations aux dérivées partielles.
- [8] Gaëlle Degrez. Des étudiants de paris-sud obtiennent la médaille d'or à l'iGEM, Dec 10, 2014. <http://www.actu.u-psud.fr/fr/formations/actualites-2014/des-etudiants-de-paris-sud-obtiennent-la-medaille-d-or-a-l-igem.html>.
- [22] iGEM Paris-Saclay. Wiki de l'équipe – this is not a lemon, 2014. http://2014.igem.org/Team:Paris_Saclay.
- [29] D.R Miller and H.D Cox. *The Theory of Stochastic Processes*. Methuen & Co, 1965.
- [32] Eric Renshaw. *Modelling biological populations in space and time*. 1993.
- [39] Eric J. Stewart, Richard Madden, Gregory Paul, and François Taddei. Aging and death in an organism that reproduces by morphologically symmetric division. *PLoS Biology*, 3(2), February 2005.

2. Agrégation de mathématiques

A la suite de mon M1 et donc en partie à cheval avec la fin de l'iGEM, j'ai décidé comme il nous est fortement conseillé à Orsay de réaliser une année de césure dans mon cursus pour passer l'Agrégation de Mathématiques.

Comme j'avais apprécié mon MAO de M1, j'ai choisi de préparer l'option *Calcul Scientifique*. Cette option m'a permis d'acquérir des connaissances en programmation – et notamment en langage python – qui me sont aujourd'hui utiles. De plus, c'est une option tournée vers les autres sciences avec de nombreuses problématiques issues de la physique et des sciences du vivant, ce qui m'a forcée à me remettre à la physique et devrait m'être utile pour ma thèse.

J'ai choisi lors de cette année de travailler pour renforcer mes connaissances mathématiques générales et me préparer pour le master recherche – notamment en renforçant mes connaissances en probabilité – plus que d'optimiser mon classement, mon but n'étant *a priori* pas d'enseigner à terme... J'ai cependant décidé de réaliser un stage enseignement cette année-là et non recherche pour pouvoir écarter la voie de l'enseignement "en connaissance de cause".

IV. M2 Mathématiques pour les Sciences du Vivant

L'année de M2 MathSV – comme pour la plupart des master recherche – se découpe en trimestre, le dernier étant réservé à la réalisation d'un stage.

1. Premier trimestre

Les cours du premier semestre permettent d'avoir un aperçu solide de la diversité des manières d'appliquer les mathématiques à la biologie :

- Statistiques en grande dimension,
- Processus stochastiques,
- Modélisation déterministe,
- Optimisation et simulation numérique,
- Concept fondamentaux de la biologie et de l'écologie,
- Séminaire hebdomadaire de biologie, d'écologie et de médecine.

Les deux premiers m'ont permis de me former en probabilité et en statistiques n'ayant pas suivi les matières éponymes en M1 ; le dernier m'a tout particulièrement intéressée : il s'agissait de conférences courtes (moins d'une heure en générale) où un chercheur travaillant sur des problématiques à l'interface entre les sciences du vivant et les mathématiques venaient nous présenter ses travaux.

A contrario l'unité d'enseignement de sciences du vivant m'a déçu : le niveau était à peine plus élevé que celui du lycée et, je n'y ai essentiellement rien appris de plus que mes connaissances acquises en autodidacte par mes lectures personnelles...

A cheval sur les deux premiers trimestres, mon "projet" de M2, dont le rapport se trouve à la page 137 portait sur l'application de la géométrie riemannienne à l'imagerie médicale. J'ai pour cela travaillé en binôme avec Pierre Roux et nous avons été encadrés par Joan Glaunès qui m'a également encadrée en stage. Ce TER fait échos au cours *Géométrie et Espaces de formes* du second trimestre ; je dirais même que sans ce TER, je n'aurais probablement pas compris ce cours et n'aurais pas poursuivi en thèse dans ce domaine. En effet, ce TER a été l'occasion pour mon binôme et moi-même d'acquérir les connaissances en géométrie riemannienne nécessaires à la compréhension du cours d'Alain Trounev et des différents modèles mis en jeu.

2. Deuxième trimestre

Le deuxième trimestre est un trimestre de spécialisation. Nous devons choisir trois unités d'enseignements parmi la douzaine proposée, plus une pour le magistère. J'ai personnellement suivi les cinq cours suivants :

- Modèles dynamiques en biologie cellulaire,
- Méthodes à noyaux pour l'apprentissage,
- Imagerie fonctionnelle cérébrale et interface cerveau ordinateur,
- Géométrie et espace de formes,
- Modélisation en neurosciences et ailleurs.

Le premier était le seul cours du master orienté système dynamique et modélisation via les équations différentielles ordinaires ; je l'ai choisi car j'avais particulièrement apprécié mon TER de L3. Les quatre suivants sont mutualisés avec le master Mathématiques Vision et Apprentissage de l'ENS Cachan, le cours de *Modélisation en Neurosciences et ailleurs* n'étant d'ailleurs pas officiellement dispensé par la master MathSV...

J'ai choisis de suivre le cours de *Méthodes à noyaux pour l'apprentissage* afin de découvrir le machine-learning et car c'est un domaine en pleine expansion et utilisant fortement le cours de Christophe Giraud, *Statistiques en grande dimension*. Il devrait m'être utile en thèse. Les deux cours suivants sont des cours de mathématiques appliquées à l'imagerie médicale. Le premier, plus appliqué, présente notamment comment analyser et traiter des données d'imagerie "brutes". Enfin, j'ai choisis de suivre le cours de Jean-Pierre Nadal *Modélisation en neurosciences et ailleurs* car les neurosciences m'ont toujours intéressées et que je voulais m'initier à la physique statistique et aux techniques de modélisation via l'entropie d'un système donnée.

Le cours d'Alain Trounev *Géométrie et Espaces de formes* m'a beaucoup plu, au même titre que mon TER de M2, aussi ai-je donc réalisé mon stage – et devrais-je réaliser ma thèse – sur un sujet connexe.

3. Troisième trimestre

Le troisième trimestre est consacré à la réalisation d'un stage orienté recherche. J'ai effectué le miens d'avril à août 2016 au laboratoire de Mathématiques Appliquées de Paris 5, MAP 5, de l'université Paris-Descartes, sous la direction de Joan Glaunès. Celui-ci portait sur l'appariement de formes par difféomorphisme, qui est une des problématiques de l'anatomie computationnelle. Mon rapport se trouve à la page 201.

J'ai beaucoup aimé cet exercice, ma principale frustration résidant dans le fait de n'avoir pu aboutir...

V. Conclusion

Mon cursus à Orsay de par son caractère généraliste m'a formée à pouvoir appréhender de nombreux domaines des maths, contrairement à des formations très spécifiques :

Cursus au sein du Magistère

ainsi, pour ma thèse, devrais-je être à même de recourir à de nombreux outils mathématiques et ainsi résoudre avec plus de facilité la problématique qui m'est posée, à savoir arriver à une meilleure compréhension du cancer du rein à partir de données médicales et ce par le biais de la géométrie riemannienne et de l'apprentissage statistique.

Déjà lors de mon M2, j'ai pu apprécier la qualité de ma formation par la compréhension, au moins partielle, de conférences que j'ai décidé de suivre par moi-même et parfois éloignées des cours que j'ai pu suivre lors de ma scolarité. De même en stage où j'ai dû me former seule aux opérateurs pseudo-différentiels – au moins sur les définitions et résultats élémentaires – ce qui n'aurait été possible sans un niveau initial correct en analyse fonctionnelle et notamment le suivi des cours *Problèmes d'évolution* et *Compléments de théorie spectrale et d'analyse harmonique* en première année de master.

Enfin, les oraux spécifiques au magistère devraient m'aider à la préparation de mes – je l'espère ! – futurs exposés.

– Deuxième partie –

Espaces de formes et recalage difféomorphique

Imagerie médicale et Anatomie computationnelle.

Sommaire – Partie II

Espaces de formes et recalage diffeomorphique	19
	22
I Généralités sur les espaces de formes	22
I.1 Encoder les formes	23
I.2 L'appariement de formes en pratique	25
II Appariement et modèles de déformations	26
II.1 Modèles de petites et grandes déformations	26
II.2 Plusieurs types d'appariements	29
III Appariement et métamorphoses	32
IV Étude transversale et bases de données longitudinales	33
V Références	34

Du fait de l'amélioration des techniques d'imagerie médicale, l'anatomie computationnelle² connaît un essor considérable ces dernières années. On entend par anatomie computationnelle l'étude des formes anatomique et de la morphologie à l'échelle macroscopique : c'est donc une discipline à l'interface entre les mathématiques comme outil de modélisation, l'informatique et l'analyse statistique.

Je m'intéresse pour l'instant essentiellement au premier point, autrement dit à produire des modèles mathématiques permettant l'interprétation statistique des données brutes fournies par les médecins. Comme décrit dans ce qui suit, la géométrie riemannienne est au cœur de cet exercice de modélisation.

Plus précisément, le modèle proposé doit permettre de comparer différentes formes anatomiques entre elles afin de pouvoir quantifier les variations inter et intra-individuelles³ de la dite forme et c'est là l'objet de la théorie des espaces de formes. En effet, outre la grande variabilité des données d'acquisition qui justifie la mise en place d'une théorie générale et non organe-dépendante, il faut bien prendre conscience que la nature nous a fait tous semblables mais loin d'être identiques. Ainsi, la forme d'un même organe peut varier très fortement d'un individu à l'autre et il apparait nécessaire de pouvoir quantifier cette variabilité afin de déterminer si deux organes diffèrent de par leur nature propre ou bien parce que l'un est malade et l'autre non.

I. Généralités sur les espaces de formes

La théorie des espaces de formes apparait comme une tentative de modélisation mathématique de cette variabilité. On introduit pour cela un espace, dit *espace de formes* dans lequel chacun des points est en fait un objet géométrique et que l'on va munir d'une métrique à même donc de quantifier l'écart entre deux objets géométriques. Pour ce faire, on définira la distance entre deux formes comme "la transformation de l'espace la moins coûteuse permettant d'amener une forme sur l'autre" : aussi s'attachera-t-on pour définir la distance à minimiser une fonctionnelle d'énergie (cf paragraphe II.1), tout cela dans un sens à préciser. Cette distance est dite géodésique et l'espace de formes ainsi défini est en particulier une variété riemannienne.

2. On s'autorise le néologisme "computationnel(le)" qui semble plus approprié sémantiquement parlant qu'une traduction plus brutale en numérique ou informatique...

3. Afin de suivre l'évolution d'une maladie par exemple...

1. Encoder les formes

Comme annoncé, l'idée des espaces de formes est de définir un objet mathématique – une variété riemannienne dans les faits – dont chacun des points représentera une forme anatomique. Il existe pour cela plusieurs manières de procéder et on en explicite ici trois, que l'on peut considérer comme des améliorations successives d'une même idée.

Points de repère : L'idée la plus naturelle pour encoder une forme est peut-être de "simplement" la discrétiser en interpolant sa surface avec un ensemble de points. Ainsi, une forme de dimension d sera encodée par un ensemble de n points de \mathbb{R}^d avec $n \in \mathbb{N}$ fixé. On parle alors de l'espace des points de repère ou landmarks en anglais, qui n'est autre que \mathbb{R}^{nd} en imposant que chacun des points soit distinct des autres :

$$\mathcal{L}^n(\mathbb{R}^d) = \{ (x_1, \dots, x_n) \in (\mathbb{R}^d)^n \mid \forall i, j \in \llbracket 1, n \rrbracket, i \neq j \implies x_i \neq x_j \}.$$

Dans un tel formalisme, apparier deux formes est équivalent à apparier n points deux à deux. L'intérêt d'un tel modèle réside principalement dans la simplicité de la mise en œuvre théorique et algorithmique de l'appariement de formes. Cependant, ce formalisme souffre de limitations importantes : la première est que ne sont appariables que des formes interpolées par le même nombre de points et à condition que l'on ait pris soin de spécifier quel point devait être apparié avec tel autre ; la seconde est que l'interpolation détruit la structure géométrique de la forme : une courbe n'est pas distinguable d'un nuage de points, etc.

Mesures discrètes : Pour pallier la première limitation, une amélioration possible est de ne plus considérer des n -uplets de points comme précédemment mais des sommes de mesures de Dirac, enracinées en chacun des dits-points, [17]. Ainsi, une forme (anatomique) sera représentée par une mesure de la forme :

$$\mu = \sum_{i=1}^{n_\mu} \delta_{x_i} \quad \text{où} \quad \forall i \in \llbracket 1, n_\mu \rrbracket, x_i \in \mathbb{R}^d.$$

En particulier, le nombre de points n_μ peut varier d'une forme à l'autre ce qui permet de coller au mieux aux données que l'on cherche à étudier. De plus, lors de l'appariement, l'utilisateur n'est pas obligé de spécifier quels points apparier deux à deux. Sur ce dernier point néanmoins, on aurait pu s'affranchir de cette contrainte dans le cadre des points de repère en décidant de réaliser les appariements pour l'ensemble des configurations possibles puis en minimisant la fonctionnelle correspondant au coût d'appariement sur l'ensemble de ces configurations ; mais, ce choix s'avère peu judicieux en pratique car alourdissant les équations géodésiques et le temps de calcul.

Éventuellement, on peut considérer une somme pondérée de mesures de Dirac, ce qui permet de mettre l'accent sur certaines zones anatomiques relativement à d'autres.

Espaces de formes et recalage diffeomorphique

Ce ne sont bien sûr pas les seuls espaces de formes existants. On peut par exemple recourir aux *courants*. Ils permettent en effet de pallier le second problème évoqué. En fait, on peut voir les courants comme une généralisation des sous-variétés : si N est une sous-variété compacte orientée de dimension p de M , variété lisse, – typiquement \mathbb{R}^d – alors on définit un p -courant sur M en posant :

$$\forall \omega \in \Omega_0^p(M), \quad T_N(\omega) := \int_N \omega$$

où $\Omega_0^p(M)$ désigne l'espace des p -formes différentielles sur M à support compact.

Ainsi, en appariant deux courants [16], on appairera deux structures géométriques, sans détruire les propriétés qualitatives de celles-ci. Cependant, les courants eux-mêmes souffrent de leur propre limitation : principalement, ils demandent de pouvoir orienter les formes que l'on cherche à appairer, ce qui n'est pas toujours possible en pratique, notamment dès lors que l'on cherche à étudier de la matière blanche comme l'illustre la figure 3.

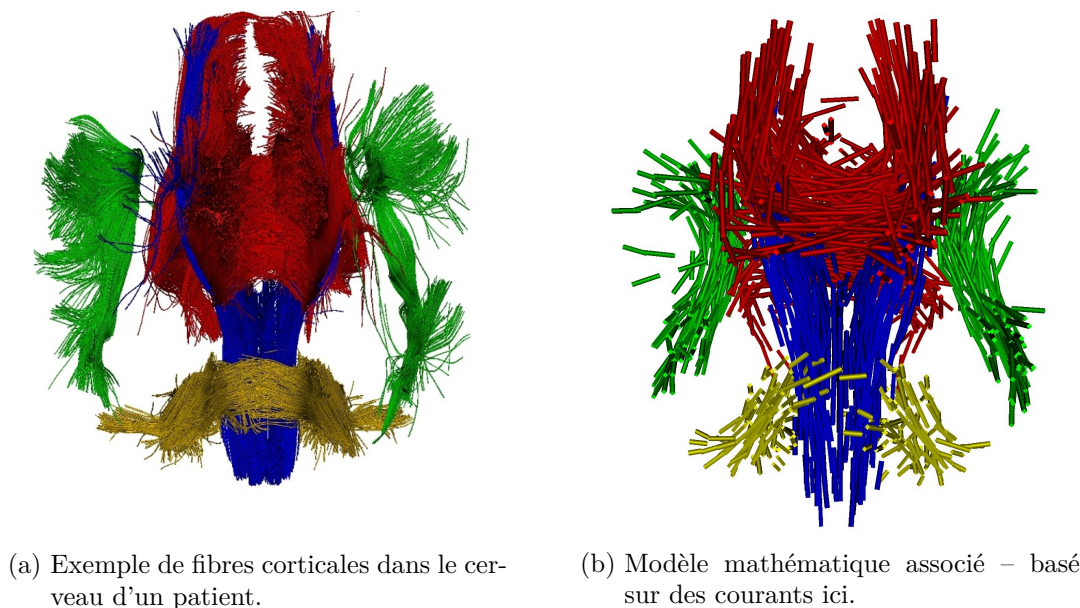


FIGURE 3. – Des faisceaux de fibres dans le cerveau.

Crédit : *D. Ducreux, Hôpital Kremlin-Bicêtre et P. Fillard, CEA, Neurospin.*

La matière blanche est principalement constituée d'axones, d'où son aspect "fibreuse". Elle est responsable de la propagation des informations dans le système nerveux et relie notamment les différentes aires de la matière grise, située en périphérie du cerveau. Sont représentés ici : en bleu, le faisceau cortico-spinal ; en jaune, le faisceau cortico-nucléaire ;

en rouge, le corps calleux et en vert les faisceaux arqués droit et gauche.

En médecine, il y a un réel intérêt à pouvoir observer ces fibres en tant que fibres justement car cela permet de détecter des maladies telles que la sclérose en plaque, qui n'est autre que la destruction de la myéline protégeant les axones de la matière blanche, la maladie d'Alzheimer et d'autres maladies neurodégénératives. Cependant, l'orientation des différentes fibres corticales est loin d'être claire d'où l'utilité d'un formalisme ne nécessitant aucune orientation. Les *varifolds* que nous ne présentons pas non-plus ici permettent par exemple cela, [6].

Précisons cependant que les varifolds reposent également sur la notion de mesure de Dirac. Finalement, on constate qu'une grande partie de l'appariement anatomique repose donc sur l'appariement de mesures discrètes, via celui de sommes de Dirac.

2. L'appariement de formes en pratique

L'intérêt pratique de l'appariement de formes est grand, notamment dans le domaine de l'imagerie médicale. En effet, comme l'avait déjà compris D'Arcy Thompson en 1917, dès lors que l'on s'intéresse à la morphologie de nos différents organes, le plus important n'est pas d'être à même de fournir une description détaillée de chacun d'entre eux mais plutôt de pouvoir quantifier les déformations des uns par rapport aux autres, et c'est là l'idée principale du modèle dit *de grandes déformations* et faisant l'objet du paragraphe suivant.

In a very large part of morphology, our essential task lies in the comparison of related forms rather than in the precise definition of each... This process of comparison, of recognizing in one form a definite permutation or deformation of another,...lies within the immediate province of mathematics and finds its solution in...the Theory of Transformations... I learnt of it from Henri Poincaré.

D'Arcy Thompson – *On Growth and Form*, 1917, [42].

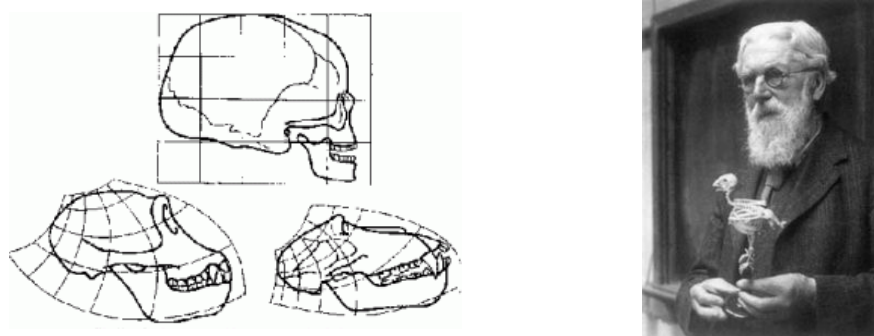


FIGURE 4. – Crânes d'humain, de chimpanzé et de babouin. – D'Arcy Thompson.

Comme expliqué en introduction, pour tenir compte de la variabilité intrinsèque des formes anatomiques, on mesure leur similitude par le biais d'une fonctionnelle encodant le coût de déformation nécessaire à la "superposition" de la première forme sur la seconde, et qui ne dépend bien évidemment pas du choix de l'organe à déformer, par symétrie du problème. Il peut également être judicieux de considérer que toutes les déformations ne se valent pas, notamment pour le diagnostique précoce de maladie où certaines déformations spécifiques d'un organe (hippocampe dans le cas d'Alzheimer par exemple) vont être symptomatiques alors que d'autres sont juste le fruit de la diversité anatomique. Dans ce cas, on imposera un poids plus important sur ces déformations spécifiques dans le coût total de déformation.

Compte-tenu de l'intérêt pratique de l'appariement de formes en anatomie computationnelle, il est nécessaire de concevoir des méthodes implémentables et donc utilisables *in extenso* par des praticiens, [14].

II. Appariement et modèles de déformations

Dans l'approche classique, les déformations sont encodées via un groupe agissant sur l'espace de formes. En fait de groupe, on va considérer une "collection" de transformations de l'espace. Historiquement, on a d'abord considéré des groupes de transformations "simples" : rotation, translation, homothétie... Mais, cela ne permettait pas de modéliser tout type de déformation. D'autre part, on a envie que les déformations soient réversibles pour des questions de symétrie évidentes et donc que les transformations mathématiques associées soient inversibles. Ainsi, considérer un groupe de diffeomorphismes – avec éventuellement des restrictions – est une idée naturelle.

1. Modèles de petites et grandes déformations

Un premier modèle dit de *petites déformations* reposait sur la minimisation d'un coût de déformation pour le champ de déplacement $\phi - Id$, avec l'idée que le morphisme Id correspond à la situation idéale où aucune déformation n'est nécessaire. Cependant, dès lors que la déformation devient trop importante, *i.e.* dès lors que ϕ devient grand devant Id en norme, un tel morphisme n'est plus inversible.

Aussi, plutôt que de considérer le résultat des déformations, le modèle de *grandes déformations* consiste à utiliser la même approche que précédemment, mais de manière infinitésimale. Plus précisément, on se donne un espace (de Hilbert) V de champs de vecteurs dont la norme définit le coût de déformation et les diffeomorphismes sont obtenus par intégration d'une famille $(v_t)_{t \in [0,1]}$ de champs de vecteurs selon l'équation :

$$\partial_t \phi_t^v = v_t \circ \phi_t^v \quad ; \quad \phi_0^v = Id.$$

Ainsi, en minimisant la norme des différents champs de vecteurs v_t , on minimise à chaque instant la déformation et donc la déformation globale, qui est donnée par ϕ_1 .

En fait, certaines limitations sont nécessaires. On se contente d'énoncer ici les théorèmes généraux qui soutiennent la suite, sans démonstration. On en trouve une démonstration au chapitre 1 de la thèse de Joan Glaunès, *Transport par difféomorphismes de points, de mesures et de courants pour la comparaison de formes et l'anatomie numérique* [16], ainsi que des résultats complémentaires.

Tout d'abord, il faut remarquer que l'équation, telle qu'écrite ci-dessus, n'est pas à valeur dans \mathbb{R}^d mais dans un espace fonctionnel. Par ailleurs, $t \mapsto \phi_t^v$ n'est à priori pas à valeurs dans \mathbb{R}^d mais dans la variété riemannienne de l'espace des formes. Aussi lui préférera-t-on le plus souvent son écriture intégrale où \int désigne l'intégrale de Bochner qui généralise aux espaces de Banach la notion d'intégration :

$$\phi_t^v(x) = x + \int_0^t v_s \circ \phi_s^v(x) ds.$$

Enfin, une telle équation est dite *équation de flot* et, compte-tenu des limitations suscitées, les résultats classiques d'existence et d'unicité ne s'appliquent donc pas en l'état. En fait, à condition d'introduire une restriction sur les champs de vecteurs v_t autorisés, on "récupère" l'existence et l'unicité à l'équation de flot.

Définition 1. *Un espace vectoriel V de champs de vecteurs de \mathbb{R}^d est dit admissible si et seulement si :*

- i. il est muni d'une structure d'espace de Hilbert. Dans ce cas, on note $\|\cdot\|_V$ sa norme.*
- ii. il s'injecte continûment dans l'espace $\mathcal{C}_{\rightarrow 0}^1(\mathbb{R}^d)$ des champs de vecteurs de classe \mathcal{C}^1 sur \mathbb{R}^d s'annulant à l'infini, ainsi que leurs dérivées partielles. En particulier, il existe une constante C telle que pour tout $v \in V$,*

$$\|v\|_\infty + \|\mathcal{D}v\|_\infty \leq C \|v\|_V \quad \text{où} \quad \mathcal{D}v: x \mapsto \mathcal{D}_x v.$$

Dorénavant, V désignera un espace de Hilbert *admissible*.

On note $L_V^1 := \mathbb{L}^1([0, 1], V)$ et $L_V^2 := \mathbb{L}^2([0, 1], V)$ les espaces des classes d'équivalences des applications mesurables $t \mapsto v_t$ de $[0, 1]$ dans V telles que les quantités respectives $\|v\|_{L_V^1} := \int_0^1 \|v_t\|_V dt$ et $\|v\|_{L_V^2}^2 := \int_0^1 \|v_t\|_V^2 dt$ soient finies. Munis de leur norme usuelle, L_V^1 est un espace de Banach et L_V^2 un espace de Hilbert. De plus, d'après l'inégalité de Cauchy-Schwarz, pour tout champ de vecteur v , $\|v\|_{L_V^1} \leq \|v\|_{L_V^2}$ et donc L_V^2 s'injecte continûment dans L_V^1 .

Théorème 2 (Cauchy-Lipschitz, [16])

Pour tout champ de vecteur $v \in \mathbb{L}_V^1$ et tout point $x \in \mathbb{R}^d$, il existe une unique application continue $t \mapsto \phi_t^v(x)$ de $[0, 1]$ dans \mathbb{R}^d et vérifiant l'équation de flot.

Ainsi, on est assuré de l'existence d'un morphisme traduisant l'effet des déformations infinitésimales. On a également un contrôle des variations du flot ϕ_t^v en temps, en espace et vis-à-vis du champ v , de manière quasiment automatique via le *lemme de Gronwall*. Un résultat plus important :

Proposition 3

Pour tout $v \in \mathbb{L}_V^1$, le flot ϕ_t^v est un difféomorphisme de classe \mathcal{C}^1 en chaque instant t de $[0, 1]$.

En particulier, au temps $t = 1$, ϕ_1^v est régulier et inversible, ce qui valide l'approche. De plus, si on impose à V de s'injecter continûment dans \mathcal{C}_0^p pour $p \geq 1$, les déformations ϕ_t^v induites seront d'autant plus régulières.

Finalement, on construit le groupe (cf proposition 4) des déformations comme l'ensemble des difféomorphismes ϕ_t^v obtenus à partir des différents champs de vecteurs $v \in \mathbb{L}_V^1$ en posant :

$$\mathcal{G}_V := \left\{ \phi_t^v \mid v \in \mathbb{L}_V^1, t \in [0, 1] \right\}.$$

En fait, il suffit de considérer les flots au temps $t = 1$. En effet, pour tout $t \in [0, 1]$ et tout $v \in \mathbb{L}_V^1$, $\phi_1^{\tilde{v}} = \phi_t^v$ où on définit \tilde{v} par : $\forall s \leq t, \tilde{v}_s = v_s$ et $\tilde{v}_s = 0$ sinon. Finalement, on écrira plus simplement :

$$\mathcal{G}_V = \left\{ \phi_1^v \mid v \in \mathbb{L}_V^1 \right\}$$

et la proposition suivante permet de transporter le coût de déformation encodé dans la norme de V sur le groupe \mathcal{G}_V :

Proposition 4

L'ensemble \mathcal{G}_V est un groupe et un espace complet pour la métrique d_V définie par :

$$\forall \phi \in \mathcal{G}_V, \quad d_V(\text{Id}, \phi) := \inf \left\{ \|v\|_{\mathbb{L}_V^1} \mid v \in \mathbb{L}_V^1, \phi_1^v = \phi \right\}$$

et que l'on étend par invariance à droite :

$$\forall \phi, \psi \in \mathcal{G}_V, \quad d_V(\phi, \psi) = d_V(\text{Id}, \psi \circ \phi^{-1}).$$

Remarque : La propriété de groupe se vérifie en remarquant que pour tout champ de vecteurs $u, v \in \mathbb{L}_V^1$, $\phi_1^u \circ \phi_1^v = \phi_1^{u \star v}$ où l'opération \star est l'opération de *concaténation*

définie par :

$$\forall t \in [0, 1], \quad (u \star v)_t = \begin{cases} \frac{1}{2}u_{2t} & \text{si } t \geq \frac{1}{2} \\ \frac{1}{2}u_{2t-1} & \text{si } t \leq \frac{1}{2} \end{cases}$$

et en remarquant que $(\phi_1^v)^{-1} = \phi_1^{\hat{v}}$ où pour tout temps $t \in [0, 1]$, $\hat{v}_t = v_{1-t}$.

En fait, la construction du groupe \mathcal{G}_V via des champs de vecteurs \mathbb{L}_V^1 reste vraie avec des champs de vecteurs \mathbb{L}_V^2 , ce qui simplifie souvent le coût algorithmique :

Proposition 5

D'une part,

$$\mathcal{G}_V = \left\{ \phi_1^v \mid v \in \mathbb{L}_V^2 \right\}$$

et, d'autre part,

$$\forall \phi \in \mathcal{G}_V, \quad d_V(Id, \phi) = \inf \left\{ \|v\|_{\mathbb{L}_V^2} \mid v \in \mathbb{L}_V^2, \phi_1^v = \phi \right\}.$$

Enfin, le théorème suivant assure l'existence d'au moins une géodésique, *i.e.* d'un champ de vecteur v permettant de transporter une configuration de l'espace de formes sur une autre et pour lequel le coût de déformation est minimal. Ainsi, le problème d'appariement admet (au moins) une solution.

Théorème 6 (*Existence de géodésique(s)*, [16])

Pour tous les difféomorphismes $\phi, \psi \in \mathcal{G}_V$, il existe un champ de vecteur $v \in \mathbb{L}_V^2$ tel que :

$$\phi_1^v = \psi \circ \phi^{-1} \quad \text{et} \quad d(\phi, \psi) = \|v\|_{\mathbb{L}_V^2} = \|v\|_{\mathbb{L}_V^1}.$$

De plus, pour un tel champ v , la norme $\|v_t\|_V$ est constante pour presque tout temps $t \in [0, 1]$.

2. Plusieurs types d'appariements

Depuis le début de cette partie, on parle d'appariement de formes. En fait, plusieurs modèles ont été développés pour permettre un tel appariement ; aussi résume-t-on d'abord les modèles principaux ainsi que leurs avantages et limitations respectives.

Très schématiquement, l'appariement de deux formes – dans notre cas d'étude mais cela reste également vrai pour des images – consiste à déterminer le difféomorphisme optimal $\phi = \phi_1^v$ tel que la première des deux formes, transportée par ϕ , soit la plus "proche" possible – en un sens à préciser – de la seconde.

Par la suite, on désigne par y et z la première et seconde forme respectivement.

Appariement exact : En accord avec les notations du paragraphe précédent, on note M l'espace de formes considéré et \mathcal{G}_V le groupe des diffeomorphismes agissant sur M . Pour tout $x \in M$ et $\phi \in \mathcal{G}_V$, $\phi(x)$ désigne donc l'image de x transportée par ϕ .

Définition 7. Le problème d'appariement exact consiste à minimiser, sur \mathcal{G}_V , la fonctionnelle suivante

$$J(\phi) := d_V(\text{Id}, \phi) \quad \text{sous contrainte que} \quad \phi(y) = z.$$

Concrètement, on force l'appariement à être exact, *i.e.* que $\phi(y) = z$ et on minimise la déformation induite par ϕ : le morphisme Id correspondant à la situation idéale où aucune déformation n'est nécessaire pour réaliser l'appariement. De manière équivalente, et d'après le théorème 6, il s'agit de minimiser sur \mathbb{L}_V^2 , et avec la même contrainte, la fonctionnelle

$$J(v) := \int_0^1 \|v_t\|_V^2 dt.$$

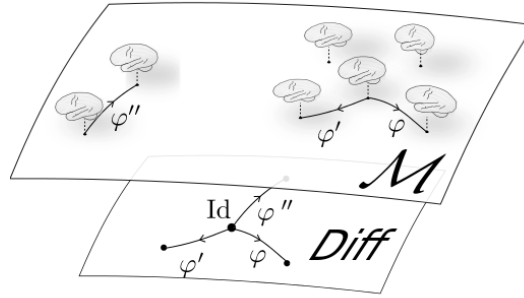


FIGURE 5. – Équivalence entre distance anatomique et coût de déformation

De manière évidente, le problème d'appariement exact admet au moins une solution : la fonctionnelle J atteint en particulier son minimum – en ne tenant pas compte des contraintes – pour $\phi \equiv \text{Id}$.

Le principal intérêt de l'appariement exact est justement le côté "exact" de celui-ci. Néanmoins, il peut détruire la structure des formes que l'on cherche à appairer, surtout dans le cas d'appariement de points de repères si ceux-ci ne sont pas choisis judicieusement.

Appariement inexact : Pour pallier ce problème, on a souvent recourt à l'appariement dit inexact, où la contrainte d'appariement est intégrée dans la fonctionnelle à minimiser, [15]. Pour cela, on doit d'abord définir une distance d_M sur l'espace de formes

afin de mesurer l'appariement entre les objets. On définit alors un terme que l'on appelle *attache aux données* et qui mesure l'écart entre la forme initiale y transportée par le morphisme ϕ et la forme cible z en posant (par exemple) :

$$\forall \phi \in \mathcal{G}_V, \quad A(\phi) := d_M(\phi(y), z).$$

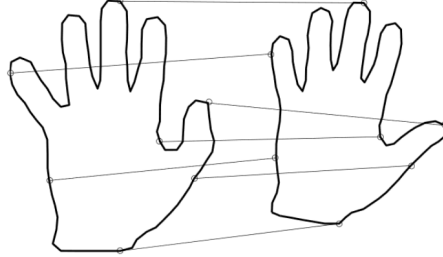


FIGURE 6. – Appariement de points de repères et choix des cibles.

Définition 8. Soit $\gamma > 0$. Le problème d'appariement inexact associé à la fonctionnelle d'attache aux données A consiste à minimiser, sur \mathcal{G}_V , la fonctionnelle suivante

$$J_\gamma(\phi) := \gamma d_V(\text{Id}, \phi) + A(\phi).$$

De même, ceci est équivalent à minimiser sur \mathbb{L}_V^2 la fonctionnelle

$$J_\gamma(v) := \gamma \int_0^1 \|v_t\|_V^2 dt + A(\phi_1^v)$$

et le théorème suivant nous assure de l'existence d'un tel minimum :

Théorème 9 (Appariement inexact, [16])

Soit $A: \mathcal{G}_V \rightarrow [0, +\infty[$ une fonctionnelle telle que $v \mapsto A(\phi_1^v)$ soit faiblement continue de \mathbb{L}_V^2 dans \mathbb{R} . Alors, pour tout $\gamma > 0$, il existe un minimum pour le problème d'appariement correspondant à A .

En pratique, on préfère le plus souvent l'appariement inexact à l'appariement exact comme expliqué ci-dessus. Cependant si l'appariement est numériquement bon avec un tel modèle, la fonctionnelle que l'on minimise et qui sert à mesurer le coût de déformation n'est *pas* une distance du fait de la présence du terme d'attache aux données. Ceci peut notamment poser un problème dès lors que l'on veut recourir à un modèle statistique, ce qui est le cas si on désire faire de l'estimation multi-modèle par exemple : on applique alors des résultats établis mathématiquement pour des distances à une quantité qui n'en est pas une et donc sans réelle garantie. Une des motivations du modèle d'appariement que l'on décrit au paragraphe suivant, et dit de *métamorphoses*, est de justement fournir une métrique sur le groupe \mathcal{G}_V ne souffrant pas des limitations de l'appariement exact.

III. Appariement et métamorphoses

Une autre motivation est de pouvoir quantifier les déformations dues à l'action du groupe \mathcal{G}_V et à l'évolution de la forme. Plus concrètement et pour paraphraser Alain Trouvé et Laurent Younes, *Metamorphoses through Lie group action* [44], si on se donne une forme $x = (x^1, \dots, x^n) \in M \subset \mathbb{R}^{nd}$ et n vecteurs unitaires u^1, u^2, \dots, u^n de \mathbb{R}^d , alors on a envie de dire que $(x^1 + u^1, \dots, x^n + u^n)$ est "plus proche" de x que $(x^1 + u^1, \dots, x^n + u^n)$ en tant que "simple" translaté de x . Pourtant, avec le modèle précédent, le terme d'attache aux données sera identique, les vecteurs u^i étant choisis unitaires. Autrement dit, on aimerait pouvoir dire que le coût de déformation est plus faible si cette déformation est due – au moins en partie – à l'action du groupe \mathcal{G}_V sur M , ce que rend possible le modèle de métamorphose.

Définition 10 (Métamorphose, [20, 44]). *On appelle métamorphose un couple de courbes $(g_t, \mu_t) \in \mathcal{G}_V \times M$ paramétré par le temps $t \in [0, 1]$ et tel que $g_0 = Id$. Son image est la courbe $m_t \in M$ définie par $m_t = g_t(\mu_t)$. On appelle partie relative à la déformation – deformation part – de la métamorphose g_t et partie relative à l'évolution du gabarit – template part – de la métamorphose μ_t . Enfin, on dit que la métamorphose est une déformation pure lorsque μ_t est constant.*

Notons v_t la vitesse du morphisme g_t au temps t , i.e. le champ de vecteur tel que

$$\forall t \in [0, 1], \quad \partial_t g_t = v_t \circ g_t.$$

Alors, l'évolution de l'image m_t de la métamorphose est donnée par :

$$\forall t \in [0, 1], \quad \partial_t m_t = \mathcal{D}_{\mu_t} g_t \cdot \partial_t \mu_t + \partial_t g_t(\mu_t) = \mathcal{D}_{\mu_t} g_t \cdot \partial_t \mu_t + v_t(m_t).$$

En particulier, pour $t = 0$,

$$\partial_t m_t|_{t=0} = \partial_t \mu_t|_{t=0} + v_0(m_0)$$

et on peut donc voir tout vecteur tangent $\zeta \in T_m M$ comme "image" d'une forme de métamorphose infinitésimale $(v, \delta) \in V \times T_m M$.

On montre alors (par un argument de projection – voir l'article d'Alain Trouvé et Laurent Younes [44]) que, pour tout $\sigma^2 > 0$,

$$|\mu|_m^2 := \inf \left\{ \|v\|_V^2 + \frac{1}{\sigma^2} \|\delta\|_m^2 \mid \mu = v(m) + \delta ; v \in V, \delta \in T_m M \right\}$$

définit une distance (riemannienne) sur M . L'énergie de la courbe image m_t est alors donnée par :

$$E(m_t) = \int_0^1 |\partial_t m_t|_{m_t}^2 dt = \inf \left\{ \int_0^1 \|v_t\|_V^2 dt + \frac{1}{\sigma^2} \int_0^1 \|\partial m_t - v_t(m_t)\|_{m_t}^2 dt \mid v_t \in V \right\}.$$

Finalement, on mesure la distance entre deux points m_y et m_z de M en minimisant sur l'ensemble des courbes $(g_t, m_t) \in V \times M$ la fonctionnelle

$$J(v_t, m_t) = \int_0^1 \|v_t\|_V^2 dt + \frac{1}{\sigma^2} \int_0^1 \|\partial m_t - v_t(m_t)\|_{m_t} dt$$

sous la contrainte $m_0 = m_y$ et $m_1 = m_z$. De manière équivalente, en notant g_t la partie déformation associée à m_t et $\mu_t := g_t^{-1}(m_t)$, la fonctionnelle à minimiser se ré-écrit :

$$J(m_t) = \int_0^1 \|v_t\|_V^2 dt + \frac{1}{\sigma^2} \int_0^1 \|\mathcal{D}_{\mu_t} g_t \cdot \partial_t \mu_t\|_{m_t} dt.$$

Du fait du caractère "métrique" de la fonctionnelle d'énergie, on espère que l'appariement par métamorphoses soit plus "robuste" numériquement et surtout compatible avec les modèles statistiques d'estimation. Par ailleurs, contrairement aux modèles difféomorphiques, l'appariement par métamorphoses permet de s'affranchir de contraintes topologiques, telle que la connexité, du fait de la présence d'un gabarit μ_t évoluant au cours du temps et que l'on transporte *a posteriori* par difféomorphisme pour avoir l'image. Cela peut avoir un intérêt notable si l'on cherche à faire de la détection/comparaison de tumeurs, qui peuvent se scinder au cours du temps ou métastaser par exemple...

Cependant, ce type d'appariement s'avère lourd, et ce même dans le cas "simple" des mesures discrètes, comme décrit dans l'article *Computing metamorphosis between discrete measures* de Casey L. Richardson et Laurent Younes [33].

IV. Étude transversale et bases de données longitudinales

Dans tout ce qui précède, on explique comment comparer deux formes anatomiques à un instant t donné ; cependant, notamment dès lors que l'on s'intéresse à l'évolution d'une maladie, on sera plus intéressé à pouvoir comparer l'évolution des formes anatomiques au cours du temps, plutôt que ces formes à un âge fixé ou stade déterminé de la maladie. Cette étude requiert la définition de modèles statistiques pour l'analyse des bases de données longitudinales dans lesquelles les observations de plusieurs individus sont acquises à plusieurs instants.

Pour des données scalaires, comme le volume d'une tumeur par exemple, l'étude de données longitudinales se fait par l'intermédiaire de modèles linéaires à effets mixtes. Cependant, ces modèles font souvent l'hypothèse que l'origine temporelle du processus observé est connue. Or, notamment dans le cas des maladies, nous sommes amenés à observer des sujets qui sont à des stades de développement différents, et pour lesquels le processus biologique sous-jacent a commencé à des âges variables et inconnus. Cela

n'a donc aucun sens de comparer la valeur du volume chez deux sujets aux mêmes âges, comme cela est fait dans le cas des modèles linéaires à effets mixtes "classiques". À l'inverse, nous devons introduire le concept de recalage temporel comme variable cachée du modèle, qui permet de comparer des données à des instants qui correspondent au même stade d'avancement du processus normal ou pathologique.

Jean-Baptiste Schiratti, lors de sa thèse, a proposé un premier modèle non linéaire à effets mixtes qui répond à ces contraintes pour des pathologies de type maladies neuro-dégénératives qui ne font que s'aggraver (comme la maladie d'Alzheimer pour laquelle les traitements ne font que ralentir partiellement les effets) [35, 36]. Ce modèle prend également en compte des observations multidimensionnelles vivant dans une variété riemannienne ce qui permet d'appliquer le même modèle à des données de type scalaires, images ou tenseurs. Le but de ma thèse est, en se basant sur les travaux de Jean-Baptiste Schiratti *et al* de proposer un modèle similaire mais adapté au cas de données longitudinales de patients traités atteints du cancer du rein : le modèle actuel ne prenant en compte que des évolutions géodésiques, elle n'est en effet pas adaptée à la dynamique d'évolution de mesures non unidirectionnelle.

V. Références

- [6] Nicolas Charon and Alain Trouvé. The varifold representation of non-oriented shapes for diffeomorphic registration. *SIAM Journal on Imaging Sciences*, 6(4), 2013.
- [9] Albrecht Dürer. *Vier bücher von menschlicher proportion*. 1528.
- [14] Joan Glaunès. Code Matlab pour l'appariement de points, courbes et surfaces. <http://www.mi.parisdescartes.fr/~glaunes/machine.zip>.
- [15] Joan Glaunès. LDDMM and equations for discrete matching functionals. Unpublished, 2015.
- [16] Joan Glaunès. *Transport par diffeomorphismes de points, de mesures et de courants pour la comparaison de formes et l'anatomie numérique*. PhD thesis, Université Paris 13, 2005.
- [17] Joan Glaunès, Alain Trouvé, and Laurent Younes. Diffeomorphic matching of distributions : a new approach for unlabelled point-sets and sub-manifolds matching. *Computer Vision and Pattern Recognition*, 2004.
- [20] Darryl D. Holm, Alain Trouvé, and Laurent Younes. The euler-poincaré theory of metamorphosis. *Quarterly of Applied Mathematics*, 67(4) :661–685, 2008.
- [33] Casey L. Richardson and Laurent Younes. Computing metamorphoses between discrete measures. *Journal of geometric mechanics*, 5(1), 2013.

-
- [35] Jean-Baptiste Schiratti, Stéphanie Allasonnière, Oliver Colliot, and Stanley Durrleman. Learning spatiotemporal trajectories from manifold-valued longitudinal data. *Neural Information Processing Systems (NIPS)*, 2015.
 - [36] Jean-Baptiste Schiratti, Stéphanie Allasonnière, Oliver Colliot, and Stanley Durrleman. A mixed effect model with time reparametrization for longitudinal univariate manifold valued data. *Information Processing in Medical Imaging (IPMI)*, 2015.
 - [42] D’Arcy Thompson. *On growth and form*. Cambridge University Press, 1917.
 - [43] Alain Trouvé. Géométrie et espaces de formes. Cours, 2016.
 - [44] Alain Trouvé and Laurent Younes. Metamorphoses through lie group action. *Foundations of computational mathematics*, pages 173–198, 2013.

– Troisième partie –

TER de L3

**Modélisation mathématiques d'une
infection par le virus du VIH**

Encadré par Hervé Lemeur *

En binôme avec Perinne Rouxel

*. LMO, Université Paris-Sud

Sommaire – Partie III

TER de L3 : Modélisation mathématiques d’une infection par le virus du VIH	37
	40
I Présentation du modèle	40
I.1 Lymphocytes non-infectés T	40
I.2 Lymphocytes infectés U	41
I.3 Virus infectants V & virus non-infectants W	41
II Admissibilité du système	42
Lemme 1.1	43
Lemme 1.2	45
Lemme 1.3	46
Lemme 1.4	47
Lemme 1.5	48
Lemme 1.6	49
III Existence globale de solutions	50
Conclusion	52
Références	52

I. Présentation du modèle

Le but de notre projet est de modéliser de la façon la plus exacte possible une infection par le virus du VIH, en nous intéressant plus particulièrement à la pluralité des antigénicités possibles pour les différents virus et lymphocytes. Nous nous restreignons pour des raisons de commodité à N antigénicités possibles, où N est une constante. Nous n'avons cependant pas eue le temps d'étudier le passage de N à $(N + 1)$ antigénicités, ce qui aurait pût nous informer sur la vraisemblance de notre modèle à la réalité et ainsi attester plus en profondeur de sa véracité, ou non.

Nous avons choisit de distinguer deux types de lymphocytes : les lymphocytes non-infectés par le virus du VIH que nous notons T et les virus infectés pas le virus du VIH, notés U . Sont également distingués les virus infectants, notés V , et les virus non-infectants, notés W .

Les différentes antigénicités possibles sont symbolisées par l'indice j qui varie donc entre 1 et N . Ainsi T_j désigne le nombre de lymphocytes non-infectés par le virus du VIH, d'antigénicités j . Et ainsi de suite...

Ce TER se base principalement sur l'article *Mathematical modeling of antigenicity for HIV dynamics* de F. Dubois, H. Lemeur et C. Reiss [10] et propose une extension du modèle qui y est présenté.

1. Lymphocytes non-infectés T

Croissance naturelle : En permanence des lymphocytes meurent (modélisé par $-\beta_j T_j$) et naissent ($+\gamma_j$). On a donc :

$$\forall j, \quad \frac{dT_j}{dt} \Big|_{naturel} = -\beta_j T_j + \gamma_j .$$

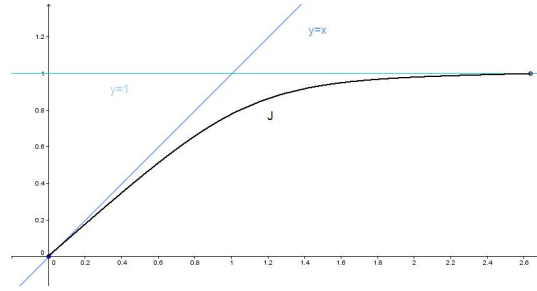
Réponse à l'infection : Les lymphocytes d'une antigénicité donnée vont se multiplier au contact d'un virus de cette même antigénicité. On modélise cela par :

$$\forall j, \quad \frac{dT_j}{dt} \Big|_{attaque} = +c_j (V_j + W_j) T_j .$$

Destruction des lymphocytes par les virus : Les virus du VIH ont besoin de pénétrer dans des lymphocytes, indépendamment de leur antigénicité, entraînant ainsi leur destruction. Il y a donc de moins en moins de lymphocytes non-infectés. Ce terme est modélisé par :

$$\forall j, \quad \frac{dT_j}{dt} \Big|_{infection} = -\frac{1}{\tau_j} J \left(\frac{V}{T} \right) T_j$$

où J est une fonction nulle sur $] -\infty, 0[$ et bornée par 1 sur $]0, +\infty[$.

FIGURE 7. – La fonction J

Finalement, pour tout antigénicité donnée j , on a :

$$\frac{dT_j}{dt} = -\beta_j + \gamma_j + c_j (V_j + W_j) T_j - \frac{1}{\tau_j} J\left(\frac{V}{T}\right) T_j.$$

2. Lymphocytes infectés U

Croissance naturelle : Contrairement, au cas des lymphocytes sains, les lymphocytes infectés ne se multiplient plus. Il n'y a donc plus de terme de création. En revanche, ils subissent une mort naturelle. (leur mort est par ailleurs rendue beaucoup plus précoce de par leur état infecté). On a donc :

$$\forall j, \quad \left. \frac{dU_j}{dt} \right|_{naturel} = -\alpha_j U_j$$

Infection par le virus : Complémentaire ment au cas des lymphocytes non-infectés, il y a de plus en plus de lymphocytes infectés et de plus :

$$\forall j, \quad \left. \frac{dU_j}{dt} \right|_{infection} = +\frac{1}{\tau_j} J\left(\frac{V}{T}\right) T_j.$$

Finalement, pour tout antigénicité donnée j , on a :

$$\frac{dU_j}{dt} = \frac{1}{\tau_j} J\left(\frac{V}{T}\right) T_j - \alpha_j U_j.$$

3. Virus infectants V & virus non-infectants W

Défense du corps à l'infection : Les lymphocytes de même antigénicité attaquent les virus, qu'ils soient infectés ou non :

$$\forall j, \quad \left. \frac{dV_j}{dt} \right|_{naturel} = -\xi_j V_j (T_j + U_j);$$

$$\forall j, \quad \left. \frac{dW_j}{dt} \right|_{naturel} = -\xi_j W_j (T_j + U_j).$$

Modélisation d'une infection par le VIH

Multiplication : Pour se multiplier, les virus ont besoin d'infecter un lymphocyte. De par leur fort taux de mutation, les virus peuvent rester infectants ou devenir non-infectants. On note θ la probabilité de muter d'un type infectant à non-infectant. On a :

$$\frac{dV}{dt} \Big|_{\text{croissance}} = a\theta U \quad \text{et} \quad \frac{dW}{dt} \Big|_{\text{croissance}} = a(1 - \theta)U.$$

Mutation : on note $S_{k,j}$ la probabilité pour un virus de muter de l'antigénicité k vers l'antigénicité j .

Donc $V_k \rightarrow V_j$ avec la probabilité $S_{k,j}\theta$ et, $V_k \rightarrow W_j$ avec la probabilité $S_{k,j}(1 - \theta)$.

Et pour obtenir l'ensemble des V_k mutant vers V_j , on fait : $\sum_k S_{k,j}\theta V_k$.

Et idem pour les V_k mutant vers W_j

On obtient donc un terme de croissance de la forme :

$$\forall j, \quad \frac{dV_j}{dt} \Big|_{\text{croissance}} = U \sum_k S_{k,j}\theta V_k A(t).$$

En sommant sur j , et en identifiant avec $\frac{dV}{dt} \Big|_{\text{croissance}} = a\theta U$, on obtient $A(t) = \frac{a}{V}$.

Finalement, pour tout antigénicité j , on a :

$$\begin{aligned} \frac{dV_j}{dt} &= \frac{a\theta U}{V} \sum_k S_{kj} V_k - \xi_j V_j (T_j + U_j); \\ \frac{dW_j}{dt} &= \frac{a(1 - \theta)U}{V} \sum_k S_{kj} V_k - \xi_j V_j (T_j + U_j). \end{aligned}$$

II. Admissibilité du système

Le but de cette section est de prouver que notre système a un sens biologiquement parlant : à savoir, qu'il ne génère pas de solutions à valeurs (tout ou en partie) négatives, et ce, quelque soit les conditions initiales.

Nous considérons le système suivant

$$\left\{ \begin{array}{l} \frac{dT_j}{dt} = -\beta_j + \gamma_j + c_j (V_j + W_j) T_j - \frac{1}{\tau_j} J\left(\frac{V}{T}\right) T_j \quad (1) \\ \frac{dU_j}{dt} = \frac{1}{\tau_j} J\left(\frac{V}{T}\right) T_j - \alpha_j U_j \quad (2) \\ \frac{dV_j}{dt} = \frac{a\theta U}{V} \sum_k S_{kj} V_k - \xi_j V_j (T_j + U_j) \quad (3) \\ \frac{dW_j}{dt} = \frac{a(1-\theta)U}{V} \sum_k S_{k,j} V_k - \xi_j V_j (T_j + U_j) \quad (4) \end{array} \right.$$

avec pour conditions initiales :

- (i) $\forall j, T_j(0) \geq 0$ et $\exists \tilde{j}$ tel que $T_{\tilde{j}}(0) > 0$,
- (ii) $\forall j, U_j(0) \geq 0$,
- (iii) $\forall j, V_j(0) \geq 0$ et $\exists \check{j}$ tel que $V_{\check{j}}(0) > 0$,
- (iv) $\forall j, W_j(0) \geq 0$ et $\exists \hat{j}$ tel que $W_{\hat{j}}(0) > 0$,

dans lequel nous supposons que, pour tout j :

- (v) $\alpha_j, \beta_j, \gamma_j, c_j, \tau_j$ et ξ_j sont des nombres réels positifs,
- (vi) $0 < \theta < 1$,
- (vii) J est une fonction concave sur $[0, +\infty]$, $J(0) = 0$ et $J'(0) = 1$,
- (viii) J est bornée sur \mathbb{R}^- et $\lim_{x \rightarrow +\infty} J(x) = 1$.

Théorème 11

Le système est biologiquement admissible. À savoir :

$$\forall j, \forall t, \quad T_j, U_j, V_j, W_j \geq 0.$$

Plusieurs lemmes vont être utiles à la démonstration de notre théorème.

Dans la suite, nous noterons \mathcal{D}_{T_j} l'ensemble de définition de T_j et de même pour U_j, V_j, W_j et nous savons que chacune de ces fonctions sont continues sur leur domaine de définition en tant que solution d'un système d'équations différentielles.

Lemme 1.1. Tant que cette quantité existe et dès lors qu'on se place en des temps strictement positifs, le nombre de lymphocytes non-infectés est, pour chaque antigénicité, strictement positif :

$$\forall j, \forall t \in \mathcal{D}_{T_j}, t > 0, \quad T_j(t) > 0.$$

Modélisation d'une infection par le VIH

Démonstration :

On considère l'équation :

$$(E) : \frac{dT_j}{dt} = -\beta_j + \gamma_j + c_j (V_j + W_j) T_j - \frac{1}{\tau_j} J \left(\frac{V}{T} \right) T_j.$$

D'après le théorème de Cauchy-Lipschitz, cette équation admet une unique solution. En effet, l'équation est linéaire en T_j . Soit j fixé.

- Si $T_j(0) > 0$. Alors, par continuité de T_j ,

$$\exists t_\varepsilon > 0 \quad \text{tel que} \quad T_j([0, t_\varepsilon]) > 0.$$

Si on suppose que $\exists t_2$ tel que $T_j(t_2) < 0$ alors nécessairement par continuité de T_j ,

$$\exists t_\varepsilon < t_1 < t_2 \quad \text{tel que} \quad T_j(t_1) = 0.$$

Notons $t_0 := \inf\{t \mid T_j(t) = 0\}$. En particulier, on sait que $t_0 > 0$. Or,

$$\frac{dT_j}{dt}(t_0) = \gamma_j > 0.$$

Donc T_j est strictement croissante au voisinage de t_0 :

$$\exists t_\eta > 0 \quad \text{tel que} \quad T_j([t_0 - t_\eta, t_0]) < 0.$$

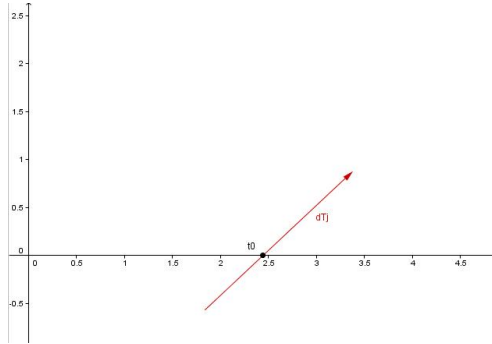


FIGURE 8. – La fonction T_j au voisinage de t_0

Or $T_j(0) > 0$, donc d'après le théorème des valeurs intermédiaires,

$$\exists \tilde{t}_0 \quad \text{tel que} \quad T_j(\tilde{t}_0) = 0$$

ce qui contredit la définition de t_0 . Finalement,

$$\forall t \in \mathcal{D}_{T_j}, \quad T_j(t) > 0.$$

- **Sinon, $\mathbf{T}_j(\mathbf{0}) = \mathbf{0}$.** Dans ce cas,

$$\frac{dT_j}{dt}(t=0) = \gamma_j > 0$$

et donc,

$$\exists t_\varepsilon > 0 \quad \text{tel que} \quad T_j(]0, t_\varepsilon]) > 0.$$

On réapplique alors le résultat précédent en prenant pour instant initial $\frac{t_\varepsilon}{2}$. □

Lemme 1.2. Tant que cette quantité existe et dès lors qu'on se place en des temps strictement positifs, le nombre de lymphocytes infectés est, pour chaque antigénicité, strictement positif :

$$\forall j, \quad \forall t \in \mathcal{D}_{T_j}, \quad t > 0, \quad U_j(t) > 0.$$

Démonstration :

Pour ce faire, nous allons procéder en deux temps. On considère l'équation :

$$(E) : \frac{dU_j}{dt} = +\frac{1}{\tau_j} J\left(\frac{V}{T}\right) T_j - \alpha_j U_j.$$

De même, cette équation admet une unique solution.

- i. **Montrons tout d'abord que $\exists t_\varepsilon > 0$ tel que $\forall j, \mathbf{U}_j(]0, t_\varepsilon]) > \mathbf{0}$.**

Soit j fixé.

- **Si $\mathbf{U}_j(\mathbf{0}) > \mathbf{0}$.** Alors la continuité de U_j nous assure que

$$\exists t_\varepsilon > 0 \quad \text{tel que} \quad U_j(]0, t_\varepsilon]) > 0.$$

- **Sinon, $\mathbf{U}_j(\mathbf{0}) = \mathbf{0}$.** Par hypothèse, on a $T(0) > 0$ et $V(0) > 0$.

- **Si $\mathbf{T}_j(\mathbf{0}) > \mathbf{0}$.** Alors $\frac{dU_j}{dt}(0) > 0$ et donc, par stricte croissance de U_j ,

$$\exists t_\varepsilon > 0 \quad \text{tel que} \quad U_j(]0, t_\varepsilon]) > 0.$$

- **Sinon, $\mathbf{T}_j(\mathbf{0}) = \mathbf{0}$.** Alors $\frac{dU_j}{dt}(0) = 0$ et on regarde donc ce qui se passe à l'ordre deux :

$$\frac{d^2U_j}{dt^2}(0) = \frac{1}{\tau_j} J\left(\frac{V(0)}{T(0)}\right) \frac{dT_j}{dt}(0) = \frac{\gamma_j}{\tau_j} J\left(\frac{V(0)}{T(0)}\right) > 0.$$

D'où U_j également strictement croissante et

$$\exists t_\varepsilon > 0 \quad \text{tel que} \quad U_j(]0, t_\varepsilon]) > 0.$$

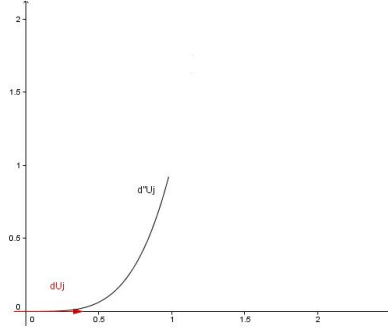


FIGURE 9. – La fonction U_j au voisinage de $t = 0$

ii. **Montrons à présent que $\forall j$, $t_\varepsilon = \sup(\mathcal{D}_{U_j})$.**

Soit j fixé. On sait que pour tout temps $t > 0$, $\frac{1}{\tau_j} J\left(\frac{V}{T}\right) \geq 0$ donc

$$\frac{1}{\tau_j} J\left(\frac{V}{T}\right) - \alpha_j U_j \geq -\alpha_j U_j \quad \text{i.e.} \quad \frac{dU_j}{dt} \geq -\alpha_j U_j.$$

En multipliant par $e^{\alpha_j t}$, on a :

$$\left(\frac{dU_j}{dt} + \alpha_j U_j\right) e^{\alpha_j t} \geq 0 \quad \text{i.e.} \quad \frac{d}{dt} (U_j e^{\alpha_j t}) \geq 0.$$

D'où, en intégrant entre $\frac{t_\varepsilon}{2}$ et $t \in \mathcal{D}_{U_j}$ tel que $t > \frac{t_\varepsilon}{2}$:

$$\int_{\frac{t_\varepsilon}{2}}^t \frac{d}{dt} (U_j e^{\alpha_j t}) dt \geq 0.$$

Autrement dit,

$$\forall t \in \mathcal{D}_{U_j}, t > \frac{t_\varepsilon}{2}, \quad U_j(t) \geq U_j\left(\frac{t_\varepsilon}{2}\right) e^{\frac{t_\varepsilon}{2} - t} > 0.$$

Finalement, comme $U_j(]0, t_\varepsilon[) > 0$ on a donc $\forall t \in \mathcal{D}_{U_j}$, $U_j(t) > 0$.

Et ce, pour tout j , ce qui achève la preuve du lemme. □

Lemme 1.3. Pour toute antigénicité et pour des temps arbitrairement petits non-nuls, le nombre de virus infectants est strictement positif :

$$\exists t_\varepsilon > 0 \quad \text{tel que} \quad \forall j, V_j(]0, t_\varepsilon[) > 0.$$

Démonstration :

On considère un j fixé.

- **Si j est tel que $V_j(\mathbf{0}) > \mathbf{0}$.** Alors, par continuité de V_j , on a le résultat.
- **Si j est tel que $V_j(\mathbf{0}) = \mathbf{0}$.** Alors, d'après les conditions initiales,

$$\frac{dV_j}{dt}(\mathbf{0}) = \frac{a\theta U(\mathbf{0})}{V(\mathbf{0})} \sum_k S_{k,j} V_k(\mathbf{0}) \geq 0.$$

- **Si $U(\mathbf{0}) > \mathbf{0}$.** Alors $\frac{dV_j}{dt}(\mathbf{0}) > 0$, c'est à dire que V_j est strictement croissante au voisinage de $\mathbf{0}$. Et donc, par continuité de V_j

$$\exists t_\varepsilon > 0 \quad \text{tel que} \quad V_j(]0, t_\varepsilon]) > 0.$$

- **Sinon $U(\mathbf{0}) = \mathbf{0}$, ou encore $\forall \mathbf{k} V_{\mathbf{k}}(\mathbf{0}) = \mathbf{0}$.** Alors $\frac{dV_j}{dt}(\mathbf{0}) = 0$. On regarde donc ce qu'il se passe à l'ordre 2 :

$$\frac{d^2V_j}{dt^2}(\mathbf{0}) = \frac{a\theta U'(\mathbf{0})}{V(\mathbf{0})} \sum_k S_{k,j} V_k(\mathbf{0}).$$

Or, comme

$$\frac{dU}{dt}(\mathbf{0}) = \sum_j \frac{T_j(\mathbf{0})}{\tau_j} J \left(\frac{V(\mathbf{0})}{T(\mathbf{0})} \right) > 0, \quad \frac{d^2V_j}{dt^2}(\mathbf{0}) > 0.$$

Dans ce cas $\exists t_\eta > 0$ tel que $\frac{dV_j}{dt}(]0, t_\eta]) > 0$. Et, par conséquent,

$$\exists t_\varepsilon > 0 \quad \text{tel que} \quad V_j(]0, t_\varepsilon]) > 0.$$

□

Lemme 1.4. Le nombre total de virus infectants est nul pour la première fois s'il ne subsiste aucun virus pour chaque antigénicité donnée

Si $\exists t$ tel que $V(]0, t]) > 0$ et $V(t) = 0$ alors $\forall j, V_j(t) = 0$.

Démonstration :

Définissons l'instant t^* comme suit :

$$t^* := \sup\{ t \mid \forall j, V_j(]0, t]) > 0 \}.$$

Modélisation d'une infection par le VIH

- **Si t^* est fini.** On pose $J = \{ j \mid V_j(t^*) = 0 \}$ et, en particulier,

$$1 \leq n \leq N \quad \text{où} \quad n := \text{Card}(J).$$

Montrons par *l'absurde* que $n = N$. Supposons pour cela que $n < N$.
Soit $j \in J$ fixé. On a :

$$\frac{dV_j}{dt}(t^*) = \frac{a\theta U(t^*)}{V(t^*)} \sum_k S_{k,j} V_k(t^*).$$

Or $V(t^*) > 0$ car $n < N$ par hypothèse et $U(t^*) > 0$ d'après le lemme 1.2.
Donc $\frac{dV_j}{dt}(t^*) > 0$ et on contredit alors la définition de t^* avec le théorème des valeurs intermédiaires (cf preuve du lemme 1.1 de la page 43). D'où $n = N$.

- **Si non, t^* est infini.** Alors pour tout temps $t \in]0, +\infty[$, $V_j(t) > 0$, ce que l'on cherche à prouver à terme.

□

Lemme 1.5. Tant que cette quantité existe et dès lors qu'on se place en des temps strictement positifs, le nombre de virus infectants est, pour chaque antigénicité, strictement positif :

$$\forall j, \quad \forall t \in \mathcal{D}_{V_j}, \quad t > 0, \quad V_j(t) > 0.$$

Démonstration :

Commençons tout d'abord par remarquer que le lemme 1.4 nous donne en fait un résultat plus fort :

$$\exists t \text{ tel que } V(t) = 0 \quad \text{si et seulement si} \quad \forall j, V_j(t) = 0.$$

En effet, $V = 0$ si et seulement si :

- Soit $\forall j, V_j = 0$;
- Soit $\exists \tilde{j}, \check{j}$ tel que $V_{\tilde{j}} > 0, V_{\check{j}} < 0$ et $\sum_j V_j = V = 0$.

Or, comme on vient de voir qu'il ne pouvait y avoir $n < N$ fonction V_j toutes nulles en même temps (Lemme 1.4, page 47) et que $\forall j, \exists t_\varepsilon$ tel que $V_j(]0, t_\varepsilon]) > 0$ (Lemme 1.3, page 46) donc d'après le théorème des valeurs intermédiaires, on ne peut pas être dans la configuration ii. sans avoir été dans la configuration i.. D'où le résultat.

De plus, par définition de V et V_j ,

$$\frac{dV}{dt} = \sum_j \frac{dV_j}{dt} = a\theta U - \sum_j \xi_j V_j (T_j + U_j).$$

Soit t^* , définis comme ci-dessus :

$$t^* = \sup\{ t \mid \forall j, V_j([0, t]) > 0 \}.$$

Alors $\frac{dV}{dt}(t^*) = a\theta U(t^*) > 0$ d'après le lemme 1.2 (page 45). C'est à dire que V est strictement croissante au voisinage de t^* . On applique alors le théorème des valeurs intermédiaires pour contredire la définition de t^* . Finalement,

$$\forall j, \quad \forall t \in \mathcal{D}_{V_j}, \quad t > 0, \quad V_j(t) > 0.$$

□

Lemme 1.6. Tant que cette quantité existe et dès lors qu'on se place en des temps strictement positifs, le nombre de virus non-infectants est, pour chaque antigénicité, strictement positif :

$$\forall j, \quad \forall t \in \mathcal{D}_{W_j}, \quad t > 0, \quad W_j(t) > 0.$$

Démonstration :

On peut réappliquer exactement le même raisonnement que celui qui a servis à montrer que

$$\forall j, \quad \forall t \in \mathcal{D}_{V_j}, \quad t > 0, \quad V_j(t) > 0$$

à savoir, les lemmes 1.3, 1.4 et 1.5 (aux pages respectives 46, 47 et 48) en posant $\tilde{\theta} = 1 - \theta$ qui satisfait aux mêmes hypothèses que θ . En particulier, on a bien $0 < \tilde{\theta} < 1$. □

Démonstration du Théorème 11 :

On a en fait montré que :

$$\left\{ \begin{array}{ll} \forall j, \quad \forall t \in \mathcal{D}_{T_j}, \quad t > 0, \quad T_j(t) > 0 & (\text{Lemme 1.1}) \\ \forall j, \quad \forall t \in \mathcal{D}_{U_j}, \quad t > 0, \quad U_j(t) > 0 & (\text{Lemme 1.2}) \\ \forall j, \quad \forall t \in \mathcal{D}_{V_j}, \quad t > 0, \quad V_j(t) > 0 & (\text{Lemme 1.5}) \\ \forall j, \quad \forall t \in \mathcal{D}_{W_j}, \quad t > 0, \quad W_j(t) > 0 & (\text{Lemme 1.6}) \end{array} \right.$$

Ce qui prouve le théorème énoncé en page 43 □

III. Existence globale de solutions

Le but de cette section est de prouver que les solutions de notre système n'explosent pas en temps fini : c'est à dire qu'une solution maximale est globale.

Nous supposons par ailleurs (*comme cela était suggéré dans l'article originel [10]*) que c_j et ξ_j ne dépendent en fait pas de j et nous les notons simplement c et ξ par la suite.

Théorème 12

La solution maximale de condition initiale (T_0, U_0, V_0, W_0) satisfaisant les conditions énoncées à la page 43 est finie pour tout temps t de $[0, +\infty[$. Plus précisément, la solution (T, U, V, W) va vérifier : $\forall t > 0$,

$$T(t) + U(t) + \frac{c}{\xi} (V(t) + W(t)) \leq \left(T_0 + U_0 + \frac{c}{\xi} (V_0 + W_0) \right) e^{\frac{ac}{\xi}t} + \frac{\gamma}{\lambda} \left(e^{\frac{ac}{\xi}t} - 1 \right).$$

Démonstration :

L'idée est de se débarrasser des termes non-linéaires du système. On remarque que :

$$\begin{aligned} & \frac{d}{dt} \left[T + U + \frac{c}{\xi} (V + W) \right] \\ &= - \sum_j \beta_j T_j + \sum_j \gamma_j - \sum_j \alpha_j U_j + \sum_j c_j (V_j + W_j) T_j \\ & \quad + \frac{c}{\xi} \sum_j \frac{aU}{V} \sum_k S_{k,j} V_k - c \sum_j (V_j + W_j) (T_j + U_j) \\ &= - \sum_j \beta_j T_j + \sum_j \gamma_j - \sum_j \alpha_j U_j - c \sum_j (V_j + W_j) U_j + \frac{ac}{\xi} U \\ &\leq \sum_j \gamma_j + \frac{ac}{\xi} U. \quad (*) \end{aligned}$$

En effet,

$$\begin{cases} \forall j, \alpha_j, \beta_j > 0 \\ \forall j, T_j, U_j, V_j, W_j \geq 0 \end{cases}.$$

On pose : $\gamma := \sum_j \gamma_j$ et $\lambda := \frac{ac}{\xi}$. (*) devient alors :

$$\frac{d}{dt} \left[T + U + \frac{c}{\xi} (V + W) \right] \leq \gamma + \lambda U.$$

Ce qui donne, après intégration, $\forall t > 0$,

$$T(t) + U(t) + \frac{c}{\xi}(V(t) + W(t)) - \left(T_0 + U_0 + \frac{c}{\xi}(V_0 + W_0)\right) \leq \gamma t + \lambda \int_0^t U(s) ds.$$

Notons $\delta_0 := T_0 + U_0 + \frac{c}{\xi}(V_0 + W_0)$ et $\phi(t) := \int_0^t U(s) ds$. Alors, $\forall t > 0$, on a :

$$\begin{aligned} T(t) + U(t) + \frac{c}{\xi}(V(t) + W(t)) &\leq \delta_0 + \gamma t + \lambda \phi(t) & (\star\star) \\ U(t) &\leq \delta_0 + \gamma t + \lambda \phi(t) - T(t) - \frac{c}{\xi}(V(t) + W(t)) \\ U(t) &\leq \delta_0 + \gamma t + \lambda \phi(t) \\ \frac{d\phi}{dt}(t) - \lambda \phi(t) &\leq \delta_0 + \gamma t. \end{aligned}$$

En multipliant par $e^{-\lambda t}$, on obtient :

$$\begin{aligned} \left(\frac{d\phi}{dt}(t) - \lambda \phi(t)\right) e^{-\lambda t} &\leq (\delta_0 + \gamma t) e^{-\lambda t} \\ \frac{d}{dt}(\phi(t) e^{-\lambda t}) &\leq (\delta_0 + \gamma t) e^{-\lambda t}. \end{aligned}$$

D'où, en intégrant l'expression précédente, et en remarquant que $\phi(0) = 0$, on a : $\forall t > 0$,

$$\begin{aligned} \phi(t) e^{-\lambda t} &\leq \int_0^t (\delta_0 + \gamma s) e^{-\lambda s} ds \\ &\leq \left[-\frac{\delta_0 + \gamma s}{\lambda} e^{-\lambda s}\right]_0^t + \frac{\gamma}{\lambda} \int_0^t e^{-\lambda s} ds \\ &\leq \frac{\delta_0}{\lambda} - \frac{\delta_0 + \gamma t}{\lambda} e^{-\lambda t} - \frac{\gamma}{\lambda} (e^{-\lambda t} - 1) \\ &\leq -\frac{1}{\lambda} \left(\delta_0 + \gamma t + \frac{\gamma}{\lambda}\right) e^{-\lambda t} + \frac{1}{\lambda} \left(\delta_0 + \frac{\gamma}{\lambda}\right). \end{aligned}$$

D'où :

$$\lambda \phi(t) \leq -\left(\delta_0 + \gamma t + \frac{\gamma}{\lambda}\right) + \left(\delta_0 + \frac{\gamma}{\lambda}\right) e^{\lambda t}.$$

En reportant dans $(\star\star)$, on a donc : $\forall t > 0$,

$$\begin{aligned} T(t) + U(t) + \frac{c}{\xi}(V(t) + W(t)) &\leq \delta_0 + \gamma t + \lambda \phi(t) \\ &\leq \delta_0 + \gamma t - \gamma t - \delta_0 - \frac{\gamma}{\lambda} + \left(\delta_0 + \frac{\gamma}{\lambda}\right) e^{\lambda t} \\ &\leq \left(\delta_0 + \frac{\gamma}{\lambda}\right) e^{\lambda t} - \frac{\gamma}{\lambda}. \end{aligned}$$

Ou encore : $\forall t > 0$,

$$T(t) + U(t) + \frac{c}{\xi}(V(t) + W(t)) \leq \left(T_0 + U_0 + \frac{c}{\xi}(V_0 + W_0)\right) e^{\frac{ac}{\xi}t} + \frac{\gamma}{\lambda} \left(e^{\frac{ac}{\xi}t} - 1\right)$$

ce qui montre que la solution n'explose pas en temps finis et achève donc de démontrer le théorème. \square

Conclusion

En conclusion, notre système est donc biologiquement admissible, dans le sens où on ne peut avoir un nombre de lymphocytes/virus négatif (Théorème de la page 43) et la solution maximale a une durée de vie infinie (Théorème de la page 50). En ce sens il présente donc certaines similitudes avec la réalité.

Cependant, nous pouvons d'ores et déjà soulever une limite à notre modélisation ; limite qui ne pourra être améliorée car intrinsèque aux équations différentielles ordinaires : on ne peut envisager une solution qui deviendrait nulle au bout d'un certain temps, ce qui a lieu dans la réalité et correspond à une "disparition" des lymphocytes ou des virus pour une antigénicité j donnée.

Une étude plus poussée de la modélisation pourrait permettre de valider ou non ce modèle.

Références

- [10] François Dubois, Hervé Lemeur, and Claude Reiss. Mathematical modeling of antigenicity for HIV dynamics. *MathS In Action*, 3(1) :1–35, 2010.

– Quatrième partie –

TER de M1
Représentations linéaires du
groupe symétrique

Encadré par Pierre-Guy Plamondon *

En binôme avec Gabriel Laumosne

*. LMO, Université Paris-Sud

Sommaire – Partie IV

TER de M1 : Représentations linéaires du groupe symétrique	53
Chapitre I :	
Représentations linéaires	59
I Quelques généralités	59
I.1 Définitions et premiers exemples	59
I.2 Théorème de complète réductibilité	61
I.3 Caractères	62
I.4 Table des caractères	64
II Représentations induites	65
III Le groupe symétrique $\mathfrak{S}_n, n \in \mathbb{N}$	66
III.1 Partitions de l'entier n	66
III.2 Représentation standard	67
III.3 Exemple 1 : \mathfrak{S}_3	70
III.4 Exemple 2 : \mathfrak{S}_4	70
Chapitre II :	
Outils pour étudier les représentations du groupe \mathfrak{S}_n	73
I Conventions d'écriture	73
I.1 Diagrammes de Young	73
I.2 Algèbre de groupe $\mathbb{C}\mathfrak{S}_n$	75
I.3 Notations pour la formule de Frobenius	77
II Présentation des résultats	78
II.1 Formule de Frobenius	78
II.2 Formule des crochets	79
II.3 Liste exhaustive des représentations du groupe \mathfrak{S}_n	79
II.4 Exemple 2 : \mathfrak{S}_4 - suite et fin	79
III Démonstration des résultats	81
III.1 Formule des crochets	81
III.2 Liste exhaustive des représentations du groupe \mathfrak{S}_n	85
Références	93

Le but de notre TER est d'étudier le groupe symétrique \mathfrak{S}_n en utilisant les outils fournis par la théorie des représentations linéaires.

Pour cela nous rappellerons dans un premier temps quelques généralités sur les représentations, en se restreignant au cas des groupes finis sur des \mathbb{C} -espaces vectoriels de dimension finie. Ensuite, nous nous attacherons à développer des outils pour étudier les représentations du groupe \mathfrak{S}_n et les appliquerons dans une troisième partie sur un exemple concret.

Pour cela, on se base principalement sur le livre de Jean-Pierre Serre, *Représentations linéaires des groupes finis* [37] et le cours éponyme de Pierre-Guy Plamondon [31] en ce qui concerne la première partie. La seconde partie consiste essentiellement en une réécriture détaillée de la preuve proposée par Fulton et Harris dans leur ouvrage intitulé *Representation Theory – A first course* [12], au chapitre 4 : une grande partie de notre TER a en effet consisté en l'étude de cette preuve.

– Chapitre I –

Représentations linéaires

On rappelle dans cette partie quelques définitions et résultats qui nous serviront par la suite, sans les démontrer.

I. Quelques généralités

1. Définitions et premiers exemples

On peut définir la notion de représentation à partir de n'importe quel groupe (y compris à partir d'un groupe infini). De même, on peut définir des représentations avec des espaces vectoriels de dimension quelconque, sur des corps quelconques.

Nous nous limiterons cependant dans le cadre de notre TER aux groupes finis et \mathbb{C} -espaces vectoriels de dimensions finies.

Définition I.1. Soit V un \mathbb{C} -espace-vectoriel. Une représentation linéaire du groupe G est un morphisme de groupe ρ de G dans le groupe des automorphismes du \mathbb{C} -espace vectoriel $V : \mathcal{GL}(V)$.

Si V est de dimension finie, l'entier $\dim(V) = n$ est appelé degré de la représentation ρ .

Un abus de langage fréquent est d'employer le terme "représentation" pour désigner aussi bien le morphisme de groupes ρ que l'espace vectoriel V .

Ainsi, on parlera de la représentation V pour désigner $\rho : G \rightarrow \mathcal{GL}(V)$.

Remarque : Une représentation $\rho : G \rightarrow \mathcal{GL}(V)$ peut être vue comme une action du groupe G sur l'espace-vectoriel V . Il s'agit de l'action définie par :

$$\forall g \in G, \quad \forall v \in V, \quad g \cdot v = \rho(g)(v).$$

On vérifie facilement qu'il s'agit d'une action de groupe en utilisant le fait que ρ est un morphisme de groupe :

- i. $\forall v \in V, e \cdot v = \rho(e)(v) = Id_V(v) = v$;
- ii. $\forall g, h \in G, \forall v \in V, g \cdot (h \cdot v) = \rho(g)(\rho(h)(v)) = (\rho(g) \circ \rho(h))(v) = \rho(gh)(v) = gh \cdot v$.

De plus, cette action vérifie la propriété de linéarité suivante :

$$\forall g \in G, \quad \forall v, w \in V, \quad \forall \lambda \in \mathbb{C}, \quad g \cdot (v + \lambda w) = g \cdot v + \lambda g \cdot w. \quad (I.1)$$

Représentations linéaires

Réciproquement, toute action du groupe G sur un espace vectoriel V et vérifiant la propriété de linéarité (1) permet de définir une représentation de G (par la même formule).

Définition I.2. Soit V un espace vectoriel et G un groupe. On dit que V est un G -module s'il existe un homomorphisme de groupes $\rho : G \rightarrow \mathcal{GL}(V)$.

De manière équivalente, se donner un G -module, c'est se donner une multiplication des éléments de V par les éléments de G . En notant $g \cdot v$ le produit de $g \in G$ par $v \in V$, cette multiplication doit vérifier les quatres points suivants :

- i. $\forall g \in G, \forall v \in V, g \cdot v \in V$;
- ii. $\forall g \in G, \forall v, w \in V, \forall \lambda \in \mathbb{C}, g \cdot (v + \lambda w) = g \cdot v + \lambda g \cdot w$;
- iii. $\forall g, h \in G, \forall v \in V, (gh) \cdot v = g \cdot (h \cdot v)$;
- iv. $\forall v \in V, e \cdot v = v$ où e désigne le neutre pour la loi de G .

Lorsqu'il n'y aura pas de confusion possible, on s'autorisera à parler de module, pour G -module.

La notation $g \cdot v$ désigne l'image de v par l'endomorphisme $\rho(g)$. Autrement dit, on identifie l'élément g avec l'endomorphisme $\rho(g)$.

Proposition I.3

Il est équivalent de se donner une représentation $\rho : G \rightarrow \mathcal{GL}(V)$ du groupe G ou une structure de G -module (à droite) sur V .

Démonstration :

On note V l'espace vectoriel \mathbb{C}^d .

Soit ρ une représentation du groupe G , de degré d .

On peut alors définir une multiplication (à droite) sur V : si $g \in G$ et $v \in V$, on pose $g \cdot v := \rho(g)(v)$ et on s'assure que les critères de la définition 2 sont vérifiés :

- i. $\forall g \in G, \forall v \in V, \rho(g)(v) \in V$;
- ii. $\forall g \in G, \forall v, w \in V, \forall c, d \in \mathbb{C}, \rho(g)(cv + dw) = c\rho(g)(v) + d\rho(g)(w)$;
- iii. $\forall g, h \in G, \forall v \in V, \rho(gh)(v) = \rho(g)(\rho(h)(v))$;
- iv. $\forall v \in V, \rho(e)(v) = v$.

Réciproquement, si V est un G -module alors $\rho(g) : v \in V \mapsto v$ convient. □

Par la suite, on confondra donc G -modules et représentations linéaires.

Avant de poursuivre, voyons dès à présent quelques exemples de représentations :

Exemple : La *représentation triviale* est la représentation de degré 1 correspondant à l'homomorphisme :

$$\rho : \begin{array}{ccc} G & \longrightarrow & \mathbb{C}^* \\ g & \longmapsto & 1 \end{array} .$$

Exemple : Soit V un \mathbb{C} -espace vectoriel de dimension $|G|$, muni d'une base $(e_g)_{g \in G}$. La *représentation régulière* du groupe G est la représentation de degré $|G|$ définie comme suit :

$$\rho : \begin{array}{ccc} G & \longrightarrow & \mathcal{GL}(V) \\ g & \longmapsto & \left(\begin{array}{ccc} \rho(g) : V & \longrightarrow & V \\ e_h & \longmapsto & e_{gh} \end{array} \right) . \end{array}$$

Exemple : La signature $\varepsilon : \mathfrak{S}_n \rightarrow \{-1, 1\}$ est une représentation de degré 1 sur \mathbb{R}, \mathbb{C} , ou même $\mathbb{Z}/3\mathbb{Z}$. On l'appelle la *représentation alternée*.

A partir de représentations connues du groupe G , on peut définir de nouvelles représentations :

Définition I.4. Une sous-représentation d'une représentation ρ est la restriction de $\rho(g)$ à un sous-espace vectoriel de V stable par l'action de G , soit à un $W \subseteq V$ tel que :

$$\forall g \in G, \quad \forall w \in W, \quad \rho(g)(w) \in W .$$

On peut montrer qu'en particulier, $\rho^W : G \rightarrow \mathcal{GL}(W)$ est bien une représentation du groupe G .

Définition I.5. Soit $\rho : G \rightarrow \mathcal{GL}(V)$ et $\rho' : G \rightarrow \mathcal{GL}(V')$ deux représentations de G . Leur somme directe est la représentation $\rho \oplus \rho' : G \rightarrow \mathcal{GL}(V \oplus V')$ définie par :

$$\forall g \in G, \quad \forall (v, v') \in V \oplus V', \quad \rho \oplus \rho'(g)(v, v') = \left(\rho(g)(v), \rho'(g)(v') \right) .$$

C'est également une représentation du groupe G .

2. Théorème de complète réductibilité

Définition I.6. Deux représentations $\rho : G \rightarrow \mathcal{GL}(V)$ et $\rho' : G \rightarrow \mathcal{GL}(V')$ sont dites isomorphes s'il existe un isomorphisme d'espaces vectoriels τ tel que :

$$\forall g \in G, \quad \tau \circ \rho(g) = \rho'(g) \circ \tau . \quad \begin{array}{ccc} V & \xrightarrow{\tau} & V' \\ \rho(g) \downarrow & & \downarrow \rho'(g) \\ V & \xrightarrow{\tau} & V' \end{array}$$

Définition I.7. Une représentation ρ de G est dite irréductible si $V \neq \{0\}$ et si pour toute sous-représentation $\rho^W : W \rightarrow \mathcal{GL}(W)$ de ρ on a que $W = V$ ou $W = \{0\}$. On dit que l'espace V est minimal pour la stabilité de G .

Théorème I.8 (de complète réductibilité)

Soit G un groupe fini, V un \mathbb{C} -espace vectoriel de dimension finie et ρ la représentation qui lui est associée.

Alors ρ peut se décomposer en somme directe de sous-représentations irréductibles : il existe $r \geq 1$ et $(\rho^{W_1}, W_1), \dots, (\rho^{W_r}, W_r)$ des sous-représentations irréductibles non triviales de (ρ, V) , telles que :

$$\rho \cong \rho^{W_1} \oplus \rho^{W_2} \oplus \dots \oplus \rho^{W_r} \quad \text{ou encore} \quad V \cong W_1 \oplus W_2 \oplus \dots \oplus W_r.$$

De plus, si $\rho \cong \rho_1 \oplus \dots \oplus \rho_r$ et $\rho'_1 \oplus \dots \oplus \rho'_s$ sont deux décompositions de ρ en représentations irréductibles, alors $r = s$ et :

$$\exists \sigma \in \mathfrak{S}_r, \quad \forall i \in \{1, \dots, r\}, \quad \rho_i = \rho'_{\sigma(i)}.$$

En particulier, pour toute représentation V d'un groupe fini G , il existe une décomposition

$$V = V_1^{\oplus a_1} \oplus V_2^{\oplus a_2} \oplus \dots \oplus V_k^{\oplus a_k}$$

où les V_i sont des représentations irréductibles non-isomorphes. De plus, cette décomposition de V en somme directe est unique, ainsi que les a_i et les V_i .

3. Caractères

Définition I.9. Soit $\rho : G \rightarrow \mathcal{GL}(V)$ une représentation de G . Le caractère de ρ est l'application :

$$\chi_\rho : \begin{array}{ccc} G & \longrightarrow & \mathbb{C} \\ g & \longmapsto & \text{Tr}(\rho(g)) \end{array}.$$

En particulier, on a immédiatement :

- i. $\chi_\rho(e) = n$, si ρ est de degré n et en notant e le neutre de G ;
- ii. $\forall g \in G, \chi_\rho(g^{-1}) = \overline{\chi_\rho(g)}$;
- iii. $\forall g, h \in G, \chi_\rho(ghg^{-1}) = \chi_\rho(h)$, i.e χ_ρ est une fonction de classe.

Définition I.10. Une fonction de classe est une application $\chi : G \rightarrow \mathbb{C}$ telle que :

$$\forall g, h \in G, \quad \chi(ghg^{-1}) = \chi(h).$$

Le nom vient du fait que χ est constante sur chacune des classes de conjugaisons de G .

Proposition I.11

Si ρ et ρ' sont deux représentations de G telles que $\chi_\rho = \chi_{\rho'}$, alors $\rho \cong \rho'$.

En d'autres termes, les caractères caractérisent réellement les représentations linéaires. Du plus, le résultat suivant (et le théorème de complète réductibilité) nous permet d'affirmer que pour étudier l'ensemble des représentations d'un groupe G , il suffit en fait de se restreindre à l'étude de ses représentations irréductibles.

Proposition I.12

Si $\rho : G \rightarrow \mathcal{GL}(V)$ et $\rho' : G \rightarrow \mathcal{GL}(V')$ sont deux représentations de G , alors :

$$\chi_{\rho \oplus \rho'} = \chi_{\rho} + \chi_{\rho'} .$$

Démonstration :

Soit $g \in G$. Le résultat est évident si on considère les formes matricielles de $\rho(g)$ et $\rho'(g)$: soit (e_i) une base de V et (e'_j) une base de V' :

$$\mathcal{M}_{(e_i, e'_j)}(\rho \oplus \rho'(g)) = \left(\begin{array}{c|c} \mathcal{M}_{(e_i)}(\rho(g)) & 0 \\ \hline 0 & \mathcal{M}_{(e'_j)}(\rho'(g)) \end{array} \right) .$$

□

Exemple : Soit G un groupe fini, et soit $\rho : G \rightarrow \mathcal{GL}(V)$ sa représentation régulière. Alors :

$$\chi_{\rho}(g) = \begin{cases} |G| & \text{si } g = e \\ 0 & \text{sinon} \end{cases} .$$

En effet, comme ρ est de degré $|G|$, on a bien que $\chi_{\rho}(e) = |G|$.
Et, si $g \neq e$, alors l'action de g sur la base $(e_h)_{h \in G}$ est définie par :

$$\forall h \in G, \quad \rho(g)(e_h) = e_{gh} .$$

Et, comme, $e \neq gh$, la matrice de $\rho(g)$ dans la base $(e_h)_{h \in G}$ ne possède que des 0 sur sa diagonale. D'où $\chi_{\rho}(g) = \text{Tr}(\rho(g)) = 0$.

Proposition I.13

Si ρ_1, \dots, ρ_r est une liste de toutes les représentations irréductibles de G (à isomorphisme près) et si n_1, \dots, n_r sont leurs degrés respectifs

- i. $\sum_{i=1}^r n_i^2 = |G|$,
- ii. Si $g \in G \setminus \{e\}$, alors $\sum_{i=1}^r n_i \chi_{\rho_i}(g) = 0$.

En travaillant sur l'ensemble des fonctions de classe, on peut montrer le résultat suivant (et qui sera la base de ce qui va suivre) :

Représentations linéaires

Proposition I.14

Le nombre de représentations irréductibles de G (à isomorphisme près) est égal au nombre de classes de conjugaison de G .

Remarque : On appelle *caractère irréductible* toute application $\chi : G \rightarrow \mathbb{C}$ qui est le caractère d'une représentation irréductible du groupe G .

4. Table des caractères

Soit G un groupe fini, dont on notera r le nombre de classes de conjugaison. On sait qu'il y a aussi r caractères irréductibles différents. On note c_1, c_2, \dots, c_r les classes de conjugaisons de G et $\chi_1, \chi_2, \dots, \chi_r$ les caractères irréductibles.

On choisira toujours $c_1 = \{e\}$ et $\chi_1 = 1$ (le caractère irréductible correspondant à la représentation triviale $\rho : g \mapsto Id_{\mathbb{C}}$).

Définition I.15. La table des caractères du groupe G est la matrice de taille $r \times r$ définie par $(\chi_i(c_j))_{1 \leq i, j \leq r}$:

$$\begin{pmatrix} \chi_1(c_1) & \cdots & \chi_1(c_r) \\ \chi_2(c_1) & \cdots & \chi_2(c_r) \\ \vdots & & \vdots \\ \chi_r(c_1) & \cdots & \chi_r(c_r) \end{pmatrix}.$$

Comme on a choisi $\chi_1 = 1$, on sait que la première ligne ne comportera que des 1. De même, comme on a choisi $c_1 = e$, on sait que la première colonne sera constituée des représentations $\chi_i(c_1) = \chi_i(e) = \dim V_i$, degré de la représentation ρ_i associée au caractère χ_i .

On présentera la table des caractères de la façon suivante :

G	c_1	\cdots	c_j	\cdots	c_r
χ_1	1	\cdots	1	\cdots	1
\vdots	\vdots		\vdots		
χ_i	$\deg(\rho_i)$	\cdots	$\chi_i(c_j)$		
\vdots	\vdots				
χ_r	$\deg(\rho_r)$				

II. Représentations induites

On introduit ici la notion de représentation induite dans l'optique de démontrer la *formule de Frobenius*, ce qui fera l'objet d'un paragraphe dans le II.3.

Définition I.16. Soit V une représentation du groupe G et $H \subseteq G$, un sous-groupe de G . On note alors $\text{Res}_H^G(V)$ et on appelle restriction de la représentation V de G à H la restriction de la représentation G au sous-groupe H de G .

C'est bien une représentation de H .

On s'intéresse ici à une construction inverse : on veut définir une représentation de G à partir de son sous-groupe H .

Remarque : Il ne s'agira pas de l'opération inverse au sens stricte : on ne retrouvera généralement pas notre représentation initiale en considérant l'induction de sa restriction ou alternativement la restriction de son induction.

Soit H un sous-groupe de G . Soit V une représentation de G et $W \subseteq V$ un sous-espace H -invariant :

$$\forall h \in H, \quad \forall w \in W, \quad h \cdot w \in w.$$

Alors, pour tout $g \in G$, le sous-espace $g \cdot W := \{g \cdot w \mid w \in W\}$ ne dépend que de la classe (à gauche) de g modulo H , gH . En effet :

$$\forall g \in G, \quad \forall h \in H, \quad gh \cdot W = g \cdot (h \cdot W) = g \cdot W.$$

Notons $\sigma \cdot W$ ce sous-espace de V , où σ est un représentant d'une classe de G/H .

Définition I.17. On dit que V est induit par W si tout élément de V peut être écrit de façon unique comme somme de tels éléments translatés de W :

$$V = \bigoplus_{\sigma \in G/H} \sigma \cdot W.$$

On note alors :

$$V = \text{Ind}_H^G(W).$$

Remarque : Si une telle représentation existe, alors elle est unique, à isomorphisme près.

On note W la représentation du sous-groupe H de G et on s'intéresse à $V = \text{Ind}_H^G(W)$. Pour chaque $\sigma \in G/H$, on se donne un élément $g_\sigma \in G$. Soit $v \in V$. On sait qu'il existe des éléments $w_\sigma \in W$ tels que l'on puisse écrire, de façon unique, v comme :

$$v = \sum_{\sigma \in G/H} g_\sigma w_\sigma.$$

Représentations linéaires

Pour obtenir le caractère correspondant à une telle représentation, remarquons d'abord que la multiplication à gauche par $g \in G$ change $\sigma \cdot W$ en $g\sigma \cdot W$. Ainsi, la trace de V est calculée à partir des $\sigma \in G/H$ tels que $g\sigma = \sigma$, donc à partir de

$$gs = s \iff s^{-1}gs \in H \text{ pour } s \in \sigma.$$

Ainsi, et en choisissant s arbitrairement dans la classe de σ , on a :

$$\chi_{\text{Ind}_H^G(W)}(g) = \sum_{g\sigma=\sigma} \chi_W(s^{-1}gs).$$

III. Le groupe symétrique $\mathfrak{S}_n, n \in \mathbb{N}$

Définition I.18. On appelle groupe symétrique et on note \mathfrak{S}_n l'ensemble des permutations sur les éléments de $1, \dots, n$, c'est à dire l'ensemble des bijections de $\{1, \dots, n\}$ dans lui-même, pour un $n \in \mathbb{N}^*$.

Il s'agit bien d'un groupe pour la loi de composition. Il est d'ordre $n!$.

Remarque : On travaille dans ce TER uniquement sur les groupes de la forme \mathfrak{S}_n , c'est à dire sur les groupes des permutations des entiers $\{1, \dots, n\}$ pour un n donné. Mais, on sait en fait que tout groupe de permutation (pas forcément d'entiers donc) est en bijection avec un \mathfrak{S}_n à condition de choisir le n convenablement.

1. Partitions de l'entier n

Nous venons de voir que pour étudier l'ensemble des représentations de \mathfrak{S}_n , il suffisait en fait de se restreindre à l'étude de ses représentations irréductibles et que celles-ci étaient au nombre de celui des classes de conjugaison de \mathfrak{S}_n . Nous allons donc nous intéresser dans un premier temps à l'ensemble des classes de conjugaisons de \mathfrak{S}_n .

Définition I.19. Soit $n \in \mathbb{N}$. Une partition λ de n est une suite d'entiers

$$\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k)$$

telle que $\lambda_1 + \lambda_2 + \dots + \lambda_k = n$. On impose aussi généralement que la suite (λ_i) soit décroissante.

On autorise parfois de telles suites à comporter des termes finaux nuls. On dira alors de deux suites qu'elles définissent une même partition si elles ne diffèrent que par leur nombre de zéros en fin de liste.

Proposition I.20

Il y a autant de classes de conjugaison de \mathfrak{S}_n que de partitions de n .

Démonstration :

Soit $\sigma \in \mathfrak{S}_n$. σ est décomposable en cycles à supports disjoints.
 Soit $c = (a_1, a_2, \dots, a_m)$ l'un d'entre ces cycles.
 On remarque que :

$$\forall \tau \in \mathfrak{S}_n, \quad \tau c \tau^{-1} = (\tau(a_1), \tau(a_2), \dots, \tau(a_m)).$$

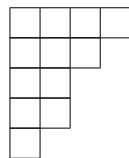
Ainsi $\tau c \tau^{-1}$ et c ont la même longueur et il y a donc autant de m -cycles dans σ que dans $\tau \sigma \tau^{-1}$. Or la somme des longueurs des cycles d'une permutation σ de \mathfrak{S}_n forme une partition de l'entier n (avec la convention que si $i \notin \text{Supp}(\sigma)$ alors $\text{longueur}(i) = 1$).

On associe donc une partition de n à chaque classe de conjugaison de \mathfrak{S}_n .
 Réciproquement, si $(\lambda_1, \lambda_2, \dots, \lambda_k) \in \mathbb{N}^k$ est une partition de n on peut lui associer la classe de conjugaison de $\sigma \in \mathfrak{S}_n$ où σ a pour décomposition en cycles à supports disjoints le produit d'un λ_1 -cycle, d'un λ_2 -cycle, ... □

Une façon pratique de représenter les partitions d'un entier n donné est d'avoir recourt au *tableaux de Young* :

Soit $n \in \mathbb{N}^*$. Si $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k) \in \mathbb{N}^k$ est une partition de l'entier n , on lui associe le tableau de Young T constitué de k lignes, la $i^{\text{ème}}$ ligne de T étant constituée de λ_i blocs. Ainsi, en tout, un tel tableau compte n blocs.

Exemple : Par exemple, pour $n = 12$, et $\lambda = (4, 3, 2, 2, 1)$ partition de 12, on a :



2. Représentation standard

Définition I.21. Soit un groupe G agissant par permutation sur un ensemble fini X . L'action de G sur X induit donc une représentation : on l'appelle la représentation par permutation.

Proposition I.22 (Formule du point fixe)

Si V est la représentation par permutation associée à l'action d'un groupe G sur un ensemble fini X , alors pour tout élément g de G , $\chi_V(g)$ est le nombre

Représentations linéaires

d'élément(s) de X fixé(s) par g .

Démonstration :

Soit M la matrice associée à l'action de g . M est une matrice de permutation. Par exemple, si on suppose que $X = x_1, x_2, x_3, x_4$ et que $\rho(g)$ est l'application définie comme suit :

$$\begin{aligned} \rho(g) : V &\longrightarrow V \\ e_{x_1} &\longmapsto e_{x_3} \\ e_{x_2} &\longmapsto e_{x_2} \\ e_{x_3} &\longmapsto e_{x_1} \\ e_{x_4} &\longmapsto e_{x_4} \end{aligned}$$

où on a noté $e_{x_1}, e_{x_2}, e_{x_3}$ et e_{x_4} les vecteurs de base de V .

En particulier, dans cette base, on a :

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

Et donc $\chi_V(g) = 2$, ce qui est le résultat que l'on cherche à prouver dans le cas général.

Plus généralement, $\exists n \in \mathbb{N}$ tel que $X = \{x_1, x_2, \dots, x_n\}$ et on note $M = (m_{k,l})_{1 \leq k, l \leq n}$. Alors, si $\exists x_i, x_j \in X$ tel que $g \cdot e_{x_i} = e_{x_j}$,

$$m_{j,i} = 1 \quad \text{et} \quad \forall k \in \{1, 2, \dots, i-1, i+1, \dots, n\}, m_{k,i} = 0.$$

En particulier, si $\exists x_i \in X$ tel que $g \cdot e_{x_i} = e_{x_i}$, alors $m_{i,i} = 1$. La trace de M est égale au nombre de 1 présents sur sa diagonale, soit exactement au nombre de points laissés fixes par g . \square

Dans le cas du groupe \mathfrak{S}_n , on note cette représentation $\rho_{\mathbb{C}^n}$ ou plus simplement \mathbb{C}^n .

Cependant, la représentation que l'on vient de définir n'est pas irréductible.

En effet, si on note (e_1, \dots, e_n) la base canonique de \mathbb{C}^n , $x := e_1 + e_2 + \dots + e_n$ et σ la permutation qui permet de définir la représentation par permutation

$$\forall j \in \{1, \dots, n\}, \quad \sigma \cdot e_j = e_{\sigma(j)}.$$

On remarque que $\sigma \cdot x = x$, autrement dit la droite engendrée par x , $\mathbb{C}(x)$ est laissée stable par σ et est donc une sous-représentation de la représentation \mathbb{C}^n . Elle est de degré 1 et il s'agit en fait de la représentation triviale ...

Mais, on sait alors que son orthogonal est également une sous-représentation de \mathbb{C}^n , de degré $n-1$, et c'est à cette représentation que l'on va s'intéresser.

Proposition I.23

Avec la notation que l'on vient d'introduire et en notant U la représentation triviale de \mathfrak{S}_n , il existe V , sous-représentation de \mathbb{C}^n , telle que $\mathbb{C}^n = U \oplus V$. De plus, cette représentation est irréductible.

Démonstration :

On note V la représentation que l'on vient de définir. On sait déjà que c'est une sous-représentation de \mathbb{C}^n et il reste donc à voir qu'elle est irréductible.

Soit $W \neq \{0\}$ une sous-représentation de V . On veut voir que $W = V$, c'est-à-dire que $V \subseteq W$.

Pour ça, on se donne v un vecteur non nul de W , on note (e_1, \dots, e_n) la base canonique de \mathbb{C}^n et $x := e_1 + e_2 + \dots + e_n$.

Comme $v \in V$, on a en particulier que $v \cdot x = 0$.

En notant

$$v = v_1 e_1 + v_2 e_2 + \dots + v_n e_n$$

on obtient que

$$v_1 + v_2 + \dots + v_n = 0.$$

Il existe donc i, j tels que $v_i \neq v_j$.

Soit τ la transposition $(i \ j)$. Comme W est une sous-représentation, W est stable par τ donc $\tau \cdot v \in W$ puis $\tau \cdot v - v \in W$. Or

$$\tau \cdot v = (v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_{j-1}, v_i, v_{j+1}, \dots, v_n).$$

Donc

$$\tau \cdot v - v = (v_j - v_i)(e_i - e_j).$$

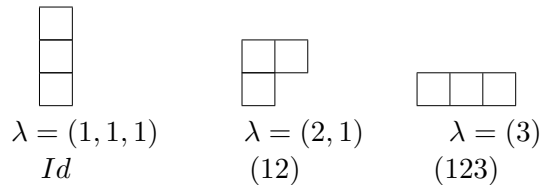
Et, comme $v_i - v_j \neq 0$, cela prouve que $e_i - e_j \in W$.

Maintenant, pour tous $1 \leq k, l \leq n$ tels que $k \neq l$ il existe une permutation $\sigma \in \mathfrak{S}_n$ envoyant i sur k et j sur l (par exemple le produit de deux transpositions, ou un 4-cycle). On a alors $e_k - e_l = \sigma \cdot (e_i - e_j) \in W$ car W est stable par σ . Ainsi W contient l'ensemble $\{e_k - e_l | k \neq l\}$, qui est une partie génératrice de V . Donc $V \subseteq W$ et, par suite $V = W$. □

On appelle cette représentation la *représentation standard*.

3. Exemple 1 : \mathfrak{S}_3

\mathfrak{S}_3 possède trois classes de conjugaison : Id , (12) et (123) , de cardinal respectif 1, 3 et 2. En effet,



Il y a donc au plus 3 représentations irréductibles différentes pour \mathfrak{S}_3 et en fait il y en a exactement 3 : la représentation triviale, la représentation alternée et la représentation standard.

On obtient donc la table des caractères suivante :

		1	3	2
\mathfrak{S}_3	Id	(12)	(123)	
triviale U	1	1	1	
alternée U'	1	-1	1	
standard V	2	0	-1	

Les deux premières lignes sont données par définition des représentations triviale et alternée. Pour remplir la troisième ligne, on peut utiliser les règles établies dans la partie I.1 qui permettent de compléter une ligne lorsque l'on connaît le reste du tableau ou bien se rappeler que $\mathbb{C}^3 = U \oplus V$ où \mathbb{C}^3 désigne la représentation par permutation, V la standard et U la triviale. On a alors :

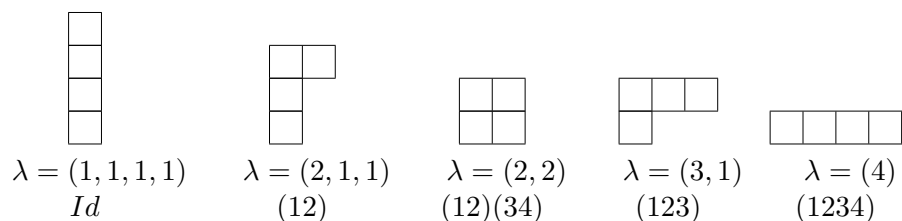
$$\chi_V = \chi_{\mathbb{C}^3} - \chi_U = (3, 1, 0) - (1, 1, 1) = (2, 0, -1)$$

en utilisant la formule du point fixe.

Par la suite, on notera toujours U la représentation triviale, U' la représentation alternée et V la représentation standard.

4. Exemple 2 : \mathfrak{S}_4

Cette fois-ci, \mathfrak{S}_4 possède cinq classes de conjugaison : Id , (12) , $(12)(34)$, (123) et (1234) de cardinaux respectifs 1, 6, 3, 8 et 6. En effet,



On peut remplir facilement les 3 premières lignes (en utilisant la formule du point fixe) :

$$\chi_V = \chi_{\mathbb{C}^4} - \chi_U = (4, 2, 0, 1, 0) - (1, 1, 1, 1, 1) = (3, 1, -1, 0, -1).$$

\mathfrak{S}_4	1	6	3	8	6
Id	(12)	(12)(34)	(123)	(1234)	
triviale U	1	1	1	1	1
alternée U'	1	-1	1	1	-1
standard V	3	1	-1	0	-1

Par ailleurs, comme il y a au plus 5 caractères irréductibles distincts, en notant α et β les degrés des représentations ρ_4 et ρ_5 et en appliquant la formule $|G| = \sum_i \dim(V_i)^2$, on a :

$$\begin{aligned} 1^2 + 1^2 + 3^2 + \alpha^2 + \beta^2 &= 24 \\ \alpha^2 + \beta^2 &= 11 \\ \Rightarrow (\alpha, \beta) &= (2, 3). \end{aligned}$$

Pour terminer de remplir la table des caractères, on peut "remarquer" (!) qu'une représentation de degré 2 est donnée par :

$$\rho_4 : G \rightarrow \mathfrak{S}_4 / \{1, (12)(34), (13)(24), (14)(23)\} \xrightarrow{\rho'_2} \mathcal{GL}(\mathbb{C})$$

où ρ'_2 est la représentation régulière de $\mathfrak{S}_4 / \dots \cong \mathfrak{S}_3$.

Une fois le caractère de ρ_4 calculé, on peut en déduire celui de ρ_5 .

Mais, cette méthode n'est pas généralisable aux dimensions supérieures, c'est pourquoi nous allons développer des outils dans notre seconde partie afin de résoudre ce problème.

– Chapitre II –

Outils pour étudier les représentations du groupe \mathfrak{S}_n

Le but de cette section est donc de développer des outils mathématiques qui vont nous permettre d'étudier les représentations du groupe symétrique \mathfrak{S}_n pour tout entier n , et donc *a fortiori* de n'importe quel groupe de permutations.

Pour cela, nous allons d'abord énoncer l'ensemble des résultats que nous avons pour objectif de démontrer afin de se donner une vision d'ensemble. On pourra alors terminer de remplir la table des caractères de \mathfrak{S}_4 que nous avons laissée en suspens.

I. Conventions d'écriture

Dans tout ce paragraphe, n sera un entier non-nul fixé.

On se donne une partition $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_r)$ de n ordonnée, c'est-à-dire telle que la suite des (λ_i) soit décroissante, et le tableau de Young T correspondant.

1. Diagrammes de Young

On numérote le tableau de Young en attribuant un entier compris entre 1 et n à chacune des n cases. Classiquement, on numérote de gauche à droite et de haut en bas.

Exemple : Pour $n = 9$ et pour la partition $\lambda = (4, 3, 2)$ de 9, on a :

1	2	3	4
5	6	7	
8	9		

Comme nous allons le voir juste après, la numérotation choisie n'aura pas d'influence pour le problème que l'on considère.

Définition II.1. On définit :

$$P_\lambda = \{g \in \mathfrak{S}_n \mid g \text{ préserve chaque ligne de } T\}$$

et

$$Q_\lambda = \{g \in \mathfrak{S}_n \mid g \text{ préserve chaque colonne de } T\}.$$

Proposition II.2

P_λ et Q_λ ne dépendent en fait pas du choix de la numérotation, dans le sens où les différents P_λ et Q_λ correspondant à différentes numérotations possibles sont conjugués les uns aux autres.

Démonstration :

On le montre pour P_λ :

Soit D et Δ deux numérotations possibles pour le tableau de Young T que l'on considère. En complétant le tableau de Young par des zéros, on obtient donc deux matrices de taille $k \times k$.

Par exemple, si on reprend la partition $\lambda = (4, 3, 2)$ de 9 vue à la page 73 et en choisissant deux numérotations : la canonique D à gauche et une autre Δ à droite :

$$\begin{array}{|c|c|c|c|} \hline 1 & 2 & 3 & 4 \\ \hline 5 & 6 & 7 & \\ \hline 8 & 9 & & \\ \hline \end{array} \implies \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 0 \\ 8 & 9 & 0 & 0 \end{pmatrix} \quad \Bigg| \quad \begin{array}{|c|c|c|c|} \hline 1 & 4 & 7 & 9 \\ \hline 2 & 5 & 8 & \\ \hline 3 & 6 & & \\ \hline \end{array} \implies \begin{pmatrix} 1 & 4 & 7 & 9 \\ 2 & 5 & 8 & 0 \\ 3 & 6 & 0 & 0 \end{pmatrix}$$

On note $D = (d_{i,j})_{1 \leq i,j \leq k}$ et $\Delta = (\delta_{i,j})_{1 \leq i,j \leq k}$ et on définit l'application σ qui permet de passer d'une numérotation D à une autre comme suit :

$$\forall i, j \in \{1, \dots, k\} \quad \sigma : \begin{array}{l} \{1, \dots, n\} \longrightarrow \{1, \dots, n\} \\ d_{i,j} \in D \longmapsto \delta_{i,j} \in \Delta \end{array} .$$

En particulier, on a que $\sigma \in \mathfrak{S}_n$.

Soit P_λ associé à la numérotation D et Π_λ associé à la numérotation Δ .

On note

$$\begin{aligned} P_\lambda &= \{g \in \mathfrak{S}_n \mid g \text{ préserve chaque ligne pour la numérotation } D\} \\ &= \{p_1, p_2, p_3, \dots\} \end{aligned}$$

et

$$\begin{aligned} \Pi_\lambda &= \{g \in \mathfrak{S}_n \mid g \text{ préserve chaque ligne pour la numérotation } \Delta\} \\ &= \{\pi_1, \pi_2, \pi_3, \dots\} . \end{aligned}$$

On conclut alors en remarquant que

$$\begin{aligned} \Pi_\lambda &= \{\sigma \circ p_1 \circ \sigma^{-1}, \sigma \circ p_2 \circ \sigma^{-1}, \sigma \circ p_3 \circ \sigma^{-1}, \dots\} \\ &\cong P_\lambda . \end{aligned}$$

En effet, pour tout k -cycle $c = (a_1 a_2 \dots a_k)$ et toute permutation $\sigma \in \mathfrak{S}_n$, on sait que :

$$\sigma \circ c \circ \sigma^{-1} = (\sigma(a_1) \sigma(a_2) \dots \sigma(a_k))$$

et les éléments de P_λ (et donc à fortiori de Π_λ) sont tous soit des k -cycles pour un $k \leq n$, soit un produit de cycles à support disjoints et, dans tous les cas on peut appliquer ce qui précède car $\sigma \circ c_1 \circ c_2 \circ \sigma^{-1} = \sigma \circ c_1 \circ \sigma^{-1} \circ \sigma \circ c_2 \circ \sigma^{-1}$. \square

Proposition II.3

P_λ et Q_λ sont des sous-groupes de \mathfrak{S}_n .

Démonstration :

On montre que P_λ est un sous-groupe de \mathfrak{S}_n :

- $Id_{\mathfrak{S}_n}$ préserve chaque case du tableau de Young donc en particulier $Id_{\mathfrak{S}_n} \in P_\lambda$.
- Soit $g, g' \in P_\lambda$. On a alors immédiatement que gg' préserve chaque ligne, i.e $gg' \in P_\lambda$.
- Soit $g \in P_\lambda$. Alors $g \in \mathfrak{S}_n$ et $g^{-1} \in \mathfrak{S}_n$ existe et préserve chaque ligne. En effet, soit $i \in \{1, \dots, n\}$ et $j \in \{1, \dots, k\}$ tel que i se trouve sur la $j^{\text{ème}}$ ligne. $g^{-1}g = Id$ et g préserve les lignes donc $g(i)$ est sur la $j^{\text{ème}}$ ligne. Id préserve les lignes donc $g^{-1}(g(i))$ restera sur la $j^{\text{ème}}$ ligne, i.e g^{-1} préserve chaque ligne et donc $g^{-1} \in P_\lambda$.

D'où le résultat. □

2. Algèbre de groupe $\mathbb{C}\mathfrak{S}_n$

Par sa qualité structurelle, nous allons transférer le problème posé à l'algèbre de groupe $\mathbb{C}\mathfrak{S}_n$. En effet, l'algèbre de groupe $\mathbb{C}\mathfrak{S}_n$ possède une structure d'espace vectoriel sous-jacente, engendrée par les $(e_g)_{g \in \mathfrak{S}_n}$ correspondant aux éléments de \mathfrak{S}_n . On lui donne la structure d'algèbre grâce à l'opération $e_g \cdot e_h = e_{gh}, \forall g, h \in \mathfrak{S}_n$.

Ce changement de point de vue n'affecte en rien la recherche des représentations de \mathfrak{S}_n car si $\rho : \mathfrak{S}_n \rightarrow \mathcal{GL}(V)$ en est une, elle s'étend par linéarité en une représentation $\tilde{\rho} : \mathbb{C}\mathfrak{S}_n \rightarrow \mathcal{L}(V)$ et réciproquement.

Nous conserverons cette notation e_g .

Définition II.4. La notion d'algèbre de groupe est généralisable à tout groupe fini G : On définit l'algèbre de groupe G , sur le corps \mathbb{K} par la formule :

$$\mathbb{K}G = \left\{ \sum_{g \in G} \lambda_g \cdot g \mid \forall g \in G, \lambda_g \in \mathbb{K} \right\}.$$

Définition II.5. On peut alors définir

$$a_\lambda = \sum_{p \in P_\lambda} e_p \quad \text{et} \quad b_\lambda = \sum_{q \in Q_\lambda} \text{sgn}(q)e_q.$$

qui sont tous deux des éléments de l'algèbre de groupe $\mathbb{C}\mathfrak{S}_n$.

On définit le symétriseur de Young par :

$$c_\lambda = a_\lambda \cdot b_\lambda \in \mathbb{C}\mathfrak{S}_n.$$

Outils pour étudier les représentations du groupe \mathfrak{S}_n

Définition II.6. On définit V_λ comme l'image de c_λ par multiplication à droite sur $\mathbb{C}\mathfrak{S}_n$:

$$V_\lambda = \mathbb{C}\mathfrak{S}_n \cdot c_\lambda.$$


V_λ est une représentation de $\mathbb{C}\mathfrak{S}_n$.

Proposition II.7

Pour toute partition λ de n ,

$$\mathbb{C}\mathfrak{S}_n \cdot a_\lambda b_\lambda \cong \mathbb{C}\mathfrak{S}_n \cdot b_\lambda a_\lambda.$$


On aurait donc pu définir V_λ comme étant $\mathbb{C}\mathfrak{S}_n \cdot b_\lambda a_\lambda$ et on aurait eu les mêmes résultats.

Exemple : Soit $\lambda = (n)$. 

Alors tout σ de \mathfrak{S}_n préserve $\{1, \dots, n\}$, soit l'ensemble des lignes donc $\mathfrak{S}_n \subseteq P_\lambda$. Alternative-ment, si $\sigma \in \mathfrak{S}_n$ préserve chaque colonne, alors pour tout $i \in \{1, \dots, n\}$, $\sigma(i) = i$, i.e $\sigma = Id_{\mathfrak{S}_n}$. On a donc $P_\lambda = \mathfrak{S}_n$ et $Q_\lambda = \{Id_{\mathfrak{S}_n}\}$. Donc :

$$V_\lambda = \mathbb{C}\mathfrak{S}_n \cdot \sum_{g \in \mathfrak{S}_n} e_g = \mathbb{C} \sum_{g \in \mathfrak{S}_n} e_g.$$


Et, on reconnaît la *représentation triviale*.

Exemple : Soit $\lambda = (1, 1, \dots, 1)$. 

Par un raisonnement similaire, on montre que $P_\lambda = \{Id_{\mathfrak{S}_n}\}$ et que $Q_\lambda = \mathfrak{S}_n$ et donc :

$$V_\lambda = \mathbb{C}\mathfrak{S}_n \cdot \sum_{g \in \mathfrak{S}_n} sgn(g)e_g = \mathbb{C} \sum_{g \in \mathfrak{S}_n} sgn(g)e_g.$$

On retrouve la *représentation alternée*.

Exemple : Soit $\lambda = (2, 1)$ dans \mathfrak{S}_3 . 

On a alors $a_\lambda = e_{Id} + e_{(12)}$ et $b_\lambda = e_{Id} - e_{(13)}$.

D'où :

$$c_\lambda = a_\lambda b_\lambda = e_{Id} - e_{(13)} + e_{(12)} - e_{(132)}.$$

En utilisant la notation $\langle \dots; \dots \rangle$ pour "engendré par", on trouve :

$$\begin{aligned} V_\lambda &= \mathbb{C}\mathfrak{S}_n \cdot c_\lambda \\ &= \langle e_{Id}; e_{(12)}; e_{(13)}; e_{(23)}; e_{(123)}; e_{(132)} \rangle \cdot c_\lambda \\ &= \langle c_\lambda; c_\lambda; e_{(13)} - e_{Id} + e_{(123)} - e_{(23)}; e_{(23)} - e_{(123)} + e_{(132)} - e_{(12)} \\ &\quad ; e_{(123)} - e_{(23)} + e_{(13)} - e_{Id}; e_{(132)} - e_{(12)} + e_{(23)} - e_{(123)} \rangle \\ &= \langle c_\lambda; e_{(13)} \cdot c_\lambda \rangle. \end{aligned}$$

V_λ est donc la représentation engendrée par c_λ et $e_{(13)} \cdot c_\lambda$. En particulier, c'est une représentation de degré 2 dans \mathfrak{S}_3 donc c'est la *représentation standard*.

Plus généralement, on pourrait montrer que, pour tout n dans \mathbb{N} , la représentation standard de \mathfrak{S}_n est V_λ avec $\lambda = (n - 1, 1)$.

3. Notations pour la formule de Frobenius

Les notations qui suivent serviront à la compréhension de la formule clé de ce dossier : *la formule de Frobenius*. Celle-ci donne une formule explicite pour tous les caractères irréductibles, en toute dimension.

On note C_i la classe de conjugaison de \mathfrak{S}_n qui consiste en l'ensemble des permutations possédant i_1 1-cycles, i_2 2-cycles, \dots , i_n n -cycles. On associe à cette classe un n -uplet $i = (i_1, i_2, \dots, i_n)$. Ce n -uplet doit en particulier vérifier que $\sum_\alpha \alpha i_\alpha = n$.

On note r le nombre de lignes dans le diagramme de Young associé à la partition λ de n et on se donne un entier k de $\{1, \dots, r\}$. On introduit alors k variables indépendantes : x_1, x_2, \dots et x_k .

Définition II.8. *Pour tout j de $\{1, \dots, n\}$ on définit*

$$P_j(x) = x_1^j + x_2^j + \dots + x_k^j$$

et le discriminant

$$\Delta(x) = \prod_{i < j} (x_i - x_j).$$

Si $f(x) = f(x_1, x_2, \dots, x_n)$ est une série polynomiale formelle et si (l_1, l_2, \dots, l_k) est un k -uplet d'entiers naturels alors $[f(x)]_{(l_1, \dots, l_k)}$ désigne le coefficient de $x_1^{l_1} x_2^{l_2} \dots x_k^{l_k}$ dans f .

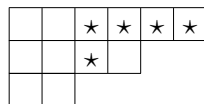
De plus, à toute partition $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k)$ de n , on associe une suite $(l_i)_{i \in \{1, \dots, k\}} \in \mathbb{N}^k$ en posant

$$\forall i \in \{1, \dots, k\}, \quad l_i := \lambda_i + k - i.$$

Cette suite $(l_i)_{i \in \{1, \dots, k\}}$ est en particulier strictement décroissante.

Enfin, on introduit la notion de longueur de crochet qui nous sera utile pour énoncer la *formule des crochets*.

Définition II.9. *Soit un diagramme de Young T . On définit la longueur de crochet d'une case donnée comme étant le nombre de case située à sa droite et sous elle en y ajoutant la case elle-même.*



Un exemple de crochet.

Outils pour étudier les représentations du groupe \mathfrak{S}_n

Exemple : On se place dans \mathfrak{S}_8 et on considère la partition $\lambda = (4, 3, 1)$ de 8.

6	4	3	1
4	2	1	
1			

Le diagramme de Young ci-dessus contient en chacune de ses cases la longueur de crochet de cette même case.

II. Présentation des résultats

1. Formule de Frobenius

La formule de Frobenius donne une formulation explicite de n'importe quel caractère associé à une représentation V_λ , en toute dimension :

Théorème II.10 (Formule de Frobenius)

Avec les notations que nous avons introduites dans le paragraphe précédent, le caractère χ_λ de V_λ évalué en $g \in C_i$ est donné par :

$$\chi_\lambda(C_i) = \left[\Delta(x) \cdot \prod_{j=1}^k P_j(x)^{i_j} \right]_{(l_1, \dots, l_k)}.$$

Exemple : On se place dans \mathfrak{S}_5 et on considère la partition $\lambda = (3, 2)$.

Alors $k = 2$, $l_1 = 2 + 3 - 1 = 4$ et $l_2 = 2 + 2 - 2 = 2$.

— On considère $C_i = [(12)(345)] \quad i = (0, 1, 1, 0, 0)$.

$$\begin{aligned} \chi_\lambda(C_i) &= [(x_1 - x_2)(x_1^2 + x_2^2)(x_1^3 + x_2^3)]_{(4,2)} \\ &= [(x_1 - x_2)(x_1^5 + x_1^2x_2^3 + x_1^3x_2^2 + x_2^5)]_{(4,2)} \\ &= [x_1^6 + x_1^3x_2^3 + \mathbf{x_1^4x_2^2} + x_1x_2^5 - x_1^5x_2 - x_1^2x_2^4 - x_1^3x_2^3 - x_2^6]_{(4,2)} \\ &= 1. \end{aligned}$$

— On considère $C_j = [Id] \quad j = (5, 0, 0, 0, 0)$.

$$\begin{aligned} \chi_\lambda(C_j) &= [(x_1 - x_2)(x_1 + x_2)^5]_{(4,2)} \\ &= [(x_1 - x_2)(x_1^5 + 5x_1^4x_2 + 10x_1^3x_2^2 + 10x_1^2x_2^3 + 5x_1x_2^4 + x_2^5)]_{(4,2)} \\ &= [x_1^6 + 5x_1^5x_2 + \mathbf{10x_1^4x_2^2} + 10x_1^3x_2^3 + 5x_1^2x_2^4 + x_1x_2^5 - x_1^5x_2 - \mathbf{5x_1^4x_2^2} \\ &\quad - 10x_1^3x_2^3 - 10x_1^2x_2^4 - 5x_1x_2^5 - x_2^6]_{(4,2)} \\ &= 5. \end{aligned}$$

2. Formule des crochets

La formule des crochets donne quant à elle la dimension de V_λ , pour toute partition λ de n , avec n quelconque.

Théorème II.11 (Formule des crochets)

$$\dim(V_\lambda) = \frac{n!}{\prod \text{longueur de crochet}}.$$

Exemple : Dans \mathfrak{S}_5 , avec la partition $\lambda = (3, 2)$.



On peut ré-emplir le diagramme de Young ci-dessus en inscrivant les longueurs de crochet de chacune des cases au sein de celle-ci (comme on avait fait à la page 78) :



On en déduit

$$\dim(V_\lambda) = \frac{n!}{\prod \text{longueur de crochet}} = \frac{5 \times 4 \times 3 \times 2 \times 1}{4 \times 3 \times 1 \times 2 \times 1} = 5.$$

En particulier, on retrouve bien que $\dim(V_\lambda) = \chi_\lambda(\text{Id})$.

3. Liste exhaustive des représentations du groupe \mathfrak{S}_n

L'intérêt de ces représentations V_λ sera finalement justifié par le théorème suivant :

Théorème II.12

L'ensemble des représentations irréductibles de \mathfrak{S}_n est exactement l'ensemble des V_λ où λ est une partition ordonnée de n .

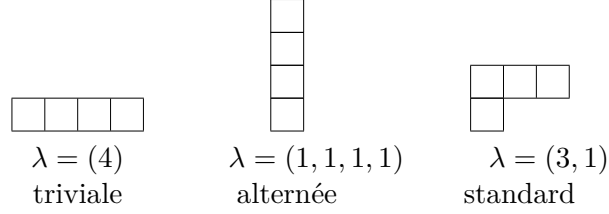
Il existe donc une bijection entre l'ensemble des partitions d'un entier n donné et l'ensemble des représentations du groupe symétrique \mathfrak{S}_n associé.

4. Exemple 2 : \mathfrak{S}_4 - suite et fin

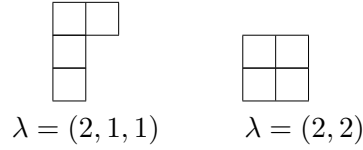
Appliquons ces théorèmes au cas de \mathfrak{S}_4 avant de se lancer dans leur démonstration :

Dans la partie I.3, à la page 70, nous avons réussi à remplir les 3 premières lignes de la table des caractères. Il en reste donc 2 à remplir. De plus, on a vu à la page 76 que les partitions associées aux représentations triviale, alternée et standard étaient

respectivement :



Il reste donc à étudier les représentations associées aux partitions :



Appliquons la formule de Frobenius à la partition $\lambda = (2, 2)$: cela nous permettra de remplir une ligne supplémentaire dans la table des caractères et on pourra alors en déduire la dernière.

On a $k = 2$, $l_1 = 2 + 2 - 1 = 3$ et $l_2 = 2 + 2 - 2 = 2$.

— on considère $C_{i_1} = [Id]$, $i_1 = (4, 0, 0, 0)$.

$$\begin{aligned}
 \chi_\lambda(C_{i_1}) &= [(x_1 - x_2)(x_1 + x_2)^4]_{(3,2)} \\
 &= [(x_1 - x_2)(x_1^4 + 4x_1^3x_2 + 6x_1^2x_2^2 + 4x_1x_2^3 + x_2^4)]_{(3,2)} \\
 &= [x_1^5 + 4x_1^4x_2 + 6x_1^3x_2^2 + 4x_1^2x_2^3 + x_1x_2^4 \\
 &\quad - x_1^4x_2 - 4x_1^3x_2^2 - 6x_1^2x_2^3 - 4x_1x_2^4 - x_2^5]_{(3,2)} \\
 &= 2.
 \end{aligned}$$

— on considère $C_{i_2} = [(12)]$, $i_1 = (2, 1, 0, 0)$.

$$\begin{aligned}
 \chi_\lambda(C_{i_2}) &= [(x_1 - x_2)(x_1 + x_2)^2(x_1^2 + x_2^2)]_{(3,2)} \\
 &= [(x_1 - x_2)(x_1^4 + 2x_1^3x_2 + 2x_1^2x_2^2 + 2x_1x_2^3 + x_2^4)]_{(3,2)} \\
 &= [x_1^5 + 2x_1^4x_2 + 2x_1^3x_2^2 + 2x_1^2x_2^3 + x_1x_2^4 \\
 &\quad - x_1^4x_2 - 2x_1^3x_2^2 - 2x_1^2x_2^3 - 2x_1x_2^4 - x_2^5]_{(3,2)} \\
 &= 0.
 \end{aligned}$$

— on considère $C_{i_3} = [(12)(34)]$, $i_1 = (0, 2, 0, 0)$.

$$\begin{aligned}
 \chi_\lambda(C_{i_3}) &= [(x_1 - x_2)(x_1^2 + x_2^2)^2]_{(3,2)} \\
 &= [(x_1 - x_2)(x_1^4 + 2x_1^2x_2^2 + x_2^4)]_{(3,2)} \\
 &= [x_1^5 + 2x_1^3x_2^2 + x_1x_2^4 - x_1^4x_2 - 2x_1^2x_2^3 - x_2^5]_{(3,2)} \\
 &= 2.
 \end{aligned}$$

— on considère $C_{i_4} = [(123)]$, $i_1 = (1, 0, 1, 0)$.

$$\begin{aligned}\chi_\lambda(C_{i_4}) &= [(x_1 - x_2)(x_1 + x_2)(x_1^3 + x_2^3)]_{(3,2)} \\ &= [(x_1 - x_2)(x_1^4 + x_1^3x_2 + x_1x_2^3 + x_2^4)]_{(3,2)} \\ &= [x_1^5 + x_1^4x_2 + x_1^2x_2^3 + x_1x_2^4 - x_1^4x_2 - x_1^3x_2^2 - x_1x_2^4 - x_2^5]_{(3,2)} \\ &= -1.\end{aligned}$$

— on considère $C_{i_5} = [(1234)]$, $i_1 = (0, 0, 0, 1)$.

$$\begin{aligned}\chi_\lambda(C_{i_5}) &= [(x_1 - x_2)(x_1^4 + x_2^4)]_{(3,2)} \\ &= [x_1^5 + x_1x_2^4 - 2x_1^4x_2 - x_2^5]_{(3,2)} \\ &= 0.\end{aligned}$$

On peut donc terminer de remplir la table des caractères :

\mathfrak{S}_4	1	6	3	8	6
	Id	(12)	(12)(34)	(123)	(1234)
triviale U	1	1	1	1	1
alternée U'	1	-1	1	1	-1
standard V	3	1	-1	0	-1
$V_{(2,2)}$	2	0	2	-1	0
$V_{(2,1,1)}$	3	-1	-1	0	1

III. Démonstration des résultats

1. Formule des crochets

Le but de cette section est d'arriver à trouver une formule explicite, et valable en toute dimension $n \in \mathbb{N}^*$, fournissant la dimension de V_λ , pour toute partition $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k)$ de l'entier n .

Par propriété des caractères, on sait que

$$\dim(V_\lambda) = \chi_\lambda([Id_{\mathfrak{S}_n}]).$$

On va donc utiliser la formule de Frobenius que l'on vient de démontrer pour calculer la quantité ci-dessus.

Avec les notations propres à la formule de Frobenius, la classe de conjugaison correspondant à $Id_{\mathfrak{S}_n}$ est $i = (n, 0, 0, \dots)$ de sorte que

$$\dim(V_\lambda) = \chi_\lambda(C_{(d)}) = [\Delta(x) \cdot (x_1 + x_2 + \dots + x_k)^n]_{(l_1, l_2, \dots, l_k)}.$$

Outils pour étudier les représentations du groupe \mathfrak{S}_n

On remarque d'une part que $\Delta(x)$ est en fait égal au *déterminant de Vandermonde* suivant :

$$\begin{aligned} \Delta(x) &= \begin{vmatrix} 1 & x_k & \cdots & x_k^{k-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_1 & \cdots & x_1^{k-1} \end{vmatrix} \\ &= \det(M) \text{ où } M = (m_{i,j}) \\ &\quad \text{et } \forall i \in \{1, \dots, k\}, \forall j \in \{2, \dots, k\} \begin{cases} m_{i,1} = 1 \\ m_{i,j} = x_{k-(i-1)}^{j-1} \end{cases} . \end{aligned}$$

Or, par définition du déterminant, on a :

$$\begin{aligned} \text{Det}(M) &= \sum_{\sigma \in \mathfrak{S}_k} \text{sgn}(\sigma) \times m_{1,\sigma(1)} \times m_{2,\sigma(2)} \times \dots \times m_{k,\sigma(k)} \\ &= \sum_{\sigma \in \mathfrak{S}_k} \text{sgn}(\sigma) \times 1 \times x_k^{\sigma(1)-1} \times \dots \times x_1^{\sigma(k)-1} . \end{aligned}$$

Et donc :

$$\Delta(x) = \sum_{\sigma \in \mathfrak{S}_k} \text{sgn}(\sigma) \times 1 \times x_k^{\sigma(1)-1} \times \dots \times x_1^{\sigma(k)-1} .$$

D'autre part, la *formule du multinôme* donne que :

$$\begin{aligned} (x_1 + x_2 + \dots + x_k)^n &= \sum_{r_1+r_2+\dots+r_k=n} \binom{n}{r_1, r_2, \dots, r_k} x_1^{r_1} x_2^{r_2} \dots x_k^{r_k} \\ \text{avec } \binom{n}{r_1, r_2, \dots, r_k} &= \frac{n!}{r_1! r_2! \dots r_k!} . \end{aligned}$$

On a donc :

$$\begin{aligned} \dim(V_\lambda) &= \left[\begin{aligned} &\left(\sum_{\sigma \in \mathfrak{S}_k} \text{sgn}(\sigma) \times x_k^{\sigma(1)-1} \times \dots \times x_1^{\sigma(k)-1} \right) \\ &\left(\sum_{r_1+r_2+\dots+r_k=n} \frac{n!}{r_1! r_2! \dots r_k!} \times x_1^{r_1} \times x_2^{r_2} \times \dots \times x_k^{r_k} \right) \end{aligned} \right]_{(l_1, l_2, \dots, l_k)} \\ &= \left[\begin{aligned} &\sum_{\sigma \in \mathfrak{S}_k} \sum_{r_1+r_2+\dots+r_k=n} \text{sgn}(\sigma) \times \frac{n!}{r_1! r_2! \dots r_k!} \\ &\times x_1^{r_1+\sigma(k)-1} \times x_2^{r_2+\sigma(k-1)-1} \times \dots \times x_k^{r_k+\sigma(1)-1} \end{aligned} \right]_{(l_1, l_2, \dots, l_k)} \\ &= \sum_{\sigma \in \mathfrak{S}_k} \text{sgn}(\sigma) \times \frac{n!}{(l_1 - \sigma(k) + 1)! (l_2 - \sigma(k-1) + 1)! \dots (l_k - \sigma(1) + 1)!} . \end{aligned}$$

En effet $l_1 = r_1 + \sigma(k) - 1 \Leftrightarrow r_1 = l_1 - \sigma(k) + 1$.

$$\begin{aligned}
 \dim(V_\lambda) &= n! \sum_{\sigma \in \mathfrak{S}_k} \operatorname{sgn}(\sigma) \prod_{j=1}^k \frac{1}{(l_j - \sigma(k-j+1) + 1)!} \quad (\star) \\
 &= n! \sum_{\sigma \in \mathfrak{S}_k} \operatorname{sgn}(\sigma) \prod_{j=1}^k \frac{(l_j - \sigma(k-j+1) + 2) \times \dots \times (l_j - 1) \times l_j}{l_j!} \\
 &= \frac{n!}{l_1! \dots l_k!} \sum_{\sigma \in \mathfrak{S}_k} \operatorname{sgn}(\sigma) \prod_{j=1}^k (l_j - \sigma(k-j+1) + 2) \times \dots \times (l_j - 1) \times l_j.
 \end{aligned}$$

En (\star) , la somme porte en fait sur les $\sigma \in \mathfrak{S}_k$ telles que

$$\forall j \in \{1, \dots, k\} \quad (r_j =) \quad l_{k-j+1} - \sigma(j) + 1 \leq 0.$$

D'autre part,

$$\begin{aligned}
 &\left| \begin{array}{cccccc}
 1 & l_k & l_k(l_k - 1) & l_k(l_k - 1)(l_k - 2) & \dots & l_k(l_k - 1) \dots (l_k - (k - 2)) \\
 \vdots & \vdots & \vdots & \vdots & & \vdots \\
 1 & l_2 & l_2(l_2 - 1) & l_2(l_2 - 1)(l_2 - 2) & \dots & l_2(l_2 - 1) \dots (l_2 - (k - 2)) \\
 1 & l_1 & l_1(l_1 - 1) & l_1(l_1 - 1)(l_1 - 2) & \dots & l_1(l_1 - 1) \dots (l_1 - (k - 2))
 \end{array} \right| \\
 &= \left| \begin{array}{cccccc}
 1 & l_1 & l_1(l_1 - 1) & l_1(l_1 - 1)(l_1 - 2) & \dots & l_1(l_1 - 1) \dots (l_1 - (k - 2)) \\
 1 & l_2 & l_2(l_2 - 1) & l_2(l_2 - 1)(l_2 - 2) & \dots & l_2(l_2 - 1) \dots (l_2 - (k - 2)) \\
 \vdots & \vdots & \vdots & \vdots & & \vdots \\
 1 & l_k & l_k(l_k - 1) & l_k(l_k - 1)(l_k - 2) & \dots & l_k(l_k - 1) \dots (l_k - (k - 2))
 \end{array} \right| \\
 &= \sum_{\sigma \in \mathfrak{S}_k} \operatorname{sgn}(\sigma) \times a_{1,\sigma(1)} \times a_{2,\sigma(2)} \times \dots \times a_{k,\sigma(k)} \\
 &= \sum_{\sigma \in \mathfrak{S}_k} \operatorname{sgn}(\sigma) \times l_1(l_1 - 1) \dots (l_1 - (\sigma(1) - 2)) \times l_2(l_2 - 1) \dots (l_2 - (\sigma(2) - 2)) \\
 &\quad \times \dots \times l_k(l_k - 1) \dots (l_k - (\sigma(k) - 2)) \\
 &= \sum_{\sigma \in \mathfrak{S}_k} \operatorname{sgn}(\sigma) \prod_{i=1}^k l_i(l_i - 1) \dots (l_i - \sigma(i) + 2) \\
 &= \sum_{\sigma \in \mathfrak{S}_k} \operatorname{sgn}(\sigma) \prod_{j=1}^k l_{k-j+1}(l_{k-j+1} - 1) \dots (l_{k-j+1} - \sigma(k - j + 1) + 2) \\
 &= \sum_{\sigma \in \mathfrak{S}_k} \operatorname{sgn}(\sigma) \prod_{j=1}^k l_j(l_j - 1) \dots (l_j - \sigma(k - j + 1) + 2)
 \end{aligned}$$

Et, comme on a construit la matrice de telle façon à avoir la condition de validité (\star)

Outils pour étudier les représentations du groupe \mathfrak{S}_n

vérifiée, on a :

$$\dim(V_\lambda) = \frac{n!}{l_1! \dots l_k!} \begin{vmatrix} 1 & l_k & l_k(l_k - 1) & l_k(l_k - 1)(l_k - 2) & \dots & l_k(l_k - 1) \dots (l_k - (k - 2)) \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & l_2 & l_2(l_2 - 1) & l_2(l_2 - 1)(l_2 - 2) & \dots & l_2(l_2 - 1) \dots (l_2 - (k - 2)) \\ 1 & l_1 & l_1(l_1 - 1) & l_1(l_1 - 1)(l_1 - 2) & \dots & l_1(l_1 - 1) \dots (l_1 - (k - 2)) \end{vmatrix}.$$

En effectuant des combinaisons entre les colonnes du déterminant ci-dessus, on peut se ramener au *déterminant de Vandermonde*. Par exemple,

$$\begin{aligned} C_1 &\leftarrow C_1 \\ C_2 &\leftarrow C_2 \\ C_3 &\leftarrow C_2 + C_3 \\ C_4 &\leftarrow -2C_2 + 3C_3 + C_4 \dots \end{aligned}$$

Finalement, on trouve donc :

$$\dim(V_\lambda) = \frac{n!}{l_1! \dots l_k!} \prod_{i < j} (l_i - l_j).$$

Théorème II.13 (Formule des crochets)

$$\dim(V_\lambda) = \frac{n!}{\prod \text{longueur de crochet}}.$$

Démonstration :

Soit le tableau de Young suivant, associée à la partition $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k)$ de l'entier k , dans lequel on a inscrit dans chacune de ses cases sa propre longueur de crochet :

On pose $\forall i \in \{1, \dots, k\}$, $l_i = \lambda_i + k - i$.

l_1											
l_2											
\vdots											
l_{k-2}	l_{k-2}	l_{k-2}	\dots	$l_{k-2} - (l_k - 1)$	$l_{k-2} - (l_k + 1)$	$l_{k-2} - (l_k + 2)$	\dots	$l_{k-2} - (l_{k-1} - 1)$	$l_{k-2} - (l_{k-1} + 1)$	\dots	1
l_{k-1}	l_{k-1}	l_{k-1}	\dots	$l_{k-1} - (l_k - 1)$	$l_{k-1} - (l_k + 1)$	$l_{k-1} - (l_k + 2)$	\dots	1			
l_k	$l_k - 1$	$l_k - 2$	\dots	1							

Pour le remplir, on peut procéder comme suit :

- La dernière ligne du tableau de Young comporte λ_k cases par définition de ce dernier. On a donc facilement les longueurs de crochet de la dernière ligne (les crochets étant quelque peu "amputés").
- On va chercher à remplir l'avant dernière ligne en avançant de la gauche vers la droite. La case la plus à gauche a pour longueur de crochet $\lambda_{k-1} + 1$ car cette ligne comporte λ_{k-1} cases et on ajoute 1 pour compter la case tout en bas à gauche. Or $\lambda_{k-1} + 1 = l_{k-1}$ par définition de l_{k-1} . On procède de même pour remplir les $l_{k-1}^{\text{ème}}$ cases suivantes, i.e pour remplir le tableau jusqu'au premier décrocher – surligné en gris sur notre diagramme Là, on remarque que lorsque l'on veut repérer le crochet correspondant à la case suivante, on perd en fait deux cases par rapport au précédent, d'où le résultat.
- On procède de même pour remplir la fin du tableau.

Si l'on effectue le produit des longueurs de crochet de la dernière ligne, au vu de ce qui précède, on trouve donc $l_k !$

Si l'on réitère l'opération sur l'avant-dernière, on trouve :

$$l_{k-1} \times (l_{k-1} - 1) \times \dots \times (l_{k-1} - l_k + 1) \times (l_{k-1} - l_k - 1) \times \dots \times 1 = \frac{l_{k-1} !}{l_{k-1} - l_k}$$

et sur l'antépénultième ligne :

$$l_{k-2} \times (l_{k-2} - 1) \times \dots \times (l_{k-2} - l_{k-1} + 1) \times (l_{k-2} - l_{k-1} - 1) \times \dots \times (l_{k-2} - l_k + 1) \times (l_{k-2} - l_k - 1) \times \dots \times 1 = \frac{l_{k-2} !}{(l_{k-2} - l_{k-1})(l_{k-2} - l_k)}.$$

Par suite, on a :

$$\prod \text{longueur de crochet} = \prod_{i < j} \frac{l_i ! \dots l_k !}{(l_i - l_j)}.$$

Et donc

$$\dim(V_\lambda) = \frac{n !}{l_1 ! \dots l_k !} \prod_{i < j} (l_i - l_j) = \frac{n !}{\prod \text{longueur de crochet}}$$

□

2. Liste exhaustive des représentations du groupe \mathfrak{S}_n

On garde les mêmes notations que jusqu'à présent et on note $A := \mathbb{C}\mathfrak{S}_n$ où n est un entier non-nul fixé dorénavant. De plus, on utilisera alternativement la notation e_g et g pour noter un élément de l'algèbre de groupe $\mathbb{C}\mathfrak{S}_n$ lorsqu'aucune confusion n'est possible.

L'objet de cette section est de démontrer le théorème suivant :

Théorème II.14

L'ensemble des représentations irréductibles de \mathfrak{S}_n est exactement l'ensemble des V_λ où λ est une partition ordonnée de n .

Pour ce faire, nous allons tout d'abord démontrer un certain nombre de lemmes.

Lemme II.15.1. Il existe au plus une façon d'écrire un élément de \mathfrak{S}_n sous la forme $p \cdot q$, avec $p \in P_\lambda$ et $q \in Q_\lambda$.

Démonstration :

On le montre par l'absurde : supposons qu'il existe $p, p' \in P_\lambda$ et $q, q' \in Q_\lambda$ tels que $p \cdot q = p' \cdot q'$. Alors, comme P_λ et Q_λ sont tous deux des groupes, on peut en particulier définir p^{-1} et $(q')^{-1}$. On a alors

$$q \cdot (q')^{-1} = p^{-1} \cdot p' \in P_\lambda \cap Q_\lambda = \{Id_{\mathfrak{S}_n}\}.$$

Et donc $p = p'$ et $q = q'$. □

En particulier, il existe une unique façon d'écrire $c_\lambda = a_\lambda \cdot b_\lambda$:

$$c_\lambda = \sum_{p \in P_\lambda, q \in Q_\lambda} sgn(q) e_{p \cdot q}$$

et le coefficient devant e_{Id} est toujours égal à 1.

Lemme II.15.2. Avec les notations introduites,

- i. $\forall p \in P_\lambda, p \cdot a_\lambda = a_\lambda \cdot p = a_\lambda$;
- ii. $\forall q \in Q_\lambda, (sgn(q)q) \cdot b_\lambda = b_\lambda \cdot (sgn(q)q) = b_\lambda$;
- iii. $\forall p \in P_\lambda, \forall q \in Q_\lambda, p \cdot c_\lambda(sgn(q)q) = c_\lambda$
et c_λ est le seul élément de \mathfrak{S}_n , à multiplication par un scalaire près, à vérifier cette propriété.

Démonstration :

- i. Soit $e_x \in P_\lambda$. Alors :

$$e_x \cdot a_\lambda = e_x \cdot \sum_{p \in P_\lambda} e_p = \sum_{p \in P_\lambda} e_{xp} = \sum_{p \in P_\lambda} e_p = a_\lambda.$$

- ii. On procède de même.

iii. Avec les points i. et ii., on a en particulier que

$$\forall p \in P_\lambda, \quad \forall q \in Q_\lambda, \quad p \cdot c_\lambda(\text{sgn}(q)q) = c_\lambda. \quad (\star)$$

Il reste à montrer que c'est le seul, à multiplication par un scalaire près. Soit $d_\lambda = \sum_{g \in \mathfrak{S}_n} n_g e_g \neq c_\lambda$ vérifiant la propriété (\star) .

On veut voir que $d_\lambda \in \mathbb{C}(c_\lambda)$. Soit $p \in P_\lambda$ et $q \in Q_\lambda$.

d_λ satisfait à la condition (\star) donc en particulier

$$\begin{aligned} d_\lambda &= \text{sgn}(q)p \cdot d_\lambda \cdot q \\ \sum_{g \in \mathfrak{S}_n} n_g e_g &= \text{sgn}(q) \sum_{g \in \mathfrak{S}_n} n_g e_{pgq}. \end{aligned}$$

On fixe $g \in \mathfrak{S}_n$ dans la somme de droite. Comme la somme de gauche porte également sur tout \mathfrak{S}_n , et qu'il existe $h \in \mathfrak{S}_n$ tel que $h = p \cdot g \cdot q$, alors par unicité du développement, on a $n_{pgq} = \text{sgn}(q)n_g$.

En particulier, $n_{pq} = \text{sgn}(q)n_1$.

Si on arrive à montrer que $\forall g \notin P_\lambda Q_\lambda, n_g = 0$ alors on a le résultat. En effet, dans ce cas :

$$d_\lambda = \sum_{p \in P_\lambda, q \in Q_\lambda} \text{sgn}(q) n_1 e_{pq} = n_1 \sum_{p \in P_\lambda, q \in Q_\lambda} \text{sgn}(q) e_{pq} = n_1 c_\lambda.$$

On va en fait montrer que pour de tels $g \in \mathfrak{S}_n; n_g = -n_g$. Pour cela, on va exhiber une transposition $t \in P_\lambda$ telle que $g^{-1}tg$ soit dans Q_λ car alors, en notant $p = t$ et $q = g^{-1}tg$:

$$pgq = tg(g^{-1}tg) = t(gg^{-1})tg = g.$$

Or :

$$\begin{aligned} p \cdot d_\lambda \cdot q &= \sum_{g \in \mathfrak{S}_n} \text{sgn}(q) n_g e_{pgq} \\ \sum_{g \in P_\lambda Q_\lambda} n_g e_{pgq} + \sum_{g \notin P_\lambda Q_\lambda} n_g e_{pgq} &= \sum_{g \in P_\lambda Q_\lambda} \text{sgn}(q) n_g e_{pgq} + \sum_{g \notin P_\lambda Q_\lambda} \text{sgn}(q) n_g e_{pgq}. \end{aligned}$$

Donc,

$$\forall g \notin P_\lambda Q_\lambda, n_g = \text{sgn}(q) n_g.$$

De plus,

$$\text{sgn}(q) = \text{sgn}(g^{-1}tg) = \text{sgn}(g)^2 \times \text{sgn}(t) = -1$$

d'où $\forall g \notin P_\lambda Q_\lambda, n_g = 0$.

Outils pour étudier les représentations du groupe \mathfrak{S}_n

Soit gT le tableau de Young obtenu à partir de T en remplaçant chacune des entrées du tableau T par leur image par g . On notera de même σT le tableau de Young obtenu à partir de T en remplaçant l'entrée i par $\sigma(i)$ pour toute permutation $\sigma \in \mathfrak{S}_n$. On remarque en particulier que si μ est une permutation de T , par exemple le cycle $(a_1 a_2 \dots a_k)$, alors $\sigma \circ \mu \circ \sigma^{-1}$ aura le même effet sur σT que μ sur T .

Si on arrive à montrer qu'il existe deux entiers qui se trouvent sur une même ligne pour le tableau T et sur une même colonne pour le tableau gT alors on aura le résultat avec ce qui précède : on appelle t la transposition qui échange ces deux entiers. On a alors $t \in P_\lambda$ car les deux entiers sont par hypothèse cités sur une même ligne dans le tableau T et $t \in gQ_\lambda g^{-1}$, i.e $g^{-1}tg \in Q_\lambda$, car les deux entiers sont situés sur une même colonne de gT et donc $\forall g \notin P_\lambda Q_\lambda, n_g = 0$.

On procède par contraposée et suppose donc que tous les éléments d'une même ligne de T se retrouve dans des colonnes différentes de gT .

$\forall p \in P_\lambda$, le tableau de Young pT contient sur chacune de ses lignes les mêmes éléments que T mais dans un ordre potentiellement différent. De même, $\forall q \in Q_\lambda$, le tableau de Young qT contient les mêmes éléments que T sur chacune de ses colonnes mais dans un ordre qui peut être différent.

Comme tous les éléments d'une même ligne de T se retrouvent sur des colonnes différentes, en permutant convenablement sur les lignes de T d'une part et sur les colonnes de gT d'autre part, on peut créer deux tableaux de Young qui auront la même première ligne : pratiquement, on permute les éléments de gT , colonne par colonnes, pour ramener tous les éléments qui étaient initialement présents sur la première ligne de T sur la première ligne de " gT ". Puis, on permute les éléments de T au sein de cette même première ligne pour reconnaître la première ligne de notre nouveau tableau de Young.

On note p_1 l'élément de P_λ qui agit sur la première ligne de T et q'_1 l'élément de $gQ_\lambda g^{-1}$ qui agit sur chacune des colonnes de gT .

Par exemple, pour $n = 5$ et avec la partition $\lambda = (3, 2)$. On considère la permutation $g \in \mathfrak{S}_5$ définie par $g = (2\ 4\ 5\ 3)$. On est alors dans le cadre de notre démonstration :

$$\begin{array}{ccc}
 \boxed{1} & \boxed{2} & \boxed{3} \\
 \boxed{4} & \boxed{5} & \\
 \end{array}
 \qquad
 \begin{array}{ccc}
 \boxed{1} & \boxed{4} & \boxed{2} \\
 \boxed{5} & \boxed{3} & \\
 \end{array}$$

T
 gT

Soit $p_1 = (2\ 3) \in P_\lambda$ et $q'_1 = (3\ 4) \in Q_\lambda$. Alors :

$$\begin{array}{ccc} \boxed{1} & \boxed{3} & \boxed{2} \\ \boxed{4} & \boxed{5} & \\ p_1 T & & q'_1(gT) \end{array}$$

On procède de même sur le reste du tableau et finalement on construit deux tableaux de Young identiques : pT avec $p \in P_\lambda$ et $q'(gT)$ avec $q' \in gQ_\lambda g^{-1}$.

Dans notre exemple, il suffit de prendre $p = (2\ 3)(4\ 5)$ et $q' = (3\ 4)$.

Dans ce cas, on a en particulier que $p = q'g$. Or $q' \in gQ_\lambda g^{-1}$ donc il existe $q \in Q_\lambda$ tel que $q' = gqg^{-1}$. D'où :

$$p = q'g = (gqg^{-1})g = gq.$$

Et, comme $q \in Q_\lambda$, qui est un sous-groupe de \mathfrak{S}_n , on peut définir q^{-1} et $q^{-1} \in Q_\lambda$ et alors on a construit deux éléments $p \in P_\lambda$ et $\tilde{q} := q^{-1} \in Q_\lambda$ tels que $g = p\tilde{q}$, ce qui conclut la preuve. □

On ordonne à présent les partitions de l'entier n en les munissant de l'ordre lexicographique.

Définition II.15. Soit $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k)$ et $\mu = (\mu_1, \mu_2, \dots, \mu_l)$ deux partitions de n . On dit que $\lambda > \mu$ si et seulement si :

$$\exists j \in \{1, \dots, \min(k, l)\}, \begin{cases} \lambda_i = \mu_i \ \forall i < j \\ \lambda_j > \mu_j \end{cases}.$$

Lemme II.16.3. Soit $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k)$ et $\mu = (\mu_1, \mu_2, \dots, \mu_l)$ deux partitions de n .

- i. Si $\lambda > \mu$, alors pour tout $x \in \mathbb{C}\mathfrak{S}_n$, $a_\lambda \cdot x \cdot b_\mu = 0$.
En particulier, si $\lambda > \mu$, alors $c_\lambda \cdot c_\mu = 0$.
- ii. $\forall x \in \mathbb{C}\mathfrak{S}_n$, $c_\lambda \cdot x \cdot c_\lambda \in \mathbb{C}(c_\lambda)$.
En particulier, $\exists n_\lambda \in \mathbb{C}$, $c_\lambda \cdot c_\lambda = n_\lambda c_\lambda$.

Démonstration :

On note T_λ le tableau de Young associé à la partition λ et T_μ celui associé à la partition μ .

- i. Soit $x \in \mathbb{C}\mathfrak{S}_n$. Alors, il existe des éléments $n_g \in \mathbb{C}$ et $e_g \in \mathfrak{S}_n$ tels que $x = \sum_{g \in \mathfrak{S}_n} n_g e_g$, de sorte que :

$$a_\lambda \cdot x \cdot b_\mu = a_\lambda \cdot \left(\sum_{g \in \mathfrak{S}_n} n_g e_g \right) \cdot b_\mu = \sum_{g \in \mathfrak{S}_n} n_g a_\lambda e_g b_\mu.$$

En particulier, les λ_j éléments de la $j^{\text{ème}}$ ligne de T_λ se retrouvent dans des colonnes différentes de T_μ . Il existe donc au moins un élément qui se trouvait dans la $j^{\text{ème}}$ ligne de T_λ qui va se retrouver dans l'une des $j-1$ premières lignes de T_μ , les lignes suivantes n'ayant pas un nombre de colonnes suffisant... On note cet élément a .

On considère alors les λ_{j-1} éléments de la $(j-1)^{\text{ème}}$ ligne de T_λ . Si la ligne $j-1$ de T_μ contient l'élément a alors elle ne peut nécessairement pas contenir tous les éléments de la $(j-1)^{\text{ème}}$ ligne de T_λ et donc il y a un élément qui est dans l'une des $j-2$ premières lignes. Sinon, c'est que a se trouvait dans l'une des $j-2$ premières lignes. Dans tous les cas il y a un élément qui appartient à ces $(j-2)$ lignes et on le note a .

On remonte de même les deux tableaux de Young jusqu'à atteindre le premier indice tel que $\lambda_1 = \lambda_i = \mu_i = \mu_1$. On note k_1 le nombre de lignes qui satisfont à cette condition. Les éléments de ces k_1 premières lignes de T_λ ne peuvent appartenir qu'au k_1 première lignes de T_μ , les autres lignes du tableau T_μ ne contenant pas un nombre de colonnes suffisant. Mais, les k_1 premières lignes de T_μ contiennent déjà l'élément a : on cherche donc à faire entrer $k_1 \times \lambda_1$ données dans $k_1 \times \lambda_1 - 1$ cases, c'est impossible !

On applique ce résultat à $x = b_\lambda \cdot a_\mu$:

$$\begin{aligned} a_\lambda \cdot x \cdot b_\mu &= 0 \\ a_\lambda \cdot (b_\lambda \cdot a_\mu) \cdot b_\mu &= 0 \\ c_\lambda \cdot c_\mu &= 0. \end{aligned}$$

ii. On a le résultat en appliquant le *lemme 2*. Soit $x \in A$,

$$\begin{aligned} p \cdot [c_\lambda \cdot x \cdot c_\lambda] \cdot (\text{sgn}(q) q) &= p \cdot (a_\lambda \cdot b_\lambda) \cdot x \cdot (a_\lambda \cdot b_\lambda) \cdot (\text{sgn}(q) q) \\ &= (p \cdot a_\lambda) \cdot b_\lambda \cdot x \cdot a_\lambda \cdot (b_\lambda \cdot (\text{sgn}(q) q)) \\ &= a_\lambda \cdot b_\lambda \cdot x \cdot a_\lambda \cdot b_\lambda && \text{(i) et (ii)} \\ &= c_\lambda \cdot x \cdot c_\lambda \\ &\in \mathbb{C}(c_\lambda). && \text{(iii)} \end{aligned}$$

En particulier, si on applique ce qui précède à $x = Id$,

$$\exists n_\lambda \in \mathbb{C}, \quad c_\lambda \cdot c_\lambda = n_\lambda \cdot c_\lambda.$$

□

Lemme II.16.4. i. Pour toute partition λ de n , V_λ est une représentation irréductible de \mathfrak{S}_n .

Outils pour étudier les représentations du groupe \mathfrak{S}_n

ii. Si $\lambda \neq \mu$, alors V_λ et V_μ ne sont pas isomorphes.

Démonstration :

i. On a déjà vu que V_λ est une représentation de $\mathbb{C}\mathfrak{S}_n$: c'est en effet un sous-espace de $\mathbb{C}\mathfrak{S}_n$, invariant par l'action que définit la multiplication à droite par c_λ sur $\mathbb{C}\mathfrak{S}_n$.

Il reste donc à voir qu'elle est irréductible.

D'après le *lemme 3*, on a que

$$c_\lambda V_\lambda = c_\lambda \mathbb{C}\mathfrak{S}_n \cdot c_\lambda \subseteq \mathbb{C}(c_\lambda) \quad \text{car} \quad c_\lambda \cdot \lambda \subseteq \mathbb{C}(c_\lambda).$$

Soit W une sous-représentation de V_λ . W est en particulier un sous-espace de V_λ et donc $c_\lambda W$, un sous-espace de $c_\lambda V_\lambda$.

Or $c_\lambda V_\lambda \subseteq \mathbb{C}(c_\lambda)$ et donc

- soit $c_\lambda \in c_\lambda W$ et dans ce cas $c_\lambda W = \mathbb{C}(c_\lambda)$,
- soit $c_\lambda \notin c_\lambda W$ et alors, $c_\lambda W = \{0\}$.

Dans le *premier cas*, comme W est une sous-représentation de V_λ , alors $c_\lambda W \subseteq W$, i.e $\mathbb{C}(c_\lambda) \subseteq W$.

En particulier, on a donc que $c_\lambda \in W$ et on en déduit que

$$V_\lambda = \mathbb{C}\mathfrak{S}_n \cdot c_\lambda \subseteq W.$$

Par suite, on a en fait montré que $W = V_\lambda$.

Dans le *second cas*, on commence par remarquer que

$$\begin{aligned} W \cdot W &\subseteq V_\lambda \cdot W && (W \text{ sous représentation de } V_\lambda) \\ W \cdot W &\subseteq (\mathbb{C}\mathfrak{S}_n \cdot c_\lambda) \cdot W \\ W \cdot W &\subseteq (\mathbb{C}\mathfrak{S}_n) \cdot (c_\lambda W) \\ W \cdot W &\subseteq \{0\}. && (c_\lambda W = \{0\}) \end{aligned}$$

Mais, dans ce cas, on a alors que $W = 0$.

En effet : comme W est une sous représentation de V_λ , d'après le *théorème de complète réductibilité*,

$$\text{il existe } W' \text{ sous - représentation de } V_\lambda, \quad \mathbb{C}\mathfrak{S}_n = W \oplus W'.$$

On peut alors définir une projection de \mathfrak{S}_n dans W comme suit :

$$\begin{aligned} \mathbb{C}\mathfrak{S}_n &= W \oplus W' \rightarrow W \\ a &\longmapsto a\phi \end{aligned}$$

où ϕ est un élément de $\mathbb{C}\mathfrak{S}_n$ tel que $\phi = \phi^2 \in W \cdot W = \{0\}$. Un tel élément existe :

$$\begin{aligned}\mathbb{C}\mathfrak{S}_n &\cong \text{End}(\mathbb{C}\mathfrak{S}_n) \\ &= \text{Hom}(W \oplus W'; W \oplus W') \\ &\cong \begin{pmatrix} \text{Hom}(W, W) & \text{Hom}(W, W') \\ \text{Hom}(W', W) & \text{Hom}(W', W') \end{pmatrix}.\end{aligned}$$

Il suffit alors de choisir

$$\Phi = \begin{pmatrix} Id & 0 \\ 0 & 0 \end{pmatrix}.$$

Dans ce cas, $\Phi^2 = \Phi$.

- ii. On peut supposer, sans perte de généralité, que $\lambda > \mu$, ce qui nous permet d'appliquer le *lemme 3*. On sait donc d'après ce lemme que :

$$c_\lambda V_\mu = c_\lambda \mathbb{C}\mathfrak{S}_n \cdot c_\mu = 0 \quad \text{car} \quad c_\lambda \cdot c_\mu = \{0\}.$$

Mais, toujours par le même lemme,

$$c_\lambda V_\lambda = \mathbb{C}(c_\lambda) \neq \{0\}.$$

Et donc, V_λ et V_μ ne peuvent pas être isomorphes en tant que module sur $\mathbb{C}\mathfrak{S}_n$. □

On est donc maintenant en mesure de prouver le théorème qui faisait l'objet de cette section et que l'on rappelle ici pour mémoire :

Théorème II.16

L'ensemble des représentations irréductibles de \mathfrak{S}_n est exactement l'ensemble des V_λ où λ est une partition ordonnée de n .

Démonstration :

Le *lemme 4* nous assure que l'ensemble des V_λ , où λ est une partition de n , est inclus dans l'ensemble des représentations irréductibles de \mathfrak{S}_n .

Or on a construit par ce procédé autant de représentations irréductibles qu'il n'y a de classes de conjugaison de \mathfrak{S}_n , (à savoir qu'aussi bien les classes de conjugaisons de \mathfrak{S}_n que l'ensemble des V_λ , pour λ partition de n sont en bijection avec l'ensemble des partitions de n).

C'est donc qu'on les a toutes trouvées. □

Références

- [12] François Dubois, Hervé Lemeur, and Claude Reiss. Mathematical modeling of antigenicity for HIV dynamics. *MathS In Action*, 3(1) :1–35, 2010.
- [31] Pierre-Guy Plamondon. Représentations linéaires des groupes finis. Cours, 2014.
- [37] Jean-Pierre Serre. *Représentations linéaires des groupes finis*. Méthodes. Hermann, 5 edition, 1998.

– Cinquième partie –

Lecture d'article – M2
Autour du problème de sériation

Encadré par Filippo Santambrogio *

En binôme avec Pierre Roux

*. LMO, Université Paris-Sud

Sommaire – Partie V

Lecture d'article : Autour du problème de sériation	95
	100
I Le problème de sériation	100
I.1 De l'usage de la théorie spectrale pour l'optimisation	101
I.2 Une solution exacte au problème de sériation	102
I.2.a Notations	103
I.2.b Principe de la résolution	104
I.3 Une approche heuristique pour les problèmes mal posés	108
I.4 Compléments	109
II Programmation de l'algorithme	112
II.1 Un algorithme naïf	112
II.2 L'algorithme spectral	116
II.2.a Implémentation de l'algorithme en Python	116
II.2.b Résultats et complexité	123
II.3 Comparaison des performances	130
III Le problème C1P	131
Références	135

Le présent document est le compte rendu d'un travail effectué sur l'article *A spectral algorithm for seriation and the consecutive ones problem* de Jonathan E. ATKINS, Erik G. BOMAN et Bruce HENDRICKSON [3] dans le cadre du cours *Optimisation et Simulation numérique* du Master Mathématiques pour les Sciences du Vivant donné par Filippo Santambrogio, Estelle Kuhn et Sylvain Faure.

On y trouve une description du contenu de l'article et des implémentations en langage python des algorithmes qui y sont décrits. Enfin, on teste ces algorithmes sur des exemples issus de l'archéologie et de la phylogénie et on les compare à des algorithmes naïfs pour appréhender le gain de temps et leur complexité respective.

I. Le problème de sériation

Nous étudions le problème dit de sériation qui consiste, étant donné un ensemble fini $X = \{x_1, \dots, x_n\}$ muni d'une relation d'ordre \leq et une fonction de corrélation symétrique

$$f : X \times X \rightarrow \mathbb{R} \\ (x, y) \mapsto f(x, y) ,$$

à trouver toutes les permutations π de X telles que

$$\forall x_i, x_j, x_k \in X, \\ \pi(x_i) < \pi(x_j) < \pi(x_k) \implies f(x_i, x_j) \geq f(x_i, x_k) \wedge f(x_j, x_k) \geq f(x_i, x_k) .$$

Le problème de sériation correspond à de nombreuses problématiques, mathématiques (étude de graphes, réduction de l'enveloppe d'une matrice creuse) ou non (séquençage génétique, datation archéologique, anthropologie, sociologie, recherche opérationnelle, visualisation de données, *etc*). Le principal intérêt de la résolution du problème de sériation en dehors des mathématiques réside dans la visualisation des relations entre divers objets. En effet, étant donné des objets (l'ensemble X) et la force du lien qui unit chaque paire (i, j) d'objets (la corrélation $f(i, j)$), résoudre le problème de sériation permet de visualiser sur une ligne quels sont pour chaque objet i_0 les autres objets avec lesquels il a les liens (la corrélation) les plus forts. La conscience humaine étant particulièrement mauvaise pour identifier des règles dans un graphe possédant beaucoup d'arrêtes, par simple incapacité de prendre en compte toutes les arrêtes à la fois, présenter les données de manière linéaire, en ne gardant visible pour chaque objet que ses liens les plus forts, permet à l'observateur d'avoir l'intuition de la règle qui régit les données fournies.

Si de plus on sait *a priori* que les données ont une nature linéaire, la visualisation que permet la résolution du problème de sériation a même des chances de traduire directement la nature du jeu de données. Par exemple, si on considère des vases produits par la même civilisation antique à des époques différentes (voir figure 1), et si l'on suppose que les changements de forme pour les vases sont graduels, alors on peut associer à chaque paire de vases une corrélation d'autant plus forte que les vases se ressemblent et résoudre le problème de sériation correspondant. La solution sera une bonne approximation de la chronologie relative de la fabrication des vases.

Contrairement à d'autres méthodes utilisées en data mining, la sériation ne cause aucune déperdition d'information puisqu'elle ne modifie pas les données. Les méthodes de sériation peuvent être étendues à des problématiques de réorganisation de matrices, particulièrement utiles par exemple pour tirer des informations des matrices sociales (en sociologie on peut représenter les liens sociaux entre des individus par une matrice sociale dont le coefficient a_{ij} est la pondération du lien ressenti par i envers j). Le résultat de la sériation peut alors donner des indications robustes sur la dynamique sociale du groupe.

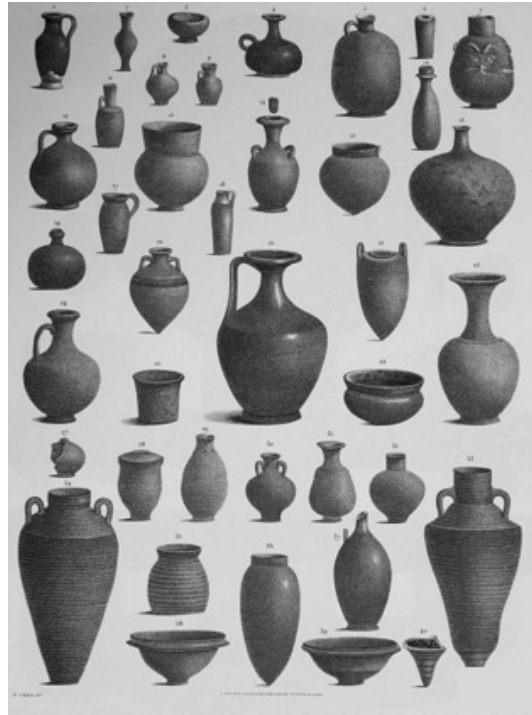


FIGURE 1. – Pots en argile égyptiens, publié en 1800, auteur inconnu.

1. De l'usage de la théorie spectrale pour l'optimisation

Les méthodes spectrales sont utiles dans de nombreux domaines du calcul scientifique et de l'informatique (résolution des équations aux dérivées partielles, modes propres du laplacien, acoustique, caractérisation des graphes bipartis, étude des graphes finis par leurs valeurs propres,...). Elles sont également liées aux problématiques de la programmation SDP (semi-définie positive) et donc à une grande variété de problèmes combinatoires. De nombreuses tentatives d'appliquer la théorie spectrale des graphes à des problèmes NP-complets ou NP-difficiles avaient déjà eu lieu, mais il s'agissait pour la plupart d'approches heuristiques ou d'algorithmes de majoration.

La piste des méthodes spectrales avait peu été explorée avant l'article que nous étudions ici. Comme nous allons le voir, un algorithme naïf pour notre problème se trouve affligé d'une complexité en $\mathcal{O}(n^3n!)$. Il est également difficile de savoir *a priori* si le problème posé admet une solution. Une méthode exacte présentant un coût en temps de calcul raisonnable n'est donc pas une chose commune dans l'utilisation de la théorie spectrale en optimisation combinatoire.

Autour du problème de sériation

L'intérêt immédiat de l'utilisation de la théorie spectrale pour résoudre un tel problème est que la méthode bénéficie à la fois des progrès réalisés dans les algorithmes de calcul matriciel et des propriétés de régularité que l'on trouve partout dans l'algèbre linéaire en dimension finie. L'algèbre linéaire regorge également d'un véritable arsenal de théorèmes d'existence, d'unicité, de décomposition (comme par exemple le théorème de Perron-Frobenius que nous utilisons par la suite).

2. Une solution exacte au problème de sériation

Précisons qu'une solution n'existe pas forcément, comme le montre l'exemple suivant : $X = \{a, b, c, d\}$, $f(a, b) = 100$, $f(a, d) = 40$, $f(b, d) = 30$, $f(a, c) = 20$, $f(c, d) = 15$, $f(b, c) = 10$ exprimé sous la forme d'un graphe sur la figure 2.

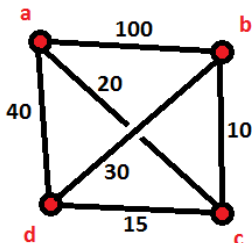


FIGURE 2. – Un problème de sériation sans solution

En effet, dans la solution finale, a et b , qui ont la plus grande corrélation, doivent être consécutifs après permutation (ab ou ba en adoptant la notation compacte sous-entendue). Mais ensuite, d "veut" fortement être proche de a , bien plus que de b , ce qui donne dab ou bad . Or,

- $cdab$ ne convient pas car c "préfère" a à d ;
- $dabc$ non plus car c préfère a à b ;
- $cbad$ non plus car c préfère a à b ;
- $badc$ non plus car c préfère a à d .

Pour résumer de manière informelle, le problème vient du fait que a préfère b et d , les met autour de lui, alors que c préfère a à tous les autres et est donc forcément insatisfait.

Par la suite, on réutilisera parfois cette terminologie informelle : " a préfère b à c " pour dire $f(a, b) > f(a, c)$.

En revanche, tous les problèmes à trois éléments peuvent être résolus. En effet, si $X = \{a, b, c\}$, on prend la plus grande valeur de la fonction de corrélation, disons $f(a, b)$,

la deuxième plus grande, disons $f(a, c)$, et l'arrangement cab convient forcément.

Notre méthode de résolution exacte doit donc être capable de discriminer de manière simple les cas où une solution existe de ceux où il n'en existe pas. La question d'une heuristique pour les cas sans solution sera étudiée dans un second temps.

a. Notations

Par la suite, l'ensemble X sera identifié à $\{1, \dots, n\}$ et on décrira la fonction de corrélation f par une matrice symétrique $A \in \mathcal{M}_n(\mathbb{R})$ dont les coefficients a_{ij} vérifient

$$\forall i, j \in \llbracket 1, n \rrbracket, \quad a_{ij} = f(i, j),$$

et dont les termes diagonaux sont arbitraires.

Le problème de sériation associé à A est dit résolu si A vérifie

$$\forall i, j, k \in \llbracket 1, n \rrbracket, \quad i < j < k \implies a_{ij} \geq a_{ik} \wedge a_{kj} \geq a_{ki}.$$

Une matrice vérifiant cette propriété est appelée *R-matrice*.

Le problème de sériation associé à A admet une solution s'il existe une permutation symétrique de A (c'est à dire une permutation sur X) telle que, en notant A^π la permutation symétrique de A par π , $A^\pi = (a_{\pi(i)\pi(j)})_{1 \leq i, j \leq n}$ est une R-matrice. Si une telle permutation existe, A est appelée *pré-R-matrice*.

Etant donnée une matrice A , on appelle *laplacien* de A et on note L_A la matrice $D_A - A$, où D_A est la matrice diagonale dont les coefficients diagonaux vérifient : $d_{ii} = \sum_{j=1}^n a_{ij}$.

Le laplacien de A est une matrice symétrique dont toutes les valeurs propres sont réelles et dont les sous-espaces propres sont deux à deux orthogonaux. Le vecteur dont toutes les coordonnées sont égales à 1, que nous noterons dorénavant e , est toujours vecteur propre de L_A pour la valeur propre 0, par construction.

La plus petite valeur propre de L_A associée à un vecteur propre orthogonal à e est appelée *valeur de Fiedler* de L_A (on dira par extension que c'est la valeur de Fiedler de A) et un vecteur propre orthogonal à e associé sera appelé *vecteur de Fiedler* de L_A (ou de A par extension).

On peut montrer que la valeur de Fiedler réalise le minimum suivant, et est réalisée pour un vecteur de Fiedler :

$$\min_{\substack{t_x e = 0, \\ t_x x = 1}} t_x L_A x$$

Autour du problème de sériation

où ${}^t x$ est la transposée du vecteur x .

Une matrice A est dite *réductible*, si il existe une permutation π telle que

$$A^\pi = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix},$$

et elle dite *irréductible* s'il n'existe pas de telle permutation.

On notera $Sp(M) = \{\lambda_1^{(i_1)}, \dots, \lambda_r^{(i_r)}\}$ le spectre d'une matrice M où les termes en exposant sont les multiplicités algébriques des valeurs propres. On écrira $M \geq 0$ le fait que tous les coefficients de la matrice M soient supérieurs ou égaux à 0 et $x \geq 0$ le fait que toutes les coordonnées du vecteur x soient supérieures ou égales à 0.

On dira que le vecteur x est monotone si ses coordonnées sont croissantes ou décroissantes, c'est à dire si x vérifie la propriété :

$$(\forall i, j \in \llbracket 1, n \rrbracket, i < j \implies x_i \leq x_j) \vee (\forall i, j \in \llbracket 1, n \rrbracket, i < j \implies x_i \geq x_j).$$

Étant donné une permutation π de $\llbracket 1, n \rrbracket$, notée $\pi = (i_1, \dots, i_n)$ si elle envoie 1 sur i_1, \dots, n sur i_n , on appelle *permutation miroir* de π la permutation $\nu = (i_n, \dots, i_1)$.

Si J est un sous-ensemble de $\llbracket 1, n \rrbracket$, et si $A \in \mathcal{M}_n(\mathbb{R})$, on notera $A(J, J)$ la matrice obtenue en retirant de A toutes les lignes et toutes les colonnes dont les indices ne sont pas dans J .

b. Principe de la résolution

Forts de ces notations et définitions, nous pouvons énoncer les théorèmes qui permettent de trouver la solution au problème de sériation à partir du vecteur de Fiedler de la matrice de corrélation.

Tout ce que nous allons faire par la suite repose sur un résultat classique d'algèbre linéaire, le théorème de Perron-Frobenius :

Théorème 17 (Perron-Frobenius)

Soit M une matrice à coefficients positifs de rayon spectral

$$\rho(M) = \max_{\lambda_i \in Sp(M)} |\lambda_i|.$$

Alors $\rho(M)$ est valeur propre et il existe un vecteur $x \geq 0$ tel que

$$Mx = \rho(M)x.$$

Le théorème de Perron-Frobenius permet la démonstration du théorème suivant :

Théorème 18

Toute R-matrice A admet un vecteur de Fiedler monotone.

La preuve de ce résultat consiste à remarquer que $x \in \mathbb{R}^n$ est un vecteur de Fiedler de A si et seulement si le vecteur $y = {}^t(x_2 - x_1, \dots, x_n - x_{n-1}) \in \mathbb{R}^{n-1}$ est vecteur propre, pour la plus grande valeur propre, d'une matrice $\tilde{M}_A \geq 0$ bien choisie.

On peut alors appliquer le théorème de Perron-Frobenius à \tilde{M}_A et obtenir un vecteur propre $y \geq 0$ de \tilde{M}_A pour la plus grande valeur propre, ce qui revient à trouver un vecteur de Fiedler monotone.

Les principales difficultés sont de construire la matrice \tilde{M}_A et de vérifier la correspondance entre ses vecteurs propres et ceux de L_A .

On tire ensuite de ce théorème une méthode de résolution des problèmes bien posés. Cependant, des conditions drastiques font que le résultat suivant est difficilement applicable en l'état. D'autres théorèmes seront nécessaires pour contourner ces difficultés, et l'algorithme final exploitera à la fois le résultat limité et les contournements possibles.

Théorème 19

Soit A une pré-R-matrice dont la valeur de Fiedler est une valeur propre simple et dont le vecteur de Fiedler n'a pas deux valeurs répétées dans ses coordonnées (*i.e.* $i < j \implies x_i \neq x_j$). Soit π_1 (respectivement π_2) la permutation qui, appliquée à x , en fait un vecteur croissant (respectivement décroissant).

Alors A^{π_1} et A^{π_2} sont des R-matrices et il n'existe pas d'autre permutation π telle que A^π soit une R-matrice.

Les hypothèses nécessaires pour résoudre un problème bien posé avec ce théorème, qui sont en passant une condition suffisante pour le fait d'être bien posé, sont le caractère simple de la valeur de Fiedler et la non répétition des coordonnées du vecteur de Fiedler. Notre objectif est à présent de comprendre ce qui se passe dans les autres cas en analysant les conséquences d'une répétition de coordonnées dans le vecteur de Fiedler.

Pour commencer, le lemme suivant montre qu'on peut se ramener au cas d'une matrice A positive ayant 0 pour plus petit coefficient hors de la diagonale :

Lemme 4.1. Soit A une matrice symétrique, α un réel et $\bar{A} = A - \alpha e {}^t e$. Alors x est un vecteur de Fiedler de A si et seulement si x est un vecteur de Fiedler de \bar{A} .

L'étape suivante est de décrire ce qui arrive à une pré-R-matrice lorsque ses blocs

Autour du problème de sériation

irréductibles sont soumis à des permutations, afin de traiter ensuite spécifiquement le cas d'un bloc irréductible.

Lemme 4.2. Soit A une pré-R-matrice, dont les blocs irréductibles sont notés A_i , $i = 1, \dots, k$. Soit, pour tout $i \in \llbracket 1, k \rrbracket$, π_i une permutation du bloc A_i qui en fait une R-matrice. Alors la permutation π obtenue en concaténant dans n'importe quel ordre les permutations π_i fait de A une R-matrice.

On peut donc bien traiter les blocs irréductibles indépendamment. En terme d'archéologie, cela revient à dire que des groupes d'objets tels qu'il n'y a aucune corrélation identifiable entre des éléments de groupes différents ne peuvent pas être ordonnés chronologiquement, mais que les éléments peuvent être ordonnés à l'intérieur des groupes s'ils ont des corrélations non nulles. La datation relative exige une certaine continuité des données, qui s'assimile mathématiquement à la propriété d'irréductibilité.

On étudie enfin ce qu'impliquent pour la matrice A les répétitions dans un vecteur de Fiedler :

Lemme 4.3. Soit $A \in \mathcal{M}_n(\mathbb{R})$ une R-matrice et x un vecteur de Fiedler monotone associé. Supposons que $J = \llbracket r, s \rrbracket$ est un intervalle maximal tel que $x_r = x_s$. Alors,

$$\forall k \notin J, a_{r,k} = a_{r+1,k} = \dots = a_{s,k}.$$

Et on énonce une forme de réciproque pour les matrices irréductibles, ce qui commence de justifier notre travail précédent sur la décomposition en blocs irréductibles :

Lemme 4.4. Soit $A \in \mathcal{M}_n(\mathbb{R})$ une R-matrice irréductible telle que $a_{n,1} = 0$. Supposons que $J = \llbracket r, s \rrbracket$ est un intervalle différent de $\llbracket 1, n \rrbracket$ tel que pour tout $k \notin J$, $a_{rk} = a_{sk}$. Alors, pour tout vecteur de Fiedler x ,

$$x_r = x_{r+1} = \dots = x_s.$$

Et en combinant les deux derniers lemmes, on obtient de manière presque immédiate :

Lemme 4.5. Soit $A \in \mathcal{M}_n(\mathbb{R})$ une R-matrice irréductible telle que $a_{n,1} = 0$ et soit x un vecteur de Fiedler monotone de A . Supposons que $J = \llbracket r, s \rrbracket$ est un intervalle tel que $x_r = x_{r+1} = \dots = x_s$. Alors, pour tout autre vecteur de Fiedler y de A ,

$$y_r = y_{r+1} = \dots = y_s.$$

Ce dernier lemme permet la démonstration d'un théorème vraiment utile pour la résolution de notre problème de sériation :

Théorème 20

Soit $A \in \mathcal{M}_n(\mathbb{R})$ une R-matrice irréductible telle que $a_{n,1} = 0$, alors sa valeur de Fiedler est une valeur propre simple.

On peut donc traiter les blocs irréductibles de A séparément et combiner les permutations obtenues dans un ordre quelconque, sous réserve que les vecteurs de Fiedler, uniques à multiplication par une constante près, soient sans répétition de coordonnées. Ce dernier écueil nécessite un résultat supplémentaire, que voici :

Théorème 21

Soit A une pré-R-matrice dont la valeur de Fiedler est simple, soit x un vecteur de Fiedler de A . Supposons qu'il y a une valeur β qui est présente plus d'une fois dans les coordonnées de x et soit I, J et K des intervalles vérifiant :

- $\forall i \in I, x_i < \beta,$
- $\forall j \in J, x_j = \beta,$
- $\forall k \in K, x_k > \beta.$

Alors π est une permutation telle que A^π est une R-matrice si et seulement si π ou la permutation miroir de π peut s'écrire (π_i, π_j, π_k) , où π_j est une permutation telle que $A(J, J)^{\pi_j}$ est une R-matrice, π_i est la restriction à $A(I, I)$ d'une permutation qui fait de A une R-matrice et π_k est la restriction à $A(K, K)$ d'une permutation qui fait de A une R-matrice.

Ce dernier théorème dit simplement que si des valeurs sont répétées dans le vecteur de Fiedler, il suffit de résoudre le problème en se restreignant successivement à chaque valeur présente dans ledit vecteur. S'il n'y a pas de valeur répétée, cela revient à ordonner le vecteur de Fiedler, puis se restreindre n fois à un singleton, c'est à dire à ordonner la matrice en suivant l'ordre du vecteur de Fiedler, terme par terme. C'est ce qui motive l'algorithme présenté dans l'article (et codé en python par nos soins!).

Remarquons que l'on a trouvé un moyen de déterminer si une matrice est pré-R. Il suffit de la décomposer en blocs irréductibles, et de tester sur chaque bloc les conclusions des théorèmes ci-dessus. Un moyen plus réaliste est d'implémenter l'algorithme présenté dans l'article, de générer une permutation et de la tester pour voir si la matrice devient une R-matrice.

Si la permutation générée ne marche pas, on est en droit de se poser des questions sur la qualité de la réorganisation que nous avons obtenue, ce que la sous-section ?? suivante explore.

3. Une approche heuristique pour les problèmes mal posés

La procédure algorithmique qui se dégage de la théorie résumée en sous-sous-section I.2.b et dont nous décrivons l'implémentation plus bas est exacte pour les problèmes bien posés, mais nous allons voir qu'elle n'est pas mauvaise pour les problèmes mal posés.

La première question à se poser est : qu'est-ce qu'une bonne solution à un problème mal posé ? La réponse informelle qui nous vient est qu'il faut que lorsque

$$\left(\pi(x_i) < \pi(x_j) < \pi(x_k)\right) \wedge \left(f(x_i, x_j) \leq f(x_i, x_k)\right),$$

ce qui est une entorse à la définition de solution, la quantité $f(x_i, x_j)$ doit avoir le moins d'écart possible avec $f(x_i, x_k)$. Comme il faut tenir compte simultanément de tous les triplets (i, j, k) pour lesquels cela pourrait "mal se passer" avec notre solution heuristique, nous introduisons la fonction de pénalité

$$\begin{aligned} g : \mathfrak{S}(X) &\rightarrow \mathbb{R} \\ \pi &\mapsto \sum_{i=1}^n \sum_{j=1}^n f(x_i, x_j) (\pi(x_i) - \pi(x_j))^2 \end{aligned}$$

qui donne un poids d'autant plus grand au terme (i, j) , pour un même $f(x_i, x_j)$, que $|\pi(i) - \pi(j)|$ est grand. La fonction g est donc minimale pour une permutation qui a tendance à rapprocher les éléments qui ont une forte corrélation, sans le faire au complet détriment de leurs corrélations avec des tiers.

Le problème, c'est qu'il n'est pas raisonnable de choisir la permutation en minimisant g naïvement, car cela requiert $n!$ étapes comprenant chacune $2n^2$ multiplications, soit une complexité en multiplication équivalente à $2n^2n!$, sans compter la procédure de génération de l'ensemble des permutations, très coûteuse elle aussi.

Une solution plus viable consiste à en faire un problème continu dans un espace de dimension n en transformant la fonction discrète g en la fonction continue

$$\begin{aligned} h : \mathbb{R}^n &\rightarrow \mathbb{R} \\ y &\mapsto \sum_{i=1}^n \sum_{j=1}^n f(x_i, x_j) (y_i - y_j)^2 \end{aligned}$$

qui a le bon goût d'être de classe \mathcal{C}^∞ et convexe. On peut donc minimiser h en un temps raisonnable à l'aide des techniques classiques de l'optimisation convexe.

Cependant, son minimum n'a aucune chance d'être atteint en un unique point puisque ajouter une constante à toutes les coordonnées de y ne change pas la valeur de h . Cette valeur minimale n'a par ailleurs aucun intérêt puisqu'il s'agit de 0, atteint pour le vecteur

nul $y = 0$. Nous devons donc ajouter des contraintes afin d'avoir un minimum qui a une signification et d'essayer d'obtenir l'unicité.

La contrainte $\sum_{i=1}^n y_i^2 = 1$ permet d'éviter les solutions de petite norme (ou de norme nulle...) qui baissent artificiellement la valeur de h en réduisant simultanément tous les y_i . Elle se reformule en ${}^t y y = 1$.

La contrainte $\sum_{i=1}^n y_i = 0$ empêche l'ajout d'une constante à y . Elle se reformule en ${}^t x e = 0$.

On remarque que si on associe la matrice $F = (f_{ij} = f(x_i, x_j))_{i,j=1,\dots,n}$ à f ,

$$h(y) = \sum_{i=1}^n \sum_{j=1}^n f_{ij} (y_i^2 - 2y_i y_j + y_j^2) = \sum_{i=1}^n \sum_{j=1}^n (y_i f_{ij} (y_i - y_j) + y_j f_{ij} (y_j - y_i)) =: \theta.$$

Or,

$$\begin{aligned} \theta &= \sum_{i,j=1}^n y_i f_{ij} (y_i - y_j) + \sum_{i,j=1}^n y_j f_{ij} (y_j - y_i) \\ &= 2 \sum_{i,j=1}^n y_i f_{ij} (y_i - y_j) \\ &= \left\langle y \mid {}^t (f_{12}(y_1 - y_2) + \dots + f_{1n}(y_1 - y_n), \dots, f_{n1}(y_n - y_2) + \dots + f_{n,n-1}(y_n - y_{n-1})) \right\rangle \\ &= {}^t y L_F y. \end{aligned}$$

La minimisation de h sous les contraintes évoquées revient donc à trouver un vecteur de Fiedler de F , ce qui implique que notre approche exacte donne de bons résultats pour les problèmes mal posés.

4. Compléments

Nous ajoutons ici des théorèmes et ou des preuves qui ne sont pas faites dans l'article afin d'éclairer certains concepts. Nous commençons par rappeler sans démonstration le théorème dit de Gershgorin afin de démontrer une propriété qui joue un rôle crucial dans la manière dont nous implémentons l'algorithme :

Théorème 22 (Gershgorin)

Soit $A \in \mathcal{M}_n(\mathbb{R})$, pour tout $i \in \llbracket 1, n \rrbracket$, on définit le $i^{\text{ième}}$ disque de Gershgorin par

$$\mathcal{D}_i = \mathcal{B}(a_{ii}, \sum_{j \neq i} |a_{ij}|)$$

où $\mathcal{B}(x, r)$ est la boule de centre x et de rayon r dans le plan complexe. Alors,

Autour du problème de sériation

pour toute valeur propre λ de A , on a

$$\lambda \in \bigcup_{i=1}^n \mathcal{D}_i.$$

Théorème 23

Si A est une R-matrice irréductible, non nulle, dont tous les coefficients sont positifs et telle que $a_{n,1} = 0$, alors sa valeur de Fiedler est la plus petite valeur propre non nulle de L_A .

Démonstration :

Les coefficients de A étant tous positifs, les termes diagonaux de L_A sont positifs et égaux en valeur absolue à la somme des valeurs absolues des coefficients non diagonaux présents sur la même ligne. Donc, les disques de Gershgorin associés à L_A ne contiennent jamais de nombre complexe ayant une partie réelle strictement négative. Il s'ensuit que le théorème de Gershgorin nous assure que les valeurs propres de L_A , réelles au demeurant, ne peuvent pas être strictement négatives.

D'après le théorème 20, la valeur de Fiedler est simple. Si elle est égale à 0, alors, puisque $e = {}^t(1, \dots, 1)$ vérifie $L_A e = 0$, e est à multiplication par une constante près le seul vecteur propre associé à la valeur de Fiedler, ce qui est impossible par définition de cette dernière.

La valeur de Fiedler est donc non nulle, c'est à dire strictement positive. C'est donc la plus petite valeur propre strictement positive de L_A . \square

Nous démontrons également le théorème de Perron-Frobenius, dont nous rappelons préalablement l'énoncé.

Théorème 24 (Perron-Frobenius)

Soit M une matrice à coefficients positifs de rayon spectral

$$\rho(M) = \max_{\lambda_i \in Sp(M)} |\lambda_i|.$$

Alors $\rho(M)$ est valeur propre et il existe un vecteur $x \geq 0$ tel que

$$Mx = \rho(M)x.$$

Démonstration :

Dans toute la démonstration, on notera, pour des vecteurs x et y , $x \leq y$ le fait que y soit supérieur terme à terme à x et $|x|$ le vecteur obtenu en remplaçant chaque coordonnée de x par sa valeur absolue (ou son module si x est complexe). On notera également $\|x\|_\infty = \max_{i=1,\dots,n} |x_i|$ la norme infinie de x .

Soit $C = \{x \in (\mathbb{R}_+)^n \mid \|x\|_\infty = 1, \rho(M)x \leq Ax\}$.

- C est convexe : soit $x, y \in C$, et soit $t \in [0, 1]$; comme x et y sont à coordonnées positives,

$$\|tx + (1-t)y\|_\infty = t\|x\|_\infty + (1-t)\|y\|_\infty = t + (1-t) = 1 ;$$

de plus,

$$\rho(M)(tx + (1-t)y) = t\rho(M)x + (1-t)\rho(M)y \leq tAx + (1-t)Ay = A(tx + (1-t)y),$$

donc $tx + (1-t)y \in C$.

- C est compact : $C \subset [0, 1]^n$ donc C est borné ; C est fermé comme image réciproque d'un fermé par une application continue ; donc, comme on est en dimension finie, C est compact.
- C est non vide : soit λ une valeur propre de A telle que $|\lambda| = \rho(M)$ et soit x un vecteur propre associé tel que $\|x\|_\infty = 1$; on a $|Ax| = |\lambda x| = \rho(M)|x|$, et par inégalité triangulaire sur chaque coordonnée, utilisant au passage la positivité des coefficients de A , on a $|Ax| \leq A|x|$; on a donc

$$\rho(M)|x| \leq A|x|,$$

et cela permet d'affirmer que $|x| \in C$.

Supposons qu'il existe $x \in C$ tel que $Ax = 0$, alors $\rho x \leq Ax = 0$, et comme x est non nul, il vient $\rho = 0$. Dans ce cas $A = 0$ et le théorème est évident.

Supposons à présent qu'il n'existe pas de tel x .

Soit la fonction

$$f : C \rightarrow C \\ x \mapsto \frac{Ax}{\|Ax\|_\infty}$$

f est bien définie car si $\rho(M)x \leq Ax$, alors $\rho(M)\frac{x}{\|Ax\|_\infty} \leq \frac{Ax}{\|Ax\|_\infty}$ et par linéarité

$$\rho(M) \left(\frac{Ax}{\|Ax\|_\infty} \right) \leq A \left(\frac{Ax}{\|Ax\|_\infty} \right),$$

d'où $f(x) \in C$. f est une fonction continue sur un compact convexe de \mathbb{R}^n à valeur dans ce même compact convexe, on peut donc lui appliquer le théorème du point fixe de Brouwer : il existe $x \in C$ tel que $f(x) = x$. Donc

$$\exists x \in (\mathbb{R}_+)^n, \quad (Ax = \|Ax\|_\infty x \wedge Ax \geq \rho(M)x),$$

ce dont on conclut à la fois que $\|Ax\|_\infty$ est une valeur propre de A admettant un vecteur propre à coordonnées positives et que $\rho(M) = \|Ax\|_\infty$.

□

II. Programmation de l'algorithme

Dans cette section, nous présentons les implémentations que nous avons faites à partir de l'article. Tous les programmes sont codés en langage Python et exécutés dans l'environnement de développement Pyzo (<http://www.pyzo.org/>).

1. Un algorithme naïf

Présentation du programme. Afin de mieux mesurer l'avantage de l'utilisation de la théorie spectrale, nous avons programmé un algorithme naïf, brutalement combinatoire, afin de faire des comparaisons de temps de calcul et de corroborer les résultats de l'autre programme dans les cas simples. Le programme utilise le package *numpy* naturellement présent dans la distribution Pyzo.

```
1 | import numpy as np
```

Puis on utilise une fonction `Permutations` qui prend en entrée une liste X et rend en sortie une liste P contenant toutes les versions permutées de la liste X . Comme le cardinal de \mathfrak{S}_n , le groupe symétrique d'ordre n , vaut $n!$, appeler cette fonction a déjà un fort coût en espace de stockage, et on verra qu'il en va de même pour le nombre d'opérations à effectuer.

```
1 | def Permutations(X):
2 |     P=[]
3 |     for i in range(len(X)): # On choisit le premier element de la
4 |         permutation
5 |         L=X[:]
6 |         p=[L[i]]
7 |         L.remove(L[i])
8 |         Q=[]
9 |         if L != []:
10 |             Q=Permutations(L) # On trouve toutes les permutations des
11 |                 elements restants
12 |         else:
13 |             P.append(p)
14 |             for q in Q:
15 |                 P.append(p+q) # On obtient (n-1)! permutations et on les
16 |                     stocke
17 |     return P
```


Le procédé est simple : on fait une boucle sur X afin de parcourir tous les choix possibles pour le premier élément de la liste permutée, puis on appelle récursivement la fonction `Permutations` sur X privé de l'élément qu'on a choisi pour figurer en premier dans la liste permutée. À chaque étape de la boucle, on obtient, pour X de taille n , $(n - 1)!$ permutations; boucle comprenant n étapes, on obtient les $n!$ permutations de X . Un test vérifie à chaque étape si la liste est vide afin de savoir quand arrêter l'appel récursif.

Cette fonction nous permet de résoudre facilement le problème de sériation, grâce à la fonction `Seriations` suivante :

```

1 def Seriations(X,F):
2     N=[]
3     P=Permutations(X)      # On genere toutes les permutations et on les
        stocke
4     for p in P:
5         Bool=True
6         for i in range(len(X)-2):
7             for j in range (i+1,len(X)-1):
8                 for k in range(j+1,len(X)):
9                     if (F[ p[i] , p[j] ] < F[ p[i] , p[k] ]) :      # test
                            de la premiere hypothese
10                        Bool = False
11                    if (F[ p[j] , p[k] ] < F[ p[i] , p[k] ]) :      # test
                            de la deuxieme hypothese
12                        Bool=False
13                if Bool:
14                    N.append(p)      # On stocke la permutation si elle est bonne
15    return N

```

Ce programme prend en entrée une liste d'objets X et la matrice de corrélation F correspondant au problème de sériation que l'on veut résoudre sur X et il rend en sortie une liste N qui contient toutes les permutations qui résolvent le problème de sériation défini par (X, F) . Pour cela, on stocke dans une liste P toutes les permutations de X possibles, obtenues en appelant la fonction `Permutations`, puis on teste une par une les $n!$ permutations de la manière la plus brutale qui soit : en testant les propriétés voulues sur chaque triplet $(i, j, k) \in \llbracket 0, n - 1 \rrbracket$ vérifiant $i < j < k$.

Complexité. Pour la sous-fonction `Permutations`, nous évaluons la complexité en temps avec le test $L \neq []$ pour opération élémentaire, car chaque autre opération est en nombre proportionnel au nombre de ces tests.

Pour une entrée de taille n , on fait d'abord n tests aux n étapes de la boucle "pour", et on fait n appels récursifs, ce qui correspond, en notant C_n le nombre de tests pour un appel de `Permutations` sur une entrée de taille n , à la formule

$$C_n = n + n \times C_{n-1},$$

Autour du problème de sériation

ce qui donne

$$C_n = n + n(n-1) + n(n-1)(n-2) + \dots + n(n-1)(n-2) \dots 5.4.3 + n!.$$

La complexité de **Permutations** en tests est donc équivalente à

$$\sum_{k=1}^{n-1} \frac{n!}{k!}.$$

La complexité en espace est tout aussi énorme, en $\mathcal{O}(n!)$ pour une entrée de taille n .

On garde le test comme opération élémentaire pour le calcul de la complexité en temps de la fonction **Sériations**. Cette dernière appelle une fois la fonction **Permutations** puis, pour chacune des $n!$ permutations (entrée de taille n), effectue trois boucles imbriquées, la troisième conduisant à faire deux tests. La première boucle, sur i , comprend $n-2$ étapes; la deuxième boucle, sur j , fait $n-1-i$ étapes; la troisième boucle fait $n-j$ étapes, ce qui donne pour la complexité C de **Sériations** :

$$C = \sum_{k=1}^{n-1} \frac{n!}{k!} + n! \sum_{i=1}^{n-2} \sum_{j=1}^{n-i-1} 2(n-j).$$

Or,

$$\begin{aligned} \sum_{i=1}^{n-2} \sum_{j=1}^{n-i-1} 2(n-j) &= 2n \sum_{i=1}^{n-2} \sum_{j=1}^{n-i-1} 1 - 2 \sum_{i=1}^{n-2} \sum_{j=1}^{n-i-1} j \\ &= 2n \sum_{i=1}^{n-2} (n-i-1) - 2 \sum_{i=1}^{n-2} \frac{(n-i-1)(n-i)}{2} \\ &= 2n(n-1) \sum_{i=1}^{n-2} 1 - 2n \sum_{i=1}^{n-2} i - 2 \sum_{i=1}^{n-2} \frac{n^2 - n - i(2n+1) + i^2}{2} \\ &= n(n-1)(n-2) - 2 \sum_{i=1}^{n-2} \frac{n^2 - n - i(2n-1) + i^2}{2} \\ &= n(n-1)(n-2) - (n^2 - n)(n-2) + (2n-1) \sum_{i=1}^{n-2} i - \sum_{i=1}^{n-2} i^2 \\ &= (2n-1) \frac{(n-2)(n-1)}{2} - \sum_{i=1}^{n-2} i^2 \\ &= (2n-1) \frac{(n-2)(n-1)}{2} - \frac{(n-2)(n-2+1)(2(n-2)+1)}{6} \\ &= \frac{1}{2}(2n-1)(n-2)(n-1) - \frac{1}{6}(n-2)(n-1)(2n-3) \\ &= (n-2)(n-1) \frac{2}{3}n \\ &= \frac{2}{3}n^3 - 2n^2 + \frac{4}{3}n. \end{aligned}$$

La complexité en tests de l'algorithme naïf est donc

$$C = \frac{2}{3}n^3.n! - 2n^2.n! + \frac{4}{3}n.n! + \sum_{k=1}^{n-1} \frac{n!}{k!}$$

soit une complexité équivalente à $\frac{2}{3}n!.n^3$.

Ce serait un euphémisme de dire que le temps de calcul risque d'être long pour toute entrée de taille supérieure à 10. La complexité en espace devient également très vite prohibitive puisque le programme utilise d'une part une liste contenant $n!$ éléments et rend d'autre part une liste qui peut contenir jusqu'à $n!$ éléments (cas d'un problème de sériation où $f(x_i, x_j)$ vaut toujours 0).

Tests de vérification. Maintenant que nous savons que notre algorithme naïf est inapplicable pour des moyennes ou grandes entrées, vérifions qu'il est au moins capable de donner la bonne réponse pour de petites entrées, afin de pouvoir contrôler les résultats de l'algorithme décrit dans l'article.

Nous testons d'abord sur une matrice 3×3 simple, celle dont les solutions ont été décrites formellement partie 1 :

$$\begin{pmatrix} 1 & 10 & 5 \\ 10 & 1 & 2 \\ 5 & 2 & 1 \end{pmatrix}.$$

Ici, $X=\{0,1,2\}$ pour coller à la numérotation de python. Nous lançons donc `Seriations` sur l'entrée suivante :

```
1 | X=[0,1,2]
2 | F=np.array( [ [1,10,5], [10,1,2], [5,2,1] ] )
3 |
4 | print( Seriations(X,F) )
```

et le programme rend la bonne réponse (calculée en partie 1), c'est à dire :

$$[[1, 0, 2], [2, 0, 1]] .$$

Nous vérifions ensuite que le programme fonctionne sur le cas dont nous avons montré précédemment qu'il n'admettait pas de solution ; son graphe est représenté partie 1, sur la figure 2, et sa matrice est :

$$\begin{pmatrix} 1 & 100 & 20 & 40 \\ 100 & 1 & 10 & 30 \\ 20 & 10 & 1 & 15 \\ 40 & 30 & 15 & 1 \end{pmatrix}.$$

Nous entrons le code

Autour du problème de sériation

```
1 | X=[0,1,2,3]
2 | F=np.array( [ [1,100,20,40], [100,1,10,30], [20,10,1,15], [40,30,15,1] ] )
3 |
4 | print( Seriations(X,F) )
```

et le programme rend [], ce qui est plutôt encourageant.

Nous vérifions enfin que pour un problème où toutes les corrélations sont nulles, c'est à dire une matrice identité, le problème rend l'ensemble des permutations. Plutôt que d'afficher la sortie du programme, on peut vérifier le résultat en affichant le nombre d'éléments de la sortie afin de s'assurer qu'on obtient $n!$ permutations :

```
1 | X=[0,1,2,3,4,5]
2 | F=np.eye(len(X)) # matrice identite de taille (longueur de X) x (
   |     longueur de X)
3 |
4 | print( len( Seriations(X,F) ) )
```

Le programme rend 720, qui a le bon goût d'être égal à $6!$.

2. L'algorithme spectral

Dans cette partie nous nous proposons d'implémenter en Python – et en utilisant librement les packages *numpy* et *scipy.linalg* – l'algorithme proposé par l'article. Pour ce faire, nous développons d'abord un certain nombre de programmes auxiliaires dont nous détaillons au paragraphe suivant les entrées, sorties et étapes principales avant de présenter l'algorithme dit "Spectral-Sort".

Dans un second paragraphe, nous présentons quelques résultats ainsi que la complexité de notre algorithme.

a. Implémentation de l'algorithme en Python

Préliminaires. On commence par importer les bibliothèques *numpy* et *scipy.linalg* comme annoncé en préambule

```
1 | import numpy as np
2 | import scipy.linalg as slin
```

et on définit une fonction `reverse` qui va nous servir par la suite. Elle sert, étant donnée une liste U quelconque de taille n , en imaginant que l'on soit capable d'effectuer des opérations sur la liste $ind = \llbracket 0, n-1 \rrbracket$ ¹, à transposer les opérations effectuées sur la liste ind à la liste U . Par exemple, si on est capable de permuter la liste ind alors `reverse` induit la même permutation à U .

```
1 | def reverse(ind,U) : # U liste a modifier selon la liste de reference ind
2 |     res = []
```

1. On adopte ici la numérotation $0, 1, \dots, n-1$ pour les n premiers entiers afin de coller à Python.

```

3 |     for u in U :
4 |         res.append(ind[u])
5 |     return res

```

Enfin, on définit une fonction `distinct` qui prend en argument un vecteur x , le parcourt, repère les valeurs répétées et stocke les valeurs correspondantes dans une liste de listes V . Par exemple, si $x = [0\ 6\ 4\ 2\ 9\ 9\ 2\ 6\ 3]$ alors, `distinct` renvoie `[[0], [3,6], [8], [2], [1,7], [4,5]]`; ce qu'il faut comprendre comme : il n'y a qu'un seul élément semblable au 0^{ème}, et c'est le plus petit, les éléments en position 3 et 6 sont identiques (ils valent ici 2) et sont les 2^{ème} plus petits, *etc.* Pour prendre en compte les erreurs d'arrondis dues à la mémoire finie de la machine, on utilise comme test `np.allclose` qui rend `true` dès lors que les deux grandeurs sont égales, plus ou moins un ε de l'ordre de l'erreur d'arrondi. Typiquement, on a envie de dire que $10^{-18} = 0 \dots$

```

1 | def distinct(x,U,n) : # n : taille du vecteur x
2 |     ind = np.argsort(x,kind='mergesort')
3 |     y = x[ind]
4 |
5 |     V, v = [], [ind[0]] # indice valeurs distinctes, ordre croissant.
6 |     for i in range(1,n) :
7 |         if np.allclose(y[i],x[v[-1]]) :
8 |             v.append(ind[i])
9 |         else :
10 |            V.append(reverse(U,v))
11 |            v = [ind[i]]
12 |        V.append(reverse(U,v))
13 |    return (V)

```

La seule subtilité réside dans le fait que, du fait de l'indexation des matrices/vecteurs par Python, i ne peut excéder la taille n du vecteur x et donc, `distinct` ne peut *a priori* retourner que des listes d'indices $i \leq n - 1$. Or, on sera amené lors de la mise en œuvre de notre algorithme final `Spectral_Sort` à appliquer `distinct` à des sous-vecteurs d'un vecteur de Fiedler de notre matrice de corrélation et on aimerait que `distinct` rende pour indices ceux correspondant au vecteur initial et non au sous-vecteur. Pour cela, on "ruse" en travaillant avec la liste `[[0, n - 1]]` puis en transposant le résultat à notre liste d'intérêt *via* la fonction `reverse` prévue à cet effet (lignes 10 et 12).

Mise sous forme irréductible. La mise sous forme de blocs d'irréductibles pour une matrice A donnée est assurée le code suivant :

```

1 | def irred(A,ll,ind,n) : # ll = liste indices a considerer
2 |     l = ll.copy()
3 |     if l == [] :
4 |         return []
5 |
6 |     u = [l[0]]
7 |     for i in u :

```

Autour du problème de sériation

```

8 |         if l != [] :
9 |             l.remove(i)
10 |         for j in l :
11 |             if not np.allclose(A[i,j],0) and j not in u :
12 |                 u.append(j)
13 |
14 |     U = [reverse(ind,u)]
15 |     U += irred(A,l,ind,n)
16 |     return U

```

La fonction `irred` construit une liste de listes, U , correspondant à la liste des indices des différents blocs dans la forme irréductible de A . Par exemple, si `irred` prend en entrée la liste $l = \llbracket 0, n - 1 \rrbracket$ des n premiers entiers et la matrice A définie par

$$A := \begin{pmatrix} 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{\text{Forme irréductible}} \left(\begin{array}{cccc|c} 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array} \right)$$

et dont la mise sous forme de blocs irréductibles est obtenue en permutant les colonnes et les lignes 3 et 4, *i.e* les deux dernières, elle rend la liste $\llbracket [0, 1, 4, 2], [3] \rrbracket$, ce qui revient à dire que le premier bloc est formé des colonnes et lignes 0, 1, 2, 4 et le deuxième uniquement de la colonne/ligne 3 et ce qui est bien le résultat escompté. On exploite à la ligne 12 du programme la structure particulière de Python qui permet d'ajouter, au sein d'une boucle `for`, des éléments à la liste sur laquelle porte cette boucle. La ligne 9 permet, d'une part de s'assurer que l'on ne va pas mettre deux fois un même élément dans la liste u et de ne pas avoir ainsi à la relire pour enlever les doublons, d'autre part de réduire le nombre de calculs à chaque incrémentation : on gagne donc doublement en complexité. Afin de gagner encore (légèrement) en complexité, on a fait le choix de ne pas trier les listes rendues car cela ne s'avère pas nécessaire. Enfin, l'idée de cette fonction est de travailler sur une liste l d'indices à considérer, dont la taille diminue à chaque étape, ce qui permet un algorithme récursif. À noter que cette liste l ne peut contenir des éléments supérieurs à la taille de la matrice A (ligne 11). Or, de même que pour la fonction `distinct`, on aura besoin d'appliquer `irred` à une liste ne respectant pas cette condition et la fonction `reverse` pallie ce problème (ligne 14).

Déterminer un vecteur de Fiedler. On rappelle que si A est irréductible, la valeur de Fiedler est la plus petite valeur propre non-nulle de la matrice du laplacien L_A associée, elle est unique et un vecteur de Fiedler est donné par n'importe quel vecteur propre qui lui est associée, *cf* théorème 23 pour plus de détails. Le code suivant retourne la Fiedler-value et un Fiedler-vecteur associé pour le cas d'une matrice irréductible :

```

1 | def Fiedler(A,n) :

```

```

2 |     d = [] # Matrice du laplacien
3 |     for i in range(n) :
4 |         d.append( np.sum(A[i,:]) )
5 |     L = np.diag(d)
6 |     L -= A
7 |
8 |     l, v = slin.eigh(L) # Vp dans l'ordre croissant
9 |     ind = np.argsort(l,kind='mergesort')
10 |    l, v = l[ind], v[:,ind]
11 |
12 |    i = 0
13 |    while l[i] < 10e-6 :
14 |        i += 1
15 |    return(l[i],v[:,i])

```

Pour cela, on utilise la fonction `slin.eigh` qui permet de retourner l'ensemble des valeurs (réelles) et vecteurs propres d'une matrice (ligne 8) hermitienne et après avoir rangé les valeurs propres dans l'ordre croissant (lignes 9 et 10) on trouve la première non-nulle, plus précisément la première supérieure à 10^{-6} pour tenir compte des erreurs d'arrondis. Cette fonction s'applique donc uniquement à des matrices irréductibles.

La structure d'arbre PQ. Afin de stocker de manière plus compacte les permutations, on introduit la structure d'arbre PQ. Un arbre PQ est un arbre qui comporte deux types de nœuds : de manière très surprenante, des nœuds de type P et d'autres de type Q. Plus précisément, un arbre PQ sur un ensemble $U = \{u_1, u_2, \dots, u_n\}$ est un arbre enraciné qui satisfait les trois propriétés suivantes :

- i. Chacun des u_i de U est une feuille et apparait une et une seule fois dans l'arbre,
- ii. Chaque nœud P a au moins deux fils,
- iii. Chaque nœud Q a au moins trois fils.

De plus, on dit que deux arbres PQ sont équivalents si on peut transformer l'un en l'autre uniquement grâce aux deux opérations suivantes :

- i. Permutation arbitraire des fils d'un nœud P,
- ii. Renversement² des fils d'un nœud Q.

Nous allons nous servir par la suite d'arbres PQ pour stocker des permutations : les nœuds Q stockeront des permutations qui peuvent uniquement être renversées tandis que les nœuds P auront pour fils des permutations que l'on peut intervertir les unes avec les autres sans restrictions. La figure 3 illustre les différents types de fils possibles que peut avoir la racine : un nœud Q, un nœud P ou une feuille. Remarquons que la définition même de nœuds P et Q nous oblige à stocker les transpositions dans des nœuds P, même si dans ce cas les seules permutations possibles sont des renversements. Par ailleurs, du

2. $(c_1 c_2 \dots c_k) \mapsto (c_k c_{k-1} \dots c_1)$.

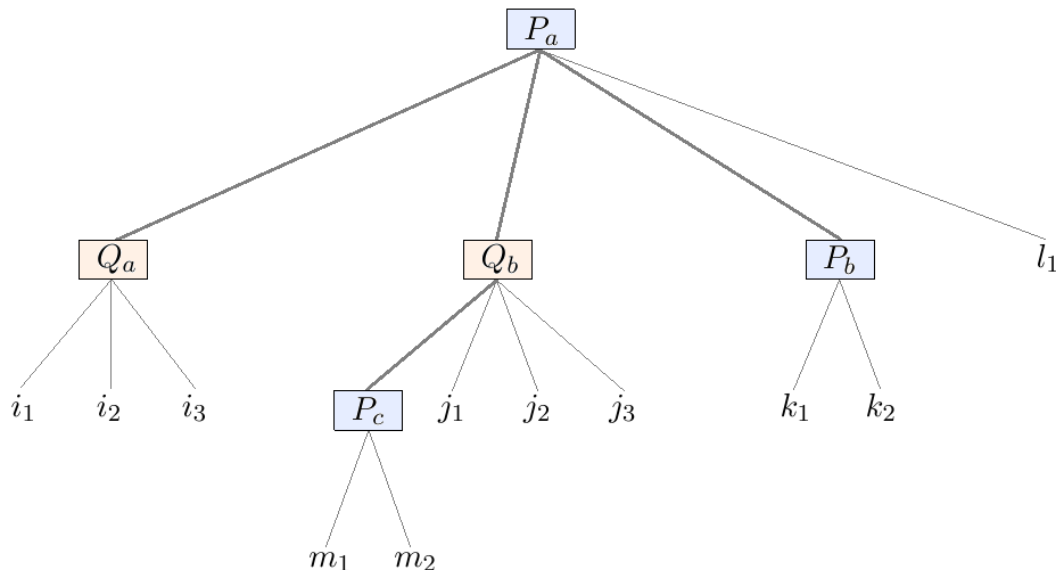


FIGURE 3. – Un arbre PQ

fait du lemme 4.2 et du théorème 21, on s’attend à ce que la racine soit de type P si la matrice est réductible et Q sinon.

Afin de mettre en œuvre l’algorithme, on introduit la structure de donnée suivante qui va permettre de stocker des arbres PQ dans Python : chaque arbre est encodé par une liste de listes³. Chacune des sous listes représente un nœud P/Q ou une feuille. Pour différencier les différents types de nœuds on introduit les convention suivantes :

P : Un nœud P est représenté par une liste à deux éléments : le premier vaut -1 et le second est une liste contenant l’ensemble de ses fils ;

Q : De même, un nœud Q est encodé par une liste à deux éléments : un label valant -2 et une liste contenant l’ensemble de ses fils ;

Feuille : Une feuille est encodée par l’unique élément u_i qu’elle contient.

Par exemple, l’arbre de la figure 3 est représenté par la liste :

$$\left[-1, \left[\underbrace{\left[-2, [i_1, i_2, i_3] \right]}_{Q_a}, \underbrace{\left[-2, [-1, [m_1, m_2]], j_1, j_2, j_3 \right]}_{Q_b}, \underbrace{\left[-1 [k_1, k_2] \right]}_{P_b}, l_1 \right] \right]$$

L’algorithme spectral de tris. On peut maintenant s’atteler à la programmation de l’algorithme proposé par l’article. L’algorithme `Spectral_Sort` prend en entrée une

3. Avec pour convention que la liste à un seul élément $[a]$ est a .

matrice B et un ensemble U qui sont les indices d'intérêt au sein de la matrice B et donne en sortie un arbre PQ stockant une permutation. Seuls les indices u de U nous intéressent au sein de B ; aussi, dès la première ligne pose-t-on $A = B[U, :][:, U]$ et on ne travaille par la suite qu'avec la matrice A . Ceci permet d'avoir accès à chaque étape à la numérotation originale des éléments et n'augmente pas significativement la complexité en temps. Par contre, ce stockage permanent a un coût en complexité spatiale mais nous ne nous intéressons pas ici à ce problème. En effet, la capacité de stockage des ordinateurs d'aujourd'hui est suffisante pour un jeu de données raisonnable et, le cas échéant il faudrait également revoir la structure de donnée adoptée précédemment en travaillant, par exemple, avec des matrices creuses.

Pour obtenir toutes les permutations, il faut alors trouver tous les arbres équivalents à l'arbre rendu par l'algorithme; ce travail est fastidieux : il faut considérer toutes les permutations possibles pour les nœuds P et renverser – ou non – les nœuds Q mais pas compliqué. Aussi, nous ne nous intéressons pas ici à énumérer toutes les permutations possibles, en dehors de cas simples, et ce pour deux raisons : la première est qu'il suffit d'exhiber une permutation pour résoudre le problème de sériation, la seconde est qu'en énumérant toutes les permutations on perd l'intérêt de la représentation compacte du résultat sous forme d'arbre PQ, en terme de stockage notamment.

L'algorithme spectral de tris utilise de manière cruciale les différents résultats mathématiques énoncés au paragraphe I.2.b et les fonctions définies ci-dessus.

```

1 | def Spectral_Sort(B,U) :
2 |     A = B[U, :][:, U]
3 |     n = A.shape[0]
4 |
5 |     alph = np.min( A+10000*np.diag(np.diag(A)) )
6 |     A -= alph*np.ones((n,n)) # translation des termes de correlation
7 |
8 |     l = [i for i in range(n)]
9 |     lU = irred(A,l,U,n) # mise sous forme irréductible
10 |         # lU : liste des indices des blocs irréductibles,
11 |     k = len(lU) # k : nombre de blocs irréductibles
12 |     T = [] # arbre PQ
13 |
14 |     if k > 1 : # si au moins deux blocs irréductibles
15 |         tmp = []
16 |         for j in range(k) :
17 |             tmp.append( Spectral_Sort(B,lU[j]) )
18 |         T = [-1,tmp] # P-node
19 |
20 |     else : # si la matrice est irréductible
21 |         if n == 1 :
22 |             T = lU[0][0] # Feuille
23 |         elif n == 2 :

```

Autour du problème de sériation

```
24         T = [-1, lU[0]] # P-node
25     else :
26         l, x = Fiedler(A,n)
27             # l : Fiedler-value ; x : Fiedler-vector
28         tmp = []
29         V = distinct(x,U,n) # des valeurs repetees ?
30         for v in V :
31             tmp.append( Spectral_Sort( B , v ) )
32         T = [-2, tmp] # Q-node
33     return(T)
```

La première étape (lignes 5 et 6) est de traduire les termes de corrélation de manière à amener le plus petit ($a_{1,n}$ dans la matrice ordonnée par définition de la corrélation) à 0, comme le justifie le lemme 4.1. Pour cela, on soustrait $\alpha = \min_{i \neq j} a_{i,j}$ terme à terme à A . Afin de ne prendre le min que sur les éléments non-diagonaux, on "gonfle" artificiellement les termes diagonaux en les multipliant par 10000 car on suppose que l'on n'étudiera pas ici des corrélations de cet ordre de grandeur. Sinon, il faut ajuster la valeur.

La seconde étape (lignes 8 et 9) est de construire la liste l_U des indices des différents blocs irréductibles de la matrice A . On peut alors travailler récursivement sur chacun des blocs (ligne 17) du fait du lemme 4.2 et ranger les arbres obtenus dans un nœud P, toutes permutations étant autorisées entre les différents blocs. Il suffit donc de traiter le cas d'une matrice irréductible pour que l'algorithme termine.

Si la matrice est irréductible, il faut distinguer trois cas, selon que A est "dégénérée" ou non. Si A est de taille 1, l'arbre associé est en fait une unique feuille, valant l'unique valeur de la matrice A (ligne 22). Si A est de taille 2, l'arbre est un nœud P contenant les deux valeurs de corrélation⁴ (ligne 24). Enfin, si A est de taille $n \geq 3$, on cherche la valeur de Fiedler et un vecteur de Fiedler associé (ligne 26). On sait de par le théorème 20 que la valeur de Fiedler est simple et on peut donc appliquer les théorèmes 19 et 21. Plus précisément, si le vecteur de Fiedler ne possède pas des valeurs répétées (ligne 29), les deux seules permutations qui conviennent sont celles qui trient le vecteur de Fiedler dans l'ordre croissant ou décroissant (théorème 19) et il convient donc de représenter cette situation par un nœud Q ayant pour fils uniquement des feuilles valant les indices triant le vecteur de Fiedler. Sinon, on travaille récursivement sur chacun des sous-blocs induits par celles-ci (ligne 31) et on stocke les arbres ainsi obtenus dans un nœud Q car seule l'opération de renversement est autorisée (ligne 32, théorème 21).

Ainsi, si A est une pré-R-matrice l'algorithme `Spectral_Sort` rend un arbre PQ qui permet de retrouver facilement l'ensemble des permutations résolvant le problème de sériation ; et la discussion précédente le prouve.

4. On ignore la diagonale et la matrice est symétrique.

Si A n'est pas une pré-R-matrice, l'algorithme rend encore un arbre PQ qui, sans être exact, donne de résultats plutôt bons d'après le paragraphe ??

b. Résultats et complexité

Complexité. On s'intéresse ici uniquement à la complexité en temps ; le type de structure de données n'étant de toute façon pas adapté à la manipulation de (très) grands objets⁵ comme annoncé plus haut.

Commençons par calculer la complexité lorsque l'on effectue une étape de la récursion, sur une matrice de taille n :

- Déterminer la taille de A (ligne 3) nécessite n opérations. De même pour définir l et calculer k aux lignes 8 et 11. Le calcul de α et la translation des termes de corrélation (ligne 6) demandent n^2 opérations.
Finalement, le préambule nécessite $\mathcal{O}(n^2)$ opérations.
- La fonction `irred` repose sur une double boucle indexée par des listes de taille au plus n et donc nécessite au plus n^2 opérations. Dans le pire des cas, c'est à dire dans le cas d'une matrice irréductible, la taille de la liste u est n assez rapidement. La fonction `reverse` nécessite n opérations mais cela ne change pas la complexité. La complexité de la fonction `irred` est donc en $\mathcal{O}(n^2)$.
- La fonction la plus gourmande en terme de complexité est `Fiedler`. En effet, elle demande de calculer l'ensemble des valeurs et vecteurs propres de A et de les trier. On note $T(n)$ la complexité de la fonction `slin.eigh` lorsqu'elle prend en entrée une matrice de taille n . Nous n'avons pas trouvé la valeur exacte de cette complexité mais on sait que la fonction `eigh` de SciPy est basée sur la bibliothèque LAPACK, dont les routines sont écrites grâce à des techniques de diviser pour régner [7, 1] ; on fait donc confiance d'une part à LAPACK, d'autre part à SciPy pour avoir écrit une fonction performante, d'autant qu'elle est optimisée pour les matrices symétriques. Une fois ces valeurs et vecteurs propres calculés, `Fiedler` les ordonne *via* la fonction `argsort`. On a choisit – grâce à l'option `kind='mergesort'` – d'utiliser un tris fusion qui a une complexité temporelle en $n \log n$, ce qui est ce qu'on peut espérer de mieux asymptotiquement. Les autres opérations effectuées par `Fiedler`, en $\mathcal{O}(n)$, sont de fait négligeables.
Finalement, on a une complexité de l'ordre de $T(n) + n \log n$.
- Pour ce qui est de la fonction `distinct`, de même que précédemment, la complexité vient du tris du vecteur x que l'on effectue grâce à un tris fusion et les autres opérations étant négligeables, il vient simplement :
La complexité de `distinct` est $\mathcal{O}(n \log n)$

5. Matrices de taille plusieurs milliers.

d'acides aminés, aussi introduit-on une matrice de corrélation C en posant simplement $C = 100e^{te} - A$. De par le caractère très conservé du cytochrome, notre matrice de corrélation est irréductible. Sur cet exemple, notre programme rend :

[-2, [19,18,17,16,1,15,0,11,9,8,10,2,6,5,4,7,3,12,13,14]]

ce qui, en terme d'arbre PQ, correspond à un unique nœud Q et qui était attendu *a priori* (matrice de corrélation irréductible). Finalement, on est en mesure de classer les différentes espèces considérées en fonction de leur distance de mutation : champignon, levure, moisissure⁷, phalène, singe, mouche, homme, pingouin, pigeon, canard, poule, chien, lapin, cochon, âne, kangourou, cheval, tortue, serpent, thon.

De tels résultats ne semblent pas complètement aberrants :

- On sait que le patrimoine génétique de la mouche et de l'homme sont assez semblables et le singe est proche de l'homme également ;
- Tous les oiseaux se retrouvent regroupés ensemble, de même pour les reptiles ;
- L'âne et le cheval se retrouvent proches, ainsi que le champignon, la levure et la moisissure.

De plus, les quelques incohérences peuvent très bien s'expliquer du fait de notre modèle très simplifié par rapport à la réalité. En premier lieu car on ne considère l'influence que d'une seule protéine sur la distance de mutation ; protéine qui peut très bien s'avérer très stable ou au contraire instable, *i.e* existant sous des formes nombreuses et variées ou non. Ensuite, on mesure ici uniquement les changements d'acides aminés dans la protéine et cela n'est pas forcément révélateur des mutations génétiques à proprement parler, ne serait-ce qu'à cause du caractère redondant du code génétique. Enfin, tous les changements d'acides aminés dans une protéine ne se valent pas : certains vont laisser la protéine peu ou prou inchangée tandis que d'autres vont radicalement modifier sa géométrie et donc son mode de fonctionnement. Or, dans notre modèle, on met un poids égal sur chacune de ces modifications.

Application en archéologie. Ici, on propose une application en "archéologie" de notre algorithme. Plus précisément, on se donne $n = 200$ objets que l'on aimerait classer par ordre chronologique. Pour cela, on simule un jeu de données en associant une datation arbitraire t_i à chacun des objets puis on définit une matrice de corrélation en posant :

- si deux objets i et j sont suffisamment proches chronologiquement, on dit qu'ils sont corrélés avec pour coefficient $c_{i,j} = \frac{1}{|t_i - t_j|}$;
- sinon, ils sont décorrélés et $c_{i,j} = 0$.

7. Il s'agit ici de *Candida* pour le champignon, *Saccharomyces* pour la levure et *Neurospora* pour la moisissure.

Autour du problème de sériation

et on espère qu'appliquer notre algorithme à notre jeu de données redonnera l'ordre pré-établi.

Pour simplifier l'implémentation, on pose $t_i = i$ pour chacun des objets et on suppose que seuls les $\frac{n}{2}$ premiers objets et $\frac{n}{2}$ derniers sont corrélés entre eux. Enfin, on multiplie la matrice de corrélation par 50 afin de limiter les erreurs d'arrondis machine. Pour résumer, la matrice de corrélation A vaut :

$$\begin{pmatrix} 0 & 50 & 25 & \dots & 0 & 0 & 0 \\ 50 & 0 & 50 & \dots & 0 & 0 & 0 \\ 25 & 50 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 50 & 25 \\ 0 & 0 & 0 & \dots & 50 & 0 & 50 \\ 0 & 0 & 0 & \dots & 25 & 50 & 0 \end{pmatrix}_{200 \times 200} .$$

On lance l'algorithme d'abord sur la matrice A telle quelle puis sur A' , obtenue à partir de A en appliquant la même permutation aléatoire sur les lignes et les colonnes de A , de manière à préserver le caractère symétrique de A .

Dans le premier cas, le programme s'exécute en environ 7 secondes et rend bien un nœud P contenant deux fils : le fils droit de type Q stocke, rangées par ordre croissant, les feuilles 100 à 199 ; le fils gauche également de type Q stocke les objets 0 à 99 :

$$[-1, [[-2, [0, 1, 2, \dots, 97, 98, 99]] , [-2, [100, 101, 102, \dots, 197, 198, 199]]]] .$$

Autrement dit, la matrice étant déjà sous forme irréductible et ordonnée, la permutation identité résout le problème de sériation.

Dans le(s) second(s) cas, le programme s'exécute en environ 3 secondes et on peut expliquer cette diminution de temps par le fait que l'on va travailler en moyenne sur des sous matrices de taille inférieure à $\frac{n}{2}$. Le résultat est le suivant :

```

[-1, [[-2,
      [122, 13, 12, 101, 170, 120, 83, 99, 136, 23, 103, 56, 113, 28, 42,
      55, 140, 173, 194, 127, 39, 199, 135, 156, 168, 160, 109, 193, 41,
      183, 124, 138, 132, 145, 54, 147, 131, 129, 71, 180, 89, 148, 188,
      110, 142, 192, 187, 44, 96, 118, 175, 174, 35, 20, 53, 50, 73, 93,
      165, 62, 70, 139, 176, 47, 16, 43, 146, 102, 25, 126, 0, 58, 172,
      91, 4, 18, 59, 153, 196, 19, 121, 112, 65, 111, 37, 158, 163, 61,
      49, 195, 105, 117, 30, 77, 1, 98, 81, 33, 63, 198]]],
[-2, [66, 152, 154, 85, 134, 57, 82, 40, 104, 90, 181, 86, 48, 94, 80,
      191, 167, 151, 31, 166, 14, 197, 67, 26, 17, 27, 92, 9, 108, 150,
      29, 38, 133, 125, 32, 95, 171, 155, 3, 116, 114, 182, 100, 179, 72,
      149, 178, 189, 115, 106, 52, 78, 7, 144, 157, 5, 159, 45, 107, 84,
      184, 22, 162, 88, 51, 21, 36, 6, 169, 119, 8, 164, 123, 75, 141,
      130, 97, 64, 68, 60, 161, 15, 11, 74, 34, 46, 185, 137, 69, 10, 2,
      186, 76, 128, 24, 177, 143, 79, 190, 87] ] ] .

```

Comme il semble compliqué d'interpréter ce résultat pour une matrice de taille 200×200 , on se propose de le faire pour une matrice de taille 10×10 :

Soit pour $n = 10$ la permutation $\sigma_0 = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 0 & 1 & 4 & 2 & 8 & 7 & 9 & 5 & 6 \end{pmatrix}$ de \mathfrak{S}_{10} que l'on applique aux lignes et colonnes de la matrice de corrélation A . L'algorithme rend dans ce cas la permutation

```
[-1, [[-2, [1, 2, 4, 0, 3]], [-2, [8, 9, 6, 5, 7]]]] .
```

On veut vérifier par le calcul que toutes les permutations engendrées par cet arbre résolvent le problème de sériation et que ce sont les seules possibles. Ainsi, cela apportera une preuve supplémentaire de la validité de notre algorithme, plus exactement que l'implémentation de l'algorithme est juste⁸. Pour cela, commençons par donner une première interprétation de ce résultat : on effectue les opérations

$$A \xrightarrow{\sigma_0} A^{\sigma_0} \xrightarrow{\sigma_T} (A^{\sigma_0})^{\sigma_T}$$

$\underbrace{\hspace{10em}}_p$

Or, $A^{\sigma_0} = (a_{\sigma_0(i), \sigma_0(j)})_{i,j}$ donc $(A^{\sigma_0})^{\sigma_T} = (a_{\sigma_0 \circ \sigma_T(i), \sigma_0 \circ \sigma_T(j)})_{i,j}$, i.e $(A^{\sigma_0})^{\sigma_T} = A^{\sigma_0 \circ \sigma_T}$ on pose donc $p = \sigma_0 \circ \sigma_T$ et en appliquant p à la matrice A de départ on connaîtra l'effet combiné du dérangement par σ_0 puis du réarrangement par σ_T .

Finalement, tous calculs faits, l'ensemble des permutations associées à cet arbre est donc, selon que l'on renverse et/ou permute les différents nœuds :

8. En théorie, l'algorithme rend le bon résultat et on l'a même prouvé; en pratique, c'est une autre question et l'expérience a montré que ce n'était pas toujours chose assurée!

Autour du problème de sériation

i. Laisser les deux nœuds à l'endroit, sans les permuter.

$$\sigma_T = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 2 & 4 & 0 & 3 & 8 & 9 & 6 & 5 & 7 \end{pmatrix}; p = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{pmatrix}$$

ii. Permuter les deux nœuds Q en les laissant à l'endroit.

$$\sigma_T = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 8 & 9 & 6 & 5 & 7 & 1 & 2 & 4 & 0 & 3 \end{pmatrix}; p = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 5 & 6 & 7 & 8 & 9 & 0 & 1 & 2 & 3 & 4 \end{pmatrix}$$

iii. Renverser le premier des deux nœuds Q , sans les permuter.

$$\sigma_T = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 0 & 4 & 2 & 1 & 8 & 9 & 6 & 5 & 7 \end{pmatrix}; p = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 3 & 2 & 1 & 0 & 5 & 6 & 7 & 8 & 9 \end{pmatrix}$$

iv. Renverser le premier des deux nœuds Q et les permuter.

$$\sigma_T = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 8 & 9 & 6 & 5 & 7 & 3 & 0 & 4 & 2 & 1 \end{pmatrix}; p = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 5 & 6 & 7 & 8 & 9 & 4 & 3 & 2 & 1 & 0 \end{pmatrix}$$

v. Renverser le deuxième des deux nœuds Q , sans les permuter.

$$\sigma_T = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 2 & 4 & 0 & 3 & 7 & 5 & 6 & 9 & 8 \end{pmatrix}; p = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 0 & 1 & 2 & 3 & 4 & 9 & 8 & 7 & 6 & 5 \end{pmatrix}$$

vi. Renverser le deuxième des deux nœuds Q et les permuter.

$$\sigma_T = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 7 & 5 & 6 & 9 & 8 & 1 & 2 & 4 & 0 & 3 \end{pmatrix}; p = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 9 & 8 & 7 & 6 & 5 & 0 & 1 & 2 & 3 & 4 \end{pmatrix}$$

vii. Renverser les deux nœuds Q , sans les permuter.

$$\sigma_T = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 0 & 4 & 2 & 1 & 7 & 5 & 6 & 9 & 8 \end{pmatrix}; p = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 3 & 2 & 1 & 0 & 9 & 8 & 7 & 6 & 5 \end{pmatrix}$$

viii. Renverser les deux nœuds Q et les permuter.

$$\sigma_T = \begin{pmatrix} 3 & 0 & 1 & 4 & 2 & 8 & 7 & 9 & 5 & 6 \\ 7 & 5 & 6 & 9 & 8 & 3 & 0 & 4 & 2 & 1 \end{pmatrix}; p = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0 \end{pmatrix}$$

en particulier, les résultats que l'on obtient par le calcul – pour p – sont cohérents avec l'idée que l'on se fait de la résolution d'un tel problème de sériation.

Le code suivant permet de visualiser le résultat annoncé pour des matrices de n'importe quelle taille, dont 200. Il prend en entrée l'arbre T rendu par l'algorithme et la permutation initiale notée s et affiche la permutation p correspondant au label demandé.


```

1 def verif(T,s,label) :
2   p1, p2 = T[1][0][1], T[1][1][1]
3   q1, q2 = p1[:, :-1], p2[:, :-1]
4
5   if label == 1 :
6     print( s[p1+p2] )
7   elif label == 2 :
8     print( s[p2+p1] )
9   elif label == 3 :
10    print( s[q1+p2] )
11   elif label == 4 :
12    print( s[p2+q1] )
13   elif label == 5 :
14    print( s[p1+q2] )
15   elif label == 6 :
16    print( s[q2+p1] )
17   elif label == 7 :
18    print( s[q1+q2] )
19   elif label == 8 :
20    print( s[q2+q1] )

```

Si on cherchait à réellement dater de vrais objets issus de l'archéologie grâce à des corrélations établies entre ces objets, on ne pourrait néanmoins sans informations supplémentaires conclure quant à la datation de ceux-ci. Cette méthode peut donc servir à estimer la datation de certains objets en établissant une chronologie relative avec des objets dont la datation est connue. Plus fréquemment, on se sert de la sériation en archéologie pour dégager des "groupes" d'objets et affiner par la suite.

Il y a néanmoins une critique assez forte que l'on peut faire à l'usage de la sériation en archéologie : l'hypothèse d'une croissance linéaire du temps de la corrélation des objets. Si on reprend l'exemple des vases présentés en figure 1, page 101, on ne peut affirmer brutalement que leur profil évolue nécessairement de manière linéaire : la mode peut très bien être d'allonger les encolures puis de les rétrécir et on aura tendance sans autres informations à corréler les deux vases de type encolure longue alors qu'il n'en est rien en réalité. Il faut donc nuancer les résultats donnés par les techniques de sériations – qui n'offrent de toute façon qu'une chronologie relative – et choisir avec soin ses indices de corrélation.

Enfin, cet exemple permet de mettre en emphase l'idée de pire des cas pour ce qui est de la complexité de notre algorithme. En effet, juste en mélangeant la matrice et donc en la "ré-équilibrant" en quelques sorte on a obtenu une réduction notable du temps de calcul. Il peut donc être intéressant d'une manière générale pour augmenter la performance de notre algorithme, et ce quelques soit la donnée en entrée, de la déranger avant d'effectuer le `Spectral_Sort...`

3. Comparaison des performances

On se propose dans cette partie de comparer les temps d'exécution des deux algorithmes développés ci-dessus sur la matrice identité de taille n croissante. Les résultats obtenus sont résumés dans le tableau 2.

Taille de la matrice	Algorithme naïf	Algo naïf théorique	Spectral_Sort
1	0.0	0.0	0.0
2	0.0	4.5373×10^{-6}	0.0
3	0.0	7.4865×10^{-5}	0.015600
4	0.0	0.00096190	0.0
5	0.015600	0.011354	0.0
6	0.17160	0.13347	0.015600
7	1.8564	1.6204	0.031200
8	21.684	20.646	0.031200
9	278.02	278.02	0.031200
10	3806.4	3965.7	0.046800
11		59923	0.046800
12		958153	0.046800
13		16185511	0.078000
14		288303702	0.062400
15		5.4044×10^9	0.062400

TABLE 2. – Temps d'exécutions en secondes

Le tableau 2 met notamment en évidence l'écrasante complexité de l'algorithme naïf et nous n'effectuons plus pour cela de tests au delà de $n = 10$, qui a déjà mis non loin d'une heure à produire un résultat !

Ne pouvant tester notre algorithme sur des valeurs de n grandes, on se propose ici d'en donner une estimation grâce au calcul exact de la complexité effectué au paragraphe II.1.2, dont on rappelle ici la formule :

$$C_{naïf} = \frac{2}{3}n^3.n! - 2n^2.n! + \frac{4}{3}n.n! + \sum_{k=1}^{n-1} \frac{n!}{k!}.$$

Pour cela, on interpole à partir de la formule théorique pour avoir la valeur exacte au point $n = 9$ en re-normalisant par 440790. Les résultats ainsi obtenus sont regroupés dans la colonne 3 du tableau 2. Le choix d'une telle renormalisation donne sur nos valeurs exactes les erreurs relatives présentées au tableau 3, ce qui a motivé notre choix

d'interpoler en $n = 9$ ⁹.

L'estimation produite par un tel procédé est un peu plus haute mais donne de bons résultats sur les premières valeurs de $n \geq 10$. Notamment, il faudrait plus de 170 ans pour calculer le terme $n = 15$!!

n	5	6	7	8	9	10
ε_{rel}	0.27167	0.22210	0.12712	0.04782	9.6805×10^{-6}	0.04186

TABLE 3. – Erreurs relatives pour le temps de calcul.

On cherche dorénavant à tester les limites de notre algorithme `Spectral_Sort` en le lançant sur des données de taille plus importante et on résume les différents temps de calcul obtenus dans le tableau 4.

n	25	50	100	200	300	400	500
t	0.14040	0.51480	2.0280	7.9092	17.956	31.559	70.672
n	600	700	800	900	950	975	982
t	49.145	96.602	125.76	158.87	178.61	187.22	189.15

TABLE 4. – Temps de calcul pour `Spectral_Sort`, en secondes.

Si notre algorithme peut encore être optimisé, il donne donc de bons résultats sur des matrices de taille "raisonnable". En revanche, on ne peut exécuter le programme pour une matrice de taille supérieure à 982 car on atteint la saturation de la pile de récursion¹⁰. On pourrait éventuellement remédier à cela en stockant des étapes intermédiaires pour limiter les appels récursifs...

III. Le problème C1P

Le problème C1P (consecutive ones problem), se révèle être un cas particulier du problème de sériation. Nous proposons donc ici d'implémenter, toujours sur Python, l'algorithme décrit à la fin de l'article pour résoudre le problème C1P.

9. $n = 8$ rendait également de très bons résultats mais estimait un peu moins bien les grandes valeurs de n , d'où le choix de $n = 9$ pour approcher au mieux le temps de calcul pour $n = 11$...

10. Maximum recursion depth exceeded.

Autour du problème de sériation

Description du problème. Le problème C1P consiste à se donner une matrice C de taille $n \times m$, non nécessairement carrée cette fois, à coefficients dans l'ensemble $\{0, 1\}$ ¹¹, et à se demander s'il existe une matrice de permutation Π telle que ΠC vérifie la propriété suivante :

$$\forall j \in \llbracket 1, m \rrbracket, \exists r, s \in \llbracket 1, n \rrbracket, \forall i \in \llbracket 1, n \rrbracket, i \in \llbracket r, s \rrbracket \iff (\Pi C)_{i,j} = 1,$$

ce qui revient à demander qu'une fois les lignes de C judicieusement permutées les 1 de chaque colonne soient consécutifs.

Une matrice C pour laquelle il existe une solution au problème C1P sera appelée une pré-P-matrice et une matrice qui vérifie la propriété sera une P-matrice. Le problème se reformule ainsi : Pour C une pré-P-matrice, trouver une matrice de permutation Π telle que ΠC soit une P-matrice.

Solution théorique. Les résultats suivant, dus à D.G. Kendall, permettent de résoudre facilement le problème C1P à partir de ce que nous avons déjà fait ; nous admettons les lemmes.

Lemme 9.1. Si C est une P-matrice, alors $C {}^t C$ est une R-matrice.

Lemme 9.2. Si C est une pré-P-matrice et si $C {}^t C$ est une R-matrice, alors C est une P-matrice.

Théorème 25

Soit C une pré-P-matrice, soit $A = C {}^t C$. Si Π est une matrice de permutation, alors ΠC est une P-matrice si et seulement si $\Pi.A. {}^t \Pi$ est une R-matrice.

Démonstration :

Si ΠC est une P-matrice, alors d'après le premier lemme $(\Pi C) {}^t (\Pi C)$ est une R-matrice. Or,

$$(\Pi C) {}^t (\Pi C) = \Pi.C {}^t C {}^t \Pi = \Pi.A. {}^t \Pi.$$

Donc $\Pi.A. {}^t \Pi$ est une R-matrice. Réciproquement, si $\Pi.A. {}^t \Pi$ est une R-matrice, alors $(\Pi C) {}^t (\Pi C)$ est une R-matrice et d'après le second lemme ΠC est une P-matrice. \square

11. Il est à noter que pour des raisons algorithmiques, voir C comme une matrice de $\mathcal{M}_n(\mathbb{Z}/2\mathbb{Z})$ est malvenu, car cela revient à prendre le modulo 2 dans les produits matriciels qui vont suivre.

On peut dire à partir de ce résultat qu'une fois le problème de sériation résolu pour la matrice $C^t C$, la permutation trouvée transforme C en P-matrice si on ne l'applique qu'aux lignes.

Programme python. Pour implémenter l'algorithme qui se dégage du travail théorique, nous chargeons de nouveau les package numpy et scipy.linalg

```
1 | import numpy as np
2 | import scipy.linalg as slin
```

et nous avons également besoin de récupérer à partir d'un arbre PQ une permutation correspondante. Nous utilisons donc la fonction `Feuilles` suivante, nommée ainsi parce qu'elle obtient la permutation en parcourant toutes les feuilles de gauche à droite. Son fonctionnement est transparent :

```
1 | def Feuilles(T):
2 |     F=[]
3 |     for t in T:
4 |         if type(t)==int :
5 |             if t >= 0:
6 |                 F.append(t)
7 |         else :
8 |             F+=Feuilles(t)
9 |     return F
```

Avec cette fonction, il devient facile de construire une fonction `C1P` qui prend en entrée une matrice à coefficients dans $\{0, 1\}$ et rend en sortie une version permutée de cette matrice qui vérifie la propriété des 1 consécutifs. En voici le code :

```
1 | def C1P(C):
2 |     A=np.dot(C,C.T)
3 |     T=Spectral_Sort(A, [i for i in range( A.shape[0])] )
4 |     p=Feuilles(T)      # On recupere une permutation a partir de l'
                        # arbre T
5 |     return C[p,: ]  # On l'applique aux lignes de C seulement
```

Le programme fait un appel à `Spectral_Sort` qui a une complexité polynomiale, et ne fait par ailleurs que des opérations de complexité polynomiale, avec des polynômes d'exposants raisonnables ; on gage donc sans calcul que la complexité de `C1P` est raisonnable et qu'elle est dans la classe de complexité P des algorithmes polynomiaux en temps.

Notons enfin que si le problème est mal posé, le programme rend quand même un résultat, et que, puisque la matrice fournie par `Spectral_Sort` tend à ressembler à une R-matrice au sens où nous avons décrit l'heuristique en Partie 1, nous avons des raisons de penser que la solution heuristique fournie sera une matrice "proche" d'être une P-matrice. La question est de donner un sens à ce "proche", et nous ne l'avons pas

Autour du problème de sériation

vraiment résolue, mais on peut d'ores et déjà constater que si une matrice extraite de $A = C {}^t C$ est une pré-R-matrice, alors l'heuristique lancée sur C produira une matrice qui est une P-matrice en dehors des colonnes qui étaient retirées à A , car retirer des colonnes à C est équivalent à retirer les mêmes colonnes et leurs lignes associées dans A . On peut donc supposer que pour C non pré-P, le programme rendra une matrice qui sera une P-matrice à quelques colonnes près pour la plupart des entrées raisonnables.

Tests. Nous vérifions ici le fonctionnement du programme pour un cas simple et testons ensuite sa robustesse en tant qu'heuristique sur un cas pour lequel il n'y a pas de solution (trouvé empiriquement et confirmé par le fait que le programme ne renvoie pas une P-matrice).

Nous testons la matrice

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

associée à la pré-R-matrice

$$A = \begin{pmatrix} 2 & 1 & 1 & 2 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 2 & 1 & 1 & 3 \end{pmatrix}$$

ce qui correspond à exécuter le code

```
1 | C=np.array([[1,0,0,0],[0,0,0,0],[1,0,0,0],[0,0,0,0],[0,0,0,0],[0,0,0,0]])
2 | print( C1P(C) )
```

et produit le résultat :

```
[[1 0 0 1]
 [1 1 0 1]
 [0 0 0 1]
 [0 0 0 1]]
```

où les 1 sont sagement consécutifs.

On obtient des résultats plus impressionnants, dont nous ne détaillons pas les entrées pour ne pas surcharger le compte rendu :

$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$	devient	$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$
--	---------	--

ou encore

$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$	devient	$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$
--	---------	--

Pour l'entrée mal posée suivante, on peut observer que le résultat tend à éviter les colonnes sans "1 consécutif" :

$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$	devient	$\begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$
--	---------	--

ce qui corrobore notre supposition.

Références

- [1] LAPACK : Linear Algebra PACKage 3.6.0. Symmetric matrix – eigenvalue, Nov 15, 2015. http://www.netlib.org/lapack/explore-html/d3/d88/group__real_s_yeigen.html.
- [3] Jonathan E. Atkins, Erik G. Boman, and Bruce Hendrickson. A spectral algorithm for seriation and the consecutive ones problem. *SIAM Journal of Computing*, 28(1) :297–310, 1998.
- [7] The Scipy Community. Scipy.org – numpy.linalg.eigh, Oct 18, 2015. <http://docs.scipy.org/doc/numpy/reference/generated/numpy.linalg.eigh>.

Autour du problème de sériation

[html](#).

- [11] Walter M. Fitch and Emanuel Margoliash. Construction of phylogenetic trees. *Science*, 155(3760) :279–284, January 1967.
- [27] Innar Lilv. Seriation and matrix reordering methods : an historical overview. *Statistical Analysis and Data Mining*, 3(2) :70–91, April 2010.

– Sixième partie –

TER de M2
Espaces de formes et courbure

Encadré par Joan Alexis Glaunès *

En binôme avec Pierre Roux

*. MAP 5, Université Paris-Descartes

Sommaire – Partie VI

TER de M2 : Espaces de formes et courbure	137
Chapitre I :	
Géométrie riemannienne et imagerie médicale	141
I Des mathématiques pour l'analyse des données médicales	141
II Modéliser la variabilité	142
III Un exemple : surfaces corticales d'amygdales cérébrales	143
Chapitre II :	
Variétés différentielles	145
I Appariement de formes	145
II Vecteurs tangents et différentielle	146
II.1 Fonction différentiable	146
II.2 Espace tangent	148
II.3 Fibré tangent	150
III Dérivation et champ de vecteurs	152
III.1 Dérivation ponctuelle	152
III.2 Dérivation globale	155
Chapitre III :	
Cadre riemannien	161
I Métrique riemannienne	161
II Connexion de Levi-Civita	164
II.1 Approche intuitive et cas euclidien	164
II.2 Cas général	164
II.3 Dérivée covariante le long d'une courbe	168
III Géodésiques	169
Chapitre IV :	
Courbure	173
I courbure d'un arc	173

II	courbure d'une surface	179
III	courbure d'une variété riemannienne	181
IV	Illustration dans le cas des amygdales	184
Chapitre V :		
	La formule de Mario	187
I	La variété riemannienne des landmarks	187
	I.1 Le groupe des difféomorphismes	187
	I.2 Fonctionnelle d'énergie	188
II	Calcul de la courbure sectionnelle	190
	II.1 Implémentation de la formule	193
Annexe A :		
	"Rappels" sur les fibrés	197
I	Espaces fibrés	197
II	Espaces fibrés vectoriels	199
	Références	199

– Chapitre I –

Géométrie riemannienne et imagerie médicale

Le présent rapport s'inscrit dans le cadre de notre projet de M2, se devant d'être à l'interface entre les sciences du vivant et les mathématiques. Nous avons travaillé avec Joan Glaunès sur l'utilisation de la géométrie riemannienne pour l'appariement de formes en imagerie médicale.

Il nous paraissait donc nécessaire de situer en contexte les développements mathématiques qui vont suivre dans le cadre de l'imagerie médicale et plus généralement de l'anatomie computationnelle.

I. Des mathématiques pour l'analyse des données médicales

Aujourd'hui l'acquisition d'images médicales est de plus en plus performante et importante et les outils mathématiques deviennent essentiels pour le traitement d'une telle masse d'informations. Si les statistiques sont bien évidemment nécessaires du fait du grand nombre de données, un traitement préalable est incontournable et c'est notamment tout l'intérêt de la géométrie riemannienne.

Entre autres, les domaines suivants interviennent de manière cruciale dans l'étude de l'anatomie computationnelle :

- L'*analyse fonctionnelle*, en premier lieu, est l'outil "indispensable" à l'analyse d'image. Notamment, on est souvent amené à appliquer des transformations en ondelettes et autres filtres aux images acquises afin de faire "émerger" les formes et rendre les acquisitions mathématiquement utilisables.
- La *géométrie riemannienne* comme nous allons le détailler plus loin permet d'étudier, de comparer les formes ainsi obtenues. Cette étape est indispensable pour la suite.
- Les *statistiques* s'imposent dès lors que l'on s'intéresse à une population d'individus, si on cherche à classifier/discriminer les données. Il faut néanmoins bien prendre conscience que sans outils pour "mathématiser" – au sens large – les données anatomiques, on est incapable de les étudier en tant qu'objet mathématique et d'ainsi proposer un cadre d'étude satisfaisant.

- L'apprentissage (*statistique*) permet de prédire l'évolution temporelle d'une forme à partir d'un état donné et de données d'autres patients. Cela prend notamment tout son intérêt dans l'élaboration d'outils d'aide à la décision : prévoir l'évolution de la forme de l'hippocampe dans la maladie d'Alzheimer pourrait permettre un diagnostic précoce et donc une meilleure prise en charge du patient.

Outre la grande variabilité des données d'acquisition qui justifie la mise en place d'une théorie générale et non organe-dépendante, il faut bien prendre conscience que la nature nous a fait tous semblables mais loin d'être identiques. Ainsi, la forme d'un même organe peut varier très fortement d'un individu à l'autre comme nous allons le voir au paragraphe III et il apparaît nécessaire de pouvoir quantifier cette variabilité afin de déterminer si deux organes diffèrent de par leur nature propre ou bien parce que l'un est malade et l'autre non.

II. Modéliser la variabilité

La théorie des espaces de formes apparaît comme une tentative de modélisation mathématique de cette variabilité. On introduit pour cela un espace, dit *espace de formes* dans lequel chacun des points est en fait un objet géométrique et que l'on va munir d'une métrique à même donc de quantifier l'écart entre deux objets géométriques. Pour ce faire, on définira la distance entre deux formes comme "la transformation de l'espace la moins coûteuse permettant d'amener une forme sur l'autre" : aussi s'attachera-t-on pour définir la distance à minimiser une fonctionnelle d'énergie (cf paragraphe I.2), tout cela dans un sens à préciser. Cette distance est dite géodésique¹.

L'espace de formes ainsi défini est en particulier une variété riemannienne, *i.e* un espace courbe, ce qui peut poser problèmes si on cherche à faire des statistiques sur des populations d'étude. Aussi est-il courant de linéariser le problème pour se ramener à l'espace tangent d'une forme prototype, ce qui permet de récupérer tous les outils classiques de l'analyse statistique. Le prix à payer est que le résultat ainsi obtenu n'est qu'une approximation. L'objet de notre TER était donc d'étudier la courbure des variétés riemanniennes afin de pouvoir juger de la pertinence d'une telle approximation lors d'un cas d'étude réelle.

Par exemple, la figure I.1a illustre une manière algorithmique rapide de calculer le point moyen \bar{x} d'un nuage de point $\{x_1, x_2, x_3, x_4, x_5\}$. Cet algorithme, dans notre cadre, permettrait de calculer une forme moyenne et ainsi de pouvoir identifier un "organe-prototype". L'intérêt médical est grand : si on est à même de déterminer un prototype pour chacun des stades d'une maladie, on pourrait alors juste en calculant la distance géodésique de l'organe d'un patient aux organes prototypes connaître son état de santé.

1. Littéralement qui permet de mesurer la terre, objet courbe s'il en est. . .

Dans les faits, moyenner tous les organes ne donne pas de très bons résultats, du fait de la grande variabilité des phénotypes rappelée en préambule, et il est préférable de se donner un ensemble de formes-types, ce qui complique les choses. . .

Connaitre la courbure de l'espace de forme sous-jacent permettrait donc de mesurer l'erreur commise par l'usage d'un tel algorithme et ainsi de s'en satisfaire sans avoir besoin de le corriger en considérant des distances géodésiques en lieu et place des euclidiennes si la courbure est faible. Un autre intérêt serait d'évaluer l'erreur commise en approchant la distance géodésique entre deux points x_1 et x_2 par la distance euclidienne qui s'obtient aisément par la différence des vecteurs tangents respectifs (cf figure I.1c), comme nous le faisons au paragraphe IV.

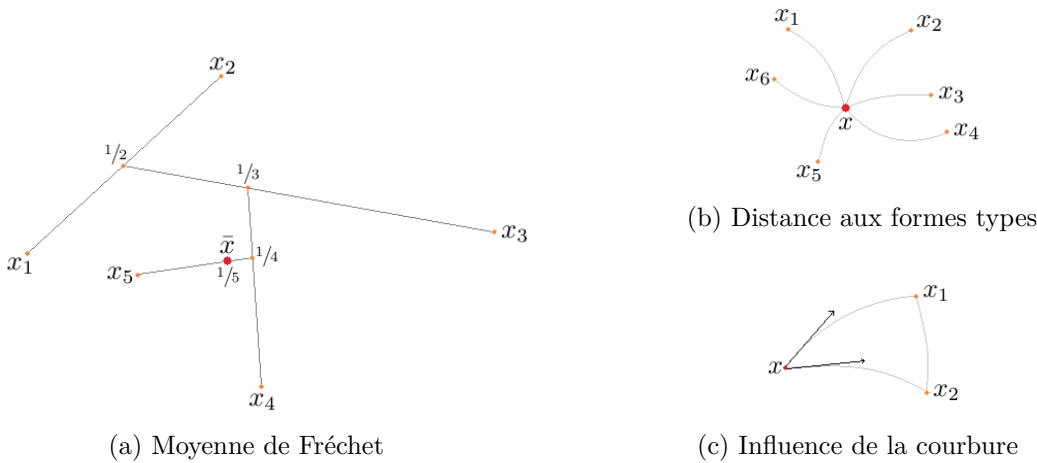


FIGURE I.1. – Distance et courbure

Dans ce TER nous nous sommes plus précisément intéressés au calcul de la courbure sectionnelle dans le cas des configurations de points, ou landmarks, et plus précisément à l'implémentation numérique d'une formule décrite dans [28] et dite de Mario calculant ladite courbure.

III. Un exemple : surfaces corticales d'amygdales cérébrales

Nous nous sommes intéressés à l'appariement de surfaces corticales d'amygdales cérébrales droites et gauches de 25 sujets. Nous avons pour cela utilisé un code fourni par M. Joan Glaunès et codé une fonction permettant un meilleur affichage graphique et une "reconstruction" de l'amygdale à partir d'un échantillon de points. L'amygdale cérébrale est une formation de matière grise située au niveau de l'hippocampe et en forme

d'amande – d'où son nom... – dont le rôle est de décoder les stimuli qui pourraient être menaçants pour l'organisme, autrement dit l'amygdale est une forme de "système d'alerte" responsable du sentiment de peur et d'anxiété.

La figure I.2 superpose l'amygdale cérébrale droite de 2 sujets et illustre notamment la variabilité morphologique sus-citée.

Pour appairer les deux amygdales présentées, on procède par appariement de landmarks (voir le paragraphe I pour plus de détails) en construisant une géodésique point par point. La figure I.3 montre l'évolution de la forme induite par les différents points lorsque ceux-ci parcourent leur géodésique propre : en $t = 0$, l'amygdale du sujet 1 et en $t = 1$, l'amygdale du sujet 2.

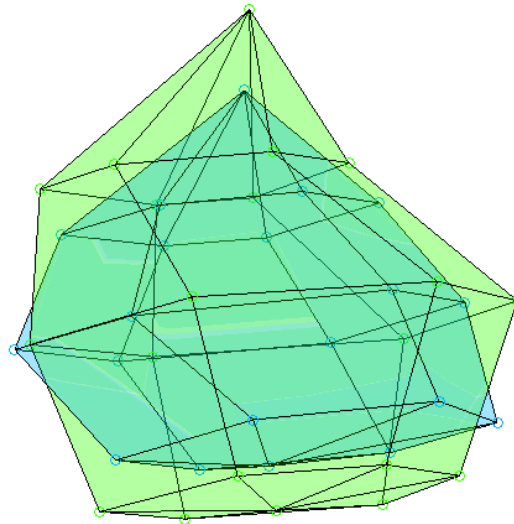


FIGURE I.2. – Amygdales cérébrales droites de deux sujets

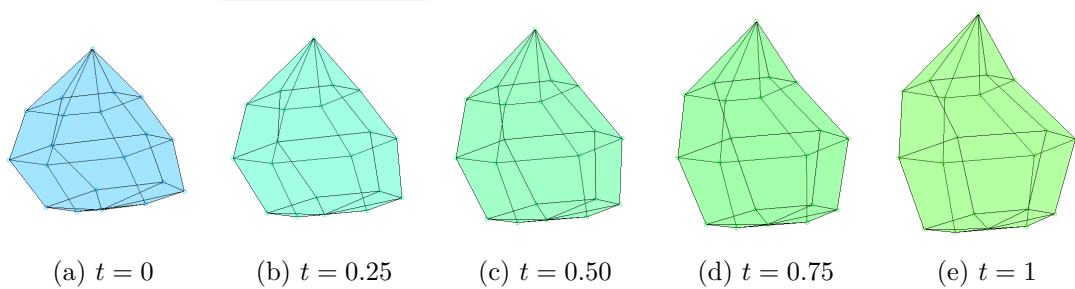


FIGURE I.3. – Appariement de deux amygdales droites

– Chapitre II –

Variétés différentielles

Une part non-négligeable de notre travail a consisté à se former aux bases de la géométrie riemannienne afin de pouvoir comprendre les articles d'intérêt.

Cela impliquait de s'assurer de "bien" maîtriser les subtilités de la géométrie différentielle et notamment les notions de dérivation et de fibré tangent, que nous n'avions pas étudiées en détail l'année passée, afin d'espérer comprendre quelque chose par la suite. A ce titre, nous répertorions ici le "fruit" de notre travail.

On suppose connue la notion de variété différentielle et on s'inspire pour beaucoup des premiers chapitres de [25] ou du premier de [13]. . .

I. Appariement de formes

Comme expliqué à la section , notre but principal est d'arriver à comparer des objets – des formes – entre eux. Pour cela on a envie d'introduire une notion de "distance" entre ces différents objets. Ainsi sera-t-on à même de donner un sens mathématique à l'assertion "l'objet A est plus semblable à l'objet B qu'à l'objet C " : il suffira pour cela de comparer les distances $d(A, B)$ et $d(A, C)$.

La manière la plus simple de procéder est d'assimiler ces différents objets à des points sur une variété M . Dans ce cas, il s'agit de construire la métrique sur une variété M pour résoudre notre problème. Si, en fait de variété, on travaille dans un espace vectoriel, la distance à considérer est simplement la distance usuelle issue du produit scalaire euclidien. A contrario, dans une variété "courbe", on ne peut relier linéairement les points : il faut, justement, tenir compte de cette courbure. Ainsi, par exemple, on a envie de dire que les deux pôles nord et sud de la sphère \mathbb{S}^2 sont éloignés d'une distance π et non 2. . .

Une façon naïve d'encoder des formes est de les interpoler avec un certain nombre de points et de chercher à apparier les points un à un. Pour préserver la structure (anatomique) de l'objet il peut néanmoins être plus judicieux de considérer plutôt des courants, etc. Nous nous intéressons ici au modèle le plus simple, à savoir : apparier des points un à un, connaissant le nombre de points et sachant qui apparier avec qui. Les situations décrites précédemment correspondent à des modèle de *landmarks*. Il existe bien évidemment d'autres façon de procéder. . .

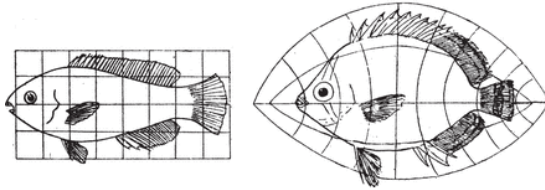


FIGURE II.1. – *D'Arcy-Thompson* : un modèle historique d'appariement de formes. Une transformation régulière du plan permet de transformer le quadrillage original "euclidien" en un quadrillage courbe.

La figure I.1 illustre une façon d'encoder différentes fibres de la matière blanche (de cerveau) mathématiquement. La matière blanche est en effet principalement constituée d'axones, d'où son aspect "fibreuse". Elle est responsable de la propagation des informations dans le système nerveux et relie notamment les différentes aires de la matière grise, située en périphérie du cerveau. Sont représentés ici : en bleu, le faisceau cortico-spinal ; en jaune, le faisceau cortico-nucléaire ; en rouge, le corps calleux et en vert les faisceaux arqués droit et gauche.

En médecine, il y a un réel intérêt à pouvoir observer ces fibres en tant que fibres justement car cela permet de détecter des maladies telles que la sclérose en plaque, qui n'est autre que la destruction de la myéline protégeant les axones de la matière blanche, la maladie d'Alzheimer et d'autres maladies neurodégénératives.

II. Vecteurs tangents et différentielle

1. Fonction différentiable

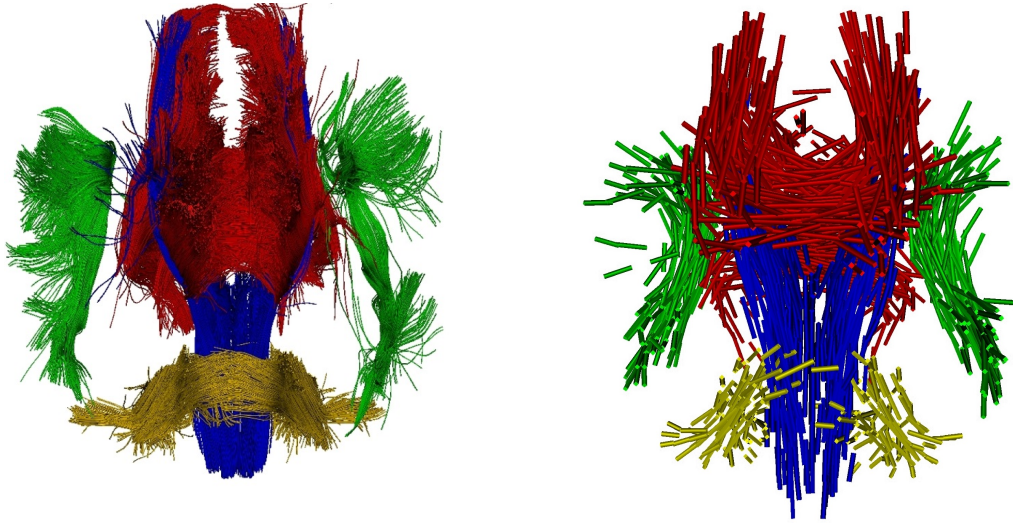
Soit M et N deux variétés, de classe \mathcal{C}^k , de dimensions respectives n_M et n_N .

Définition II.1. Une application continue $f : M \rightarrow N$ est dite de classe \mathcal{C}^k si pour tout $m \in M$ il existe une carte (\mathcal{U}, ϕ) de M avec $m \in \mathcal{U}$, une carte (\mathcal{V}, ψ) de N avec $f(m) \in \mathcal{V}$ telle que l'application

$$\psi \circ f \circ \phi^{-1} : \phi(f^{-1}(\mathcal{V}) \cap \mathcal{U}) \rightarrow \psi(\mathcal{V})$$

soit de classe \mathcal{C}^k , en tant que fonction de \mathbb{R}^{n_M} dans \mathbb{R}^{n_N} .

Autrement dit, on considère le diagramme commutatif suivant :



(a) Exemple de fibres corticales dans le cerveau d'un patient.

(b) Modèle mathématique associé – basé sur des courants ici.

FIGURE II.2. – Des faisceaux de fibres dans le cerveau.

Crédit : D. Ducreux, Hôpital Kremlin-Bicêtre et P. Fillard, CEA, Neurospin.

$$\begin{array}{ccc}
 \tilde{\mathcal{U}} & \xrightarrow{f} & \mathcal{V} \\
 \phi \downarrow & & \downarrow \psi \\
 \phi(\tilde{\mathcal{U}}) & \xrightarrow{\psi \circ f \circ \phi^{-1}} & \psi(\mathcal{V})
 \end{array}$$

où $\tilde{\mathcal{U}} = f^{-1}(\mathcal{V}) \cap \mathcal{U}$ pour que les compositions d'applications soient bien définies.

Remarque : Si $N = \mathbb{R}$ alors la définition ci-dessus est équivalente à demander que l'application continue $f \circ \phi^{-1}: \mathbb{R}^{n_M} \rightarrow \mathbb{R}$ soit de classe \mathcal{C}^k pour toute carte (\mathcal{U}, ϕ) de M . En particulier, la somme et le produit de deux fonctions numériques \mathcal{C}^k sur une variété est \mathcal{C}^k .

Proposition II.2

Toute composée d'applications \mathcal{C}^k est de classe \mathcal{C}^k .

Démonstration :

Soit P une troisième variété de dimension notée n_P , $f \in \mathcal{C}(M, N)$ et $g \in \mathcal{C}(N, P)$. Soit $m \in M$ et (\mathcal{U}, ϕ) , (\mathcal{V}, ψ) et (\mathcal{W}, χ) des cartes de M, N, P contenant respectivement $m, f(m)$ et $g \circ f(m)$. Alors la proposition découle trivialement du diagramme

suivant et de la propriété correspondante dans le cadre vectoriel :

$$\begin{array}{ccccc}
 \tilde{\mathcal{U}} & \xrightarrow{f} & \tilde{\mathcal{V}} & \xrightarrow{g} & W \\
 \phi \downarrow & & \psi \downarrow & & \chi \downarrow \\
 \phi(\tilde{\mathcal{U}}) & \xrightarrow{\psi \circ f \circ \phi^{-1}} & \psi(\tilde{\mathcal{V}}) & \xrightarrow{\chi \circ g \circ \psi^{-1}} & \chi(W)
 \end{array}$$

où $\tilde{\mathcal{U}} = \mathcal{U} \cap f^{-1}(\tilde{\mathcal{V}})$ et $\tilde{\mathcal{V}} = g^{-1}(W) \cap \mathcal{V}$ □

A partir de maintenant et sauf mention contraire, on supposera toutes les applications lisses pour ne pas s'embêter avec des questions de régularité mais on pourrait raffiner certains résultats obtenus pour des applications seulement \mathcal{C}^k , $k \in \mathbb{N}$.

2. Espace tangent

Notons \mathcal{C}_m^M l'ensemble des courbes lisses $\gamma: I \rightarrow M$ définies sur un intervalle ouvert I contenant 0 et tel que $\gamma(0) = m$.

Définition II.3. Deux courbes $\gamma_1: I_1 \rightarrow M$ et $\gamma_2: I_2 \rightarrow M$ de \mathcal{C}_m^M sont dites tangentes en m si $\gamma_1(0) = \gamma_2(0) = m$ et il existe une carte (\mathcal{U}, ϕ) telle que $m \in \mathcal{U}$ et

$$(\phi \circ \gamma_1)'(0) = (\phi \circ \gamma_2)'(0) \quad (\in \mathbb{R}^{n_M})$$

Intuitivement, cela revient à demander à ce que les courbes γ_1 et γ_2 aient la même vitesse. Par ailleurs, cette définition ne dépend pas de la carte (\mathcal{U}, ϕ) choisie. En effet si (\mathcal{V}, ψ) est une autre carte telle que $m \in \mathcal{V}$, on a :

$$(\psi \circ \gamma_i)'(0) = (\psi \circ \phi^{-1} \circ \phi \circ \gamma_i)'(0) = \underbrace{\mathcal{D}_{\phi(m)}(\psi \circ \phi^{-1})}_{\mathcal{C}^k\text{-difféo}} \cdot (\phi \circ \gamma_i)'(0) \quad (*)$$

On définit donc une relation d'équivalence sur \mathcal{C}_m^M ce qui amène à la définition :

Définition II.4. Un vecteur tangent à M en un point m de M est une classe d'équivalence de la relation ci-dessus. L'ensemble des vecteurs tangents à M en m forme l'espace tangent à M en m , on le note $T_m M$.

Exemple : Dans le cas dégénéré ou M est un espace vectoriel alors toute carte de M s'obtient par restriction de (M, Id) . En particulier, deux courbes $\gamma_1: I_1 \rightarrow M$ et $\gamma_2: I_2 \rightarrow M$ sont tangentes en un point $m := \gamma_i(0) \in M$ ssi $\gamma_1'(0) = \gamma_2'(0)$. Autrement dit, ssi les courbes γ_1 et γ_2 admettent la même tangente – même vecteur directeur $\gamma_i'(0)$ et même point d'application $\gamma_i(0)$ – au sens usuel du terme et le vecteur tangent n'est autre que le vecteur directeur de cette unique tangente.

Proposition II.5

$T_m M$ admet une structure d'espace vectoriel réel de dimension n_M .

Démonstration :

- i Commençons par montrer que $T_m M$ est en bijection avec \mathbb{R}^{n_M} .
Soit (\mathcal{U}, ϕ) une carte telle que $m \in \mathcal{U}$. Puisque la quantité $(\phi \circ \gamma)'(0)$ ne dépend que de la classe d'équivalence de γ , on peut définir l'application

$$\begin{aligned} \theta_\phi : T_m M &\rightarrow \mathbb{R}^{n_M} \\ \xi &\mapsto (\phi \circ \gamma)'(0) \end{aligned} .$$

Par définition, θ_ϕ est injective. Voyons qu'elle est surjective : tout vecteur $v \in \mathbb{R}^{n_M}$ est l'image par θ_ϕ de la classe d'équivalence de la courbe γ_v définie par $t \mapsto \phi^{-1}(tv)$. En effet :

$$(\phi \circ \gamma_v)'(0) = \frac{d}{dt} (t \mapsto tv)|_{t=0} = v$$

et donc θ_ϕ est bijective.

- ii Ensuite, on munit $T_m M$ d'une structure d'espace vectoriel.
D'après le calcul (\star) si (\mathcal{V}, ψ) est une autre carte telle que $m \in \mathcal{V}$ et $v \in \mathbb{R}^{n_M}$.

$$(\theta_\phi \circ \theta_\psi^{-1})(v) = \mathcal{D}_{\psi(m)}(\phi \circ \psi^{-1}) \cdot v .$$

Ainsi, par propriété de la différentielle, $\theta_\phi \circ \theta_\psi^{-1}$ est une application linéaire et on gagne donc "gratuitement" la compatibilité avec la structure d'espace vectoriel des opérations $+$ et \cdot définies par : $\forall \xi, \zeta \in T_m M, \forall \lambda \in \mathbb{R}$,

$$\xi + \zeta := \theta_\phi^{-1}(\theta_\phi(\xi) + \theta_\phi(\zeta)) \quad \text{et} \quad \lambda \cdot \xi := \theta_\phi^{-1}(\lambda \theta_\phi(\xi)) .$$

Par définition de ϕ , la construction ci-dessus ne dépend pas de ϕ . □

Enfin, on conclut ce paragraphe en définissant l'application tangente entre deux variétés :

Définition II.6. Soit $f : M \rightarrow N$ une application lisse et $m \in M$. On définit l'application linéaire tangente de f en m , notée $T_m f$, par passage au quotient de l'application $\gamma \mapsto f \circ \gamma$ de \mathcal{C}_m^M dans $\mathcal{C}_{f(m)}^N$.

Exemple : Reprenons l'exemple précédent en supposant que M et N sont des espaces vectoriels. Dans ce cas, l'application $T_m f$ est la différentielle $\mathcal{D}_m f$ de f en m .
En effet, dans ce cas, $(f \circ \gamma)'(0) = \mathcal{D}_m f \cdot \gamma'(0)$ d'après la formule de dérivation composée d'où le résultat en quotientant.

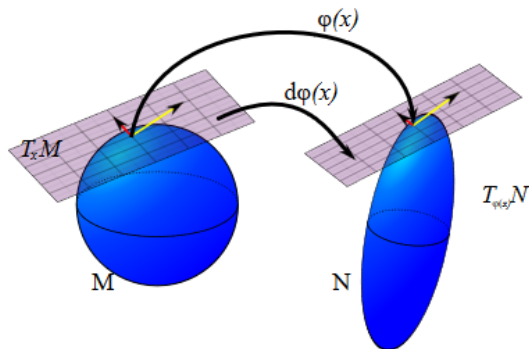


FIGURE II.3. – Vecteur tangent, plant tangent et application tangente.

Proposition II.7

Si $f: M \rightarrow N$ et $g: N \rightarrow P$ sont deux applications lisses entre variétés, alors :

$$\forall m \in M, \quad T_m(g \circ f) = T_{f(m)}g \circ T_m f .$$

Démonstration :

Ce résultat découle trivialement du résultat analogue dans des espaces vectoriels. Il suffit pour cela de se placer dans des cartes :

Soit $(U, \phi), (V, \psi)$ et (W, χ) des cartes de M, N et P respectivement et telles que $x \in U, f(x) \in V$ et $f \circ g(x) \in W$. On a le même diagramme commutatif qu'à la preuve de la proposition II.2 et le résultat s'en suit. □

3. Fibré tangent

Définition II.8. On appelle fibré tangent et on note TM l'union disjointe des espaces tangents à M en chacun de ses points.

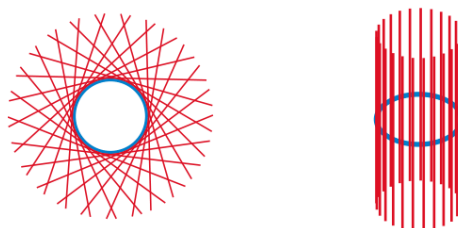


FIGURE II.4. – Deux représentations du fibré tangent du cercle.

Pour le moment, le fibré tangent ainsi défini est une somme ensembliste mais on peut le munir d'une topologie comme le précise la proposition suivante :

Proposition II.9

Si M est de classe \mathcal{C}^k , $k \geq 1$, alors son fibré tangent TM peut être muni d'une structure canonique de variété différentielle, de dimension $2n_M$ et de classe \mathcal{C}^{k-1} .

Démonstration :

On munit TM d'une topologie en procédant comme suit : Soit $(\mathcal{U}_i, \phi_i)_{i \in I}$ un atlas de M ,

- On définit pour tout $i \in I$ l'application

$$\begin{aligned} \Phi_i : \quad TU &\rightarrow \phi(\mathcal{U}_i) \times \mathbb{R}^{n_M} \\ (m, \xi) &\mapsto (\phi_i(m), T_m\phi_i \cdot \xi) \end{aligned}$$

où m est un point de \mathcal{U} et $\xi \in T_m\mathcal{U}$ un vecteur tangent associé, qui est en particulier bijective d'après la proposition II.5.

- On impose les conditions suivantes : pour tout $i \in I$,
 - i. $T\mathcal{U}_i$ est un ouvert de TM ,
 - ii. les applications Φ_i sont des homéomorphismes.

Ainsi, $\Omega \subset TM$ est ouvert ssi, pour tout $i \in I$, $\Phi_i(\Omega \cap T\mathcal{U}_i)$ est un ouvert de $\phi(\mathcal{U}_i) \times \mathbb{R}^{n_M}$.

On a alors plus précisément munit M d'un atlas $(T\mathcal{U}_i, \Phi_i)_{i \in I}$, ce qui en fait une variété différentiable de dimension $2n_M$.

En effet, puisque $(\mathcal{U}_i, \phi_i)_{i \in I}$ est un atlas et par définition de l'application tangente, pour tout i, j tels que $\mathcal{U}_i \cap \mathcal{U}_j \neq \emptyset$, l'application

$$\begin{aligned} \Phi_i \circ \Phi_j^{-1} : \quad \phi_j(\mathcal{U}_i \cap \mathcal{U}_j) \times \mathbb{R}^{n_M} &\rightarrow \phi_i(\mathcal{U}_i \cap \mathcal{U}_j) \times \mathbb{R}^{n_M} \\ (m, \xi) &\mapsto ((\phi_i \circ \phi_j^{-1})(m), T_m(\phi_i \circ \phi_j^{-1}) \cdot \xi) \end{aligned}$$

est un difféomorphisme.

Par ailleurs, le recourt à l'application tangente dans la définition des Φ_i implique que TM soit de classe \mathcal{C}^{k-1} . □

Enfin, l'appellation "fibré" est justifié par la proposition :

Proposition II.10

La projection canonique p de TM sur M est une fibration.

Démonstration :

Soit $(\mathcal{U}_i, \phi_i)_{i \in I}$ un atlas de M . On définit, pour tout i de I ,

$$\begin{aligned} \psi_i : p^{-1}(\mathcal{U}_i) = T\mathcal{U}_i &\rightarrow \mathcal{U}_i \times \mathbb{R}^{n_M} \\ \xi &\mapsto (m, T_m\phi_i \cdot \xi) \quad \text{où } m \text{ est tel que } \xi \in T_mM. \end{aligned}$$

Pour tout i de I , ψ_i est clairement un difféomorphisme et : $\forall m \in \mathcal{U}_i, \forall v \in \mathbb{R}^{n_M}$,

$$p(\psi_i^{-1}(m, v)) = p(\xi) = m \quad \text{où } \xi \in T_mM \text{ tel que } v = T_m\phi_i \cdot \xi$$

et donc (cf annexe A) TM est espace fibré, de base M et de fibre \mathbb{R}^{n_M} . \square

En fait, le fibré tangent est même un fibré vectoriel réel de rang n_M .

Définition II.11. Une variété dont le fibré tangent est trivialisable est dite parallélisable.

Exemple : Le cercle \mathbb{S}^1 est parallélisable comme l'illustre la figure II.4. A droite, on a trivialisé le fibré tangent.

III. Dérivation et champ de vecteurs

1. Dérivation ponctuelle

Soit $m \in M$. On définit une relation d'équivalence sur l'ensemble des fonctions définies sur un ouvert contenant m en posant :

$$(f: \mathcal{U} \rightarrow \mathbb{R}) \sim (g: \mathcal{V} \rightarrow \mathbb{R}) \quad \text{ssi} \quad \exists \mathcal{W} \text{ ouvert, } \mathcal{W} \subset \mathcal{U} \cap \mathcal{V}, \quad \text{tel que } f|_{\mathcal{W}} = g|_{\mathcal{W}}.$$

On appelle *germe* de la fonction f cette valeur commune, *ensemble des germes* l'espace quotient relatif à cette relation et on note \dot{f} le germe de f . L'ensemble des germes est muni d'une structure d'anneau : il suffit de définir l'addition et la multiplication de deux germes via leur représentant respectif. Enfin, on note \mathcal{F}_mM l'ensemble des germes en m des fonctions lisses sur la variété M . \mathcal{F}_mM est trivialement un sous-anneau de l'ensemble des germes.

Remarque : Soit $m \in M$. Alors, l'espace \mathcal{F}_mM est isomorphe à $\mathcal{F}_0\mathbb{R}^{n_M}$.

En effet, soit (\mathcal{U}, ϕ) une carte telle que $m \in \mathcal{U}$ et $\phi(m) = 0$. On associe à toute application f définie sur un voisinage \mathcal{V} de m une application $f \mapsto f \circ \phi^{-1}$ définie

sur $\phi(\mathcal{U} \cap \mathcal{V})$. En quotientant l'application précédente pour la relation "de germe", on obtient un isomorphisme – non canonique! – entre $\mathcal{F}_m M$ et $\mathcal{F}_0 \mathbb{R}^{n_M}$.

Définition II.12. Une dérivation ponctuelle en un point m de M est une application linéaire $\delta: \mathcal{F}_m M \rightarrow \mathbb{R}$ vérifiant l'identité de Leibniz. Autrement dit :

$$\forall f, g \in \mathcal{F}_m M, \quad \delta(f \times g) = f(m)\delta(g) + g(m)\delta(f).$$

On notera parfois $\mathcal{D}_m M$ l'ensemble des dérivations de $\mathcal{F}_m M$.

Lemme II.13.1 (Lemme de Hadamard). Soit f une fonction de classe \mathcal{C}^k , $k \in \mathbb{N} \setminus \{0\}$, définie sur \mathcal{V} , un voisinage ouvert de 0 dans \mathbb{R}^{n_M} . Alors, il existe n_M fonctions h_i de classe \mathcal{C}^{k-1} telles que

$$\forall x \in \mathcal{V}, \quad f(x) = f(0) + \sum_{i=1}^{n_M} x_i h_i(x)$$

De plus, pour tout i , $h_i(0) = \frac{\partial f}{\partial x_i}(0)$.

Démonstration :

D'après la formule de Taylor avec reste intégral,

$$\forall x \in \mathcal{V}, \quad f(x) = f(0) + \int_0^1 \frac{df}{dt}(tx) dt = f(0) + \sum_{i=1}^{n_M} x_i \int_0^1 \frac{\partial f}{\partial x_i}(tx) dt.$$

On pose alors, pour tout $i \in \llbracket 1, n_M \rrbracket$ $h_i := \int_0^1 \frac{\partial f}{\partial x_i}(t \cdot) dt$, qui est de classe \mathcal{C}^{k-1} d'après le théorème de dérivation sous le signe intégral.

Enfin, un calcul direct montre que :

$$\forall i \in \llbracket 1, n_M \rrbracket, \quad \forall x \in \mathcal{V}, \quad h_i(0) = \int_0^1 \frac{\partial f}{\partial x_i}(t \times 0) dt = \frac{\partial f}{\partial x_i}(0).$$

□

Théorème II.13

Toute dérivation δ de $\mathcal{F}_0 \mathbb{R}^{n_M}$ peut s'écrire sous la forme :

$$\forall f \in \mathcal{F}_0 \mathbb{R}^{n_M}, \quad \delta(f) = \sum_{i=1}^{n_M} \delta(x_i) \frac{\partial f}{\partial x_i}(0)$$

où (x_1, \dots, x_n) est l'écriture en coordonnées dans \mathbb{R}^{n_M} et f un représentant de f . Plus exactement, pour une telle dérivation δ , il existe un unique vecteur

$v \in \mathbb{R}^{n_M}$ tel que :

$$\forall \dot{f} \in \mathcal{F}_0 \mathbb{R}^{n_M}, \quad \delta(f) = T_0 f \cdot v = \sum_{i=1}^{n_M} \frac{\partial f}{\partial x_i}(0) v_i.$$

Ce théorème traduit en particulier l'existence d'une bijection naturelle entre \mathbb{R}^{n_M} et les dérivations de $\mathcal{F}_0 \mathbb{R}^{n_M}$.

Démonstration :

Soit $\dot{f} \in \mathcal{F}_0 \mathbb{R}^{n_M}$ et f un représentant de \dot{f} . Notons $\mathbb{1}$ l'application \mathcal{C}^∞ de \mathbb{R}^n dans $\{1\}$ définie par $x \mapsto 1$. Par définition de la dérivation,

$$\delta(\mathbb{1}) = \delta(\mathbb{1} \times \mathbb{1}) = 1 \times \delta(\mathbb{1}) + 1 \times \delta(\mathbb{1}) = 2\delta(\mathbb{1}) \quad \text{d'où} \quad \delta(\text{cste}) = 0.$$

Donc $\delta(f - f(0)) = \delta(f) - \delta(f(0)) = \delta(f)$ et, d'après le lemme de Hadamard,

$$\delta(f) = \delta\left(\sum_{i=1}^{n_M} x_i h_i\right) = \sum_{i=1}^{n_M} \delta(x_i) h_i(0) = \sum_{i=1}^{n_M} \delta(x_i) \frac{\partial f}{\partial x_i}(0)$$

car $\forall i \in \llbracket 1, n_M \rrbracket$, $\delta(x_i h_i) = 0 \times \delta(h_i) + \delta(x_i) \times h_i(0)$.

Enfin, de manière évidente, l'application $f \mapsto \frac{\partial f}{\partial x_i}(0)$ ne dépend, pour tout i , que du germe \dot{f} de f . Il s'agit également d'une dérivation. \square

Enfin, ce résultat s'étend directement aux variétés et permet de donner une nouvelle définition de l'espace tangent qui ne fait pas appel aux cartes.

Théorème II.14

Soit $m \in M$. Alors, pour toute dérivation δ de $\mathcal{F}_m M$, il existe un unique $\xi \in T_m M$ tel que

$$\delta(f) = T_m f \cdot \xi := \mathcal{L}_\xi(f)$$

où \mathcal{L}_ξ désigne la dérivée de Lie dans la direction ξ .

En particulier, l'ensemble des dérivations sur $\mathcal{F}_m M$ est isomorphe à $T_m M$ et un isomorphisme est donné par l'application dérivée de Lie $\xi \mapsto \mathcal{L}_\xi$.

Démonstration :

Soit (\mathcal{U}, ϕ) une carte de M telle que $m \in \mathcal{U}$ et $\phi(m) = 0$. D'après la remarque précédente, si δ est une dérivation de $\mathcal{F}_m M$ alors l'application $g \mapsto \delta(g \circ \phi)$ est une dérivation de $\mathcal{F}_0 \mathbb{R}^{n_M}$.

D'après le théorème II.13, il existe un unique vecteur $v \in \mathbb{R}^{n_M}$ tel que

$$\delta(g \circ \phi) = T_0 g \cdot v$$

Alors, si $f = g \circ \phi$,

$$\delta(f) = \delta(g \circ \phi) = T_0g \cdot v = T_mf \cdot T_0\phi^{-1} \cdot v.$$

Or, $T_0\phi^{-1} \cdot v = \theta_\phi^{-1}(v) = \xi$ donc $\delta(f) = T_mf \cdot \xi = \mathcal{L}_\xi$. \square

Autrement dit, à tout point m de M on peut associer, de manière équivalente, une dérivation locale et/ou un vecteur tangent.

2. Dérivation globale

Définition II.15. On appelle champ de vecteurs sur une variété une section (lisse) de son fibré tangent.

On note $\mathcal{C}^\infty(E)$ l'ensemble des sections lisses d'un fibré vectoriel E^1 (cf annexe A.II) et donc $\mathcal{C}^\infty(TM)$ – parfois, on note aussi $\Gamma(TM)$ – l'ensemble des champs de vecteurs sur M . On peut munir $\mathcal{C}^\infty(E)$ d'une structure d'espace vectoriel en posant :

$$\forall s, t \in \mathcal{C}^\infty(E), \quad \forall x \in M, \quad (s +_{\mathcal{C}^\infty(E)} t)(x) = s(x) +_{E_x} t(x)$$

où l'opération $+_{E_x}$ est bien définie de par la structure d'espace vectoriel des fibres E_x .

Exemple : Dans le cas où M est un espace vectoriel, son fibré tangent est trivial et en particulier l'ensemble des sections de TM sont de la forme $m \in M \mapsto (m, f(m)) \in M \times \mathbb{R}^{n_M}$ et s'identifient aux applications de classe \mathcal{C}^∞ de M dans \mathbb{R}^{n_M} : $\mathcal{C}^\infty(M, \mathbb{R}^{n_M})$.

La définition précédente est donc cohérente avec le cadre vectoriel.

Par la suite, si $X \in \mathcal{C}^\infty(TM)$ est un champ de vecteurs sur M , on note X_x sa valeur en un point x .

Définition II.16. On appelle dérivation – éventuellement globale – sur une variété M une application linéaire de $\mathcal{C}^\infty(TM)$ dans lui-même vérifiant l'identité de Leibniz :

$$\forall f, g \in \mathcal{C}^\infty(TM), \quad \delta(f \cdot g) = f \cdot \delta(g) + g \cdot \delta(f).$$

On note $\mathcal{D}(M)$ l'ensemble des dérivations de M .

Le but de ce paragraphe est d'établir le pendant du théorème II.14 dans le cas des dérivations globales et des champs de vecteurs. Pour cela on va avoir besoin de la notion de partition de l'unité que l'on rappelle ici :

Définition II.17. Soit $(\mathcal{U}_i)_{i \in I}$ un recouvrement ouvert localement fini de M . On appelle partition de l'unité subordonnée au recouvrement (\mathcal{U}_i) , de classe \mathcal{C}^k , une famille de fonctions $(\alpha_i)_{i \in I}$ de classe \mathcal{C}^k , définies sur M et à valeurs dans $[0, 1]$ telles que :

i Pour tout $i \in I$, le support de α_i est inclus dans \mathcal{U}_i ;

1. On rappelle que d'après la proposition II.10, TM est un espace fibré de base M et de fibre \mathbb{R}^{n_M} .

- ii Pour tout $m \in M$, il existe un voisinage \mathcal{V} de m tel que toutes les fonctions α_i soient nulles sur \mathcal{V} à l'exception d'un nombre fini d'entre elles ;
- iii Pour tout $m \in M$, la somme de toutes les valeurs prises par les fonctions α_i est égale à 1.

Théorème II.18 (Partition de l'unité)

Soit M une variété différentiable et $(\mathcal{U}_i)_{i \in I}$ un recouvrement d'ouverts de M – par exemple un atlas – localement fini. Alors, il existe une partition lisse de l'unité subordonnée au recouvrement (\mathcal{U}_i) .

Ce résultat, bien que d'une apparente simplicité, est, dès lors que M n'est pas métrisable, non-trivial et on l'admet donc. En particulier, l'existence d'une distance sur M est loin d'être évidente – c'est même ce que l'on veut/va prouver au théorème III.3 – et on ne peut recourir à cet argument ici, sauf dans le cas d'une variété dégénérée².

On associe à tout champ de vecteurs X de M l'application

$$L_X : \mathcal{C}^\infty(M) \rightarrow \mathcal{C}^\infty(M)$$

$$f \mapsto (L_x f : x \mapsto L_X f(x) = T_x f \cdot X_x)$$

dite *dérivation associée à X* . Les résultats suivants justifient cette appellation.

Lemme II.19.1 (Localisation des dérivations). Soit δ une dérivation sur M .

- i. Si f et g sont deux fonctions lisses telles que $f|_{\mathcal{U}} = g|_{\mathcal{U}}$, alors $(\delta f)|_{\mathcal{U}} = (\delta g)|_{\mathcal{U}}$.
- ii. Pour tout ouvert $\mathcal{U} \subset M$, il existe une unique dérivation sur \mathcal{U} , notée $\delta|_{\mathcal{U}}$, telle que pour tout $f \in \mathcal{C}^\infty(M)$, $\delta|_{\mathcal{U}}(f|_{\mathcal{U}}) = (\delta f)|_{\mathcal{U}}$.

Démonstration :

On commence par montrer i. utile à la démonstration de ii. :

- i. Par linéarité de δ , il suffit donc de montrer que pour toute fonction f de $\mathcal{C}^\infty(M)$, telle que $f \equiv 0$ sur \mathcal{U} alors $\delta f \equiv 0$ sur \mathcal{U} .

Soit une telle fonction f et $x \in \mathcal{U}$. Soit également un ouvert \mathcal{V}_x de \mathcal{U} tel que $x \in \mathcal{V}_x$ et $\overline{\mathcal{V}_x} \subset \mathcal{U}$ et une fonction plateau h à support dans \mathcal{U} et valant 1 sur \mathcal{V}_x . On vérifie alors aisément que $f = (1 - h)f$ et donc

$$\delta f = (1 - h)\delta f + f\delta(1 - h).$$

2. Dans ce cas, $M \simeq \mathbb{R}^n$ et il suffit donc de poser $\forall i \in I, F_i := \mathbb{R}^n \setminus \mathcal{U}_i$ et

$$\forall i \in I, \forall x \in \mathbb{R}^n, \quad \alpha_i(x) := \frac{d(x, F_i)}{\sum_{j \in I} d(x, F_j)} \quad \text{où} \quad d(x, F) := \min_{y \in F} \|x - y\|$$

Par suite, δf est nulle sur \mathcal{V}_x et ceci étant vrai pour tout x il vient que δf est nulle sur \mathcal{U} .

- ii. Soit $x \in \mathcal{U}$ et \mathcal{V}_x et h comme ci dessus : $x \in \mathcal{V}_x$, $\overline{\mathcal{V}_x} \subset \mathcal{U}$ et h à support dans \mathcal{U} valant 1 sur \mathcal{V}_x . Si $f \in \mathcal{C}^\infty(\mathcal{U})$ alors fh définit donc une fonction lisse sur tout M et valant f sur \mathcal{V}_x . On pose alors $\delta_{|\mathcal{U}}f := \delta(fh)$, défini sans ambiguïté sur \mathcal{V}_x d'après i. On procède de même pour tout $x \in \mathcal{U}$ et finalement, on a défini une dérivation $\delta_{|\mathcal{U}}$ telle que :

$$\forall f \in \mathcal{C}^\infty(M), \quad \delta_{|\mathcal{U}}(f|_{\mathcal{U}}) = \delta(f|_{\mathcal{U}}h) = (\delta f)|_{\mathcal{U}}$$

où h est une somme de fonctions plateaux et où la dernière égalité vient du fait que $\delta(\text{cste}) = 0$.

□

Théorème II.19

L'application $L : X \mapsto L_X$ est une bijection de $\mathcal{C}^\infty(TM)$ sur l'ensemble des dérivations de M .

Démonstration :

L'idée de la preuve est une idée générale lorsque l'on travaille avec des variétés et on la ré-exploitera dans la démonstration du théorème III.3 : on se ramène grâce à un atlas de la variété au cas des espaces vectoriels où les démonstrations sont généralement plus aisées puis on vérifie que le "recollement" se passe bien.

- i. Supposons que $M = \mathcal{U}$ est un ouvert de \mathbb{R}^{n_M} .

Dans ce cas, un champ de vecteurs sur \mathcal{U} est, comme vu précédemment, simplement une application \mathcal{C}^∞ de \mathcal{U} dans \mathbb{R}^{n_M} .

- L est injective : Soit X un champ de vecteurs sur \mathcal{U} et $a \in \mathcal{U}$ tel que $X(a) \neq 0$. Alors,

$$L_X Id(a) = T_x Id \cdot X_a = \sum_{i=0}^{n_M} X_i(a) \neq 0$$

et donc, par contraposée, si pour toute fonction lisse f $L_X f \equiv 0$ alors $X \equiv 0$ et donc L injective.

- L est surjective : Soit δ dérivation sur \mathcal{U} et $f \in \mathcal{C}^\infty(\mathcal{U})$. Commençons par remarquer que le lemme de Hadamard n'utilise que la convexité de \mathbb{R}^{n_M} et s'étend donc facilement à n'importe quel convexe de \mathbb{R}^{n_M} . En particulier, pour toute boule $\mathcal{B} \subset \mathcal{U}$ et tout $y \in \mathcal{B}$, le lemme de Hadamard (étendu) appliqué à la fonction $x \mapsto f(x) - f(y)$ donne :

$$\forall x \in \mathcal{B}, \quad f(x) = f(y) + \sum_{i=1}^{n_M} (x_i - y_i) \frac{\partial f}{\partial x_i}(y).$$

De même que dans le cas des dérivations locales, la dérivation d'une constante est nulle et donc, en notant f la restriction de f à \mathcal{B} :

$$\delta_{|\mathcal{B}}f(y) = \delta_{|\mathcal{B}}(f - f(y))(y) = \sum_{i=1}^{n_M} \delta_{|\mathcal{B}}(x_i - y_i) \frac{\partial f}{\partial x_i}(y) = T_y f \cdot Y_y$$

où Y est le champ de vecteurs défini par $Y_y^i := \delta_{|\mathcal{B}}(x_i)(y)$. D'où la surjectivité de L en considérant un recouvrement de \mathcal{U} et le lemme II.19.1

ii. Sinon : on se ramène au cas précédent comme annoncé.

- L est injective : Soit X un champ de vecteurs sur M et $a \in M$ tel que $X_a \neq 0$. Soit (\mathcal{U}, ϕ) une carte telle que $a \in \mathcal{U}$. Puisque $X_a \neq 0$, le vecteur $T_a \phi \cdot X_a$ est non-nul et admet en particulier au moins une coordonnée non-nulle que l'on note k et soit donc $f := \phi_k$ la $k^{\text{ème}}$ composante de ϕ . Soit $\mathcal{V} \subset \mathcal{U}$ un ouvert contenant a tel que $\bar{\mathcal{V}} \subset \mathcal{U}$ et h une fonction lisse à support dans \mathcal{U} et valant à 1 sur \mathcal{V} . Alors, $fh \in \mathcal{C}^\infty(\mathbb{R}^{n_M})$ et

$$L_X(fh)(a) = L_X f(a) = T_a f \cdot X_a = T_a \phi_k \cdot X_a = (T_a \phi \cdot X_a)_k \neq 0$$

et le résultat s'en suit par contraposée.

- L est surjective : Soit (\mathcal{U}_i, ϕ_i) un atlas de M et δ une dérivation sur M . D'après le lemme II.19.1, δ induit en particulier une dérivation sur chacun des ouverts \mathcal{U}_i . Remarquons que l'application

$$h \mapsto \phi_i \circ \left(\delta(h \circ \phi_i^{-1}) \right)$$

est une dérivation de $\phi_i(\mathcal{U}_i) \subset \mathbb{R}^{n_M}$ et on sait d'après ce qui précède qu'il existe un champ de vecteurs Y^i sur $\phi_i(\mathcal{U}_i)$ tel que

$$\phi_i \circ \left(\delta(h \circ \phi_i^{-1}) \right) = L_{Y^i} h.$$

Par ailleurs, $\tilde{\delta} : h \mapsto \phi_i^{-1} \circ \delta(h) \circ \phi_i$ est une dérivation de M telle que

$$\tilde{\delta}_{|\mathcal{U}_i}(h) = \left(L_{Y^i}(h \circ \phi_i) \right) \circ \phi_i^{-1} := L_{X^i} h$$

en posant $X^i := \phi_*^{-1} Y^i$ (cf remarque suivante).

Enfin, si $\mathcal{U}_i \cap \mathcal{U}_j \neq \emptyset$, les dérivations L_{X^i} et L_{X^j} induisent la même dérivation sur $\tilde{\delta}_{|\mathcal{U}_i \cap \mathcal{U}_j}$ et par injectivité de l'application L , les champs X^i et X^j sont donc égaux que $\mathcal{U}_i \cap \mathcal{U}_j$. Par suite, on peut définir proprement X telle que $L_X = \tilde{\delta}$; autrement dit, L est surjective. □

Remarque : Soit $\phi: \mathcal{U} \rightarrow \mathcal{V}$ un difféomorphisme entre ouverts de \mathbb{R}^{n_M} et δ une dérivation sur \mathcal{U} . Alors l'image de δ par ϕ est la dérivation sur \mathcal{V} définie par :

$$g \mapsto (\delta(g \circ \phi)) \circ \phi^{-1}.$$

Si X est le champ de vecteurs associé à δ , on note ϕ_*X le champ de vecteurs associé à l'image de δ . On dit alors que ϕ_*X est l'image de X par ϕ .

En particulier :

$$\forall f \in \mathcal{C}^\infty(\mathcal{V}), \quad L_{\phi_*X}f = (L_X(f \circ \phi)) \circ \phi^{-1}$$

et

$$\forall y \in \mathcal{V}, \quad (\phi_*X)_y = T_{\phi^{-1}(y)}\phi \cdot X_{\phi^{-1}(y)}.$$

D'après le théorème des fonctions implicites :

$$\forall y \in \mathcal{V}, \quad \delta(g \circ \phi)(\phi^{-1}(y)) = T_{\phi^{-1}(y)}(g \circ \phi) \cdot X_{\phi^{-1}(y)} = T_y g \cdot T_{\phi^{-1}(y)}\phi \cdot X_{\phi^{-1}(y)}.$$

Ainsi, de même qu'il existait un point de vue double sur les vecteurs tangents, on peut voir les champs de vecteurs comme des dérivations. Ce point de vue justifie l'abus de notation classique suivant : Soit (\mathcal{U}, ϕ) une carte de M . On notera encore $\frac{\partial}{\partial x_i}$ – ou parfois dx_i – le champ de vecteurs sur \mathcal{U} image par ϕ^{-1} du champ de vecteurs $\frac{\partial}{\partial x_i}$ sur $\phi(\mathcal{U})$. Cette écriture, choix de la carte dépendante, permet d'écrire tout champ de vecteurs X sur \mathcal{U} de manière unique comme :

$$X = \sum_{i=1}^{n_M} X^i \frac{\partial}{\partial x_i} \quad \text{où} \quad X^i \in \mathcal{C}^\infty(\mathcal{U}).$$

En particulier, on récupère une écriture en terme de coordonnées des champs de vecteurs sur une variété.

L'ensemble des dérivations $\mathcal{D}(M)$ est naturellement muni d'une structure linéaire. Néanmoins, la composée de deux dérivations n'est pas nécessairement une dérivation. En effet, soit δ et δ' deux dérivations sur M . Alors :

$$\begin{aligned} \forall f, g \in \mathcal{C}^\infty(M), \quad \delta\delta'(fg) &= \delta(f.\delta'(g) + \delta'(f).g) \\ &= \delta(f).\delta'(g) + f.\delta\delta'(g) + \delta\delta'(f).g + \delta'(f).\delta(g). \end{aligned}$$

Mais, on a le résultat suivant :

Lemme II.20.1. Si δ et δ' sont deux dérivations sur M alors $\delta \circ \delta' - \delta' \circ \delta$ est une dérivation sur M .

Démonstration :

Soit $f, g \in \mathcal{C}^\infty(M)$. Alors, d'après le calcul précédent,

$$(\delta\delta' - \delta'\delta)(fg) = f \cdot (\delta\delta' - \delta'\delta)(g) + (\delta\delta' - \delta'\delta)(f) \cdot g$$

et $\delta\delta' - \delta'\delta$ est bien évidemment linéaire. \square

Définition II.20. Le crochet de deux champs de vecteurs X et Y , noté $[X, Y]$ est le champ de vecteurs associé à la dérivation $L_X L_Y - L_Y L_X$.

Proposition II.21

Soit X et Y deux champs de vecteurs sur M s'écrivant dans des coordonnées locales

$$X = \sum_{i=1}^{n_M} X^i \frac{\partial}{\partial x_i} \quad \text{et} \quad Y = \sum_{i=1}^{n_M} Y^i \frac{\partial}{\partial x_i}.$$

Alors,

$$[X, Y] = \sum_{i=1}^{n_M} Z^i \frac{\partial}{\partial x_i} \quad \text{où} \quad Z^i = \sum_{j=1}^{n_M} \left(X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j} \right).$$

Démonstration :

Pour toute fonction lisse f ,

$$L_X L_Y f = L_X \left(\sum_{i=1}^{n_M} Y^i \frac{\partial f}{\partial x_i} \right) = \sum_{i,j=1}^{n_M} \left(X^j \frac{\partial Y^i}{\partial x_j} \frac{\partial f}{\partial x_i} + X^j \frac{\partial^2 f}{\partial x_i \partial x_j} \right)$$

et le résultat découle trivialement du lemme de Schwarz. \square

Lemme II.22.1 (Lemme de Jacobi). Pour tout champs de vecteurs X, Y et Z sur M ,

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

Démonstration :

La preuve découle d'un calcul algébrique sur les dérivations correspondantes, sans réelle difficulté mais fastidieux. \square

– Chapitre III –

Cadre riemannien

Dans tout ce qui suit M désigne une variété différentiable de dimension n .

I. Métrique riemannienne

Pour tout $m \in M$, $T_m M$ étant un espace vectoriel, on peut le munir d'un produit scalaire $\langle \cdot | \cdot \rangle_m$. Ainsi, on peut espérer transporter la structure courbe de la variété M sur l'espace linéaire $T_m M$ et récupérer la notion de distance que l'on connaît bien. La notion de variété riemannienne repose sur cette idée.

On rappelle que si $X: \mathcal{C}^\infty(M) \rightarrow \mathcal{C}^\infty(M)$ est un champ de vecteurs sur M alors l'application $X_m: \mathcal{C}^\infty(M) \rightarrow \mathbb{R}$ définie par

$$\forall \phi \in \mathcal{C}^\infty(M), \quad X_m(\phi) = (X(\phi))(m)$$

est un vecteur tangent, *i.e* $X_m \in T_m M$.

Définition III.1. Une structure riemannienne sur M est la donnée, en chaque point m de M d'un produit scalaire $\langle \cdot | \cdot \rangle_m$ de $T_m M$ tel que, pour tout champ de vecteurs X , Y de M l'application $m \mapsto \langle X_m | Y_m \rangle_m$ est une fonction de classe \mathcal{C}^k , $k \in \mathbb{N}$.

Ici, on prend $k = +\infty$ pour ne pas s'embêter avec des questions de régularité et on note classiquement $\|\cdot\|_m$ la norme issue du produit scalaire $\langle \cdot | \cdot \rangle_m$.

Il faut bien comprendre que la construction est locale. Plus précisément : soit (\mathcal{U}, ϕ) une carte de M dans laquelle tout point de \mathcal{U} se décompose en les coordonnées $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$. Comme vu précédemment, tout vecteur tangent en $m \in \mathcal{U}$ s'exprime comme combinaison linéaire des $\partial_{x_i}(m) := \phi^{-1} \left(\frac{\partial \phi}{\partial x_i}(m) \right)$ où $\frac{\partial}{\partial x_i}$ désigne la dérivée par rapport au $i^{\text{ème}}$ vecteur de base dans \mathbb{R}^n , *i.e* $\forall \xi, \zeta \in T_m M$,

$$\exists (\xi_i)_{i \in \llbracket 1, n \rrbracket}, (\zeta_i)_{i \in \llbracket 1, n \rrbracket}, \quad \xi = \sum_{i=1}^n \xi_i \partial_{x_i}(m) \quad ; \quad \zeta = \sum_{i=1}^n \zeta_i \partial_{x_i}(m).$$

La métrique en m est alors donnée par :

$$g_m(\xi, \zeta) = \sum_{i,j=1}^n g_{i,j}(m) \xi_i \zeta_j \quad \text{avec} \quad g_{i,j}(m) = g(\partial_{x_i}(m), \partial_{x_j}(m)).$$

Autrement dit :

$$g = \sum_{i,j=1}^n g_{i,j} \partial_{x_i} \otimes \partial_{x_j} \quad \text{noté abusivement par la suite} \quad \sum_{i,j=1}^n g_{i,j} \partial_{x_i} \partial_{x_j}.$$

Remarquons que la construction ci-dessus est équivalente à se donner, pour tout $m \in \mathcal{U}$, une matrice symétrique définie positive g_m ¹, dont les coefficients sont des fonctions \mathcal{C}^∞ de m telle que :

$$\forall \xi, \zeta \in T_m M, \quad \langle \xi \mid \zeta \rangle_m = {}^t \xi g_m \zeta.$$

Définition III.2. On appelle métrique riemannienne et on note g l'application

$$g : M \rightarrow \bigsqcup_{m \in M} (T_m M \times T_m M)^{\mathbb{R}}$$

$$m \mapsto g_m : (\xi, \zeta) \mapsto \langle \xi \mid \zeta \rangle_m \quad \text{tel que} \quad \xi, \zeta \in T_m M.$$

Théorème III.3

Pour toute variété M , il existe au moins une métrique riemannienne.

La preuve du théorème précédent utilise la notion de partition de l'unité rappelée au paragraphe III.2.

Démonstration :

L'idée de la preuve est simple : construire la métrique sur des cartes locales puis étendre la construction à M dans sa globalité via une partition de l'unité.

Soit $(\mathcal{U}_i, \phi_i)_{i \in I}$ un atlas de M tel que (\mathcal{U}_i) soit un recouvrement localement fini de M et $(\alpha_i)_{i \in I}$ une partition de l'unité subordonnée. Comme vu précédemment, on peut définir sur chacun des \mathcal{U}_i une métrique riemannienne que l'on note g^i .

Posons $g_m := \sum_{i \in I} \alpha_i g_m^i$ et montrons que g est bien une métrique riemannienne sur tout M . g est trivialement lisse en la variable m , bilinéaire, symétrique ; reste donc à montrer qu'elle est définie positive. Soit $m \in M$. Il existe au moins un $j \in I$ tel que $\alpha_j(m) \neq 0$ et donc, par définition de g^j :

$$\forall \xi \in T_m M \setminus \{0\}, \quad g_m(\xi, \xi) = \sum_{i \in I} \alpha_i(m) g_m^i(\xi, \xi) \geq \alpha_j(m) g_m^j(\xi, \xi) > 0.$$

□

1. En confondant matrice et application sous-jacente. . .

Enfin, la proposition suivante généralise la formule de changement de base pour les formes quadratiques dans le cas des espaces euclidiens :

Proposition III.4

Soit (\mathcal{U}, ϕ) et (\mathcal{U}', ϕ') deux cartes locales de M dans lesquelles les coordonnées locales s'écrivent respectivement (x_i) et (x'_i) et où est la métrique est donnée par $(g_{i,j})$ et $(g'_{i,j})$. Alors

$$(g'_{i,j}) = {}^t\Phi^{-1}(g_{i,j})\Phi^{-1}$$

où Φ désigne la matrice jacobienne de la fonction de transition $\phi' \circ \phi^{-1}$.

Démonstration :

Il suffit de remarquer que par définition de g et g' ,

$$g' \left(\frac{\partial}{\partial x'_i}, \frac{\partial}{\partial x'_j} \right) = \sum_{k,l=1}^n \frac{\partial x_k}{\partial x'_i} \frac{\partial x_l}{\partial x'_j} g \left(\frac{\partial}{\partial x_k}, \frac{\partial}{\partial x_l} \right).$$

D'où le résultat par changement de variable. □

On est ainsi à même de remplir notre objectif, *i.e* de mesurer des déplacements le long de la variété M :

Définition III.5. Soit $\gamma : [0, a] \rightarrow M$ une courbe continue et \mathcal{C}^1 par morceaux : soit en particulier $(0 = a_0 < a_1 < \dots < a_n = a)$ la partition de $[0, a]$ telle que γ soit \mathcal{C}^1 sur chacun des $]a_k, a_{k+1}[$. On définit alors la longueur de γ par :

$$l(\gamma) = \sum_{i=0}^{n-1} \int_{a_i}^{a_{i+1}} \left\| \frac{d\gamma}{dt}(\gamma(t)) \right\|_{\gamma(t)} dt.$$

En particulier, si γ est \mathcal{C}^1 sur $]0, a[$, $l(\gamma) = \int_0^a \left\| \frac{d\gamma}{dt}(\gamma(t)) \right\|_{\gamma(t)} dt$.

Exemple : On se place sur \mathbb{R}^2 munit de sa structure euclidienne. Soit $\gamma : [0, \pi] \rightarrow \mathbb{R}^2$ définie par $\gamma(t) = e^{it}$, autrement dit γ parcourt \mathbb{S}^1 dans le sens trigonométrique. Dans une carte évidente on peut écrire γ comme $\gamma(t) = (\cos(t), \sin(t))$ d'où :

$$\langle \gamma'(t) \mid \gamma'(t) \rangle_{\gamma(t)} = \sin^2 t + \cos^2 t = 1 \quad \text{et} \quad l(\gamma) = \int_0^\pi dt = \pi$$

qui est le résultat attendu. Cette définition semble donc, de premier abord, cohérente.

II. Connexion de Levi-Civita

Nous avons dans les sections précédentes défini des espaces tangents en chacun des points de notre variété. On aimerait à présent pouvoir les relier les uns aux autres afin de pouvoir étudier des champs de vecteurs tangents comme des objets "continus" et notamment les dériver, dans un sens à préciser.

1. Approche intuitive et cas euclidien

Dans le cas euclidien, la problématique est simple : dans ce cas, $M \equiv \mathbb{R}^n$ s'identifie à "ses" propres espaces tangents en chacun de ses points et on passe d'un espace tangent à l'autre par une simple translation (linéaire). Commençons par détailler l'affirmation précédente :

Soit Y un champ de vecteurs sur \mathbb{R}^n , autrement dit une application lisse de \mathbb{R}^n dans lui-même, m_0 un point de \mathbb{R}^n et ξ un vecteur de \mathbb{R}^n . Pour étudier la dérivée du champ Y , en m_0 et dans la direction de ξ on forme le quotient

$$\lim_{t \rightarrow 0} \frac{Y(m_0 + t\xi) - Y(m_0)}{t} = \mathcal{D}_{m_0} Y \cdot \xi.$$

Par définition du fibré tangent sur \mathbb{R}^n , pour tout réel t , $Y(m_0 + t\xi) \in T_{m_0 + t\xi} \mathbb{R}^n$ et $Y(m_0) \in T_{m_0} \mathbb{R}^n$ et en particulier ces deux quantités ne vivent pas dans le même espace, on ne devrait pas pouvoir les additionner. Mais, dans le cas euclidien, on peut identifier de manière canonique l'espace tangent $T_m \mathbb{R}^n$ à $T_0 \mathbb{R}^n \equiv \mathbb{R}^n$ pour tout m de \mathbb{R}^n . Il suffit pour cela d'envoyer le vecteur tangent $\frac{\partial}{\partial x_i}$ en x sur le vecteur tangent éponyme en 0, par translation.

Dans le cas dégénéré ou $M \equiv \mathbb{R}^n$, "passer" d'un espace tangent à l'autre ne pose pas de réel problème et on peut même compliquer légèrement la chose en considérant non plus un unique vecteur $\xi \in \mathbb{R}^n$ mais un second champ de vecteurs X . La dérivée de Y en m_0 et dans la direction X_{m_0} est alors donnée par $\mathcal{D}_{m_0} Y \cdot X_{m_0}$ et, en laissant m_0 parcourir \mathbb{R}^n on définit une application dite *dérivée directionnelle* de Y par rapport à X

$$\nabla_Y(X) : \begin{array}{ccc} \mathbb{R}^n & \rightarrow & \mathbb{R}^n \\ m & \mapsto & \mathcal{D}_m Y \cdot X_m \end{array}.$$

En revanche, dans le cas général, l'approche précédente ne fonctionne plus, ne serait-ce que parce que les différents espaces tangents ne s'identifient plus aussi facilement. On a besoin pour cela d'introduire la notion de connexion.

2. Cas général

Dans le cas général, du fait de la "courbure" de la variété, on ne peut plus se contenter de faire translater les différents espaces tangents pour les identifier les uns aux autres.

Intuitivement, il s'agit de les faire "rouler" d'un point à l'autre comme l'illustre la figure III.1.

Formellement, on a la définition suivante :

Définition III.6. Une connexion sur une variété (lisse) M est une application bilinéaire ∇ de $\mathcal{C}^\infty(TM) \times \mathcal{C}^\infty(TM)$ dans $\mathcal{C}^\infty(TM)$ telle que :
 $\forall X, Y \in \mathcal{C}^\infty(TM)$ et $f \in \mathcal{C}^\infty(M)$,

$$\nabla_{fX}Y = f\nabla_XY \quad \text{et} \quad \nabla_X(fY) = (X \cdot f)Y + f\nabla_XY$$

où $X \cdot f$ désigne la dérivée directionnelle de f dans la direction X est définie comme précédemment par :

$$\begin{aligned} X \cdot f : M &\rightarrow M \\ m &\mapsto \mathcal{D}_m f \cdot X_m \end{aligned}$$

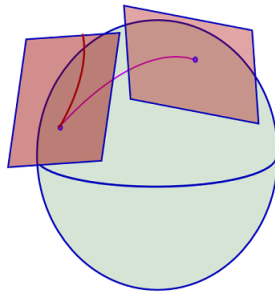


FIGURE III.1. – Connexion affine sur la sphère

On parle également de *dérivée covariante*. ∇ est en particulier un champ de vecteurs et le même argument de fonction plateau qu'au lemme II.19.1 montre que, à l'instar des dérivations, les connexions se localisent. On a :

$$(\nabla_X Y)|_{\mathcal{U}} = \nabla_{X|_{\mathcal{U}}} Y|_{\mathcal{U}}$$

pour tout ouvert \mathcal{U} . De plus, la première propriété implique que la valeur en m de $\nabla_X Y$ pour un champ donné Y ne dépend que de X_m ², tandis que la seconde est un avatar de la formule de Leibniz.

On vérifie enfin que cette définition est cohérente avec le paragraphe précédent : les deux propriétés sont satisfaites par la différentielle d'un champ de vecteurs et dans ce

2. Pour le montrer, on procède également de manière équivalente à au lemme II.19.1 avec des fonction test judicieusement choisies

cas, $\nabla_Y X$ est l'application définie précédemment.

En effet, pour toute fonction $f \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R})$ et tout $m \in \mathbb{R}^n$, par linéarité

$$\nabla_{fX} Y(m) = \mathcal{D}_m Y \cdot (fX)_m = f(m) \mathcal{D}_m Y \cdot X_m$$

et

$$\nabla_X (fY)(m) = \mathcal{D}_m (fY) \cdot X_m = \mathcal{D}_m f \cdot X_m Y_m + f(m) \mathcal{D}_m Y \cdot X_m .$$

Définition III.7. On définit la torsion de la connexion ∇ par

$$\forall X, Y \in \mathcal{C}^\infty(TM), \quad T^\nabla(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y].$$

Si ∇ est de torsion nulle, on dit qu'elle est symétrique ou sans torsion.

| **Remarque :** On peut montrer que la torsion est un tenseur.

Théorème III.8

Il existe, pour toute variété riemannienne une unique connexion symétrique consistante avec la métrique, *i.e* telle que pour tout champs de vecteurs X, Y, Z ,

$$Z \cdot g(X, Y) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y) .$$

On dit aussi que g est parallèle pour cette connexion³ et on l'appelle *connexion de Levi-Civita* ou connexion canonique pour la métrique g .

Démonstration :

Unicité : Supposons qu'il existe une telle connexion. Alors

$$X \cdot g(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) ,$$

$$Y \cdot g(Z, X) = g(\nabla_Y Z, X) + g(Z, \nabla_Y X) ,$$

$$Z \cdot g(X, Y) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y) .$$

D'où, en sommant les deux premières identités et retranchant la troisième,

$$\begin{aligned} X \cdot g(Y, Z) + Y \cdot g(Z, X) + Z \cdot g(X, Y) \\ = g(\nabla_X Y + \nabla_Y X, Z) + g(\nabla_X Z - \nabla_Z X, Y) + g(\nabla_Y Z - \nabla_Z Y, X) \end{aligned}$$

et donc, par symétrie de la connexion,

$$\begin{aligned} 2g(\nabla_X Y, Z) = X \cdot g(Y, Z) + Y \cdot g(Z, X) + Z \cdot g(X, Y) \\ + g([X, Y], Z) + g([X, Y], Z) - g([Y, Z], X) . \end{aligned} \quad (**)$$

3. cf transport parallèle...

D'où l'unicité de la connexion par non-dégénérescence de g ⁴.

Existence : Si X et Y sont deux champs de vecteurs sur M , on définit le champ $\nabla_X Y$ comme l'unique champ de vecteurs sur M vérifiant l'identité précédente et on peut vérifier par le calcul que ∇ ainsi définie est bien une connexion symétrique rendant g parallèle pour la connexion ∇ . \square

Proposition III.9

Soit g une métrique riemannienne, X et Y deux champs de vecteurs définis dans des coordonnées locales par :

$$g = \sum_{i,j=1}^n g_{i,j} \partial_{x_i} \partial_{x_j} \quad ; \quad X = \sum_{i=1}^n X_i \partial_{x_i} \quad ; \quad Y = \sum_{i=1}^n Y_i \partial_{x_i} .$$

Alors, dans ces mêmes coordonnées,

$$\nabla_X Y = \sum_{i=1}^n \left(\sum_{j=1}^n X_j \frac{\partial Y_i}{\partial x_j} + \sum_{j,k=1}^n \Gamma_{jk}^i X_j Y_k \right) \partial_{x_i}$$

où les Γ_{jk}^i sont définis par la relation $\nabla_{\partial_{x_j}} \partial_{x_k} = \sum_{i=1}^n \Gamma_{jk}^i \partial_{x_i}$.

Les Γ_{jk}^i sont appelés *symboles de Christoffel* et, de manière explicite :

$$\Gamma_{jk}^i = \frac{1}{2} \sum_{l=1}^n (g_{i,l})^{-1} \left(\frac{\partial g_{k,l}}{\partial x_j} + \frac{\partial g_{l,j}}{\partial x_k} - \frac{\partial g_{j,k}}{\partial x_l} \right) .$$

Nous verrons au paragraphe II une notation plus compacte, via la *notation d'Einstein*.

Démonstration :

La localisation de la dérivation justifie le fait que l'on puisse se placer dans des cartes afin de travailler en terme de coordonnées.

Par propriété de la connexion, $\nabla_X Y = \sum_{j=1}^n \nabla_{X_j} \partial_{x_j} Y = \sum_{j=1}^n X_j \nabla_{\partial_{x_j}} Y$.

Or, $\forall j \in \llbracket 1, n \rrbracket$,

$$\nabla_{\partial_{x_j}} Y = \sum_{k=1}^n \nabla_{\partial_{x_j}} (Y_k \partial_{x_k}) = \sum_{k=1}^n \left((\partial_{x_j} \cdot Y_k) \partial_{x_k} + Y_k \nabla_{\partial_{x_j}} \partial_{x_k} \right) .$$

4. qui est même défini positif en tant que produit scalaire.

D'où, par définition du symbole de Christoffel, $\forall j \in \llbracket 1, n \rrbracket$,

$$\nabla_{\partial_{x_j}} Y = \sum_{k=1}^n \frac{\partial Y_k}{\partial x_j} \partial_{x_k} + \sum_{i,k=1}^n Y_k \Gamma_{j,k}^i \partial_{x_i}$$

et donc
$$\nabla_X Y = \sum_{i=1}^n \left(\sum_{j=1}^n X_j \frac{\partial Y_i}{\partial x_j} + \sum_{j,k=1}^n \Gamma_{j,k}^i X_j Y_k \right) \partial_{x_i}.$$

Commençons par noter que pour tout i, j , $[\partial_{x_j}, \partial_{x_k}] = 0$ et donc, en utilisant ($\star\star$), pour tout j, k, l ,

$$\begin{aligned} 2g\left(\nabla_{\partial_{x_j}} \partial_{x_k}, \partial_{x_l}\right) &= \partial_{x_j} \cdot g(\partial_{x_k}, \partial_{x_l}) + \partial_{x_k} \cdot g(\partial_{x_l}, \partial_{x_j}) + \partial_{x_l} \cdot g(\partial_{x_j}, \partial_{x_k}) \\ &= \frac{\partial g_{k,l}}{\partial x_j} + \frac{\partial g_{l,j}}{\partial x_k} - \frac{\partial g_{j,k}}{\partial x_l}. \end{aligned}$$

D'où la définition explicite des symboles de Christoffel. □

3. Dérivée covariante le long d'une courbe

Avant d'introduire la notion de géodésique on formalise dans ce paragraphe la notion de déformation infinitésimale d'une courbe.

Définition III.10. *Un champ de vecteurs le long d'une courbe $\gamma : I \subset \mathbb{R} \rightarrow M$ est une courbe $X : I \rightarrow TM$ telle que $X(t) \in T_{\gamma(t)}M$ pour tout $t \in I$.*

Exemple : De manière naturelle, $\gamma'(t)$ est un champ de vecteurs le long de la courbe γ .

De même que dans le cas général, à courbe γ fixée, l'ensemble des champs de vecteurs le long de γ est muni d'une structure d'espace vectoriel.

Théorème III.11

Soit ∇ la connexion de Levi-Civita associée à M et $\gamma : I \rightarrow M$ une courbe sur M . Il existe un unique opérateur que l'on note \mathcal{D}_t défini sur l'espace vectoriel des champs de vecteurs le long de γ tel que, pour tout tel champ X :

- i. Pour toute fonction $f : I \rightarrow \mathbb{R}$ lisse,

$$\mathcal{D}_t(fX)(t) = f'(t)X(t) + f(t)\mathcal{D}_tX(t).$$

- ii. S'il existe un voisinage de t_0 dans I tel que X soit la restriction à γ d'un champ de vecteurs Y défini sur un voisinage de $\gamma(t_0) \subset M$, alors :

$$\mathcal{D}_tX(t_0) = \left(\nabla_{\gamma'(t_0)} Y \right)_{\gamma(t_0)}.$$

On admet le théorème précédent et la proposition suivante dont les preuves respectives sont assez techniques et similaires.

Proposition III.12

Soit X et Y deux champs de vecteurs le long de γ . Alors :

$$\frac{d}{dt}g(X(t), Y(t)) = g(\mathcal{D}_t X(t), Y(t)) + g(X(t), \mathcal{D}_t Y(t)) .$$

III. Géodésiques

Nous commençons par définir formellement la notion de "plus courte distance", ce qui permet de munir M d'une distance (proposition III.15) compatible avec sa topologie. Enfin, la proposition donne l'équivalence entre la définition suivante et la minimisation d'une fonctionnelle d'énergie comme annoncée en section II.

On se contente ici d'énoncer les résultats fondamentaux. Pour les démontrer rigoureusement il faudrait introduire un certain nombre de résultats intermédiaires, ce que nous n'avons pas (eu) le temps de faire mais on peut se reporter, par exemple, au chapitre 2.C de [13]

On reprend les notations du paragraphe précédent et ∇ désigne la connexion de Levi-Civita et \mathcal{D}_t l'opérateur défini en III.11.

Définition III.13. Une courbe paramétrée γ sur M est une géodésique si $\nabla_{\gamma'}\gamma' = 0$, autrement dit $\mathcal{D}_t\gamma' = 0$.

En particulier, d'après la proposition III.12, la norme de $\gamma'(t)$ est constante.

En effet :

$$\frac{1}{2}g(\gamma'(t), \gamma'(t)) = g(\mathcal{D}_t\gamma', \gamma') = 0.$$

Et, dans des coordonnées locales où $x(t)$ s'écrit $(x_1(t), \dots, x_n(t))$, pour tout $i \in \llbracket 1, n \rrbracket$,

$$\frac{d^2x_i}{dt^2} + \sum_{j,k=1}^n \Gamma_{jk}^i(x(t)) \frac{dx_k}{dt} \frac{dx_j}{dt} = 0$$

et on a un moyen "simple" de calculer les géodésiques en résolvant l'équation ci-dessus.

Le théorème suivant donne l'existence et l'unicité locale des géodésiques. Non seulement l'existence n'est pas gratuite au vu de la définition mais, surtout, il faut bien prendre garde que cette unicité est uniquement **locale**. Pour s'en convaincre, on peut considérer le cas d'une sphère (figure III.2) : il existe bien évidemment une infinité de géodésiques reliant deux pôles opposés. . .

Théorème III.14

Soit $m_0 \in M$. Il existe un ouvert $\mathcal{U} \subset M$, $m_0 \in \mathcal{U}$ et $\varepsilon > 0$ tel que pour tout $m \in \mathcal{U}$ et $v \in T_m M$ vérifiant $|v| < \varepsilon$ il existe une unique géodésique $\gamma_v :]-1, 1[\rightarrow M$ vérifiant $\gamma_v(0) = m$ et $\gamma'_v(0) = v$.

Exemple : Dans le cas $M \equiv \mathbb{R}^n$, une courbe γ sur M est une géodésique ssi :

$$\mathcal{D}_t \gamma'(t) = 0 \iff \gamma''(t) = 0 \iff \exists x_0, v \in \mathbb{R}^n, \quad \gamma(t) = x_0 + tv.$$

Autrement dit, les géodésiques sont les lignes droites paramétrisées à vitesse constante, ce qui est cohérent avec l'idée que l'on se fait de "plus court chemin".

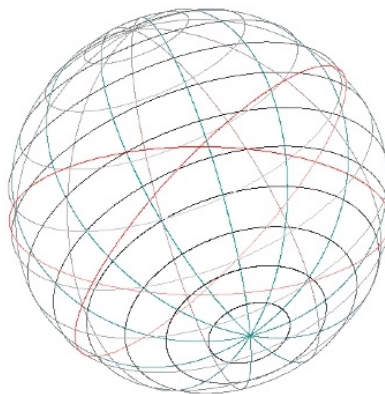


FIGURE III.2. – Géodésiques sur la sphère \mathbb{S}^2

On rappelle la définition de longueur d'une courbe vu précédemment (définition III.5) et qui, dans le cas d'une courbe $\gamma : I \rightarrow M$ de classe \mathcal{C}^1 donne :

$$l(\gamma) = \int_I \left\| \frac{d\gamma}{dt}(\gamma(t)) \right\|_{\gamma(t)} dt.$$

On suppose à présent que M est connexe, ce qui est le cas de l'espace des formes, et on définit pour tout $x, y \in M$:

$$d(x, y) := \inf \{ l(\gamma) \mid \gamma \mathcal{C}^1 \text{ par morceaux reliant } x \text{ à } y \}.$$

Proposition III.15

La quantité d ainsi définie est une distance sur M qui respecte la topologie de M .

Théorème III.16

Pour tout point m_0 de M il existe un voisinage \mathcal{U} de m_0 et un $\varepsilon > 0$ tel que pour tout $x, y \in \mathcal{U}$ il existe une unique géodésique γ de longueur au plus ε reliant x et y . De plus, $l(\gamma) = d(x, y)$.

Plus précisément, on peut montrer que les géodésiques minimisent **localement** la distance. Encore une fois, ce résultat est uniquement local.

Définition III.17. Soit $\gamma: I \rightarrow M$ une courbe \mathcal{C}^1 par morceaux sur M . On définit l'énergie de γ par la quantité

$$E(\gamma) = \frac{1}{2} \int_I \left\| \frac{d\gamma}{dt} \right\|_{\gamma(t)}^2 dt$$

qu'il faut à considérer, comme pour la distance, une subdivision sur laquelle γ est \mathcal{C}^1 ...

Proposition III.18

Soit $\gamma_0: [a, b] \rightarrow M$ une courbe \mathcal{C}^1 par morceaux.

- i. Si γ_0 minimise la distance de $\gamma_0(a)$ à $\gamma_0(b)$, i.e si $d(\gamma_0(a), \gamma_0(b)) = l(\gamma_0)$, et si γ_0 est paramétrisée proportionnellement à "arclength", alors γ_0 est une géodésique.
- ii. Si pour toute courbe γ reliant $\gamma_0(a)$ et $\gamma_0(b)$ $E(\gamma) \geq E(\gamma_0)$, alors γ_0 est une géodésique. En particulier, elle est alors paramétrisée proportionnellement à "arclength".

A la section I on revient plus en détail sur la paramétrisation des courbes.

Ce résultat est important pour l'appariement de formes car, le plus souvent, on préfère voir la distance entre deux formes A et B comme un coût de déformation pondéré par une contrainte d'attache aux données. Le choix de l'attache aux données est une difficulté en soit car un mauvais choix peut amener à des conclusions erronées. Contrairement à ce qu'on pourrait penser de prime abord, il est souvent plus judicieux de choisir une contrainte "souple", autrement dit une contrainte autorisant un appariement inexact car cela permet souvent de gagner en régularité dans l'appariement.

L'article que nous avons étudié se base sur un modèle dit de *métamorphose* où le coût d'attache aux données est intégré dans le noyau du RKHS. Tout cela est détaillé dans la section IV.

– Chapitre IV –

Courbure

L'idée naturelle de distinguer dans \mathbb{R}^n les objets "droits" des objets "courbes" amène les mathématiciens à des surprises depuis des centaines d'années, à commencer par Carl Friedrich Gauss qui démontra le résultat que l'on appelle *theorema egregium*, qui dit que la courbure d'une surface de \mathbb{R}^3 ne dépend pas de la manière dont on la plonge dans \mathbb{R}^3 mais est au contraire intrinsèque, ce que l'on reformulerait aujourd'hui ainsi :

Théorème IV.1 (*Theorema Egregium*)

La courbure de Gauss d'une surface est invariante par isométrie locale.

La question de la courbure, et l'idée d'une courbure intrinsèque à la "forme géométrique" considérée, prend toute son importance dans le cadre riemannien dans lequel nous travaillons, en particulier parce que la courbure nous permet de savoir si la géométrie construite sur notre forme est proche d'une géométrie euclidienne, ou non, ce qui est d'une importance majeure pour calculer en pratique la distance géodésique.

I. courbure d'un arc

Le premier objet dont l'on pourrait vouloir déterminer la courbure est une courbe lisse et birégulière (dérivées première et seconde du paramétrage linéairement indépendantes) du plan. Il existe plusieurs façons équivalentes de le faire, comme par exemple prendre l'inverse du rayon du cercle qui épouse la courbe au plus près au point dont on veut déterminer la courbure : c'est le *cercle osculateur* que l'on peut définir comme la limite sur ε des cercles tangents à la courbe au point d'abscisse x et passant par des points de la courbe d'abscisse x et $x + \varepsilon$.

On voit bien qu'en tout point d'une droite, le cercle osculateur a un rayon infini, ce qui correspond à une courbure nulle. En revanche, en tout point d'un cercle \mathcal{C} de rayon R , le cercle osculateur est \mathcal{C} lui-même, ce qui correspond à une courbure de $\frac{1}{R}$.

Une manière plus analytique de le voir est d'utiliser les paramétrages des courbes considérées. Introduisons à ce titre la définition suivante :

Définition IV.2. Soit I un intervalle de \mathbb{R} , soit $m \in \mathbb{N}^*$. La fonction $x : I \rightarrow \mathbb{R}^n$ est une représentation paramétrique régulière de classe \mathcal{C}^m de la courbe γ si :

- $x(I) = \gamma$ (ou $x(I) \subset \gamma$ selon que l'objectif du paramétrage est local ou global),

- x est de classe \mathcal{C}^m sur I ,
- x' ne s'annule pas sur I .

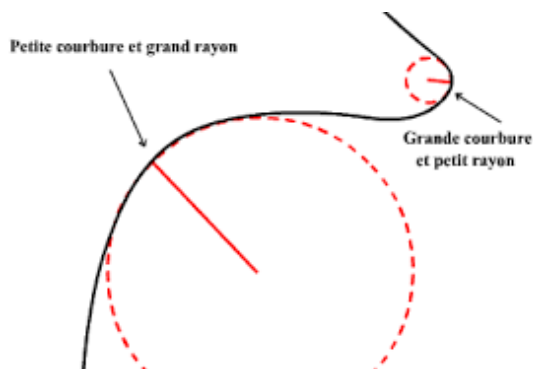


FIGURE IV.1. – Courbure à partir du cercle osculateur

Remarquons que cette définition n'interdit pas de paramétrer des courbes qui possèdent des points multiples (il peut exister t_1 et t_2 dans I tels que $x(t_1) = x(t_2)$). Cependant, on peut montrer que pour tout point $t_0 \in I$, il existe un voisinage I_0 de t_0 tel que $x(I_0)$ est une courbe simple (sans point multiple).

Il est évident qu'en "accélération" le temps t , on parcourt la même courbe, ce qui nous amène à la définition suivante :

Définition IV.3. Soit I, J des intervalles de \mathbb{R} et $x : I \rightarrow \mathbb{R}^n$ une représentation paramétrique régulière de classe \mathcal{C}^m de la courbe γ . La fonction $\phi : J \rightarrow I$ est un changement de paramètre admissible pour la courbe γ si ϕ est un \mathcal{C}^m -difféomorphisme de J dans I .

Remarque : Grâce au théorème de la bijection, il suffit de vérifier que ϕ est de classe \mathcal{C}^m , surjective et strictement monotone de I dans J .

Remarque : Sous ces conditions, $y = x \circ \phi$ est une représentation paramétrique régulière de la courbe γ , car $y(J) = x(\phi(J)) = x(I) = \gamma$ et

$$\forall t \in J, \quad y'(t) = \phi'(t)x'(\phi(t)) > 0.$$

Comme la composition de \mathcal{C}^m difféomorphismes $\psi : K \rightarrow J$ et $\phi : J \rightarrow I$ est un \mathcal{C}^m -difféomorphisme $\phi \circ \psi : K \rightarrow I$, on peut facilement achever de vérifier que les changements de paramètre admissibles définissent une relation d'équivalence.

Définition IV.4. *Etant donnée une représentation paramétrique régulière x , on appelle arc géométrique associé à x la classe d'équivalence de x définie par tous les changements de paramètres admissibles.*

Cette classe d'équivalence peut-être vue comme la "vraie" courbe γ associée au paramétrage x . Dans ce cadre, on peut revoir d'une autre manière la définition III.5, avec dans l'idée que la correspondance entre les arcs paramétrés de \mathbb{R}^n et les variétés de dimension n doit être naturelle, en substance au moins.

Définition IV.5. *Soit $x : [a, b] \rightarrow \mathbb{R}^n$ une représentation paramétrique régulière d'une courbe γ . On appelle longueur de la courbe paramétrée par x la quantité*

$$l(\gamma, x) = \sup_{k \in \mathbb{N}^*} \left(\sup_{\alpha_0=a < \alpha_1 < \dots < \alpha_k < \alpha_{k+1}=b} \left(\sum_{i=0}^k \|x(\alpha_{i+1}) - x(\alpha_i)\| \right) \right).$$

Si cette quantité est finie, on dira que la courbe γ est rectifiable.

Théorème IV.6

Si x et y sont deux représentations paramétriques régulières d'une courbe γ ,

$$l(\gamma, x) = l(\gamma, y).$$

Démonstration :

Supposons un paramétrage $y = x \circ \phi$ pour la courbe γ , avec $\phi : [c, d] \rightarrow [a, b]$. ϕ étant une bijection, pour toute subdivision $a, \alpha_1, \dots, \alpha_k, b$ de $[a, b]$, il existe une subdivision $c, \beta_1, \dots, \beta_k, d$ de $[c, d]$ telle qu'on ait

$$\phi(c) = a, \phi(\beta_1) = \alpha_1, \dots, \phi(\beta_k) = \alpha_k, \phi(d) = b$$

dans le cas où ϕ est croissant, ou

$$\phi(c) = b, \phi(\beta_1) = \alpha_k, \dots, \phi(\beta_k) = \alpha_1, \phi(d) = a$$

dans le cas où ϕ est décroissant. On a donc :

$$\sum_{i=0}^k \|y(\beta_{i+1}) - x(\beta_i)\| = \sum_{i=0}^k \|x(\phi(\beta_{i+1})) - x(\phi(\beta_i))\| = \sum_{i=0}^k \|x(\alpha_{i+1}) - x(\alpha_i)\|$$

D'où par double passage au sup,

$$l(\gamma, y) = l(\gamma, x).$$

□

On peut donc noter $l(\gamma) := l(\gamma, x)$ en calculant pour n'importe quel paramétrage. Nous pouvons enfin recoller les morceaux avec la définition III.5 :

Théorème IV.7

Si $x : [a, b] \rightarrow \mathbb{R}^n$ est une représentation paramétrique régulière de classe \mathcal{C}^1 d'une courbe γ , alors cette courbe est rectifiable et

$$l(\gamma) = \int_a^b \|x'(t)\| dt.$$

Démonstration :

Soit une représentation paramétrique $x : [a, b] \rightarrow \mathbb{R}^n$ de classe \mathcal{C}^1 , notons $x(t) = x_1(t)e_1 + \dots + x_n(t)e_n$ et $x'(t) = x'_1(t)e_1 + \dots + x'_n(t)e_n$. Comme x est de classe \mathcal{C}^1 , sa dérivée est bornée sur $[a, b]$, et donc il existe un réel positif M tel que pour tout j , pour tout t , $|x'_j(t)| \leq M$. Alors, pour toute subdivision $a < \alpha_1 < \dots < \alpha_k < b$ de $[a, b]$, on a par inégalité triangulaire et inégalité des accroissements finis :

$$\begin{aligned} \sum_{i=0}^k \|x(\alpha_{i+1}) - x(\alpha_i)\| &\leq \sum_{i=0}^k \sum_{j=0}^n |x_j(\alpha_{i+1}) - x_j(\alpha_i)| \\ &\leq \sum_{i=0}^k \sum_{j=0}^n M |\alpha_{i+1} - \alpha_i| \\ &\leq M \times n \times \sum_{i=0}^k |\alpha_{i+1} - \alpha_i| \\ &\leq M(b-a)n. \end{aligned}$$

Donc la courbe γ est rectifiable.

Montrons à présent la formule intégrale. Pour tout $\varepsilon \in \mathbb{R}_+^*$, pour tout $\delta \in \mathbb{R}_+^*$, on peut trouver une subdivision $a < \alpha_1 < \dots < \alpha_k < b$ de $[a, b]$ telle que d'une part pour tout i on ait $|\alpha_i - \alpha_{i+1}| \leq \delta$ et d'autre part,

$$\left| l(\gamma) - \sum_{i=0}^k \|x(\alpha_{i+1}) - x(\alpha_i)\| \right| \leq \varepsilon$$

ce qui se justifie par le fait que pour tout subdivision ε -proche de $l(\gamma)$, une subdivision plus précise donnera une somme plus grande, et donc plus proche de $l(\gamma)$.

Ensuite, en voyant l'intégrale comme une intégrale de Riemann, on a que pour tout $\varepsilon \in \mathbb{R}_+^*$, il existe $\delta \in \mathbb{R}_+^*$ tel que pour toute subdivision $a < \alpha_1 < \dots < \alpha_k < b$ vérifiant, pour tout i , $|\alpha_i - \alpha_{i+1}| \leq \delta$,

$$\left| \int_a^b \|x'(t)\| dt - \sum_{i=0}^k \|x'(\eta_i)\| \cdot |\alpha_{i+1} - \alpha_i| \right| \leq \varepsilon \quad \text{avec } \alpha_i \leq \eta_i \leq \alpha_{i+1}.$$

Or, d'après le théorème des accroissements finis, pour tout j , pour tout i , il existe $\eta_{i,j}$ tel que $\alpha_i \leq \eta_{i,j} \leq \alpha_{i+1}$ et

$$x_j(\alpha_{i+1}) - x_j(\alpha_i) = x'_j(\eta_{i,j})(\alpha_{i+1} - \alpha_i).$$

On peut décomposer le problème de la sorte :

$$\begin{aligned} \left| l(\gamma) - \int_a^b \|x'(t)\| dt \right| &\leq \left| l(\gamma) - \sum_{i=0}^k \|x(\alpha_{i+1}) - x(\alpha_i)\| \right| \\ &\quad + \left| \int_a^b \|x'(t)\| dt - \sum_{i=0}^k \|x(\alpha_{i+1}) - x(\alpha_i)\| \right| \\ &\leq \varepsilon + \left| \int_a^b \|x'(t)\| dt - \sum_{i=0}^k \|x(\alpha_{i+1}) - x(\alpha_i)\| \right| \\ &\leq \varepsilon + \left| \int_a^b \|x'(t)\| dt - \sum_{i=0}^k \left\| \sum_{j=1}^n x'_j(\eta_{i,j})(\alpha_{i+1} - \alpha_i) e_j \right\| \right|. \end{aligned}$$

Or,

$$\begin{aligned} &\left| \int_a^b \|x'(t)\| dt - \sum_{i=0}^k \left\| \sum_{j=1}^n x'_j(\eta_{i,j})(\alpha_{i+1} - \alpha_i) e_j \right\| \right| \\ &\leq \left| \int_a^b \|x'(t)\| dt - \sum_{i=0}^k \|x'(\alpha_i)\| \cdot |\alpha_{i+1} - \alpha_i| \right| \\ &\quad + \left| \sum_{i=0}^k \|x'(\alpha_i)\| \cdot |\alpha_{i+1} - \alpha_i| - \sum_{i=0}^k \left\| \sum_{j=1}^n x'_j(\eta_{i,j})(\alpha_{i+1} - \alpha_i) e_j \right\| \right| \\ &\leq \varepsilon + \left| \sum_{i=0}^k \left(\|x'(\alpha_i)\| \cdot |\alpha_{i+1} - \alpha_i| - \left\| \sum_{j=1}^n x'_j(\eta_{i,j})(\alpha_{i+1} - \alpha_i) e_j \right\| \right) \right| \\ &\leq \varepsilon + \left| \sum_{i=0}^k \|x'(\alpha_i)(\alpha_{i+1} - \alpha_i) - \sum_{j=1}^n x'_j(\eta_{i,j})(\alpha_{i+1} - \alpha_i) e_j \right| \\ &\leq \varepsilon + \left| \sum_{i=0}^k \left\| \sum_{j=1}^n (x'_j(\alpha_i) - x'_j(\eta_{i,j})) e_j \cdot (\alpha_{i+1} - \alpha_i) \right\| \right| \end{aligned}$$

Le deuxième terme est aussi petit que l'on veut par uniforme continuité de x' sur $[a, b]$ (théorème de Heine). On en déduit le résultat. \square

Remarquons à présent que pour tout $t_0 \in I$, l'application

$$s : t \mapsto \int_{t_0}^t \left| \frac{dx(u)}{du} \right| du$$

est régulière quand x est de classe \mathcal{C}^2 a une dérivée qui ne s'annule pas (comme c'est le cas des représentations régulières), comme le montre ce calcul :

$$\frac{ds}{dt} = \frac{d}{dt} \int_{t_0}^t \left| \frac{dx}{du}(u) \right| du = \left| \frac{dx}{du}(u) \right| \neq 0$$

et s est alors un changement de paramètre admissible, que l'on appelle *l'abscisse curviligne*.

On peut donc dériver une courbe au sens temporel, et nous noterons alors

$$x'(t) = \frac{dx}{dt} \quad \text{et} \quad x''(t) = \frac{d^2x}{dt^2},$$

ou par rapport à son abscisse curviligne,

$$\frac{dx}{ds}(s) = \lim_{\delta s \rightarrow 0} \frac{x(s + \delta s) - x(s)}{\delta s} \quad \text{et} \quad \frac{d^2x}{ds^2}(s) = \lim_{\delta s \rightarrow 0} \frac{\frac{dx}{ds}(s + \delta s) - \frac{dx}{ds}(s)}{\delta s}.$$

Définition IV.8. On appelle *vecteur unitaire tangent en s* le vecteur $\frac{dx}{ds}(s)$.

Remarque : Le vecteur unitaire tangent vérifie, par dérivées composées :

$$\frac{dx}{ds}(s(t)) = \frac{x'(t)}{|x'(t)|}$$

et c'est donc bien un vecteur unitaire et tangent (d'où son nom).

Le plan orthogonal au vecteur tangent en s et passant par $x(s)$ est appelé *plan normal* en s . La courbure qui nous intéresse est caractérisée par un vecteur de ce plan normal.

Définition IV.9. Soit x la représentation paramétrique régulière d'un arc de classe \mathcal{C}^2 d'abscisse curviligne s . On appelle *vecteur de courbure au point s* le vecteur

$$k(s) = \frac{d^2x}{ds^2}(s).$$

Sa valeur absolue $|k'(s)|$ est appelée la courbure en s et si la courbure est non nulle on appelle rayon de courbure la quantité

$$r(s) = \frac{1}{|k'(s)|}.$$

Remarque : En dérivant par rapport à s l'égalité

$$\left\| \frac{dx}{ds}(s) \right\|^2 = \left\langle \frac{dx}{ds}(s) \mid \frac{dx}{ds}(s) \right\rangle = 1$$

on obtient

$$2 \left\langle \frac{d^2x}{ds^2}(s) \mid \frac{dx}{ds}(s) \right\rangle = 0$$

ce qui nous assure que le vecteur de courbure est bien dans le plan normal.

Remarque : On peut montrer

$$\|k(s(t))\| = \frac{\|x'(t) \wedge x''(t)\|}{\|x'(t)\|^3}$$

où $x \wedge y$ est le produit vectoriel de x et y .

II. courbure d'une surface

La construction précédente pour les courbes de \mathbb{R}^n permet de définir des notions de courbure pour les surfaces et constitue une intuition robuste sur laquelle appuyer une construction abstraite de la courbure d'une variété riemannienne. Nous allons voir qu'il y a plusieurs manières d'appréhender la notion de courbure d'une surface de \mathbb{R}^3 .

Soit U un ouvert de \mathbb{R}^2 .

Définition IV.10. Une partie S de \mathbb{R}^3 est une surface régulière si pour tout point $z \in S$, il existe un voisinage ouvert V de z et un \mathcal{C}^∞ -difféomorphisme $\phi : U \rightarrow S \cap V$ (où $S \cap V$ est muni de la topologie induite) tels que pour tout $(x, y) \in U$, $\mathcal{D}_{(x,y)}\phi$ est de rang 2.

Les surfaces régulières sont des variétés différentielles de dimension 2 si on les considère intrinsèquement et munies de la topologie induite. On peut donc définir en tout point z de S un plan tangent $T_z S$. On notera (X_z, Y_z) une base du plan tangent en z .

Remarque : Les surfaces peuvent être traitées avec les mêmes outils bas-niveau que ceux que nous avons utilisés pour les courbes en les définissant comme l'image de paramétrisations régulières continues $\phi : U \rightarrow \mathbb{R}^3$ dont la différentielle est de rang 2 en tout point. Si on considère alors une base $(\frac{\partial}{\partial x}, \frac{\partial}{\partial y})$ des champs de vecteurs sur U , le plan tangent en $z = \phi(x, y)$ est le sous-espace vectoriel engendré par $X_z = \mathcal{D}_{(x,y)}\phi \cdot \frac{\partial}{\partial x}$ et $Y_z = \mathcal{D}_{(x,y)}\phi \cdot \frac{\partial}{\partial y}$.

Définition IV.11. Soit S une surface régulière de paramétrisation ϕ et de plan tangent au point z engendré par la base (X_z, Y_z) . On appelle vecteur normal unitaire en z le vecteur

$$N_z = \frac{X_z \wedge Y_z}{\|X_z \wedge Y_z\|}.$$

De plus, on appelle application de Gauss de S l'application

$$\begin{aligned} N : S &\rightarrow \mathbb{S}^2 \\ z &\mapsto N_z \end{aligned}$$

où \mathbb{S}^2 est la sphère unité de \mathbb{R}^3 .

La sphère \mathbb{S}^2 est une variété lisse, et on peut montrer que l'application de Gauss de S est lisse. Sa différentielle

$$\mathcal{D}_z N : T_z S \rightarrow T_{N_z} \mathbb{S}^2$$

induit un morphisme d'espaces vectoriels de $T_z S$ dans $T_{N_z} \mathbb{S}^2$, qui est en fait un endomorphisme puisque ces espaces sont isomorphes.

Si on muni S d'une métrique riemannienne, on peut montrer que pour tout z , $\mathcal{D}_z N$ est auto adjointe. Elle a alors deux valeurs propres $\lambda_1(z)$ et $\lambda_2(z)$, que l'on peut nommer de sorte que $\lambda_1(z) < \lambda_2(z)$. Notons alors $\lambda_1(z) = -k_1(z)$ et $\lambda_2(z) = -k_2(z)$; on a $k_1(z) > k_2(z)$.

Définition IV.12. On appelle courbures principales de S au point z les valeurs $k_1(z)$ et $k_2(z)$ définies ci-dessus. $k_1(z)$ est appelée la courbure maximale et $k_2(z)$ la courbure minimale. On appelle courbure moyenne en z la quantité

$$k_{\text{moy}}(z) = \frac{k_1(z) + k_2(z)}{2}.$$

On appelle courbure de Gauss en z le déterminant de $\mathcal{D}_z N$, que l'on note

$$k(z) = k_1(z)k_2(z).$$

Mais tout cela n'a un sens qu'algébrique, il est préférable pour en avoir une vision géométrique d'introduire la *seconde forme fondamentale* :

Définition IV.13. On appelle seconde forme fondamentale de S au point z et on note II_z l'application :

$$II_z : T_z S \rightarrow \mathbb{R}$$

$$\eta \mapsto - \left\langle D_z N \cdot \eta \mid \eta \right\rangle_z .$$

Plusieurs résultats sur la seconde forme fondamentale sont évidents à démontrer et cependant d'une importance capitale :

- La seconde forme fondamentale est une forme quadratique sur $T_z S$;
- $k_1(z)$ et $k_2(z)$ sont respectivement le maximum et le minimum de II_z sur le cercle unité de l'espace euclidien $T_z S$;
- La forme quadratique II_z a une matrice diagonale $diag(k_1, k_2)$ dans la base (e_1, e_2) . La courbure moyenne est alors la demi-trace de II_z dans cette base et la courbure de Gauss le déterminant de II_z dans cette base.

Un résultat plus surprenant encore, que nous ne démontrerons pas, est le suivant :

Théorème IV.14 (Théorème de Meusnier)

Toutes les courbes tracées sur S , passant par z et ayant la même droite tangente en z ont la même courbure.

Ce résultat vient du fait, que nous ne démontrons pas non plus, que la courbure de toute courbe passant par z , contenue dans S et dirigée par le vecteur tangent $\eta \in T_z S$ a pour courbure la valeur

$$k = II_z(\eta).$$

On termine avec la définition d'une autre forme de courbure, spéciale parce qu'elle est globale (ne dépend pas d'un point de la surface) :

Définition IV.15. On appelle courbure totale de la surface S la quantité

$$k_{tot} = \int_S k_{gauss}(z) dz .$$

La courbure totale s'interprète géométriquement comme l'aire balayée sur \mathbb{S}^2 par le vecteur normal unitaire en z lorsque z parcourt tout S .

III. courbure d'une variété riemannienne

Dans une variété riemannienne de dimension n , on ne peut pas continuer à tracer des courbes pour étudier la courbure. Il devient nécessaire d'utiliser le formalisme algébrique développé dans les sections précédentes afin de capturer au mieux et dans la plus vaste généralité l'idée d'espace "courbe". On introduit donc la définition suivante :

Définition IV.16. Soit une variété riemannienne M munie de sa connexion de Levi-Civita ∇ . On appelle tenseur de courbure de M l'application

$$R : \mathcal{C}^\infty(TM) \times \mathcal{C}^\infty(TM) \times \mathcal{C}^\infty(TM) \rightarrow \mathcal{C}^\infty(TM)$$

$$(X, Y, Z) \mapsto R(X, Y)Z := \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z \quad .$$

La définition a un sens pour n'importe quelle variété M munie d'une connexion, mais c'est avec la métrique riemannienne et la connexion de Levi-Civita que l'on a des espoirs de donner à une variété un "vrai" sens géométrique de courbure. Certains auteurs définissent R comme l'opposé de la quantité ci-dessus, c'est à dire

$$R(X, Y)Z := \nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z + \nabla_{[X, Y]} Z$$

comme on peut le trouver dans [13].

Cet objet représente en quelque sorte le défaut de commutativité de la dérivée covariante.

En effet, si $X = \frac{\partial}{\partial x_i}$ et $Y = \frac{\partial}{\partial x_j}$, on a $[X, Y] = 0$ et par conséquent

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_X \nabla_Y Z .$$

Dans le cas d'un espace euclidien muni de la métrique usuelle, en notant $X = \frac{\partial}{\partial x_i}$ et

$$Y = \frac{\partial}{\partial x_j},$$

$$(R(X, Y)Z)_m = \nabla_m X \circ \nabla_m Y \cdot Z_m - \nabla_m Y \circ \nabla_m X \cdot Z_m$$

$$= \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} Z_m - \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i} Z_m = 0$$

d'après le lemme de Schwarz. Cela traduit le fait que la courbure soit nulle dans un espace euclidien muni de la métrique usuelle. On peut donc, également, voir la courbure comme un "défaut de vérification du lemme de Schwarz", c'est à dire, si on a en tête la preuve la plus pédestre qui soit du lemme de Schwarz, une différence entre le transport parallèle selon X puis selon Y et le transport parallèle selon Y puis selon X d'un champ de vecteurs de classe \mathcal{C}^∞ .

Avec des outils d'algèbre tensorielle que nous n'avons pas eu le temps de travailler, on pourrait définir une version (encore) plus algébrique et générale de la courbure : si on considère une variété M et un fibré vectoriel E de base M , si on note $\mathcal{C}^\infty(E)$ les champs de vecteurs sur M définis comme sections lisses de E et $\Omega^p(M) = \mathcal{C}(M, \Lambda^p M)^1$, où $\Lambda^p M$ est un produit extérieur sur M , si enfin on définit $\Omega^p(E) := \mathcal{C}^\infty(E) \otimes \Omega^p(M)$, on peut alors utiliser la définition suivante :

1. C'est l'ensemble des p -formes extérieures, et il peut être défini comme l'ensemble des sections du fibré vectoriel $\Lambda^p(M)$. On peut définir une dérivée extérieure $d_p : \Omega^p(M) \rightarrow \Omega^{p+1}(M)$, vérifiant $d_{p+1} \circ d_p = 0$, ce qui ouvre sur la cohomologie de De Rham. Ce secteur de la géométrie nous intéressait mais nous n'avions ni le temps ni les raisons suffisantes de l'inclure dans le projet.

Définition IV.17. La courbure d'une connexion ∇ est l'opérateur

$$\begin{aligned} F : \Omega^0(E) &\rightarrow \Omega^2(E) \\ \mu &\mapsto \nabla \circ \nabla(\mu) \end{aligned} .$$

On notera que dans ce formalisme, $\nabla : \Omega^p(E) \rightarrow \Omega^{p+1}(E)$, mais que $\nabla \circ \nabla \neq 0$ dans le cas général; ∇ ne définit pas un complexe différentiel croissant. Si $F=0$, on dira que la connexion est *plate*. L'espace euclidien \mathbb{R}^n muni de sa connexion de Levi-Civita (voir II.1) est un exemple de courbure nulle. Pour revenir vers des choses concrètes, il faudrait utiliser les propriétés de la courbure ainsi définie, dans le cas riemannien, pour écrire F sous la forme :

$$\begin{aligned} F : \Omega^0(E) &\rightarrow \Omega^2(E) \\ \mu &\mapsto R(\cdot, \cdot)\mu \end{aligned} .$$

On obtient alors comme un théorème la définition que nous prenions au début, c'est à dire :

Théorème IV.18

Le tenseur de courbure R d'une connexion ∇ vérifie, pour tous champs de vecteurs X, Y , pour tout $\mu \in \mathcal{C}^\infty(E)$

$$R(X, Y)\mu = \nabla_X \nabla_Y \mu - \nabla_Y \nabla_X \mu - \nabla_{[X, Y]}\mu .$$

Le problème du tenseur de courbure est qu'il est en pratique très difficile à calculer en l'état, et qu'il est difficile de se représenter la courbure d'une variété à partir du seul tenseur. L'outil que nous allons utiliser en pratique, qui est présenté dans l'article et dont nous avons programmé un calcul sur python est la courbure sectionnelle.

Définition IV.19. Soit $G_m^2 M$ l'ensemble des 2-Grassmanniennes de $T_m M$ et soit

$$G^2 M = \bigcup_{m \in M} G_m^2 M .$$

La courbure sectionnelle est l'application

$$\begin{aligned} \mathcal{K} : G^2 M &\rightarrow \mathbb{R} \\ P = \mathbb{R}X + \mathbb{R}Y &\mapsto \frac{\langle R(X, Y)Y \mid X \rangle_g}{\|X\|_g^2 \|Y\|_g^2 - \langle X \mid Y \rangle_g^2} . \end{aligned}$$

Pour deux vecteurs tangents linéairement indépendants, on notera $\mathcal{K}(X, Y)$ la valeur de la courbure sectionnelle sur le plan $\mathbb{R}X + \mathbb{R}Y$. On remarquera que si X et Y sont orthonormaux, on a

$$\mathcal{K}(X, Y) = \langle R(X, Y)Y \mid X \rangle_g .$$

Théorème IV.20

La courbure sectionnelle détermine entièrement le tenseur de courbure.

Nous n'avons pas le temps de montrer ce théorème, car sa preuve est calculatoire et demande d'introduire l'identité de Bianchi. On peut également montrer que dans le cas d'une surface de \mathbb{R}^3 la courbure sectionnelle de deux vecteurs tangents linéairement indépendants est la courbure de Gauss de la surface au point considéré.

IV. Illustration dans le cas des amygdales

On revient ici sur les amygdales cérébrales présentées au paragraphe III. Pour cela, on utilise le même programme d'appariement mais en l'utilisant cette fois deux à deux sur trois sujets. Sur chaque amygdale, on a positionné 20 points labellisés que l'on cherche appairier sur les 20 points correspondants. La figure IV.2 illustre la distance (géodésique) dans l'espace de formes courbe pour chacun des points des trois amygdales. La grille en arrière plan montre la torsion que doit subir l'espace pour appairier un amygdale sur un autre de manière exacte. Comme on effectue le programme trois fois, les trois grilles se superposent...

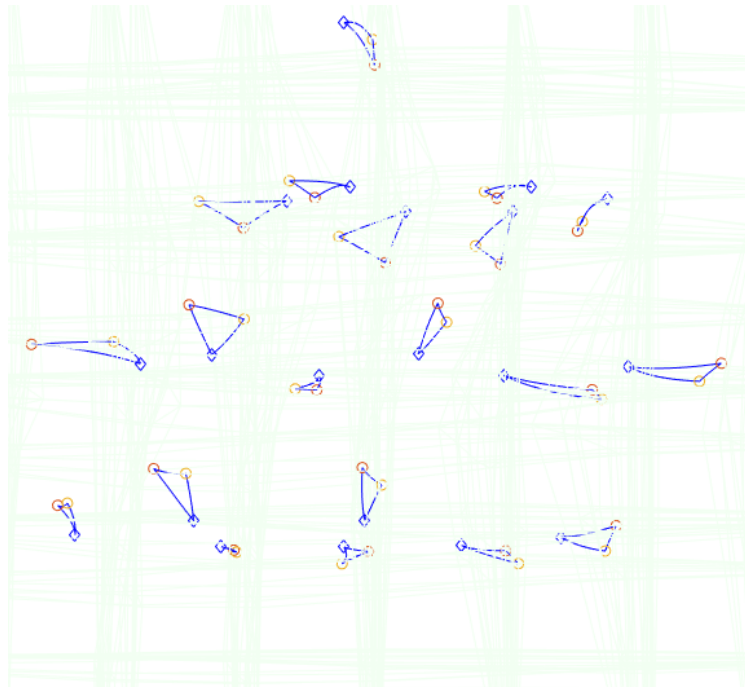


FIGURE IV.2. – Appariement de trois amygdales cérébrales gauches

On cherche à évaluer la courbure de l'espace induit par ces trois amygdales : on travaille donc dans $\mathbb{R}^{3 \times 20}$.

Commençons par s'assurer que la courbure n'est pas nulle. Pour cela, on choisit arbitrairement un des côtés des triangles ci-dessus (celui reliant les points rouges et oranges par exemple) et on calcul la distance géodésique donnée par le programme d'appariement d'une part et la distance euclidienne entre ces deux points brutalement d'autre part. On est à même de récupérer les vecteurs tangents au départ de chacune des géodésiques informatiquement : notons-les τ_{br} , τ_{bo} et τ_{or} selon qu'ils relient les points bleus→rouges, bleus→orange ou oranges→rouges.

On a vu au paragraphe III que la norme de la dérivée de toute géodésique est constante et en particulier, si on note γ_{or} une géodésique reliant le point orange au point rouge, alors :

$$l(\gamma_{or}) = \int_0^1 \|\gamma'_{or}(t)\|_{\gamma_{or}(t)} dt = \int_0^1 \|\|\tau_{or}\|_2 u_{\gamma_{or}(t)}\|_{\gamma_{or}(t)} = \|\tau_{or}\|_2$$

où $u_{\gamma_{or}(t)}$ est un vecteur tangent et unitaire pour la métrique locale (de la variété) en chacun des points $\gamma_{or}(t)$ et on peut donc calculer la distance géodésique facilement.

On résume dans le tableau suivant les distances obtenues en calculant la distance géodésique comme expliqué ci-dessus puis la distance euclidienne (différence des vecteurs τ_{br} et τ_{bo}). En particulier, la distance géodésique est le plus souvent supérieure à la distance euclidienne et on peut en déduire que la courbure n'est pas nulle : en moyenne la distance géodésique est de 1.8974 contre 1.0729.

Distance géodésique	1.2259	2.7798	1.1279	0.6511	1.5291	2.1225	1.9434
Distance euclidienne	0.3739	0.6846	1.3678	0.0051	0.6372	0.3692	0.4764
Distance géodésique	1.2124	2.2932	1.5440	1.3602	2.1148	2.8375	1.7637
Distance euclidienne	1.3061	2.4165	1.0827	1.1140	0.8670	0.8699	2.1142
Distance géodésique	1.8273	2.0922	2.4267	2.3209	1.8771	1.6015	
Distance euclidienne	1.0783	1.8696	2.4909	1.9935	1.2094	0.3878	

TABLE IV.1. – Distance euclidienne vs distance géodésique

Pour déterminer si la courbure est positive ou négative, on va chercher à calculer la somme des angles (géométriques) des triangles et regarder si elle est supérieure ou inférieure à π . On répertorie dans le tableau IV.2 les différents angles des triangles ci-dessus. An particulier, la courbure est fortement négative car la somme des angles des triangles vaut, en moyenne $1.0013 \simeq \frac{\pi}{3}$.

\widehat{obr}	1.2906	1.2167	1.1149	0.2144	1.5558	1.1196	1.0098	0.9328
\widehat{bro}	0.6461	1.6185	1.1491	1.6712	0.8920	1.1422	1.3746	1.7922
\widehat{bor}	0.9617	0.3297	0.1831	0.6634	0.0318	0.5108	0.4266	0.5469
Somme	2.0367	0.6288	1.0327	1.5403	0.6747	1.3623	1.1536	0.9229
\widehat{obr}	2.3510	2.0933	0.1449	0.5399	0.2613	0.4554	0.7448	0.9966
\widehat{bro}	0.7657	0.4335	2.6407	1.9952	2.5478	2.3653	2.0104	1.3275
\widehat{bor}	0.2638	1.6385	0.8182	0.6926	0.1848	0.3499	0.7403	0.9116
Somme	0.2812	2.0469	0.9305	1.1386	0.3221	0.5345	1.0499	1.6957
\widehat{obr}	1.7004	1.8162	1.8856	0.2280				
\widehat{obr}	0.6418	0.8028	0.7075	2.6872				
\widehat{bor}	0.4285	0.1954	0.1270	0.2734				
Somme	1.1003	0.6471	0.5774	0.3490				

TABLE IV.2. – Angles des triangles

– Chapitre V –

La formule de Mario

Le principal objectif de notre projet était d'étudier et d'implémenter la formule dite "*formule de Mario*" : il s'agit d'une expression de la courbure sectionnelle sensée simplifier le calcul numérique pour des espaces de landmarks. Ce qui suit est principalement tiré de l'article [28] et du cours *Géométrie et espaces de formes* d'Alain Trouvé.

I. La variété riemannienne des landmarks

1. Le groupe des difféomorphismes

Nous nous intéressons ici à un ensemble de landmarks

$$\mathcal{L}^n(\mathbb{R}^d) = \{(P^1, \dots, P^n) \mid \forall i \in \{1, \dots, n\}, P^i \in \mathbb{R}^d\}$$

que nous souhaitons doter d'une structure riemannienne.

On définit l'ensemble des chemins continûment différentiables dans l'espace des landmarks :

$$\mathcal{Q} = \{q : [0, 1] \rightarrow \mathcal{L}^n(\mathbb{R}^d) \mid \forall i \in \{1, \dots, n\}, q^i \in C^1([0, 1], \mathbb{R}^d)\}.$$

Soit $(V, \langle \cdot | \cdot \rangle_V)$ un espace de Hilbert de champs de vecteurs sur \mathbb{R}^d . On le suppose admissible, c'est à dire vérifiant les deux propriétés suivantes :

- i. V s'injecte continûment dans l'espace $C_b^1(\mathbb{R}^d, \mathbb{R}^d)$ des fonctions de classe \mathcal{C}^1 bornées à dérivées premières bornées ;
- ii. si pour $x_1, \dots, x_M \in \mathbb{R}^d$ et $\alpha_1, \dots, \alpha_M \in \mathbb{R}^d$ on a

$$\forall u \in V, \sum_{i=1}^M \langle \alpha_i | u(x_i) \rangle_{\mathbb{R}^d} = 0$$

alors $\alpha_1 = \dots = \alpha_M = 0$.

En outre, on suppose que V est un espace de Hilbert à noyau reproduisant, c'est à dire que pour tout $\alpha, x \in \mathbb{R}^d$, il existe $K_x^\alpha \in V$ vérifiant :

$$\forall f \in V, \langle K_x^\alpha | f \rangle_V = \langle \alpha | f(x) \rangle_{\mathbb{R}^d}.$$

Comme, d'après la définition ci-dessus, $(\alpha, \beta) \mapsto \langle K_x^\alpha \mid K_y^\beta \rangle$ est une forme bilinéaire, il existe pour la représenter une matrice $d \times d$ que nous notons $K(x, y)$. Dans la suite, nous supposons que pour tout $x, y \in \mathbb{R}^d$, la matrice $K(x, y)$ est de la forme

$$K(x, y) = K(x - y)I_d$$

où I_d est la matrice identité de taille $d \times d$.

En supposant connue la notion d'intégrale de Bochner qui généralise aux espaces de Banach la notion d'intégration, nous pouvons définir l'ensemble $L^p([0, 1], V)$ des applications de $[0, 1]$ dans V Bochner mesurables et de norme à la puissance p intégrable. On admet que pour $v \in L^1([0, 1], V)$, dont les évaluations sont notées $v(t) = v_t$ il existe une unique solution q , au système dynamique

$$q(t) = x + \int_0^t v_s(q(s)) ds$$

avec condition initiale $q(0) = x$ ¹. On note alors $q(t) = \phi_t^v(x)$.

Définition V.1. On appelle groupe des difféomorphismes engendrés par V et on note \mathcal{G}_V l'ensemble

$$\mathcal{G}_V = \{\phi_1^v \mid v \in L^1([0, 1], V)\}.$$

On montre que \mathcal{G}_V est bien un groupe, et même un groupe topologique métrisable.

2. Fonctionnelle d'énergie

Soit $\lambda \in \mathbb{R}_+^*$ un paramètre réel. On définit pour tous $v \in L^2([0, 1], V)$ et $q \in \mathcal{Q}$ la fonctionnelle d'énergie

$$E_\lambda(v, q) = \int_0^1 \left(\|v_t\|_V^2 + \lambda \sum_{i=1}^n \left\| \frac{dq^i}{dt}(t) - v_t(q^i(t)) \right\|_{\mathbb{R}^d}^2 \right) dt.$$

De manière informelle, plus le terme λ est grand, plus les défauts de vérification de $\frac{dq}{dt}(t) = v_t(q(t))$ coûtent de l'énergie. Cette notion nous permet de définir, pour $I, I' \in \mathcal{L}^n(\mathbb{R}^d)$, la quantité

$$d(I, I') = \inf_{v, q} \left\{ \sqrt{E(v, q)} \mid v \in L^2([0, 1], V), q \in \mathcal{Q}, q(0) = I, q(1) = I' \right\}.$$

1. On peut écrire de manière informelle $\frac{dq}{dt} = v(q)$, mais l'écriture intégrale, avec le soutien de la théorie de l'intégration de Bochner, a plus de sens.

On peut en fait prouver que d est une distance géodésique pour une certaine métrique riemannienne. L'infimum représente la manière la moins "coûteuse", au sens de E_λ , de transformer la forme I en la forme I' suivant la trajectoire q et le champ de vecteurs dépendant du temps v .

Soit $q \in \mathcal{L}^n(\mathbb{R}^d)$. En écrivant $q = (q^1, \dots, q^n)$ et avec le noyau $K(x, y)$ précédemment défini, on peut écrire la matrice $n \times n$, dite de Gram, associée à q :

$$\mathbf{K}(q) = \left(K(q^i, q^j) \right)_{1 \leq i, j \leq n}.$$

Par les propriétés des noyaux, la matrice $\mathbf{K}(q)$ est inversible. Comme l'ensemble des matrices inversibles est un ouvert, pour $\frac{1}{\lambda}$ assez petit, c'est à dire λ assez grand, la matrice

$$\mathbf{K}(q) + \frac{I_n}{\lambda}$$

est inversible. On introduit alors la matrice $nd \times nd$ suivante, également fonction de $q \in \mathcal{L}^n(\mathbb{R}^d)$:

$$g(q) = \begin{pmatrix} \left(\mathbf{K}(q) + \frac{I_n}{\lambda} \right)^{-1} & 0 & \cdots & 0 \\ 0 & \left(\mathbf{K}(q) + \frac{I_n}{\lambda} \right)^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \left(\mathbf{K}(q) + \frac{I_n}{\lambda} \right)^{-1} \end{pmatrix}.$$

Nous allons montrer que la matrice g est un le tenseur d'une métrique riemannienne dont les géodésiques sont données par la minimisation de la fonctionnelle E_λ . On commence par le lemme suivant :

Lemme V.2.1. Soit $\bar{q} \in \mathcal{Q}$. Il existe un unique $v^* \in L^2([0, 1], V)$ tel que

$$E_\lambda(v^*, \bar{q}) = \inf_{v \in L^2([0, 1], V)} E_\lambda(v, \bar{q})$$

et on a en outre pour tout $t \in [0, 1]$,

$$\forall x \in \mathbb{R}^d, \quad v_t^*(x) = \sum_{i=1}^n K(x, \bar{q}^i(t)) \begin{pmatrix} \sum_{j=1}^n \left(\mathbf{K}(\bar{q}(t)) + \frac{I_n}{\lambda} \right)^{-1}_{ij} \times \frac{d\bar{q}_1^j}{dt}(t) \\ \vdots \\ \sum_{j=1}^n \left(\mathbf{K}(\bar{q}(t)) + \frac{I_N}{\lambda} \right)^{-1}_{ij} \times \frac{d\bar{q}_d^j}{dt}(t) \end{pmatrix}.$$

Forts de ce résultat, nous introduisons la quantité

$$\tilde{E}_\lambda(\bar{q}) = E_\lambda(v^*(\bar{q}), \bar{q}).$$

On a alors le théorème suivant, où le vecteur $\frac{dq}{dt}(t)$ doit être vu comme le long vecteur colonne ${}^t\left(\frac{dq_1^1}{dt}(t), \dots, \frac{dq_1^n}{dt}(t), \dots, \frac{dq_d^1}{dt}(t), \dots, \frac{dq_d^n}{dt}(t)\right)$, dont la taille est nd .

Théorème V.2

Si (\hat{v}, \hat{q}) minimise $E_\lambda(v, q)$, alors \hat{q} minimise $\tilde{E}_\lambda(q)$ et $E_\lambda(\hat{v}, \hat{q}) = \tilde{E}_\lambda(\hat{q})$.
 Réciproquement, si \hat{q} minimise $\tilde{E}_\lambda(q)$ alors $(v^*(\hat{q}), \hat{q})$ minimise $E_\lambda(v, q)$ et $E_\lambda(v^*(\hat{q}), \hat{q}) = \tilde{E}_\lambda(\hat{q})$.
 Enfin, on a $\forall q \in \mathcal{Q}$,

$$\begin{aligned} \tilde{E}_\lambda(q) &= \int_0^1 \sum_{i,j=1}^n \sum_{s=1}^d \frac{dq_s^i}{dt}(t) \times \frac{dq_s^j}{dt}(t) \times \left(\mathbf{K}(\bar{q}(t)) + \frac{I_n}{\lambda}\right)^{-1}_{ij} dt \\ &= \int_0^1 {}^t\left(\frac{dq}{dt}(t)\right) \cdot g(q(t)) \cdot \frac{dq}{dt}(t) dt \end{aligned}$$

Ce résultat signifie que pour la métrique riemannienne associée à g , étant donné $I, I' \in \mathcal{L}^n(\mathbb{R}^d)$, la distance géodésique entre I et I' est la quantité $d(I, I')$ que nous avons définie plus tôt. Notons que cette quantité se réécrit, à la lumière des derniers résultats, sous la forme

$$d(I, I') = \inf_q \left\{ \sqrt{\tilde{E}(q)} \mid q \in \mathcal{Q}, q(0) = I, q(1) = I' \right\}.$$

II. Calcul de la courbure sectionnelle

Dans cette partie, nous nous plaçons sur une variété riemannienne M , de tenseur métrique $g : TM \times_M TM \rightarrow \mathbb{R}$. Pour une carte (U, ϕ) associée à la variété, de coordonnées locales (x^1, \dots, x^n) et de coordonnées pour les champs de vecteurs $(\partial_1, \dots, \partial_n) = (\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n})$, on note dx^i l'application $(x^1, \dots, x^n) \mapsto x^i$, et on écrit $g_{ij} := g(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j})$.

A partir de maintenant, nous utiliserons souvent, implicitement ou non, la convention de sommation d'Einstein :

$$a_i b^i := \sum_{i=1}^n a_i b^i$$

Avec cette convention, un champ de vecteurs X sur M s'écrira donc de manière compacte :

$$X = \sum_{i=1}^n X^i \partial_{x_i} = X^i \partial_i$$

et la métrique est donnée par :

$$g|_U = \sum_{i,j=1}^n g \left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right) dx^i \otimes dx^j = g(\partial_i, \partial_j) dx^i \otimes dx^j = g_{ij} dx^i \otimes dx^j .$$

Comme précédemment, on assimile chacun des points m de la variété M à son image par le morphisme de carte ϕ . Ainsi, par exemple, on notera pour tout $x \in \phi(U)$, $g_{ij}(x)$ pour $(g_{ij} \circ \phi^{-1})(x) \dots$

Afin de rendre les expressions plus compactes et par soucis de cohérence avec la notation d'Einstein, on introduit les notations suivantes :

- $g_{ij,k}(x) := \frac{\partial}{\partial x^k} g_{ij}(x)$ et $g_{ij,kl}(x) := \frac{\partial}{\partial x^k} \frac{\partial}{\partial x^l} g_{ij}(x)$;
- la cométrique g^{-1} est notée avec des indices en haut :

$$g^{-1}|_U = \sum_{i,j=1}^n g^{ij} \partial_i \partial_j = g^{ij} \partial_i \partial_j \quad \text{de sorte que} \quad \sum_{k=1}^n g^{ik} g_{kj} = g^{ik} g_{kj} = \delta_j^i ;$$

- $g^{ij},{}_k(x) := \frac{\partial}{\partial x^k} g^{ij}(x)$ et $g^{ij},{}_{kl}(x) := \frac{\partial}{\partial x^k} \frac{\partial}{\partial x^l} g^{ij}(x)$;
- pour tout champ de vecteurs $X^i \partial_i$, $X_i := g_{ij} X^j$.

Avec la notation d'Einstein, les symboles de Christoffel vus à la proposition III.9 sont définis par $\nabla_{\partial_i} \partial_j = \Gamma_{ij}^k \partial_k$ pour tout i, j et, de manière explicite :

$$\forall i, j, k, \quad \Gamma_{ij}^k = \frac{1}{2} g^{kl} (g_{il,j} + g_{jl,i} - g_{ij,l}) .$$

L'article que nous étudions utilisait pour la courbure sectionnelle la définition que nous avons présentée, mais l'écrivait ensuite sous la forme suivante :

$$\mathcal{K}(X, Y) = \frac{R(X, Y, Y, X)}{\|X\|_g^2 \|Y\|_g^2 - \langle X | Y \rangle_g^2} = \frac{R_{ijkl} X^i Y^j Y^k X^l}{\|X\|_g^2 \|Y\|_g^2 - \langle X | Y \rangle_g^2} .$$

C'est cette courbure sectionnelle que nous voulons faire calculer à un ordinateur.

Enfin, on introduit les notations suivantes, duales de la courbure de Riemann et des symboles de Christoffel :

$$R^{ursv} := R_{ijkl}g^{iu}g^{jr}g^{ks}g^{mv} \quad \text{et} \quad \Gamma_u^{rs} := g^{ir}g^{js}g_{ku}\Gamma_{ij}^k$$

et les symboles $g^{ij,k} := g^{ij}_{,\xi}g^{\xi k}$ et $g^{ij,kl} := g^{ij}_{,\xi\eta}g^{\xi k}g^{\eta l}$ de sorte à avoir, par analogie avec la formule rappelée ci-dessus,

$$\Gamma_u^{rs} = -\frac{1}{2}g_{u\phi} \left(g^{s\phi,r} + g^{r\phi,s} - g^{rs,\phi} \right).$$

La formule, dite de Mario ci-dessous permet de calculer la courbure d'un espace de landmarks. Initialement, Mario Micheli espérait pouvoir se débarrasser entièrement des termes dépendants de la métrique g , qui demandent de calculer l'inverse des noyaux, ce qui est coûteux algorithmiquement. Néanmoins, ça n'a pas été possible, comme le montre la présence d'un terme $g_{\lambda\mu}$ dans la formule ci-dessous.

Théorème V.3 (Formule de Mario)

Soit $m \in M$ et $X = X^i\partial_i$ et $Y = Y^i\partial_i$ deux vecteurs tangents à M en m . Alors, le numérateur de la courbure sectionnelle en m est donné par :

$$\begin{aligned} R(X, Y, Y, X) &= (X_u Y_r - Y_u X_r) \\ &\quad \times \left(\frac{1}{2}g^{su,rv} + \frac{1}{2}g^{us}_{,\rho}g^{\rho r,v} - \frac{1}{8}g^{us}_{,\sigma}g^{rv,\sigma} - \frac{3}{4}g^{\lambda u,r}g_{\lambda\mu}g^{\mu s,v} \right) \\ &\quad \times (X_s Y_v - Y_s X_v). \end{aligned}$$

Plus précisément, sans recourir à la notation d'Einstein, on a la formule suivante :

$$\begin{aligned} R(X, Y, Y, X) &= \sum_{u=1}^n \sum_{r=1}^n \sum_{s=1}^n \sum_{v=1}^n \left[\left(\sum_{i=1}^n \sum_{j=1}^n g_{ui}g_{rj}X^iY^j - \sum_{i=1}^n \sum_{j=1}^n g_{ui}g_{rj}X^jY^i \right) \right. \\ &\quad \times \left(\frac{1}{2} \sum_{\xi=1}^n \sum_{\eta=1}^n g^{su}_{,\xi\eta}g^{\xi r}g^{\eta v} + \frac{1}{2} \sum_{\rho=1}^n \left[g^{us}_{,\rho} \sum_{\xi=1}^n g^{\rho r}_{,\xi}g^{\xi v} \right] \right. \\ &\quad \left. - \frac{1}{8} \sum_{\sigma=1}^n \left[g^{us}_{,\sigma} \sum_{\xi=1}^n g^{rv}_{,\xi}g^{\xi\sigma} \right] - \frac{3}{4} \sum_{\lambda=1}^n \left[\left(\sum_{\xi=1}^n g^{\lambda u}_{,\xi}g^{\xi r} \right) g_{\lambda u} \left(\sum_{\xi=1}^n g^{us}_{,\xi}g^{\xi v} \right) \right] \right) \\ &\quad \left. \times \left(\sum_{i=1}^n \sum_{j=1}^n g_{si}g_{vj}X^iY^j - \sum_{i=1}^n \sum_{j=1}^n g_{si}g_{vj}X^jY^i \right) \right]. \end{aligned}$$

C'est là le principal écueil de notre projet. En effet, on espérait grâce à cette formule donner un moyen pratique de calculer la courbure de n'importe quel espace de landmarks et d'ainsi pouvoir déterminer *a priori* l'algorithme le mieux adapté au problème considéré. Mais, la complexité de la formule en a décidé autrement...

1. Implémentation de la formule

On a choisi d'implémenter la formule de Mario en python. On reproduit ici le code qui nous semble suffisamment limpide (il s'agit essentiellement d'imbriquer des boucles pour calculer les différentes sommes) pour se passer de commentaires au-delà de ceux présents au sein de ce-dernier. Les fonctions g et f calculent respectivement la métrique et la cométrique, les deux derniers arguments servant à dire si l'on veut les dérivées partielles de la métrique et la cométrique ou pas (-1,-1 pour ne pas les avoir, i,-1 pour l'ordre 1 et i,j pour dériver deux fois). Le calcul des dérivées partielles se fait par différences finies dans l'espace euclidien ambiant dans lequel est plongée la sous-variété, ce qui suppose un minimum de régularité des paramétrages locaux de ladite sous-variété.

```

1 import numpy as np
2 import numpy.linalg as lng
3
4
5 n=3                #
6 eps=10**-2        # Donnees initiales
7 X=np.arange(n)    #
8 Y=np.ones((n,))  #
9
10
11
12 def PlusEps(m,k):
13     # Fonction pour ajouter une petite quantite sur une coordonnee d'un
14     # vecteur m
15     h = m.copy()
16     h[k]+=eps
17     return h
18
19
20 def g(m,i,j,k,l):
21     # Calcul du tenseur metrique et de ses derivees
22     if k==-1 and l==-1:
23         # g=(1/(abs(sum(m))+1))**2*np.eye(n)
24         g=np.eye(n)
25         return g[i,j]
26     elif k!=-1 and l==-1:
27         return (g(PlusEps(m,k),i,j,-1,-1)-g(m,i,j,-1,-1))/eps
28     else :
29         return (g(PlusEps(m,l),i,j,k,-1)-g(m,i,j,k,-1))/eps
30
31
32
33 def f(m,i,j,k,l):
34     # Calcul de la cometrique et de ses derivees
35     if k==-1 and l==-1:
36         # com=(sum(m)+1)**2*np.eye(n)

```

```

37         com=np.eye(n)
38         return com[i,j]
39     elif k!=-1 and l==-1:
40         return ( f(PlusEps(m,k),i,j,-1,-1) - f(m,i,j,-1,-1))/eps
41     else :
42         return (f(PlusEps(m,l) ,i,j,k,-1)-f(m,i,j,k,-1))/eps
43
44
45 def Mario(m,X,Y):
46     # Calcul de la courbure sectionnelle
47
48     temp1,temp2,temp3=0,0,0
49     for i in range(n):
50         for j in range(n):
51             temp1+=X[i]*X[j]*g(m,i,j,-1,-1)
52             temp2+=Y[i]*Y[j]*g(m,i,j,-1,-1)
53             temp1+=X[i]*Y[j]*g(m,i,j,-1,-1)
54     Den=temp1*temp2-temp3**2      # Le denominateur
55
56     Num=0
57     for u in range(n):
58         for v in range(n):
59             for r in range(n):
60                 for s in range(n):
61
62                     Calc1,Calc2,Calc3,Calc4=0,0,0,0
63
64                     for xi in range(n):
65                         for eta in range(n):
66                             Calc1 += f(m,s,u,xi,eta) * f(m,xi,r,-1,-1) *
67                                 f(m,eta,v,-1,-1)
68
69                     for rho in range(n):
70                         for xi in range(n):
71                             Calc2 += f(m,u,s,rho,-1) * f(m,rho,r,xi,-1) *
72                                 f(m,xi,v,-1,-1)
73
74                     for sig in range(n):
75                         for xi in range(n):
76                             Calc3 += f(m,u,s,sig,-1) * f(m,r,v,xi,-1) * f
77                                 (m,xi,sig,-1,-1)
78
79                     for lamb in range(n):
80                         S1, S2 = 0, 0
81                         for xi in range(n):
82                             S1 += f(m,lamb,u,xi,-1) * f(m,xi,r,-1,-1)
83                             S2 += f(m,u,s,xi,-1) * f(m,xi,v,-1,-1)
84                         Calc4 += S1 * g(m,lamb,u,-1,-1)
85
86     Xu,Xr,Xs,Xv,Yu,Yr,Ys,Yv=0,0,0,0,0,0,0,0,0

```

```

84
85         for t in range(n):
86             Xu += g(m,u,t,-1,-1)*X[t]
87             Xr += g(m,r,t,-1,-1)*X[t]
88             Xs += g(m,s,t,-1,-1)*X[t]
89             Xv += g(m,v,t,-1,-1)*X[t]
90             Yu += g(m,u,t,-1,-1)*Y[t]
91             Yr += g(m,r,t,-1,-1)*Y[t]
92             Ys += g(m,s,t,-1,-1)*Y[t]
93             Yv += g(m,v,t,-1,-1)*Y[t]
94
95             Num += (Xu*Yr-Yu*Xr) * ( (1/2)*Calc1 + (1/2)*Calc2 -
96                   (1/8)*Calc3 - (3/4)*Calc4 ) * (Xs*Yv-Ys*Xv)
97
98         if Den!=0:
99             return Num/Den
100        else :
101            print('infinity')
102
103    print( Mario(np.ones((n,)),X,Y) )
104

```

Comme dit plus haut, la complexité de la formule, non-apparente du fait de la notation d'Einstein, ne nous a même pas permis de la tester sur notre exemple des amygdales, et ce malgré sa faible taille.

On aurait pu améliorer légèrement l'algorithme pour gagner en terme de stockage mais reste la limite dure du $O(n^6)$ opérations... Une dernière amélioration possible est de ne pas prendre uniquement les vecteurs X et Y en lesquels on veut calculer la courbure sectionnelle mais aussi les co-vecteurs définis par $X_i = g_{ij}X^j$ pour tout vecteur $X = X^i\partial_i$ que les méthodes d'appariements classiques permettent d'obtenir directement. Mais, la complexité ne serait pas réduite pour autant et on ne l'a donc pas fait, cet algorithme s'avérant inutilisable en pratique. Un énorme gain pouvait être réalisé dans la complexité en espace, en utilisant autant que possible les fonctions de numpy adaptées pour les matrices creuses, mais ce travail d'ingénierie était inutile puisque le facteur limitant était la complexité en temps.

Nous avons testé la formule, d'abord sur une métrique constante égale à l'identité, c'est à dire pour un espace euclidien classique, et le programme rend bien 0. Nous l'avons également testée sur une sphère de rayon 1 dans l'espace euclidien de dimension 3, en utilisant le code suivant :

```

1  n=2                                     #
2  eps=10**-5                             #  Donnees initiales
3  X=np.array([1,1])                      #
4  Y=np.array([1,0])                      #
5

```

```

6
7 def g(m,i,j,k,l):          # Calcul du tenseur metrique et de ses derivees
8     if k==-1 and l==-1:
9
10        metr = np.eye(2)
11        metr[1,1] = m[0]**2 + m[1]**2
12
13
14        return metr[i,j]
15     elif k!=-1 and l==-1:
16        return (g(PlusEps(m,k),i,j,-1,-1)-g(m,i,j,-1,-1))/eps
17     else :
18        return (g(PlusEps(m,l),i,j,k,-1)-g(m,i,j,k,-1))/eps
19
20 def f(m,i,j,k,l):          #Calcul de la cometrique et de ses derivees
21     if k==-1 and l==-1:
22        metr = np.eye(2)
23        metr[1,1] = m[0]**2 + m[1]**2
24        com=lng.inv(metr)
25        return com[i,j]
26     elif k!=-1 and l==-1:
27        return ( f(PlusEps(m,k),i,j,-1,-1) - f(m,i,j,-1,-1))/eps
28     else :
29        return ( f(PlusEps(m,l) ,i,j,k,-1)-f(m,i,j,k,-1))/eps
30
31 print( Mario([1,0],X,Y) )

```

et nous obtenons le résultat suivant :

```
In [1]: (executing lines 1 to 122 of "Courbure.py")
0.999959929674
```

Ce qui correspond bien à la courbure de Gauss de la sphère \mathbb{S}^2 .

– Annexe A –

"Rappels" sur les fibrés**I. Espaces fibrés**

Soit $k \in [0, +\infty]$. On définit ici la notion d'espaces fibrés, liée à celles de fibrés tangents (proposition II.10) et de champs de vecteurs sur une variété (définition II.15).

Définition A.1. Soit E et B deux variétés \mathcal{C}^k . Une application p de classe \mathcal{C}^k de E dans B est une fibration – de base B et d'espace total E – si pour tout $b \in B$ il existe un ouvert \mathcal{U} contenant b , une variété F et un difféomorphisme $\phi: \mathcal{U} \times F \rightarrow p^{-1}(\mathcal{U})$ tel que $p(\phi(y, z)) = y$ pour tout y de \mathcal{U} et z de F . On dit alors que E est un espace fibré.

Pour tout $b \in B$, $E_b := p^{-1}(b)$ est une sous-variété fermée de E , de dimension $\dim(E) - \dim(B)$ et que l'on appelle *fibres* de b . Par construction E_b est difféomorphe à F .

Intuitivement et au moins localement, cette définition signifie que l'on peut voir E comme un "empilement" indexé par F , éventuellement continu donc, d'ouverts difféomorphes à \mathcal{U} . Dans le cas où la base B est connexe, comme le formalise la proposition A.2, il existe une "unique" façon d'indexer ces ouverts; d'où le fait qu'on parle souvent de *fibres type*...

Exemple : Un exemple important d'espaces fibrés est celui des revêtements. Dans ce cas, "la" fibre F est discrète. Le revêtement du cercle \mathbb{S}^1 par une hélice est un exemple parmi d'autres...

Proposition A.2

Si $p: E \rightarrow B$ est une fibration de base connexe alors les fibres E_b sont difféomorphes.

Démonstration :

Soit $b_0 \in B$. Alors, pour tout point b_1 de B il existe un chemin continu $\gamma: [0, 1] \rightarrow B$ joignant b_0 et b_1 .

Tout $x \in \gamma([0, 1])$ satisfait en particulier la définition A.1 et il existe donc un ouvert \mathcal{U}_x contenant x tel que dans la définition. Par compacité de $\gamma([0, 1])$ on peut extraire de ce recouvrement un sous-recouvrement fini que l'on note $\mathcal{U}_1, \dots, \mathcal{U}_m$. De plus, quitte à re-numéroter les ouverts et en rajouter un nombre fini on peut supposer que

$$b_0 \in \mathcal{U}_0, \quad b_1 \in \mathcal{U}_m, \quad \forall i \in \llbracket 1, m-1 \rrbracket, \quad \mathcal{U}_i \cap \mathcal{U}_{i+1} \neq \emptyset.$$

"Rappels" sur les fibrés

Soit $b \in \mathcal{U}_i \cap \mathcal{U}_{i+1}$. Alors $b \in \mathcal{U}_i$ donc les fibres E_x pour $x \in \mathcal{U}_i$ sont diffeomorphes à E_b par définition. De même, $b \in \mathcal{U}_{i+1}$ donc les fibres E_x pour $x \in \mathcal{U}_{i+1}$ sont diffeomorphes à E_b et donc à l'ensemble des E_x pour $x \in \mathcal{U}_i$ par transitivité. Ainsi, de proche en proche, E_{b_0} est diffeomorphe à E_{b_1} . \square

Définition A.3. La fibration triviale est celle pour laquelle $E = B \times F$ et $p = pr_1$. On dit qu'une fibration est trivialisable s'il existe un diffeomorphisme $\phi: E \rightarrow B \times F$ tel que $p = pr_1 \circ \phi$. On dit alors que le fibré E est trivial (resp. trivialisable) et que ϕ est une trivialisations.

Remarque : Par définition, les fibrations sont localement trivialisables. Autrement dit, la fibration devient trivialisable au dessus d'ouverts suffisamment petits de la base.

Exemple : Un exemple non-trivial¹ de fibration est la *fibration de Hopf*. Il s'agit d'une fibration de la sphère \mathbb{S}^3 de dimension 4 sur \mathbb{S}^2 , de fibre \mathbb{S}^1 .

Une façon de définir \mathbb{S}^4 est de voir \mathbb{R}^4 comme \mathbb{C}^2 . Alors :

$$\mathbb{S}^3 = \{(z, w) \in \mathbb{C}^2 \mid |z|^2 + |w|^2 = 1\}.$$

De manière analogue, on peut voir \mathbb{R}^3 comme $\mathbb{C} \times \mathbb{R}$ et :

$$\mathbb{S}^2 = \{(z, x) \in \mathbb{C} \times \mathbb{R} \mid |z|^2 + |x|^2 = 1\}.$$

$$\begin{aligned} \text{On pose alors } p : \mathbb{S}^3 \subset \mathbb{C}^2 &\rightarrow \mathbb{S}^2 \subset \mathbb{C} \times \mathbb{R} \\ (z, w) &\mapsto (2z\bar{w}, |z|^2 - |w|^2). \end{aligned}$$

La projection stéréographique de \mathbb{S}^4 dans \mathbb{R}^3 permet de "visualiser" cette fibration comme le montre la figure A.1.

On représente ici de la même couleur un point sur l'espace de base \mathbb{S}^2 et la fibre correspondante. Ces deux images sont tirées d'une animation réalisée par Niles Johnson, accessible depuis sa page à l'adresse <http://nilesjohnson.net/hopf.html>.

$$\begin{array}{ccc} E_1 & \xrightarrow{f} & E_2 \\ & \searrow p_1 & \swarrow p_2 \\ & B & \end{array}$$

Définition A.4. Un isomorphisme entre deux fibrés E_1 et E_2 ayant pour même base B est un diffeomorphisme $f: E_1 \rightarrow E_2$ tel que $p_2 \circ f = p_1$.

Autrement dit, le diagramme ci-contre est commutatif.

Exemple : Un fibré trivialisable est isomorphe au fibré trivial.

1. Que l'on admet complètement ici car ce n'est pas le propos de notre projet mais qu'on cite parce que quand même c'est joli!

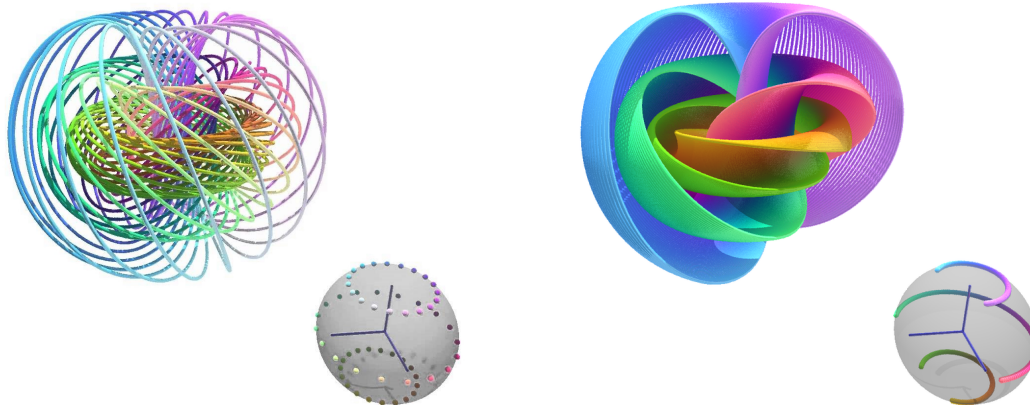


FIGURE A.1. – Fibration de Hopf – Projection stéréographique.

Crédit : Niles Johnson, The Ohio State University.

II. Espaces fibrés vectoriels

Définition A.5. Un fibré vectoriel réel (resp. complexe) de rang k sur une variété B est un espace fibré (E, p, B) tel que :

- i. la fibre type F et les fibres $p^{-1}(b)$, $b \in B$, sont des espaces vectoriels réels (resp. complexes) ;
- ii. pour toute trivialisation locale ϕ , la restriction de ϕ à $p^{-1}(b)$ induit un isomorphisme d'espaces vectoriels sur F .

De même que précédemment, on note E_b la fibre $p^{-1}(b)$ pour $b \in B$.

Définition A.6. Un morphisme entre deux fibrés vectoriels (E_1, p_1, B_1) et (E_2, p_2, B_2) est une application lisse $f: E_1 \rightarrow E_2$ telle que :

- i. f envoie chacune des fibres de E_1 sur une fibre de E_2 . Autrement dit, il existe une application $g: B_1 \rightarrow B_2$ rendant le diagramme ci-contre commutatif ;
- ii. la restriction de f à chacune des fibres $(E_1)_b$ est une application linéaire de $(E_1)_b$ dans $(E_2)_{g(b)}$.

$$\begin{array}{ccc}
 E_1 & \xrightarrow{f} & E_2 \\
 p_1 \downarrow & & \downarrow p_2 \\
 B_1 & \xrightarrow{g} & B_2
 \end{array}$$

Définition A.7. Un fibré est trivialisable s'il est isomorphe à un fibré trivial.

Définition A.8. Une section d'un fibré vectoriel E de base B est une application lisse s de B dans E telle que $p \circ s = \text{Id}_B$, autrement dit telle que $s(b) \in E_b$ pour tout b de B .

Références

- [13] Sylvestre Gallot, Dominique Hulin, and Jacques Lafontaine. *Riemannian Geometry*. Universitext. Springer, 3 edition, 2004.
- [23] Niles Johnson. A visualization of the hopf fibration, 2011.
<http://nilesjohnson.net/hopf.html>.
- [24] Jürgen Jost. *Riemannian geometry and geometry analysis*. Universitext. Springer, 3 edition, 2002.
- [25] Jacques Lafontaine. *Introduction aux variétés différentielles*. Grenoble Sciences. EDP Sciences, 2 edition, 2010.
- [28] Mario Micheli, Peter W. Michor, and David Mumford. Sectional curvature in terms of the co-metric and with applications to the riemannian manifolds of landmarks. *SIAM J. on Imaging Sciences*, 5(1), 2012.
- [34] Ricardo Sa Earp and Eric Toubiana. *Introduction à la géométrie hyperbolique et aux surfaces de Riemann*. Enseignement des mathématiques. Cassini, 2 edition, 2009.
- [43] Alain Trouvé. *Géométrie et espaces de formes*. Cours, 2016.
- [47] Laurent Younes. *Invariance, déformations et reconnaissance de formes*. Mathématiques et applications 44. Springer, 2003.

– Septième partie –

Mémoire de M2

Modèles de pseudo-métamorphoses pour le recalage diffeomorphique de courbes et surfaces

Encadré par Joan Alexis Glaunès *

*. MAP 5, Université Paris-Descartes

Sommaire – Partie VII

Mémoire de M2 : Modèle de pseudo métamorphoses pour le recalage diffeomorphique de courbes et surfaces	201
Chapitre I :	
Recalage diffeomorphique de courbes et surfaces	209
I Appariement de formes	209
I.1 Généralités sur les espaces de formes	209
I.1.a Encoder les formes	210
I.1.b L'appariement de formes en pratique	214
I.2 Appariement et modèles de déformations	215
I.2.a Modèles de petites et grandes déformations	215
I.2.b Plusieurs types d'appariements	218
I.3 Appariement et métamorphoses	220
II Appariement de mesures discrètes	221
II.1 Appariement (in)exact de mesures signées	222
II.1.a Formulation variationnelle	222
II.1.b Aspects algorithmiques	224
II.2 Modèle de métamorphoses	225
II.2.a Mise en équations du modèle	226
II.2.b Minimisation de la fonctionnelle d'appariement	230
II.3 Modèle de pseudo-métamorphose	231
Chapitre II :	
Modèle de pseudo-métamorphoses pour les mesures	233
I Modèle de pseudo-métamorphoses	233
I.1 Mise en équations du modèle	233
I.2 Équations géodésiques et reparamétrisation du modèle	236
I.2.a Recourt à la mécanique hamiltonienne	236
I.2.b Méthode de shooting	237
II Mise en œuvre algorithmique	240
II.1 Minimisation par descente de gradient	240
II.1.a Gradient de la fonctionnelle d'attache aux données	240

II.1.b	Établissement de l'équation différentielle	242
II.1.c	Établissement de l'équation adjointe	244
II.2	Pré-requis à l'implémentation numérique	246
II.3	Implémentation numérique	248
II.4	Expérimentations sur des données tests	250
II.4.a	Validation du Shooting par pseudo-métamorphoses	250
II.4.b	Validation du Matching par pseudo-métamorphoses	253
III	Métamorphoses vs pseudo-métamorphoses	260
Annexe A :		
	Résultats complémentaires sur les courants	263
I	Formes extérieures	263
II	Formes différentielles et courants	264
Annexe B :		
	Rappels de théorie des distributions	267
I	Distributions sur un ouvert de \mathbb{R}^n	267
I.1	Distributions	267
I.2	Support	269
I.3	Noyaux	270
II	Convolution	271
II.1	Convolution avec une fonction	271
II.2	Convolution de deux distributions	272
II.3	Support d'un produit de convolution	273
III	Transformation de Fourier	274
III.1	Espace de Schwartz	274
III.2	Distributions tempérées	276
IV	Solutions fondamentales	278
IV.1	La distribution de Dirac	278
IV.2	Existence et régularité des solutions	279
Annexe C :		
	Opérateurs pseudos-différentiels	281
I	Symbole d'un opérateur	281
I.1	Définitions et exemples	282
I.2	Sommes asymptotiques et symboles classiques	283
II	Opérateurs pseudo-différentiels	283

II.1	Définition et exemple	283
II.2	Régularité spatiale restreinte	284
	Références	285

Le présent document résume les travaux que j'ai effectués durant mon stage recherche de M2, *Mathématiques pour les Sciences du Vivant*, et qui conclura ainsi mes quatre années d'études au sein du Magistère de Mathématiques d'Orsay. Ce stage a été encadré par Joan Glaunès, que je remercie donc de m'avoir consacré du temps, alors même que le moment n'était pas le plus propice pour lui.

Une partie de mon travail a d'abord consisté à m'assurer – au moins pour moi – que j'avais acquis une compréhension suffisante de la modélisation induite par le recours aux "espaces de formes" en renforçant mes connaissances du cours d'Alain Trouvé. Également, avant de commencer à étudier le modèle dit de *pseudo-métamorphose*, il m'a fallu comprendre les limites du modèles de "simples" métamorphoses, au moins dans le cas des Landmarks. J'ai pour cela étudié l'article de Casey L. Richardson et Laurent Younes [33], ce qui m'a demandé de me former sur les opérateurs pseudo-différentiels, n'en n'ayant jamais étudiés auparavant. Tout ceci fait l'objet du premier chapitre.

Ensuite, j'ai pu me concentrer sur l'étude du modèle de pseudo-métamorphose proposé par Joan Glaunès et son implémentation numérique, comme détaillé au second chapitre. Néanmoins, le modèle, en l'état, ne s'est pas révélé aussi promoteur et il conviendrait de lui apporter certaines améliorations pour le rendre efficient.

Enfin, un dernier aspect de mon travail effectué lors de ce stage, et qui ne transparait pas dans ce rapport, a consisté en la traduction du code MatLab [14] réalisant le matching – exact et inexact – de deux configurations de points en langage Python. Si d'un point de vu purement mathématique, je n'en ai pas acquis de connaissances nouvelles, cela m'a permis de renforcer – assez solidement je pense – mes connaissances en programmation, et notamment en programmation orientée objets, et explique en partie le fait que je n'ai pas eu le temps d'améliorer le modèle de pseudo-métamorphose pour le rendre efficient.

Ce travail est cependant partiellement terminé à ce jour – le shooting "fonctionne" mais le matching n'est pas complètement débuggé – d'où son absence du rapport. J'espère pour autant pouvoir le terminer, d'abord à titre personnel, mais également dans l'objectif de promouvoir des langages/logiciels libres de droit tels que *numpy* et *scipy*.

– Chapitre I –

Recalage diffeomorphique de courbes et surfaces

I. Appariement de formes

Aujourd'hui l'acquisition d'images médicales est de plus en plus performante et importante et les outils mathématiques deviennent essentiels pour le traitement d'une telle masse d'informations. Si les statistiques sont bien évidemment nécessaires du fait du grand nombre de données, un traitement préalable est incontournable et c'est notamment tout l'intérêt de la géométrie riemannienne.

Plus précisément, on a envie de pouvoir comparer différentes formes anatomiques entre elles afin de pouvoir quantifier les variations inter et intra-individuelles¹ de la dite forme et c'est là l'objet de la théorie des espaces de formes. En effet, outre la grande variabilité des données d'acquisition qui justifie la mise en place d'une théorie générale et non organe-dépendante, il faut bien prendre conscience que la nature nous a fait tous semblables mais loin d'être identiques. Ainsi, la forme d'un même organe peut varier très fortement d'un individu à l'autre et il apparait nécessaire de pouvoir quantifier cette variabilité afin de déterminer si deux organes diffèrent de par leur nature propre ou bien parce que l'un est malade et l'autre non.

1. Généralités sur les espaces de formes

La théorie des espaces de formes apparait comme une tentative de modélisation mathématique de cette variabilité. On introduit pour cela un espace, dit *espace de formes* dans lequel chacun des points est en fait un objet géométrique et que l'on va munir d'une métrique à même donc de quantifier l'écart entre deux objets géométriques. Pour ce faire, on définira la distance entre deux formes comme "la transformation de l'espace la moins coûteuse permettant d'amener une forme sur l'autre" : aussi s'attachera-t-on pour définir la distance à minimiser une fonctionnelle d'énergie (cf paragraphe I.2.a), tout cela dans un sens à préciser. Cette distance est dite géodésique et l'espace de formes ainsi défini est en particulier une variété riemannienne.

1. Afin de suivre l'évolution d'une maladie par exemple...

a. Encoder les formes

Comme annoncé, l'idée des espaces de formes est de définir un objet mathématique – une variété riemannienne dans les faits – dont chacun des points représentera une forme anatomique. Il existe pour cela plusieurs manières de procéder et on en explicite ici trois, que l'on peut considérer comme des améliorations successives d'une même idée.

Landmarks : L'idée la plus naturelle pour encoder une forme est peut-être de "simplement" la discrétiser en interpolant sa surface avec un ensemble de points. Ainsi, une forme de dimension d sera encodée par un ensemble de n points de \mathbb{R}^d avec $n \in \mathbb{N}$ fixé. On parle alors de l'espace des landmarks, qui n'est autre que \mathbb{R}^{nd} en imposant que chacun des points soit distinct des autres :

$$\mathcal{L}^n(\mathbb{R}^d) = \{ (x_1, \dots, x_n) \in (\mathbb{R}^d)^n \mid \forall i, j \in \llbracket 1, n \rrbracket, i \neq j \implies x_i \neq x_j \}$$

Dans un tel formalisme, apparier deux formes est équivalent à apparier n points deux à deux. L'intérêt des landmarks réside principalement dans la simplicité de la mise en œuvre théorique et algorithmique de l'appariement de formes. Cependant, ce formalisme souffre de limitations importantes : la première est que ne sont appariables que des formes interpolées par le même nombre de points et à condition que l'on ait pris soin de spécifier quel point devait être apparié avec tel autre ; la seconde est que l'interpolation détruit la structure géométrique de la forme : une courbe n'est pas distinguable d'un nuage de points, etc.

Mesures discrètes : Pour palier la première limitation une amélioration possible est de ne plus considérer des n -uplets de points comme dans le cas des landmarks mais des sommes de mesures de dirac, enracinées en chacun des dits-points. Ainsi, une forme sera représentée par une mesure de la forme :

$$\mu = \sum_{i=1}^{n_\mu} \delta_{x_i} \quad \text{où} \quad \forall i \in \llbracket 1, n_\mu \rrbracket, x_i \in \mathbb{R}^d$$

En particulier, le nombre de points n_μ peut varier d'une forme à l'autre ce qui permet de coller au mieux aux données que l'on cherche à étudier. De plus, lors de l'appariement, l'utilisateur n'est pas obligé de spécifier quels points apparier deux à deux. Sur ce dernier point néanmoins, on aurait pu s'affranchir de cette contrainte dans le cadre des landmarks en décidant de réaliser les appariements pour l'ensemble des configurations possibles puis en minimisant la fonctionnelle correspondant au coût d'appariement sur l'ensemble de ces configurations ; mais, ce choix s'avère peut judicieux en pratique car alourdissant les équations géodésiques et le temps de calcul.

Éventuellement, on peut considérer une somme pondérée de mesures de Dirac, ce qui permet de mettre l'accent sur certaines zones anatomiques relativement à d'autres. Dans le modèle présenté au chapitre 2, on travaille surtout sur ce type d'encodage.

Courants : Enfin, les courants permettent de palier le second problème évoqué. On présente à l'annexe A, page 262, les connaissances théoriques qui soutiennent l'utilisation des courants pour l'encodage de formes. On détaille ici comment les courants peuvent servir à "encoder" des courbes et des surfaces. En fait, on peut voir les courants comme une généralisation des sous-variétés : si N est une sous-variété compacte orientée de dimension p de M , variété lisse, – typiquement \mathbb{R}^d – alors on définit un p -courant sur M en posant :

$$\forall \omega \in \Omega_0^p(M), \quad T_N(\omega) := \int_N \omega$$

et où $\Omega_0^p(M)$ désigne les p -formes différentielles sur M à support compact. On se contente ici d'expliciter l'affirmation précédente dans le cas des courbes et surfaces :

- **Cas des courbes :** Soit Γ la courbe orientée définie comme la classe d'équivalence dans $\mathcal{C}^1([0, 1], \mathbb{R}^d) / \text{Diff}_+^1([0, 1])$ d'une courbe paramétrée $\gamma : [0, 1] \rightarrow \mathbb{R}^d$ de classe \mathcal{C}^1 .
On pose :

$$\forall \omega \in \Omega_0^1(\mathbb{R}^d), \quad T_\Gamma(\omega) := \int_\Gamma \omega := \int_0^1 \omega_{\gamma(t)}(\gamma'(t)) dt$$

En particulier, on peut vérifier que l'expression ci-dessus ne dépend *que* de l'orientation de la paramétrisation γ et que T_Γ est continue pour la norme $\|\cdot\|_\infty$. Soit à présent une "généralisation" de la mesure de Dirac

$$\forall x, \alpha \in \mathbb{R}^d, \quad \forall \omega \in \Omega_0^1(\mathbb{R}^d), \quad \delta_x^\alpha(\omega) = \omega_x(\alpha)$$

que l'on peut voir comme un courant d'un élément infinitésimal de ligne, centré en $x \in \mathbb{R}^d$, orienté selon la direction α et de longueur $|\alpha|$. Cet élément bien que "basique" permet d'approcher des 1-courants plus généraux :

Proposition I.1

Posons $x_i := \gamma(\frac{i}{n})$ pour tout $i \in \llbracket 0, n \rrbracket$, $n \in \mathbb{N}$. Alors :

$$\forall \omega \in \Omega_0^1(\mathbb{R}^d), \quad \lim_{n \rightarrow +\infty} \sum_{i=1}^n \delta_{x_i}^{x_{i+1} - x_i}(\omega) = T_\Gamma(\omega)$$

Ainsi, on peut approcher toute courbe par une approximation discrète. De plus, cette approximation ne dépend pas du choix de la paramétrisation de la courbe mais de l'orientation de cette dernière.

- **Cas des surfaces :** On procède de même que précédemment. Soit Σ une surface, *i.e* une variété de dimension 2, de \mathbb{R}^d et $\sigma : \mathcal{U} \subset \mathbb{R}^2 \rightarrow \mathbb{R}^d$ une paramétrisation \mathcal{C}^1

Recalage difféomorphique de courbes et surfaces

de Σ . On suppose² que

$$|S| := \int_{\mathcal{U}} \left| \frac{\partial \sigma}{\partial s}(s, t) \wedge \frac{\partial \sigma}{\partial t}(s, t) \right| ds dt < +\infty$$

On définit donc un courant en posant :

$$\forall \omega \in \Omega_0^2(\mathbb{R}^d), \quad T_\Sigma(\omega) := \int_\Sigma \omega := \int_{\mathcal{U}} \omega_{\sigma(s,t)} \left(\frac{\partial \sigma}{\partial s}(s, t) \wedge \frac{\partial \sigma}{\partial t}(s, t) \right) ds dt$$

et on peut vérifier que, comme précédemment, l'expression précédente ne dépend que de l'orientation de Σ , via le choix de σ .

On définit également les courants "élémentaires" dans le cas des surfaces en posant :

$$\forall x, u, v \in \mathbb{R}^d, \quad \forall \omega \in \Omega_0^2(\mathbb{R}^d), \quad \delta_x^{u \wedge v}(\omega) = \omega_x(u \wedge v)$$

et que l'on peut interpréter comme un élément de surface infinitésimal, centré en x , d'aire $u \wedge v$ dans le plan engendré par (u, v) et on dispose du pendant de la proposition I.1 :

Proposition I.2

Soit une triangulation $(T_i)_{i \in I}$ de la surface Σ d'orientation consistante deux à deux, *i.e* telle que deux triangles de la triangulation adjacents ont un sens de parcours *opposés* lorsqu'ils partagent une arête commune. On note, pour tout i de I , $T_i := (a_i, b_i, c_i)$ et on se donne un point $x_i \in T_i$, typiquement le barycentre de T_i . Alors, on peut approcher le courant T_Σ par

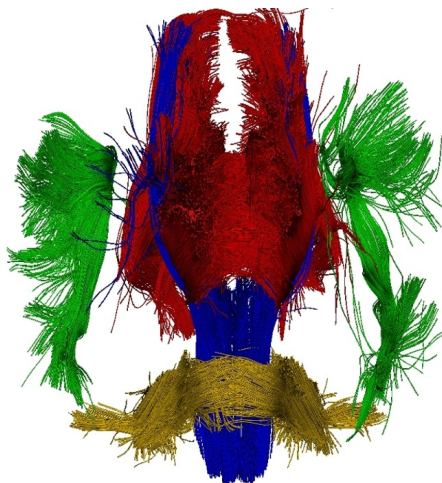
$$\bar{T}_\Sigma := \sum_{i \in I} \delta_{x_i}^{\vec{a}_i \wedge \vec{b}_i \wedge \vec{c}_i}$$

De plus, le choix de la triangulation est sans réelle conséquence sur l'approximation, à condition que la taille des différents triangles soient d'une taille cohérente vis-à-vis de la dynamique des équations de transports... Tout ceci dans un sens à préciser plus formellement, ce qui est fait au chapitre 5 de [16] par exemple.

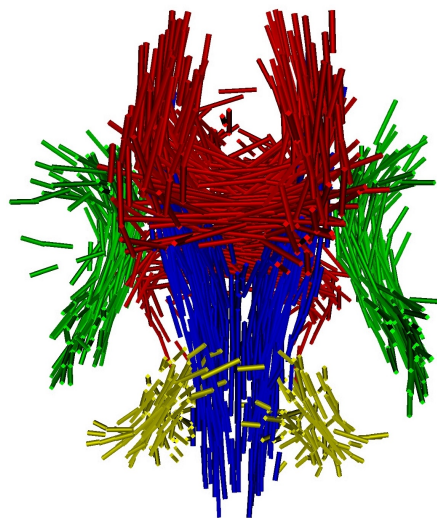
Ce ne sont bien sûr pas les seuls espaces de formes existants. Et les courants eux-mêmes souffrent de leur propre limitation : principalement, ils demandent de pouvoir orienter les formes que l'on cherche à apparier, ce qui n'est pas toujours possible en pratique, notamment dès lors que l'on cherche à étudier de la matière blanche comme l'illustre la figure I.1.

2. Cette hypothèse est en fait automatique dans le cas des courbes car γ étant supposée \mathcal{C}^1 , γ' est en particulier continue et donc $|L| = \int_0^1 |\gamma'(t)| dt < +\infty$.

Elle illustre une façon d'encoder différentes fibres de la matière blanche mathématiquement. La matière blanche est en effet principalement constituée d'axones, d'où son aspect "fibreux". Elle est responsable de la propagation des informations dans le système nerveux et relie notamment les différentes aires de la matière grise, située en périphérie du cerveau. Sont représentés ici : en bleu, le faisceau cortico-spinal ; en jaune, le faisceau cortico-nucléaire ; en rouge, le corps calleux et en vert les faisceaux arqués droit et gauche.



(a) Exemple de fibres corticales dans le cerveau d'un patient.



(b) Modèle mathématique associé – basé sur des courants ici.

FIGURE I.1. – Des faisceaux de fibres dans le cerveau.

Crédit : D. Ducreux, Hôpital Kremlin-Bicêtre et P. Fillard, CEA, Neurospin.

En médecine, il y a un réel intérêt à pouvoir observer ces fibres en tant que fibres justement car cela permet de détecter des maladies telles que la sclérose en plaque, qui n'est autre que la destruction de la myéline protégeant les axones de la matière blanche, la maladie d'Alzheimer et d'autres maladies neurodégénératives. Cependant, l'orientation des différentes fibres corticales est loin d'être claire d'où l'utilité d'un formalisme ne nécessitant aucune orientation. Les *varifolds* que nous ne présentons pas ici permettent par exemple cela.

Précisons cependant que les *varifolds* reposent également sur la notion de mesure de Dirac. Finalement, on constate qu'une grande partie de l'appariement anatomique repose donc sur l'appariement de mesures discrètes, via le matching de sommes de Dirac. D'où notre choix lors de ce stage de se concentrer sur l'appariement de mesures discrètes. On présente au paragraphe II du présent chapitre deux méthodes d'appariements (cf I.2.b

pour une description des différentes techniques d'appariement) spécifiquement pour les mesures discrètes ainsi que l'idée générale de notre modèle de pseudo-métamorphoses et qui fait l'objet du chapitre 2.

b. L'appariement de formes en pratique

L'intérêt pratique de l'appariement de formes est grand, notamment dans le domaine de l'imagerie médicale. En effet, comme l'avait déjà compris D'Arcy Thompson en 1917, dès lors que l'on s'intéresse à la morphologie de nos différents organes, le plus important n'est pas d'être à même de fournir une description détaillée de chacun d'entre eux mais plutôt de pouvoir quantifier les déformations des uns par rapport aux autres, et c'est là l'idée principale du modèle dit *de grande déformation* et faisant l'objet du paragraphe suivant.

In a very large part of morphology, our essential task lies in the comparison of related forms rather than in the precise definition of each... This process of comparison, of recognizing in one form a definite permutation or deformation of another,...lies within the immediate province of mathematics and finds its solution in...the Theory of Transformations... I learnt of it from Henri Poincaré.

D'Arcy Thompson – *On Growth and Form*, 1917, [42].

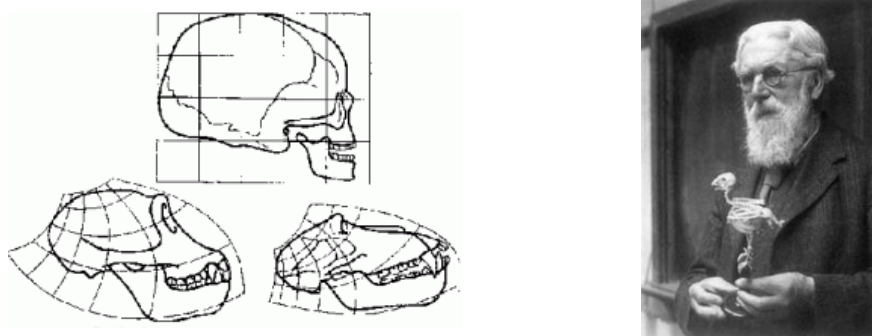


FIGURE I.2. – Crânes d'humain, de chimpanzé et de babouin. – D'Arcy Thompson.

Comme expliqué en introduction, pour tenir compte de la variabilité intrinsèque des formes anatomiques, on mesure leur similitude par le biais d'une fonctionnelle encodant le coût de déformation nécessaire à la "superposition" de la première forme sur la seconde, et qui ne dépend bien évidemment pas du choix de l'organe à déformer, par symétrie du problème. Il peut également être judicieux de considérer que toutes les déformations ne se valent pas, notamment pour le diagnostic précoce de maladie où certaines déformations spécifiques d'un organe (hippocampe dans le cas d'Alzheimer par exemple) vont être symptomatiques alors que d'autres sont juste le fruit de la diversité anatomique.

Dans ce cas, on imposera un poids plus important sur ces déformations spécifiques dans le coût total de déformation.

Compte-tenu de l'intérêt pratique de l'appariement de formes en anatomie computationnelle, il est nécessaire de concevoir des méthodes implémentables et donc utilisables *in extenso* par des praticiens. Du fait de cette "contrainte" une grande partie de mon stage a également consisté en de la programmation : notamment par la traduction du code MatLab décrit en [14] en langage Python – travail partiellement abouti à ce jour – et l'implémentation d'un code MatLab réalisant le matching de deux formes par pseudo-métamorphoses, comme décrit au chapitre 2, paragraphe II.3.

2. Appariement et modèles de déformations

Dans l'approche classique, les déformations sont encodées via un groupe agissant sur l'espace de formes. En fait de groupe, on va considérer une "collection" de transformations de l'espace. Historiquement, on a d'abord considéré des groupes de transformations "simples" : rotation, translation, homothétie... Mais, cela ne permettait pas de modéliser tout type de déformation. D'autre part, on a envie que les déformations soient réversibles pour des questions de symétrie évidentes et donc que les transformations mathématiques associées soient inversibles. Ainsi, considérer un groupe de difféomorphismes – avec éventuellement des restrictions – est une idée naturelle.

a. Modèles de petites et grandes déformations

Un premier modèle dit de *petites déformations* reposait sur la minimisation d'un coût de déformation pour le champ de déplacement $\phi - Id$, avec l'idée que le morphisme Id correspond à la situation idéale où aucune déformation n'est nécessaire. Cependant, dès lors que la déformation devient trop importante, *i.e* dès lors que ϕ devient grand devant Id en norme, un tel morphisme n'est plus inversible.

Aussi, plutôt que de considérer le résultat des déformations, le modèle de *grandes déformations* consiste à utiliser la même approche que précédemment, mais de manière infinitésimale. Plus précisément, on se donne un espace (de Hilbert) V de champs de vecteurs dont la norme définit le coût de déformation et les difféomorphismes sont obtenus par intégration d'une famille $(v_t)_{t \in [0,1]}$ de champs de vecteurs selon l'équation :

$$\partial_t \phi_t^v = v_t \circ \phi_t^v \quad ; \quad \phi_0^v = Id$$

Ainsi, en minimisant la norme des différents champs de vecteurs v_t , on minimise à chaque instant la déformation et donc la déformation globale, qui est donnée par ϕ_1 .

Recalage difféomorphique de courbes et surfaces

En fait, certaines limitations sont nécessaires. On se contente d'énoncer ici les théorèmes généraux qui soutiennent la suite, sans démonstration. On en trouve une démonstration au chapitre 1 de [16], ainsi que des résultats complémentaires.

Tout d'abord, il faut remarquer que l'équation, telle qu'écrite ci-dessus, n'est pas à valeur dans \mathbb{R}^d mais dans un espace fonctionnel. Par ailleurs, $t \mapsto \phi_t^v$ n'est à priori pas à valeurs dans \mathbb{R}^d mais dans la variété Riemannienne de l'espace des formes. Aussi lui préférera-t-on le plus souvent son écriture intégrale où \int désigne l'intégrale de Bochner qui généralise aux espaces de Banach la notion d'intégration :

$$\phi_t^v(x) = x + \int_0^t v_s \circ \phi_s^v(x) \, ds$$

Enfin, une telle équation est dite *équation de flot* et, compte-tenu des limitations suscitées, les résultats classiques d'existence et d'unicité ne s'appliquent donc pas en l'état. En fait, à condition d'introduire une restriction sur les champs de vecteurs v_t autorisés, on "récupère" l'existence et l'unicité à l'équation de flot.

Définition I.3. *Un espace vectoriel V de champs de vecteurs de \mathbb{R}^d est dit admissible si et seulement si :*

- i. *V est un espace de Hilbert. Dans ce cas, on note $\|\cdot\|_V$ sa norme.*
- ii. *V s'injecte continument dans l'espace $C_0^1(\mathbb{R}^d)$ des champs de vecteurs de classe C^1 sur \mathbb{R}^d s'annulant à l'infini, ainsi que leurs dérivés partielles. Autrement dit, il existe une constante C telle que pour tout $v \in V$,*

$$\|v\|_\infty + \|\mathcal{D}v\|_\infty \leq C \|v\|_V \quad \text{où} \quad \mathcal{D}v: x \mapsto \mathcal{D}_x v$$

Dorénavant, V désignera un espace de Hilbert *admissible*.

On note $L_V^1 := \mathbb{L}^1([0, 1], V)$ et $L_V^2 := \mathbb{L}^2([0, 1], V)$ les espaces des classes d'équivalences des applications $t \mapsto v_t$ de $[0, 1]$ dans V telles que les quantités respectives $\|v\|_{L_V^1} := \int_0^1 \|v_t\|_V \, dt$ et $\|v\|_{L_V^2}^2 := \int_0^1 \|v_t\|_V^2 \, dt$ soient finies. Munis de leur norme usuelle, L_V^1 est un espace de Banach et L_V^2 un espace de Hilbert. De plus, d'après l'inégalité de Cauchy-Schwartz, pour tout champ de vecteur v , $\|v\|_{L_V^1} \leq \|v\|_{L_V^2}$ et donc L_V^2 s'injecte continument dans L_V^1 .

Théorème I.4 (Cauchy-Lipschitz)

Pour tout champ de vecteur $v \in L_V^1$ et tout point $x \in \mathbb{R}^d$, il existe une unique application continue $t \mapsto \phi_t^v(x)$ de $[0, 1]$ dans \mathbb{R}^d et vérifiant l'équation de flot.

Ainsi, on est assuré de l'existence d'un difféomorphisme traduisant l'effet des déformations infinitésimales. On a également un contrôle des variations du flot ϕ_t^v en temps, en espace et vis-à-vis du champ v , de manière quasiment automatique via le *lemme de Gronwall*. Un résultat plus important :

Proposition I.5

Pour tout $v \in \mathbb{L}_V^1$, le flot ϕ_t^v est un difféomorphisme de classe \mathcal{C}^1 à chaque instant t .

En particulier, au temps $t = 1$, ϕ_1^v est régulier et inversible, ce qui valide l'approche. De plus, si on impose à V de s'injecter continument dans \mathcal{C}_0^p pour $p \geq 1$, les déformations ϕ_t^v induites seront d'autant plus régulières.

Finalement, on construit le groupe des déformations comme l'ensemble des difféomorphismes ϕ_t^v obtenus à partir des différents champs de vecteurs $v \in \mathbb{L}_V^1$ en posant :

$$\mathcal{G}_V := \left\{ \phi_t^v \mid v \in \mathbb{L}_V^1, t \in [0, 1] \right\}$$

En fait, il suffit de considérer les flots au temps $t = 1$. En effet, pour tout $t \in [0, 1]$ et tout $v \in \mathbb{L}_V^1$, $\phi_1^{\tilde{v}} = \phi_t^v$ où on définit \tilde{v} par : $\forall s \leq t, \tilde{v}_s := v_s$ et $\tilde{v}_s = 0$ sinon. Finalement, on écrira plus simplement :

$$\mathcal{G}_V = \left\{ \phi_1^v \mid v \in \mathbb{L}_V^1 \right\}$$

et la proposition suivante permet de transporter le coût de déformation encodé dans la norme de V sur le groupe \mathcal{G}_V :

Proposition I.6

\mathcal{G}_V est un groupe et un espace complet pour la métrique d_V définie par :

$$\forall \phi \in \mathcal{G}_V, \quad d_V(Id, \phi) := \inf \left\{ \|v\|_{\mathbb{L}_V^1} \mid v \in \mathbb{L}_V^1, \phi_1^v = \phi \right\}$$

et que l'on étend par invariance à droite :

$$\forall \phi, \psi \in \mathcal{G}_V, \quad d_V(\phi, \psi) = d_V(Id, \psi \circ \phi^{-1})$$

En fait, la construction du groupe \mathcal{G}_V via des champs de vecteurs \mathbb{L}_V^1 reste vraie avec des champs de vecteurs \mathbb{L}_V^2 , ce qui simplifie souvent le coût algorithmique :

Proposition I.7

D'une part,

$$\mathcal{G}_V = \left\{ \phi_1^v \mid v \in \mathbb{L}_V^2 \right\}$$

et, d'autre part,

$$\forall \phi \in \mathcal{G}_V, \quad d_V(Id, \phi) := \inf \left\{ \|v\|_{\mathbb{L}_V^2} \mid v \in \mathbb{L}_V^2, \phi_1^v = \phi \right\}$$

Recalage difféomorphique de courbes et surfaces

Enfin, le théorème suivant assure l'existence d'au moins une géodésique, *i.e* d'un champ de vecteur v permettant de transporter une configuration de l'espace de formes sur une autre et pour lequel le coût de déformation est minimal. Ainsi, le problème de matching admet (au moins) une solution.

Théorème I.8

Pour tout difféomorphisme $\phi, \psi \in \mathcal{G}_V$ il existe un champ de vecteur $v \in \mathbb{L}_V^2$ tel que :

$$\phi_1^v = \psi \circ \phi^{-1} \quad \text{et} \quad d(\phi, \psi) = \|v\|_{\mathbb{L}_V^2} = \|v\|_{\mathbb{L}_V^1}$$

De plus, pour un tel champ v , la norme $\|v_t\|_V$ est constante pour presque tout temps $t \in [0, 1]$.

b. Plusieurs types d'appariements

Depuis le début du chapitre, on parle d'appariement de formes. En fait, plusieurs modèles ont été développés pour permettre un tel appariement. Le but du présent rapport est de proposer un nouveau modèle ; aussi résume-t-on d'abord les modèles principaux ainsi que leurs avantages et limitations respectives.

Très schématiquement, l'appariement de deux formes – dans notre cas d'étude mais cela reste également vrai pour des images – consiste à déterminer le difféomorphisme optimal $\phi = \phi_1^v$ tel que la première des deux formes, transportée par ϕ , soit la plus "proche" possible – en un sens à préciser – de la seconde.

Par la suite, on désigne par y et z la première et seconde forme respectivement.

Appariement exact : En accord avec les notations du paragraphe précédent, on note M l'espace de formes considéré et \mathcal{G}_V le groupe des difféomorphismes agissant sur M . Pour tout $x \in M$ et $\phi \in \mathcal{G}_V$, $\phi \cdot x$ désigne donc l'image de x transportée par ϕ .

Définition I.9. *Le problème d'appariement exact consiste à minimiser, sur \mathcal{G}_V , la fonctionnelle suivante*

$$J(\phi) := d_V(Id, \phi) \quad \text{sous contrainte que} \quad \phi(y) = z$$

Concrètement, on force l'appariement à être exact, *i.e* que $\phi(y) = z$ et on minimise la déformation induite par ϕ : le morphisme Id correspondant à la situation idéale où aucune déformation n'est nécessaire pour réaliser l'appariement. De manière équivalente, et d'après le théorème I.8, il s'agit de minimiser sur \mathbb{L}_V^2 , et avec la même contrainte, la fonctionnelle

$$J(v) := \int_0^1 \|v_t\|_V^2 dt$$

De manière évidente, le problème d'appariement exact admet au moins une solution : la fonctionnelle J atteint en particulier son minimum – en ne tenant pas compte des

contraintes – pour $\phi \equiv Id$.

Le principal intérêt de l'appariement exact est justement le côté "exact" de celui-ci. Néanmoins, il peut détruire la structure des formes que l'on cherche à apparier, surtout dans le cas de matching par landmarks si les points de repère ne sont pas choisis judicieusement.

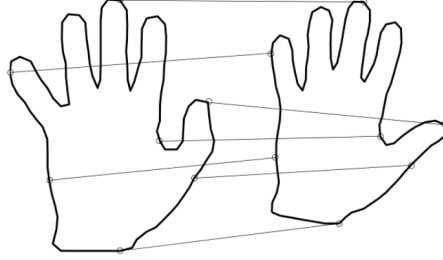


FIGURE I.3. – Matching par landmarks et targeting.

Appariement inexact : Pour palier ce problème, on a souvent recourt à l'appariement dit inexact, où la contrainte d'appariement est intégrée dans la fonctionnelle à minimiser. Pour cela, on doit d'abord définir une distance d_M sur l'espace de formes afin de mesurer l'appariement entre les objets. On définit alors un terme que l'on appelle *attache aux données* et qui mesure l'écart entre la forme initiale y transportée par le morphisme ϕ et la forme cible z en posant (par exemple) :

$$\forall \phi \in \mathcal{G}_V, \quad A(\phi) := d_M(\phi \cdot y, z)$$

Définition I.10. Soit $\gamma > 0$. Le problème d'appariement inexact associé à la fonctionnelle d'attache aux données A consiste à minimiser, sur \mathcal{G}_V , la fonctionnelle suivante

$$J_\gamma(\phi) := \gamma d_V(Id, \phi) + A(\phi)$$

De même, ceci est équivalent à minimiser sur \mathbb{L}_V^2 la fonctionnelle

$$J_\gamma(v) := \gamma \int_0^1 \|v_t\|_V^2 dt + A(\phi_1^v)$$

et le théorème suivant nous assure de l'existence d'un tel minimum :

Théorème I.11

Soit $A: \mathcal{G}_V \rightarrow [0, +\infty[$ une fonctionnelle telle que $v \mapsto A(\phi_1^v)$ soit faiblement continue de \mathbb{L}_V^2 dans \mathbb{R} . Alors, pour tout $\gamma > 0$, il existe un minimum pour

le problème d'appariement correspondant à A .

En pratique, on préfère le plus souvent l'appariement inexact à l'appariement exact comme expliqué ci-dessus. Cependant si l'appariement est numériquement bon avec un tel modèle, la fonctionnelle que l'on minimise et qui sert à mesurer le coût de déformation n'est *pas* une distance du fait de la présence du terme d'attache aux données. Ceci peut notamment poser un problème dès lors que l'on veut recourir à un modèle statistique, ce qui est le cas si on désire faire de l'estimation multi-template par exemple : on applique alors des résultats établis mathématiquement pour des distances à une quantité qui n'en est pas une et donc sans réelle garantie. Une des motivations du modèle d'appariement que l'on décrit au paragraphe suivant, et dit de *métamorphoses*, est de justement fournir une métrique sur le groupe \mathcal{G}_V ne souffrant pas des limitations de l'appariement exact.

3. Appariement et métamorphoses

Une autre motivation est de pouvoir quantifier les déformations dues à l'action du groupe \mathcal{G}_V et à l'évolution de la forme. Plus concrètement et pour paraphraser [44], si on se donne une forme $x = (x^1, \dots, x^n) \in M \subset \mathbb{R}^{nd}$ et n vecteurs unitaires de \mathbb{R}^d : u^1, u^2, \dots, u^n ; alors, on a envie de dire que $(x^1 + u^1, \dots, x^n + u^1)$ est "plus proche" de x que $(x^1 + u^1, \dots, x^n + u^n)$ en tant que "simple" translaté de x . Pourtant, avec le modèle précédent, le terme d'attache aux données sera identique, les vecteurs u^i étant choisis unitaires. Autrement dit, on aimerait pouvoir dire que le coût de déformation est plus faible si cette déformation est due – au moins en partie – à l'action du groupe \mathcal{G}_V sur M , ce que rend possible le modèle de métamorphose.

Définition I.12. *On appelle métamorphose un couple de courbes $(g_t, \mu_t) \in \mathcal{G}_V \times M$ paramétrisé par le temps et tel que $g_0 = Id$. Son image est la courbe $m_t \in M$ définie par $m_t = g_t \cdot \mu_t$. On appelle partie relative à la déformation – *deformation part* – de la métamorphose g_t et partie relative à l'évolution du gabarit – *template part* – de la métamorphose μ_t . Enfin, on dit que la métamorphose est une déformation pure lorsque μ_t est constant.*

Notons v_t la vitesse du morphisme g_t au temps t , *i.e* le champ de vecteur tel que

$$\forall t \in [0, 1], \quad \partial_t g_t = v_t \circ g_t$$

Alors, l'évolution de l'image m_t de la métamorphose est donnée par :

$$\forall t \in [0, 1], \quad \partial_t m_t = \mathcal{D}_{\mu_t} g_t \cdot \partial_t \mu_t + \partial_t g_t(\mu_t) = \mathcal{D}_{\mu_t} g_t \cdot \partial_t \mu_t + v_t(m_t)$$

En particulier, pour $t = 0$,

$$\partial_t m_t|_{t=0} = \partial_t \mu_t|_{t=0} + v_0(m_0)$$

et on peut donc voir tout vecteur tangent $\zeta \in T_m M$ comme "image" d'une forme de *métamorphose infinitésimale* $(v, \delta) \in V \times T_m M$.

On montre alors (par un argument de projection – voir l'article d'Alain Trouvé et Laurant Younes [44]) que, pour tout $\sigma^2 > 0$,

$$|\mu|_m^2 := \inf \left\{ \|v\|_V^2 + \frac{1}{\sigma^2} \|\delta\|_m^2 \mid \mu = v(m) + \delta \right\}$$

définit une distance (riemannienne) sur M . L'énergie de la courbe image m_t est alors donnée par :

$$E(m_t) = \int_0^1 |\partial_t m_t|_{m_t}^2 dt = \inf \left\{ \int_0^1 \|v_t\|_V^2 dt + \frac{1}{\sigma^2} \int_0^1 \|\partial m_t - v_t(m_t)\|_{m_t}^2 dt \mid v_t \in V \right\}$$

Finalement, on mesure la distance entre deux points m_y et m_z de M en minimisant sur l'ensemble des courbes $(g_t, m_t) \in V \times M$ la fonctionnelle

$$J(v_t, m_t) = \int_0^1 \|v_t\|_V^2 dt + \frac{1}{\sigma^2} \int_0^1 \|\partial m_t - v_t(m_t)\|_{m_t} dt$$

sous la contrainte $m_0 = m_y$ et $m_1 = m_z$. De manière équivalente, en notant g_t la partie déformation associée à m_t et $\mu_t := g_t^{-1}(m_t)$, la fonctionnelle à minimiser se ré-écrit :

$$J(m_t) = \int_0^1 \|v_t\|_V^2 dt + \frac{1}{\sigma^2} \int_0^1 \|\mathcal{D}_{\mu_t} g_t \cdot \partial_t \mu_t\|_{m_t} dt$$

Du fait du caractère "métrique" de la fonctionnelle d'énergie, on espère que l'appariement par métamorphoses soit plus "robuste" numériquement et surtout compatible avec les modèles statistiques d'estimation. Par ailleurs, contrairement aux modèles difféomorphiques, l'appariement par métamorphoses permet de s'affranchir de contraintes topologiques, telle que la connexité, du fait de la présence d'un gabarit μ_t évoluant au cours du temps et que l'on transporte *a posteriori* par difféomorphisme pour avoir l'image. Cela peut avoir un intérêt notable si l'on cherche à faire de la détection/comparaison de tumeurs, qui peuvent se scinder au cours du temps ou métastaser par exemple...

Cependant, ce type d'appariement s'avère lourd, et ce même dans le cas "simple" des mesures discrètes, comme décrit au paragraphe II.2 et dans [33].

II. Appariement de mesures discrètes

Comme nous l'avons vu au paragraphe I.1, il y a un réel intérêt à s'intéresser à l'appariement de mesures, et plus particulièrement de mesures discrètes. On propose ici trois

approches différentes : la première, plus élémentaire et basée sur [17] permet l'appariement exact et inexact de mesures signées et donc, en particulier, l'appariement de sommes de Dirac ; la seconde, tirée de [20] et [33] réalise l'appariement par métamorphoses de telles sommes ; enfin, le dernier paragraphe présente les grandes idées de notre modèle de pseudo-métamorphoses. On revient en détail sur l'étude de ce modèle au chapitre suivant et notamment sur l'aspect algorithmique de celui-ci.

1. Appariement (in)exact de mesures signées

On se place dans le cadre du modèle de grande déformation : à savoir, on se donne un espace V de champs de vecteurs et le groupe de difféomorphismes \mathcal{G}_V qui lui est naturellement associé via l'équation de flot. A savoir,

$$\forall t \in [0, 1], \quad \begin{aligned} \partial_t \phi_t &= v_t \circ \phi_t & \text{et} & & \mathcal{G}_V &= \{ \phi_1^v \mid v \in \mathbb{L}_V^1 \} \\ \phi_0 &= Id \end{aligned}$$

et on munit \mathcal{G}_V de sa distance (invariante à droite) usuelle :

$$\forall \phi, \psi \in \mathcal{G}_V, \quad d(\phi, \psi) = \inf \left\{ \int_0^1 \|v_t\|_V dt \mid \phi_1^v \circ \phi = \psi \right\}$$

Enfin, on note \mathcal{M}_s l'ensemble des mesures signées de \mathbb{R}^d , à savoir le dual de l'espace $\mathcal{C}_{\rightarrow 0}(\mathbb{R}^d)$ des fonctions continues et tendant vers 0 à l'infini.

a. Formulation variationnelle

Le groupe \mathcal{G}_V agit sur \mathcal{M}_s via l'action suivante et que l'on appelle *transport de masse* :

$$\begin{aligned} \mathcal{G}_V \times \mathcal{M}_s &\rightarrow \mathcal{M}_s & \text{où} & & \forall f \in \mathbb{L}_V^1, & \phi\mu(f) &:= \int f \circ \phi d\mu \\ (\phi, \mu) &\mapsto \phi\mu \end{aligned}$$

pour toute fonction f mesurable et bornée. En particulier, dans le cas d'une mesure discrète $\mu = \sum_i \alpha^i \delta_{x^i}$, on a $\phi\mu = \sum_i \alpha^i \delta_{\phi(x^i)}$. En effet, l'action est linéaire et pour toute masse de Dirac δ_x , et toute fonction test f mesurable,

$$\phi\delta_x(f) = \int_{\mathbb{R}^d} f \circ \phi d\delta_x = f \circ \phi(x) \quad \text{i.e.} \quad \phi \cdot \delta_x = \delta_{\phi(x)}$$

Soit μ_y et μ_z deux mesures signées et que l'on cherche à appairer. On cherche à munir l'espace des mesures signées \mathcal{M}_s d'une distance afin de pouvoir comparer $\phi \cdot \mu_y$, pour tout difféomorphisme ϕ de \mathcal{G}_V , à μ_z et ainsi évaluer l'écart entre notre mesure cible μ_z et la mesure initiale déformée $\phi \cdot \mu_y$. \mathcal{M}_s est en fait muni d'une norme naturelle, dite *norme de la variation totale* et qui en fait la norme duale de la norme de $\mathcal{C}_{\rightarrow 0}(\mathbb{R}^d)$:

$$\forall \mu \in \mathcal{M}_s, \quad \|\mu\|_{VT} = \sup \left\{ \int_{\mathbb{R}^d} f d\mu \mid f \in \mathcal{C}_{\rightarrow 0}(\mathbb{R}^d), \|f\|_{\infty} = 1 \right\}$$

Dans le but de fournir un algorithme implémentable "facilement", on va supposer que \mathcal{M}_s est en fait plongé dans le dual d'un espace de Hilbert à noyau reproduisant – et que l'on désignera par l'acronyme anglais RKHS par la suite – H ; en faisant alors hériter \mathcal{M}_s de la norme duale de la norme de H , on aura en effet un moyen efficace d'évaluer les distances entre deux mesures de \mathcal{M}_s , du fait des propriétés des noyaux reproduisants.

On suppose donc que l'ensemble des fonctions bornées de \mathbb{R}^d s'injecte continument dans un second espace de Hilbert H . Alors, en particulier $\mathcal{C}_{\rightarrow 0}(\mathbb{R}^d) \subset \mathbb{L}^\infty(\mathbb{R}^d)$ s'injecte dans H . Autrement dit, on suppose qu'il existe une constante c telle que

$$\forall f \in \mathbb{L}^\infty(\mathbb{R}^d), \quad |f|_\infty \leq c |f|_H$$

et donc l'intégration de toute fonction bornée contre une mesure signée est une forme linéaire continue pour la structure hilbertienne de H : c'est le cas pour la norme infinie ; or, la norme hilbertienne étant plus fine que cette-dernière, toute application continue pour $|\cdot|_\infty$ l'est également pour $|\cdot|_H$. Finalement, \mathcal{M}_s est en bijection avec un sous ensemble du dual H' de H et on pose :

$$\forall \mu, \nu \in \mathcal{M}_s, \quad d(\mu, \nu) = |\mu - \nu|_{H'} \quad \text{où} \quad \forall \mu \in \mathcal{M}_s, \quad |\mu|_{H'} = \sup_{|f|_H=1} \int f \, d\mu$$

La formulation variationnelle du problème d'appariement s'écrit dans ce cas :

$$\text{Minimiser } J_{\mu_y, \mu_z}(v) = \int_0^1 |v_t|_V^2 \, dt + \frac{1}{\sigma^2} |\phi_1^v \mu_y - \mu_z|_{H'}^2 \quad \text{avec } v : t \mapsto v_t$$

où le problème d'appariement exact est obtenu en faisant tendre σ vers $+\infty$ et en imposant $\phi_1^v \cdot \mu_y = \mu_z$. Notons dorénavant K_H le noyau du RKHS H , autrement dit le morphisme de H' dans H et tel que pour tout point x de \mathbb{R}^d ,

$$\forall f \in H, \quad f(x) = \delta_x(f) = \left\langle K_H \delta_x, f \right\rangle_H$$

Alors, pour toute fonction f de H et toute mesure $\mu \in \mathcal{M}_s$,

$$\int_{\mathbb{R}^d} f(x) \, d\mu(x) = \int_{\mathbb{R}^d} \left\langle K_H \delta_x, f \right\rangle_H \, d\mu(x) = \left\langle \int_{\mathbb{R}^d} K_H \delta_x \, d\mu(x), f \right\rangle_H$$

par continuité de l'intégration contre μ pour la structure hilbertienne de H . En particulier, en posant $K_H \mu = \int_{\mathbb{R}^d} K_H \delta_x \, d\mu(x)$ et d'après l'inégalité de Cauchy-Schwarz, on a

$$\forall \mu \in \mathcal{M}_s, \quad \forall f \in H, \quad \left| \int_{\mathbb{R}^d} f(x) \, d\mu(x) \right| = \left| \left\langle K_H \mu, f \right\rangle_H \right| \leq |K_H \mu|_H^2 |f|_H^2$$

avec égalité si et seulement si il existe une constante λ telle que $f = \lambda K_H \mu$.

D'où, en passant au sup pour $f \in H$ telle que $|f|_H = 1$ des deux côtés de l'inégalité, $|\mu|_{H'}^2 = |K_H \mu|_H^2$ et le maximum est atteint pour $f = K_H \mu$ donc

$$\forall \mu \in \mathcal{M}_s, \quad |\mu|_{H'}^2 = |K_H \mu|_H^2 = \int_{\mathbb{R}^d} K_H \mu(x) \, d\mu(x)$$

Recalage difféomorphique de courbes et surfaces

Soit, en posant pour tout $x, y \in \mathbb{R}^d$, $k_H(x, y) = (K_H \delta_x)(y)$,

$$|\mu|_{H'}^2 = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K_H \delta_x(y) d\mu(x) d\mu(y) = \int_{\mathbb{R}^{2d}} k_H(x, y) d\mu(x) d\mu(y)$$

En particulier, dans le cas qui nous intéresse, $\mu = \sum_i a^i \delta_{x^i}$ et

$$|\mu|_{H'}^2 = \sum_{i,j} a^i a^j \int_{\mathbb{R}^{2d}} k_H(x, y) d\delta_{x^i}(x) d\delta_{x^j}(y) = \sum_{i,j} a^i a^j k_H(x^i, x^j)$$

b. Aspects algorithmiques

Si on suppose de plus que V est admissible, alors du fait du théorème I.11, on est assuré de l'existence d'un champ de vecteurs v minimisant la fonctionnelle d'appariement.

Dans le cas où les deux mesures μ_y et μ_z sont discrètes, *i.e*

$$\mu_y := \sum_{i=1}^m b^i \delta_{y^i} \quad \text{et} \quad \mu_z := \sum_{i=1}^n c^i \delta_{z^i}$$

alors on a vu que pour tout champ de vecteur v , $\phi_1^v \cdot \mu_y$ était encore une somme de mesures de Dirac – plus précisément, $\phi_t^v \cdot \mu_y = \sum_{i=1}^m b^i \delta_{\phi_t^v(y^i)}$ pour tout temps $t \in [0, 1]$ – et il existe donc des constantes γ^i , un entier ν et des points ζ^i de telle sorte que

$$\phi_1^v \cdot \mu_y - \mu_z = \sum_{i=1}^{\nu} \gamma^i \delta_{\zeta^i} \quad \text{et donc} \quad |\phi_1^v \cdot \mu_y - \mu_z|_{H'}^2 = \sum_{i,j=1}^{\nu} \gamma^i \gamma^j k_H(\zeta^i, \zeta^j)$$

De plus, pour tout temps $t \in [0, 1]$ et à trajectoire $x^i: t \mapsto \phi_t^v(y^i)$ fixée, on sait qu'il existe des vecteurs – dépendants du temps – $\alpha^i(t) \in \mathbb{R}^d$ tels que le champ de vecteurs d'énergie minimale, *i.e* de norme L_V^2 minimale, s'écrive :

$$\forall t \in [0, 1], \quad \forall x \in \mathbb{R}^d, \quad v_t(x) = \sum_{i=1}^m k_V(x^i(t), x) \alpha^i(t)$$

En effet, on dispose du théorème suivant dans le cas du matching de Landmarks dont la preuve repose principalement sur le caractère reproduisant du noyau k_V .

Théorème I.13

Soit $v \in V$, x^1, \dots, x^n des points de \mathbb{R}^d tous distincts et ξ^1, \dots, ξ^n des vecteurs de \mathbb{R}^d . Si on suppose de plus que le noyau k_V est strictement positif alors il existe une unique solution $v_* \in V$ au problème d'interpolation

optimale :

$$\begin{cases} v(x^i) = \xi^i, & \forall i \in \llbracket 1, n \rrbracket \\ \|v\|_V \text{ minimal} \end{cases}$$

De plus, la solution s'écrit : $\forall x \in \mathbb{R}^d, v_*(x) = \sum_{i=1}^n k_V(x^i, x) \alpha^i$

où les $\alpha^i \in \mathbb{R}^d$ sont les solutions du système linéaire

$$\forall j \in \llbracket 1, n \rrbracket, \quad \sum_{i=1}^n k_V(x^i, x^j) \alpha^i = \xi^j$$

En particulier, il vient que

$$\begin{aligned} \forall t \in [0, 1], \quad \|v_t\|_V^2 &= \sum_{i,j=1}^m \left\langle k_V(x^i(t), x) \alpha^i(t), k_V(x^j(t), x) \alpha^j(t) \right\rangle_V \\ &= \sum_{i,j=1}^m \alpha^j(t) \cdot k_V(x^i(t), x^j(t)) \alpha^i(t) \end{aligned}$$

et pour toute trajectoire $x^j(t) = \phi_t^v(y^j)$,

$$\forall t \in [0, 1], \quad x^j(t) = y^j + \sum_{i=1}^m \int_0^t k_V(x^i(s), x^j(s)) \alpha^i(s) ds \quad (\star)$$

Finalement, la fonctionnelle d'appariement se ré-écrit :

$$J_{\mu_y, \mu_z}(v) = \sum_{i,j=1}^m \int_0^1 \alpha^j(t) \cdot k_V(x^i(t), x^j(t)) \alpha^i(t) dt + \frac{1}{\sigma^2} \sum_{i,j=1}^\nu \gamma^i \gamma^j k_H(\zeta^i, \zeta^j)$$

et on peut en fait faire porter la minimisation sur les vecteurs $t \mapsto \alpha^i(t)$, où les trajectoires $t \mapsto x^i(t)$ sont obtenues par résolution du système intégrale (\star) à chaque instant.

En pratique, la minimisation de la fonctionnelle d'appariement est réalisée par une descente de gradient aussi le calcul du gradient est une étape importante du processus d'appariement. On ne détaille pas ces calculs pour ce modèle-ci (mais on les trouve par exemple dans [15]) mais on les détaille pour le modèle de pseudo-métamorphoses car ils sont "originaux".

2. Modèle de métamorphoses

Le modèle de métamorphoses pour les mesures discrètes que l'on présente ici est dû à Casey L. Richardson et Laurent Younes, [33]. Ces travaux se basent principalement sur

ceux de Darryl D. Holm, Alain Trounev et Laurent Younes, [20], qui ont établis les équations d'évolution relatives aux métamorphoses pour les mesures (discrètes). On prend comme point de départ ces travaux, sans chercher à en comprendre la démonstration – ou plus exactement après un "échec" à vouloir en comprendre tous les détails, l'approche étant très physicienne (mécanique des fluides)...

Dans un premier temps, on résume donc l'article [20], autrement dit le cadre mathématique des métamorphoses de mesures ; dans un second temps on considère l'apport de [33], qui permet par exemple la mise-en-oeuvre algorithmique des métamorphoses de mesures et donc une comparaison avec le modèle de pseudo-métamorphoses décrit ci-après.

Soit donc deux mesures discrètes m_y et m_z que l'on cherche à apparier.

a. Mise en équations du modèle

Soit Ω un ouvert borné, à bords lisses, de \mathbb{R}^d et \mathcal{G}_V le groupe de difféomorphismes générés depuis un RKHS V de champs de vecteurs via l'équation de flot suivante :

$$\begin{aligned} \forall t \in [0, 1], \quad \dot{g}_t &= v_t \cdot g_t & g_t \in \mathcal{G}_V, \quad v_t \in V \\ g_0 &= Id \end{aligned}$$

On impose de plus que $v: t \mapsto v_t$ soit d'énergie cinétique finie, *i.e* que $v \in \mathbb{L}_V^2$:

$$\int_0^1 \|v_t\|_V^2 dt < +\infty$$

On impose également à V d'être admissible, de régularité p pour un certain $p \geq 1$: plus précisément, on suppose que V s'injecte continument dans l'espace $\mathcal{C}_0^p(\Omega)$ des champs de vecteurs de classe \mathcal{C}^p sur \mathbb{R}^d , s'annulant sur $\mathbb{R}^d \setminus \Omega$. On sait alors d'après le paragraphe I.2.a que l'équation de flot ci-dessus admet des solutions uniques et que, de plus,

$$\forall t \in [0, 1], \quad g_t \in \mathcal{C}^p(\Omega, \mathbb{R}^d) \subset \mathcal{H}^p(\Omega, \mathbb{R}^d)$$

On note K_V le noyau reproduisant associé à V et $L_V = K_V^{-1}$ son inverse. Enfin, pour tout temps t , on prolonge les difféomorphismes g_t à \mathbb{R}^d tout entier en posant $g_t = Id$ en dehors de Ω . Ainsi, on écrira plus simplement :

$$\forall t \in [0, 1], \quad g_t \in \mathcal{C}^p(\mathbb{R}^d) \subset \mathcal{H}^p(\mathbb{R}^d)$$

On se base de nouveau sur le formalisme introduit dans [17], paragraphe précédent, et on considère un RKHS $H \subset \mathcal{D}'(\mathbb{R}^d)$ et $M = H'$ son espace dual. De même que précédemment, on note K_H son noyau reproduisant et $L_H = K_H^{-1}$ l'inverse de celui-ci, qui est en fait l'application issue du théorème de représentation de Riesz... En particulier,

$$\forall h \in H, \quad \forall \mu \in M, \quad \langle \mu \mid h \rangle = \langle K_H \mu, h \rangle_H$$

où $\langle \cdot | \cdot \rangle$ désigne le crochet de dualité et $\langle \cdot , \cdot \rangle_H$ le produit scalaire sur H et on peut munir M d'un produit scalaire en posant :

$$\forall \mu, \tilde{\mu} \in M, \quad \langle \mu , \tilde{\mu} \rangle_M := \langle \mu | K_H \tilde{\mu} \rangle$$

Enfin, on fait agir le groupe \mathcal{G}_V sur M via l'action (à gauche) suivante :

$$\forall g \in \mathcal{G}_V, \quad \forall \mu \in M, \quad \forall h \in H, \quad \langle g \cdot \mu | h \rangle = \langle \mu | h \circ g \rangle$$

En fait, on va construire spécifiquement les espaces H et M pour avoir les régularités suffisantes : Soit $L: \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}(\mathbb{R}^d)$ un opérateur (différentiel) elliptique, à coefficients constants d'ordre $m = 2\gamma > d$, par exemple $L = (I - \Delta)^\gamma$ avec $\gamma > d/2$. Soit E la solution fondamentale de l'opérateur L , i.e la distribution telle que $LE = \delta_0$. On peut en fait montrer que la solution fondamentale E est dans l'espace de Sobolev \mathcal{H}_{loc}^s pour tout $s < m - d/2$. En particulier, $E \in \mathcal{H}_{loc}^s$ pour des $s > d/2$ et donc, par injection des espaces de Sobolev dans ceux des fonctions continues, la solution fondamentale E est en fait continue. En effet, on a les résultats suivants :

Espace de Sobolev :

Soit $s > 0$. On appelle *espace de Sobolev* d'ordre s et on note $\mathcal{H}^s(\mathbb{R}^d, \mathbb{R}^m)$ ou $\mathcal{H}^s(\mathbb{R}^d)$ si $d = m$, l'ensemble des fonctions $f \in L^2(\mathbb{R}^d, \mathbb{R}^m)$ dont la transformée de Fourier \hat{f} satisfait :

$$\|f\|_{\mathcal{H}^s}^2 = \int_{\mathbb{R}^d} (1 + |\xi|^2)^s |\hat{f}(\xi)|^2 d\xi < +\infty$$

De plus, dès lors que $s > k + d/2$, $\mathcal{H}^s(\mathbb{R}^d, \mathbb{R}^m)$ s'injecte continument dans $\mathcal{C}_0^k(\mathbb{R}^d, \mathbb{R}^m)$. En particulier, dès que $s > d/2$, $\mathcal{H}^s(\mathbb{R}^d, \mathbb{R}^m)$ s'injecte dans $\mathcal{C}_0^0(\mathbb{R}^d, \mathbb{R}^m)$ et peut donc être muni d'un noyau reproduisant.

On construit alors H comme l'extension de Friedrich du produit scalaire $\langle L\psi , \psi \rangle_{\mathbb{L}^2}$ à $H \subset \mathbb{L}^2$ et on note L_H – ou L si aucune confusion n'est possible – l'extension de L et on peut donc voir L_H comme un opérateur différentiel d'ordre m avec $L_H: H \subset \mathcal{S}' \rightarrow M \subset \mathcal{S}'$.

Extension de Friedrich :

Soit T un opérateur positif défini sur un espace de Hilbert H . On peut munir $Dom T$ d'un produit scalaire – si ce n'est pas déjà le cas – en posant :

$$\forall \mu, \tilde{\mu} \in Dom T, \quad Q(\mu, \tilde{\mu}) = \langle \mu , T\tilde{\mu} \rangle_H + \langle \mu , \tilde{\mu} \rangle_H$$

En effet, Q ainsi définie est trivialement sesqui-linéaire et, par positivité de T^3

$$\forall \mu \in Dom T, \quad Q(\mu, \mu) = \langle \mu , T\mu \rangle_H + \langle \mu , \mu \rangle_H \geq \| \mu \|_H^2$$

Soit H_1 le complété de $Dom T$ qui préserve Q . A priori, H_1 est obtenue de manière très formelle et on a aucune garantie que les éléments de H_1 qui peuvent être de nature complètement différente (des classes d'équivalence par exemple...) puissent s'identifier avec ceux de H . En fait, on peut montrer que :

Proposition I.14

L'inclusion canonique $Dom T \hookrightarrow H$ s'étend en une application *injective* de H_1 dans H . Ainsi, on peut voir H_1 comme un sous-ensemble de H .

En particulier, on peut faire hériter H_1 de la topologie de H . Soit l'ensemble

$$\mathcal{D} := \{ \tilde{\mu} \in H_1 \mid \phi_\mu : \mu \mapsto Q(\mu, \tilde{\mu}) \text{ est borné pour la topologie de } H_1 \}$$

Enfin, d'après le théorème de représentation de Riesz appliqué à ϕ_μ vu comme fonction de H , il existe un unique $A\mu \in H$ telle que

$$\forall \tilde{\mu} \in H_1, \quad Q(\mu, \tilde{\mu}) = \langle A\mu, \tilde{\mu} \rangle_H$$

et ce pour tout μ . On a ainsi défini un opérateur A de domaine de définition \mathcal{D} .

Théorème I.15

$A : \mathcal{D} \subset H_1 \rightarrow H$ est un opérateur auto-adjoint positif tel que $T_1 := A - Id$ soit une extension de T .

On appelle *extension de Freidrich* de T l'opérateur T_1 .

En dernier lieu, on peut montrer que, puisque l'opérateur L_H est elliptique, H est isomorphe à l'espace de Sobolev \mathcal{H}^γ . De plus, comme on ne considère par hypothèse que des morphisme $g \in \mathcal{C}_0^p$, alors pour tout $h \in \mathcal{H}^\gamma$, $g \circ h \in \mathcal{H}^\gamma$ et donc l'action de \mathcal{G}_V sur H est bien définie.

En accord avec le paragraphe I.3, on note $(g_t, \mu_t) \in \mathcal{G}_V \times M$ où $g_0 = Id$ et $t \in [0, 1]$ une métamorphoses. On désigne par $m_t = g_t \cdot \mu_t$ l'image de la métamorphose et puisque le flot g est entièrement déterminé par le champ de vecteur v auquel il est associé, on notera parfois abusivement $(v_t, \mu_t) \in V \times M$ pour désigner une métamorphose.

Le problème variationnelle que l'on cherche à résoudre est donc : Minimiser la fonctionnelle définie par

$$J(v, \mu) = \int_0^1 \|v_t\|_V^2 dt + \frac{1}{\sigma^2} \int_0^1 \|\dot{m}_t - v_t(m_t)\|_M^2 dt$$

lorsque (v, μ) parcourt l'ensemble des métamorphoses satisfaisant aux conditions aux bords $m_0 = m_y$ et $m_1 = m_z$ et où σ est le paramètre de matching.

Théorème I.16

Pour toute condition m_y et m_z dans \mathcal{H}^γ , le problème de minimisation sous contrainte ci-dessus admet une solution.

Le théorème ci-dessus est tiré de [20] où il est démontré dans un cadre plus général, *i.e* pour H non-nécessaire un espace de Sobolev.

Toujours dans [20], on trouve une description des équations d'évolution des métamorphoses solutions d'un tel problème. On nomme de telles équations, *équations d'Euler-Lagrange*. Elles prennent ici la forme suivante : il existe $f_0 \in H$ telle que

$$\forall t \in [0, 1], \quad \begin{cases} L_V v_t &= \nabla f_t \otimes m_t \\ \mu_t &= \mu_0 + \sigma^2 \int_0^t g_s^{-1} \cdot L_H(f_0 \circ g_s^{-1}) ds \end{cases}$$

où, pour tout temps t , $f_t := f_0 \circ g_t^{-1}$ et $u \otimes \mu$ désigne la forme linéaire définie sur V par :

$$\forall u, v \in V, \quad \forall \mu \in M, \quad \langle u \otimes \mu \mid v \rangle := \langle \mu \mid {}^t u v \rangle$$

On introduit les deux opérateurs suivants – qui sont en fait des opérateurs différentiels elliptiques, de manière non-triviale – en posant, pour toute métamorphose (g, μ) :

$$\forall f \in H, \quad L^{g_t} f := g_t^{-1} \cdot L_H(f \circ g_t^{-1}) = L(f \circ g_t^{-1}) \circ g_t$$

et $\bar{L}^g := \int_0^1 L^{g_t} dt$. Enfin, pour des raisons pratiques, on note également

$$\bar{L}_t^g := \int_0^t L^{g_s} ds \quad \text{de sorte que} \quad \bar{L}^g = \bar{L}_1^g$$

D'après [20], on peut en fait exprimer f_0 à partir des opérateurs définis ci-dessus et des formes initiales et finales, *i.e* de μ_t aux temps $t = 0$ et $t = 1$. Plus précisément, on a :

$$f_0 = (\bar{L}^g)^{-1}(\mu_1 - \mu_0)$$

et les équations d'Euler-Lagrange se ré-écrivent donc :

$$\forall t \in [0, 1], \quad \begin{cases} L_V v_t &= \nabla f_t \otimes m_t \\ \mu_t &= \mu_0 + \sigma^2 \bar{L}_t^g \circ (\bar{L}^g)^{-1}(\mu_1 - \mu_0) \end{cases}$$

La deuxième forme permet de voir l'évolution du gabarit μ_t comme une composition/intégration de la contrainte aux bords, stockée dans $\mu_1 - \mu_0$. C'est cette seconde forme qui est utile dans la démonstration du théorème I.20.

3. Un opérateur T est dit positif si, pour tout $\mu \in \text{Dom } T$, $\langle \mu, T\mu \rangle_H \geq 0$.

De même [20] garantit l'existence et l'unicité d'une solution pour le système ci-dessus, et même dans un cadre moins restrictif.

Théorème I.17

Pour toute condition initiale $m_0 \in (\mathcal{H}^{\gamma-1})'$, toute fonction $f_0 \in \mathcal{H}^{\gamma+1}$ et tout temps $T > 0$, il existe une unique solution au système d'équations.

b. Minimisation de la fonctionnelle d'appariement

Si tout ce qui précède est valable pour des mesures quelconques, on se place dorénavant dans le cas où les mesures m_y et m_z que l'on cherche à apparier sont toutes deux des sommes de Dirac, ne possédant pas nécessairement le même nombre de termes. Daryl D. Holm, Alain Trouvé et Laurent Younes proposent dans [20] une conjecture sur la forme des solutions à savoir, qu'en tout temps t , m_t puisse s'écrire comme une somme de masses de Dirac et d'une fonction intégrable. En fait, la forme des solutions est plus compliquée et moins régulière, comme énoncé dans le théorème I.20, issu de [33].

Définition I.18. • *On dit qu'une distribution $Z_0 \in \mathcal{D}'(\mathbb{R}^d)$ est de Calderón-Zygmund³ si elle est l'extension dans les distributions d'une fonction z_0 , homogène de degrés $-d$ sur $(\mathbb{R}^d)^*$. Si z_0 est également lisse sur $(\mathbb{R}^d)^*$, on dit alors que Z_0 est une distribution de Calderón-Sygmund lisse.*

• *De même, on dit que Z_x est une distribution de Calderón-Zygmund centrée en x si $Z_x \circ T_x$ est une distribution de Calderón-Zygmund, où T_x désigne l'application de translation qui envoie le point x sur 0.*

En admettant le résultat suivant, dont on trouve par exemple une démonstration dans le livre de Stein [38], utile à la preuve et concernant les distributions de Calderón-Zygmund, on est à même d'énoncer le théorème qui nous intéresse.

Proposition I.19

Si une distribution $v \in \mathcal{D}'(\mathbb{R}^d)$ est homogène de degré 0 alors sa transformée de Fourier est homogène de degré $-d$ et est de la forme

$$\hat{v} = c \delta_0 + Z_0 \quad ; \quad c \in \mathbb{R}$$

où Z_0 est une distribution de Calderón-Zygmund.
De plus, si v restreinte à $(\mathbb{R}^d)^*$ est lisse, alors \hat{v} est également lisse sur $(\mathbb{R}^d)^*$, ainsi que la distribution Z_0 de la décomposition précédente.

Finalement, le théorème qui suit invalide la conjecture évoquée précédemment.

3. En fait, le terme Calderón-Zygmund désigne généralement des généralisations de la définition I.18, plus irrégulières encore.

Théorème I.20

Supposons que V s'injecte continument dans l'espace \mathcal{C}_0^p pour $p \geq m + 1$ et que les conditions aux bords sont telles que la mesure $\mu_1 - \mu_0$ est donnée par la somme de mesures de Dirac

$$\mu := \sum_{k=1}^N \alpha_k \delta_{x_k}$$

Alors, pour tout champ de vecteurs $(v_t)_{t \in [0,1]}$ fixé, tout point stationnaire pour la fonctionnelle $J(v, \cdot)$ est de la forme :

$$\forall t \in [0, 1], \quad \mu_t = f(t) dx + \sum_{k=1}^M \tilde{\alpha}_k(t) \delta_{x_k} + Z_{x_k}(t)$$

où N désigne le nombre total de points singuliers, en regroupant mesure source m_y et mesure cible m_z ; Z_{x_k} est une distribution de Calderón-Zygmund centré en x_k et $f dx$ une mesure absolument continument par rapport à la mesure de Lebesgue, de densité $f \in \mathbb{L}_{loc}^1(\mathbb{R}^d)$.

L'image de la métamorphose est alors donnée par $g_t \cdot \mu_t$ pour tout temps t .

3. Modèle de pseudo-métamorphose

L'idée du modèle de pseudo-métamorphoses est de fournir une méthode aussi facilement implémentable que la méthode de matching de landmarks classique présentée au paragraphe II.1 mais tout en conservant, au moins en partie, les avantages des métamorphoses : notamment découpler – au moins partiellement – l'effet du transport des points et de la déformation due au groupe \mathcal{G}_V . Pour cela, contrairement aux deux modèles précédents, on part de la forme qu'on aimerait que les solutions aient et on va construire une fonctionnelle d'appariement à minimiser à partir de cette forme solution.

Soit deux mesures discrètes μ_y et μ_z que l'on écrit comme somme de masses de Dirac :

$$\mu_y = \sum_{i=1}^m b^i \delta_{y^i} \quad \text{et} \quad \mu_z = \sum_{i=1}^n c^i \delta_{z^i}$$

On désire réaliser l'appariement par *pseudo-métamorphoses* de μ_y et μ_z . Dans le modèle de matching, on réalise le matching en se plaçant sur l'espace \mathcal{M}_s des mesures signées qui contiennent en particulier les sommes de Dirac ; ici, on fait l'hypothèse qu'un modèle d'appariement est possible en se restreignant à l'espace que l'on note $E \subset \mathcal{M}_s$ des sommes (finies) de mesures de Dirac : $E = \bigcup_{n \in \mathbb{N}} E_n$ où pour tout entier $n \in \mathbb{N}$,

Recalage difféomorphique de courbes et surfaces

$$E_n := \left\{ \mu \in \mathcal{M}_s \mid \exists (a_1, \dots, a_n) \in \mathbb{R}^n, \exists (x_1, \dots, x_n) \in (\mathbb{R}^d)^n, \mu = \sum_{i=1}^n a^i \delta_{x^i} \right\}$$

On va munir E d'une métrique encodant le coût de déformation, aussi a-t-on peu d'espoir que l'espace E reste de courbure nulle. Cependant, si c'était le cas, le "plus court chemin" entre les mesures μ_y et μ_z , vues comme points de l'espace E serait donné le segment $[\mu_y, \mu_z]$, à savoir l'ensemble des points (μ_t) qui s'écrivent sous la forme :

$$\forall t \in [0, 1], \quad \mu_t := t\mu_z + (1-t)\mu_y$$

En particulier, on peut vérifier qu'en chaque instant μ_t reste dans E .

$$\forall t \in [0, 1], \quad \mu_t = t \sum_{i=1}^n c^i \delta_{z^i} + (1-t) \sum_{i=1}^m b^i \delta_{y^i} := \sum_{i=1}^{m+n} a^i(t) \delta_{x^i}$$

où on a posé, pour tout temps $t \in [0, 1]$:

$$\forall i \in \llbracket 1, n \rrbracket, \quad \begin{cases} a^i(t) &= t \times c^i \\ x^i(t) &= z^i \end{cases} \quad \text{et} \quad \forall i \in \llbracket n+1, n+m \rrbracket, \quad \begin{cases} a^i(t) &= (1-t) \times b^i \\ x^i(t) &= y^i \end{cases}$$

Plus généralement et en se basant sur cette approche "naïve", on va chercher une solution au problème d'appariement sous la forme d'un chemin dans l'espace E , *i.e*

$$\forall t \in [0, 1], \quad \mu_t := \sum_{i=1}^{m+n} a^i(t) \delta_{x^i(t)} \quad \text{où} \quad \forall i \in \llbracket 1, m+n \rrbracket, \quad a^i(t) \in \mathbb{R} \text{ et } x^i(t) \in \mathbb{R}^d$$

en imposant $\mu_0 = \mu_y$ et $\mu_1 = \mu_z$. Par suite, la variation des points x^i correspondra à la partie transport de la pseudo-métamorphose, tandis que le n -uplet (a^1, \dots, a^n) à la partie déformation.

Pour alléger les calculs par la suite, on notera pour tout indice i et temps t , $x_t^i := x^i(t)$ et $a_t^i := a^i(t)$ de sorte que

$$\forall t \in [0, 1], \quad \mu_t = \sum_{i=1}^{m+n} a_t^i \delta_{x_t^i}$$

– Chapitre II –

Modèle de pseudo-métamorphoses pour les mesures

I. Modèle de pseudo-métamorphoses

Dans ce chapitre, on détaille l'étude mathématique du modèle dit de *pseudo-métamorphoses* et dont nous venons d'introduire l'idée sous-jacente. Dans tout ce qui suit, E désigne donc l'espace des mesures discrètes et, pour tout entier $n \in \mathbb{N}$, E_n celui des sommes à n termes de mesures de Dirac. On plonge E_n dans E_{n+k} pour tout $k \geq 1$ en imposant, par exemple, un poids nul aux termes $n+1$ à $n+k$, sans contrainte pour les points.

Soit donc deux mesures, μ_y et μ_z , ne possédant pas nécessairement le même nombre de termes et que l'on cherche à appairier :

$$\mu_y := \sum_{i=1}^m b^i \delta_{y^i} \quad ; \quad \mu_z := \sum_{i=1}^n c^i \delta_{z^i} .$$

Au besoin, on considérera leur plongement dans E_{m+n} via les identifications suivantes :

$$\bar{\mu}_y = \sum_{i=1}^{m+n} \bar{b}^i \delta_{\bar{y}^i} \quad \text{où} \quad \forall i \in \llbracket 1, m \rrbracket, \begin{cases} \bar{b}^i & := b^i \\ \bar{y}^i & := y^i \end{cases} \quad \text{et} \quad \forall i \in \llbracket 1, n \rrbracket, \begin{cases} \bar{b}^{m+i} & := 0 \\ \bar{y}^{m+i} & \in \mathbb{R}^d \end{cases}$$

et

$$\bar{\mu}_z = \sum_{i=1}^{m+n} \bar{c}^i \delta_{\bar{z}^i} \quad \text{où} \quad \forall i \in \llbracket 1, m \rrbracket, \begin{cases} \bar{c}^i & := 0 \\ \bar{z}^i & \in \mathbb{R}^d \end{cases} \quad \text{et} \quad \forall i \in \llbracket 1, n \rrbracket, \begin{cases} \bar{c}^{m+i} & := c^i \\ \bar{z}^{m+i} & := z^i \end{cases} .$$

1. Mise en équations du modèle

On garde les notations du paragraphe précédent et on cherche à construire la fonctionnelle d'appariement propre à notre modèle. Pour ce faire, on reprend le cadre introduit au paragraphe I.2.a, chapitre 1, et on se donne un groupe \mathcal{G}_V de difféomorphismes associés à un RKHS V dont on note le noyau k_V via l'équation de flot classique :

$$\forall t \in [0, 1], \quad \begin{cases} \dot{\phi}_t & = v_t \circ \phi_t \\ \phi_0 & = Id \end{cases} \quad \phi_t \in \mathcal{G}_V, \quad v_t \in V .$$

Modèle de pseudo-métamorphoses pour les mesures

Afin de garantir l'existence et l'unicité à l'équation précédente, on prend les hypothèses usuelles, à savoir que V est admissible et que les applications $v: t \mapsto v_t$ sont dans \mathbb{L}_V^2 . En procédant de même que pour le matching de landmarks (paragraphe II.1 du chapitre précédent) on peut supposer que E est en bijection avec le dual d'un second RKHS H dont on note le noyau k_H , tous les calculs effectués dans le cadre sus-cité restant valides. De plus, on fait agir le groupe \mathcal{G}_V sur E par transport de masse, *i.e*

$$\text{Pour tout entier } n \in \mathbb{N}, \quad \begin{aligned} \mathcal{G}_V \times E_n &\longrightarrow E_n \\ \left(\phi, \sum_{i=1}^n a^i \delta_{x^i} \right) &\longmapsto \sum_{i=1}^n a^i \delta_{\phi(x^i)} \end{aligned} .$$

Enfin, afin d'alléger les notations, on appelle N le nombre de points à considérer, en se rappelant que dans le cas des pseudo-métamorphoses, $N = m + n$ où m est le nombre de points initiaux et n le nombre de points cibles.

Pour s'assurer que l'espace sur lequel on se place, *a priori* E_N , soit une variété riemannienne et pouvoir réutiliser les travaux relatifs aux landmarks, on impose aux points x^i de rester distincts en tout temps. En effet, l'identification suivante, valable pour tout entier $N \in \mathbb{N}$,

$$\begin{aligned} E_N &\longrightarrow \mathcal{L}^N(\mathbb{R}^d) \times \mathbb{R}^N \\ \sum_{i=1}^N a^i \delta_{x^i} &\longmapsto (x^1, \dots, x^N; a^1, \dots, a^N) \end{aligned}$$

permet d'utiliser les méthodes et algorithmes développés pour le matching de landmarks pour des sommes de mesures de Dirac. En fait, on a déjà eu recours à une telle identification, de manière tacite, lorsque l'on a utilisé le théorème I.13 pour l'appariement de mesures signées. Ainsi, on se place sur le sous-ensemble de E_N isomorphe à $\mathcal{L}^N(\mathbb{R}^d) \times \mathbb{R}^N$, que l'on continue à noter E_N par abus de notations.

On cherche à déterminer en tout temps $t \in [0, 1]$, le champ de vecteur v_t , d'énergie minimale, qui nous permettra d'avoir les formes solutions annoncées en introduction. En considérant l'extension $\bar{\mu}_y$ de μ_y dans E_N , on note, pour tout temps $t \in [0, 1]$,

$$\mu_t := \phi_t^v \cdot \bar{\mu}_y = \sum_{i=1}^N \bar{b}^i \delta_{\phi_t^v(\bar{y}^i)} := \sum_{i=1}^N a_t^i \delta_{x_t^i} .$$

Alors, pour tout temps $t \in [0, 1]$,

$$\dot{\mu}_t = \sum_{i=1}^N \dot{a}_t^i \delta_{x_t^i} + \sum_{i=1}^N a_t^i \delta_{x_t^i}^{\nabla, \dot{x}_t^i} \quad \text{où} \quad \forall \phi \in H, \quad \left\langle \delta_{x_t^i}^{\nabla, \dot{x}_t^i} \mid \phi \right\rangle := \nabla \phi(x_t^i) \cdot \dot{x}_t^i .$$

La première somme peut donc être vue comme un terme de coloration : les points x_t^i restent inchangés tandis que les poids se trouvent modifiés – et éventuellement annulés, ce qui revient à faire "disparaître" le point associé. Au contraire, la seconde somme correspond au transport des points par le flot ϕ_t^v de l'équation de flot et on utilise le formalisme du matching pour les landmarks pour traiter ce terme de transport de masse.

On cherche donc les champs de vecteurs v_t qui minimisent le terme de coloration $\sum_{i=1}^N \dot{a}_t^i \delta_{x_t^i}$ mais également, comme dans le matching pour les landmarks, tels que $\dot{x}_i = v(x_i)$ pour tout indice $i \in \llbracket 1, N \rrbracket$ et d'énergie $\|v_t\|_V$ minimale. On procède donc comme au paragraphe II.1 du chapitre 1 : Pour tout temps $t \in [0, 1]$,

$$\left\| \sum_{i=1}^N \dot{a}_t^i \delta_{x_t^i} \right\|_{H'}^2 = \sum_{i,j=1}^N \dot{a}_t^i \dot{a}_t^j k_H(x_t^i, x_t^j) = {}^t \dot{a}_t K_H(x_t) \dot{a}_t = \left\langle \dot{a}_t \mid K_H(x_t) \dot{a}_t \right\rangle$$

en notation matricielle où, de manière évidente et pour tout t , $a_t := (a_t^1, \dots, a_t^N) \in \mathbb{R}^N$, $x_t := (x_t^1, \dots, x_t^N) \in \mathbb{R}^{Nd}$ et $K_H(x)$ est la matrice de $\mathcal{M}_N \mathbb{R}$ définie par :

$$\forall i, j \in \llbracket 1, N \rrbracket, \quad K_H(x)_{i,j} := k_H(x^i, x^j).$$

On introduit de même la matrice $K_V \in \mathcal{M}_{Nd} \mathbb{R}$ définie par

$$\forall i, j \in \llbracket 1, N \rrbracket, \quad K_V(x)_{i,j} := k_V(x^i, x^j).$$

D'après le théorème I.13, on sait qu'il existe, pour tout temps t , des vecteurs $\xi_t^i \in \mathbb{R}^d$ – que l'on appellera par la suite *moments associés à x_t^i* – tels que les champs de vecteurs s'écrivent, pour tout temps $t \in [0, 1]$,

$$v_t = \sum_{i=1}^N k_V(x_t^i, \cdot) \xi_t^i \quad \text{et donc} \quad \|v_t\|_V^2 = \sum_{i,j=1}^N k_V(x_t^i, x_t^j) \left\langle \xi_t^i \mid \xi_t^j \right\rangle.$$

De plus, les vecteurs $\xi_t = (\xi_t^1, \dots, \xi_t^N) \in \mathbb{R}^{Nd}$ sont solutions des systèmes linéaires

$$\forall t \in [0, 1], \quad K_V(x_t) \xi_t = \dot{x}_t \quad \text{i.e.} \quad \forall t \in [0, 1], \quad \xi_t = K_V(x_t)^{-1} \dot{x}_t$$

d'où,

$$\forall t \in [0, 1], \quad \|v_t\|_V^2 = \left\langle {}^t \xi_t \mid K_V(x_t) \xi_t \right\rangle = \left\langle \dot{x}_t \mid K_V(x_t)^{-1} \dot{x}_t \right\rangle.$$

Finalement, on définit une fonctionnelle d'appariement en posant :

$$J(x, a) := \int_0^1 \left\langle \dot{a}_t \mid K_H(x_t) \dot{a}_t \right\rangle + \left\langle \dot{x}_t \mid K_V(x_t)^{-1} \dot{x}_t \right\rangle dt$$

où $x: t \mapsto x_t$ et $a: : t \mapsto a_t$.

2. Équations géodésiques et reparamétrisation du modèle

De manière équivalente, on cherche donc à minimiser la fonctionnelle suivante :

$$J(x, a) = \frac{1}{2} \int_0^1 \left\langle \dot{a}_t \mid K_H(x_t) \dot{a}_t \right\rangle + \left\langle \dot{x}_t \mid K_V(x_t)^{-1} \dot{x}_t \right\rangle dt$$

sous la contrainte que $\mu_0 = \bar{\mu}_y$ et $\mu_1 = \bar{\mu}_z$. En définissant la métrique (riemannienne) g suivante sur l'espace $\mathcal{K}^N(\mathbb{R}^d) := \mathcal{L}^N(\mathbb{R}^d) \times \mathbb{R}^N$ où $\mathcal{L}^N(\mathbb{R}^d)$ désigne l'espace des landmarks constitués de N points dans \mathbb{R}^d

$$\forall q, q', q'' \in \mathcal{K}^N(\mathbb{R}^d), \quad \left\langle q' \mid q'' \right\rangle_{g(q)} := \left\langle a' \mid K_H(x) a'' \right\rangle + \left\langle x' \mid K_V(x)^{-1} x'' \right\rangle$$

où $q = (x, a)$, on remarque que la fonctionnelle à minimiser se ré-écrit

$$J(x, a) = \frac{1}{2} \int_0^1 \|\dot{q}_t\|_{g(q_t)} dt$$

et on cherche donc à trouver le chemin $q: [0, 1] \rightarrow \mathcal{K}^N(\mathbb{R}^d)$ d'énergie minimale et tel que $\mu_0 = \bar{\mu}_y$ et $\mu_1 = \bar{\mu}_z$. Des résultats classiques de géométrie riemannienne nous assurent de l'existence en particulier (au moins locale) de tels chemins : il suffit de déterminer les géodésiques (minimisantes) associées. En effet, on a :

Théorème II.1

Soit M une variété riemannienne (connexe et complète) et m un point de M . Alors il existe un voisinage \mathcal{U} de m et un réel $\varepsilon > 0$ tel que pour tout point x, y de \mathcal{U} il existe une unique géodésique γ de longueur au plus ε reliant x et y . De plus, $l(\gamma) = d_M(x, y)$.

où $l(\gamma)$ désigne la longueur de la courbe γ et que l'on définit dans le cas \mathcal{C}^1 par :

$$\forall \gamma \in \mathcal{C}^1(I), \quad I \subset \mathbb{R}, \quad l(\gamma) := \int_I \left\| \frac{d\gamma}{dt}(\gamma(t)) \right\|_{\gamma(t)} dt$$

et en procédant de même sur la subdivision *ad hoc* et en sommant les longueurs des sous-intervalles dans le cas \mathcal{C}^1 par morceaux ; et où d_M désigne la distance entre x et y , obtenue par minimisation des différentes longueurs des courbes joignant x à y :

$$\forall x, y \in M, \quad d_M(x, y) := \inf \left\{ l(\gamma) \mid \gamma \mathcal{C}^1 \text{ par morceaux, reliant } x \text{ à } y \right\}.$$

a. Recourt à la mécanique hamiltonienne

Pour cela, nous allons utiliser la mécanique hamiltonienne [18, 19, 40]. En effet, nous sommes précisément en train d'étudier le mouvement d'une particule libre – *i.e* soumise

à aucune force – dans la variété riemannienne $\mathcal{K}^N(\mathbb{R}^d)$ des landmarks "augmentés", pour la métrique g . Fort du formalisme hamiltonien, on introduit le moment conjugué à q , que l'on note $p := (\xi, \alpha)^1$, et l'opérateur hamiltonien suivant :

$$\mathcal{H}(q, p) = \frac{1}{2} \|p\|_{g(q)^{-1}} = \frac{1}{2} \left\langle \alpha \mid K_H(x)^{-1} \alpha \right\rangle + \frac{1}{2} \left\langle \xi \mid K_V(x) \xi \right\rangle .$$

On sait par ailleurs que le Hamiltonien satisfait aux équations suivantes :

$$\begin{cases} \dot{q} &= \nabla_p \mathcal{H}(q, p) \\ \dot{p} &= -\nabla_q \mathcal{H}(q, p) \end{cases}$$

ce qui, dans notre cas, s'écrit :

$$\begin{cases} \dot{x} &= \nabla_\xi \mathcal{H}(x, a, \xi, \alpha) = K_V(x) \xi \\ \dot{a} &= \nabla_\alpha \mathcal{H}(x, a, \xi, \alpha) = K_H(x)^{-1} \alpha \\ \dot{\xi} &= -\nabla_x \mathcal{H}(x, a, \xi, \alpha) = -\frac{1}{2} \nabla_x \left[\left\langle \alpha \mid K_H(x)^{-1} \alpha \right\rangle + \left\langle \xi \mid K_V(x) \xi \right\rangle \right] \\ \dot{\alpha} &= -\nabla_a \mathcal{H}(x, a, \xi, \alpha) = 0 \end{cases} .$$

À condition de se donner une condition initiale raisonnable – *i.e* appartenant à l'espace $\mathcal{K}^N(\mathbb{R}^d) = \mathcal{L}^N(\mathbb{R}^d) \times \mathbb{R}^N$ – on est donc à même de déterminer les équations du mouvement. Dans les faits, nous n'avons pas de contrainte sur le moment initial mais sur les positions initiales et finales de la "particule", aussi va-t-on recourir à une méthode de shooting pour résoudre notre système.

b. Méthode de shooting

Nos contraintes d'attache aux données portant uniquement sur les positions initiales – $(x_0, a_0) = (y, b)$ – et finales – $(x_1, a_1) = (z, c)$ – du nuage de points, on ne peut résoudre "brutalement" le système d'équations ci-dessous : en effet, on aurait besoin pour cela de connaître les moments initiaux, ce qui n'est pas le cas. Au lieu de quoi, nous allons simuler la résolution de ce dernier pour une donnée initiale "arbitraire" $(x_0, a_0, \xi_0, \alpha_0)$ et considérer la "distance" entre la position finale du nuage de points induite par une telle condition et la valeur attendue. Pour ce faire, on va définir un terme dit d'attache aux données qui permettra de mesurer cette distance et la courbe géodésique reliant $\mu_y \sim (b, y)$ à $\mu_z \sim (c, z)$ sera celle minimisant ce terme d'attache que l'on note A .

En fait, dans le cadre des pseudo-métamorphoses, on travaille dans $\mathcal{K}^N(\mathbb{R}^d)$ avec $N = m + n$ alors que nos contraintes aux temps $t = 0$ ne portent que sur les m premières valeurs de (x_0, a_0) . On va donc initialiser notre système sur des grandeurs partiellement arbitraires et intégrer la condition initiale dans la fonctionnelle d'attache aux données. En pratique, et pour symétriser le problème, on initialise le système sur les entrées

1. où ξ est le vecteur que l'on a déjà défini précédemment...

Modèle de pseudo-métamorphoses pour les mesures

$x_0 = (y, z)$, $a_0 = (b, c)$, $\xi_0 = 0_{\mathbb{R}^{Nd}}$ et $\alpha_0 = 0_{\mathbb{R}^N}$; où (y, z) désigne la concaténation des vecteurs y et z et de même pour (b, c) .

Par définition, la fonctionnelle A dépend des positions initiales et finales, aussi écrira-t-on $A(x_0, a_0, x_1, a_1)$. Il faut néanmoins faire attention au fait que x_1 et a_1 dépendent de manière implicite des conditions initiales, et donc en particulier de x_0 et a_0 . . .

La "proposition" suivante justifie la méthode de shooting :

Proposition II.2

Soit $(x_0, a_0, \xi_0, \alpha_0)$ la condition initiale telle que la fonctionnelle d'attache aux données A soit minimale et $t \mapsto (x_t, a_t, \xi_t, \alpha_t)$ solution du système

$$\begin{cases} \dot{x} &= K_V(x)\xi \\ \dot{a} &= K_H(x)^{-1}\alpha \\ \dot{\xi} &= -\frac{1}{2}\nabla_x \left[\langle \alpha \mid K_H(x)^{-1}\alpha \rangle + \langle \xi \mid K_V(x)\xi \rangle \right] \\ \dot{\alpha} &= 0 \end{cases} .$$

Alors une courbe géodésique (minimisante) reliant μ_y à μ_z est donnée par

$$\forall t \in [0, 1], \quad \mu(t) = \sum_{i=1}^{m+n} a_i(t)\delta_{x_i(t)} .$$

Démonstration :

Quitte à se restreindre à un voisinage de $\mu_y \sim (y, b)$ et $\mu_z \sim (z, c)$ – dans la variété riemannienne $\mathcal{K}^N(\mathbb{R}^d)$ – on est assuré de l'existence d'une trajectoire géodésique (minimisante) reliant μ_y à μ_z .

L'ensemble des courbes géodésiques, et donc en particulier celle reliant μ_y à μ_z , sont solutions du système. De plus, l'attache aux données A est nulle lorsque l'appariement est exact, strictement positive sinon. Aussi, en minimisant l'attache aux données pour l'ensemble des courbes géodésiques possibles – générées par le système – on trouve un appariement exact. \square

Réciproquement, il est bien sûr évidemment qu'une courbe géodésique reliant μ_y et μ_z satisfera les équations géodésiques d'une part et induira un terme d'attache aux données minimal car nul d'autre part. La figure II.1 résume cette équivalence des points de vue.

Attention, il faut néanmoins prendre garde que ce "résultat" est purement théorique et ne nous assure pas de la convergence numérique vers le minimum (global) de la fonctionnelle d'attache aux données. La convergence numérique dépend en effet fortement du choix de l'algorithme. On reprend le choix de [14] et on va procéder grâce à un algorithme

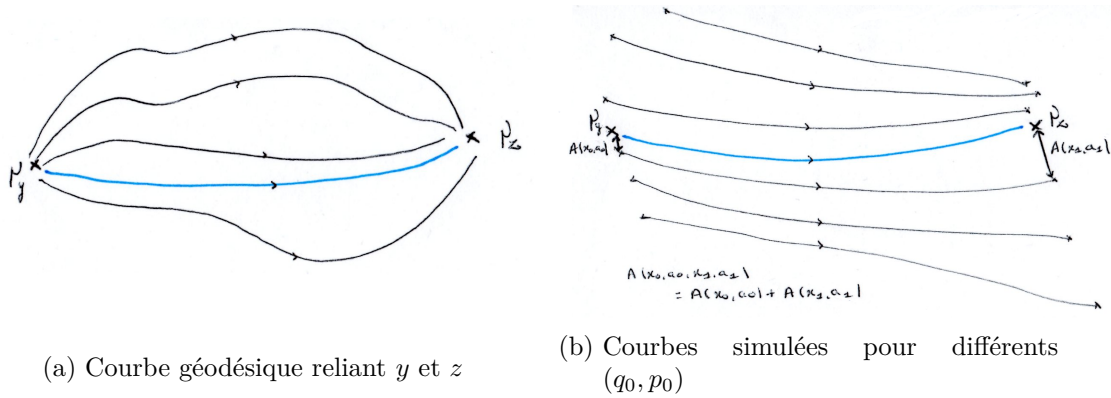


FIGURE II.1. – Méthode de shooting

de type L-BFGS – Low Broyden-Fletcher-Goldfarb-Shano² – qui est particulièrement adapté à la résolution de problèmes d'optimisation non-linéaires sans contrainte, ce qui est notre cas.

Choix de la fonctionnelle d'attache aux données : Finalement, il ne reste plus qu'à exprimer le terme d'attache aux données A associé à notre problème. On rappelle pour ce faire que les contraintes d'appariement que l'on s'est fixées sont les suivantes :

$$\mu_0 = \mu_y \quad \text{et} \quad \mu_1 = \mu_z \quad \text{soit, alternativement,} \quad \mu_0 = \bar{\mu}_y \quad \text{et} \quad \mu_1 = \bar{\mu}_z$$

où $N = m + n$. Autrement dit, pour tout $i \in \llbracket 1, m + n \rrbracket$, la condition "initiale" est donnée par :

$$a^i(0) = a_0^i = \begin{cases} b^i & \text{si } i < m \\ 0 & \text{sinon} \end{cases} \quad \text{et} \quad x^i(0) = x_0^i = y^i \quad \text{si } i < m$$

et la condition "finale" par :

$$a^i(1) = a_1^i = \begin{cases} c^{m-i} & \text{si } i > m \\ 0 & \text{sinon} \end{cases} \quad \text{et} \quad x^i(1) = x_1^i = z^{m-i} \quad \text{si } i > m.$$

En particulier, remarquons qu'aucune contrainte n'est imposée aux m premières valeurs de x_1 et n dernières de x_0 . L'espoir est que x étant solution des équations géodésiques, la valeur de x aux temps 0 et 1 soient contrôlées par ces dernières.

Le choix le plus simple d'attache aux données est de considérer la distance euclidienne entre les couples (z, c) et (y, b) attendus et ceux (x_1, a_1) et (x_0, a_0) associés au choix des

2. lui-même une amélioration de l'algorithme BFGS, permettant de limiter le stockage mémoire.

grandeurs arbitraires d'initialisation :

$$\begin{aligned}
 A(x_0, a_0, x_1, a_1) &= \left\| \left(x_0^1 - y^1, \dots, x_0^m - y^m ; \star, \dots, \star ; \right. \right. \\
 &\quad \left. \left. a_0^1 - b^1, \dots, a_0^m - b^m ; a_0^{m+1}, \dots, a_0^{m+n} \right) \right\|_{\mathbb{R}^{2N(d+1)}}^2 \\
 &+ \left\| \left(\star, \dots, \star ; x_1^{m+1} - z^1, \dots, x_1^{m+n} - z^n ; \right. \right. \\
 &\quad \left. \left. a_1^1, \dots, a_1^m ; a_1^{m+1} - c^1, \dots, a_1^{m+n} - c^n \right) \right\|_{\mathbb{R}^{2N(d+1)}}^2 \\
 &= \sum_{i=1}^m |x_0^i - y^i|^2 + \sum_{i=1}^m |a_0^i - b^i|^2 + \sum_{i=1}^n |a_0^{m+i}|^2 \\
 &\quad + \sum_{i=1}^n |x_1^{m+i} - z^i|^2 + \sum_{i=1}^m |a_1^i|^2 + \sum_{i=1}^n |a_1^{m+i} - c^i|^2 .
 \end{aligned}$$

II. Mise en œuvre algorithmique

Afin de tester la méthode sur des données "réelles" nous avons choisi de l'implémenter dans un premier temps en MatLab, en modifiant le code [14] qui réalise l'appariement (exact ou inexact) de points, courbes et surfaces par shooting.

1. Minimisation par descente de gradient

Comme énoncé à la proposition II.2, il suffit de minimiser le terme d'attache aux données pour déterminer les courbes géodésiques et en particulier celle qui nous intéresse. Pour ce faire, nous allons réaliser en s'inspirant du code sus-cité une descente de gradient pour cette fonctionnelle, vue comme fonction de la variable p_0 . Plus rigoureusement, on définit une nouvelle fonctionnelle que l'on note $B : \mathbb{R}^{Nd} \times \mathbb{R}^N \rightarrow \mathbb{R}$ par

$$\forall (\xi_0, \alpha_0) \in \mathbb{R}^{Nd} \times \mathbb{R}^N, \quad B(\xi_0, \alpha_0) := A(x_0, a_0, x_1(\xi_0, \alpha_0), a_1(\xi_0, \alpha_0))$$

et on cherche à déterminer le couple (ξ_0, α_0) qui la minimise. Pour cela, on va devoir calculer son gradient. Comme on s'intéresse uniquement en la dépendance à ξ_0 et α_0 de l'attache aux données et afin d'alléger les notations, on implice la dépendance en x_0 et a_0 de A dans les calculs suivants.

a. Gradient de la fonctionnelle d'attache aux données

En confondant applications (différentielles partielles) et représentations matricielles associées, on a :

Pour tout $(\xi_0, \alpha_0) \in \mathbb{R}^{Nd} \times \mathbb{R}^N$ et $\delta\xi_0 \in \mathbb{R}^{Nd}$,

$$\begin{aligned} \frac{\partial B}{\partial \xi_0}(\xi_0, \alpha_0) \cdot \delta\xi_0 &= \frac{\partial A}{\partial x_1}(x_1, a_1) \cdot \frac{\partial x_1}{\partial \xi_0}(\xi_0, \alpha_0) \cdot \delta\xi_0 + \frac{\partial A}{\partial a_1}(x_1, a_1) \cdot \frac{\partial a_1}{\partial \xi_0}(\xi_0, \alpha_0) \cdot \delta\xi_0 \\ &= \left\langle \nabla_{x_1} A(x_1, a_1) \left| \frac{\partial x_1}{\partial \xi_0}(\xi_0, \alpha_0) \delta\xi_0 \right. \right\rangle + \left\langle \nabla_{a_1} A(x_1, a_1) \left| \frac{\partial a_1}{\partial \xi_0}(\xi_0, \alpha_0) \delta\xi_0 \right. \right\rangle \\ &= \left\langle {}^t \left(\frac{\partial x_1}{\partial \xi_0}(\xi_0, \alpha_0) \right) \nabla_{x_1} A(x_1, a_1) + {}^t \left(\frac{\partial a_1}{\partial \xi_0}(\xi_0, \alpha_0) \right) \nabla_{a_1} A(x_1, a_1) \left| \delta\xi_0 \right. \right\rangle \end{aligned}$$

et par conséquent, pour tout $(\xi_0, \alpha_0) \in \mathbb{R}^{Nd} \times \mathbb{R}^N$,

$$\nabla_{\xi_0} B(\xi_0, \alpha_0) = {}^t \left(\frac{\partial x_1}{\partial \xi_0}(\xi_0, \alpha_0) \right) \nabla_{x_1} A(x_1, a_1) + {}^t \left(\frac{\partial a_1}{\partial \xi_0}(\xi_0, \alpha_0) \right) \nabla_{a_1} A(x_1, a_1).$$

De même, pour tout $(\xi_0, \alpha_0) \in \mathbb{R}^{Nd} \times \mathbb{R}^N$,

$$\nabla_{\alpha_0} B(\xi_0, \alpha_0) = {}^t \left(\frac{\partial x_1}{\partial \alpha_0}(\xi_0, \alpha_0) \right) \nabla_{x_1} A(x_1, a_1) + {}^t \left(\frac{\partial a_1}{\partial \alpha_0}(\xi_0, \alpha_0) \right) \nabla_{a_1} A(x_1, a_1)$$

et la principale difficulté réside dans la détermination des grandeurs suivantes :

$${}^t \left(\frac{\partial x_1}{\partial \xi_0}(\xi_0, \alpha_0) \right), \quad {}^t \left(\frac{\partial a_1}{\partial \xi_0}(\xi_0, \alpha_0) \right), \quad {}^t \left(\frac{\partial x_1}{\partial \alpha_0}(\xi_0, \alpha_0) \right) \quad \text{et} \quad {}^t \left(\frac{\partial a_1}{\partial \alpha_0}(\xi_0, \alpha_0) \right).$$

Pour cela et pour tout $\varepsilon \in \mathbb{R}$, on considère les variations des conditions initiales

$$\xi_0(\varepsilon) := \xi_0 + \varepsilon \tilde{\xi}_0 \quad \text{et} \quad \alpha_0(\varepsilon) := \alpha_0 + \varepsilon \tilde{\alpha}_0 \quad ; \quad \tilde{\xi}_0, \tilde{\alpha}_0 \in \mathbb{R}^{(m+n)d}$$

auxquelles on associe les solutions correspondantes aux équations géodésiques : $x(\varepsilon)$, $a(\varepsilon)$, $\xi(\varepsilon)$ et $\alpha(\varepsilon)$, pour la condition initiale $(y, b, \xi_0(\varepsilon), \alpha_0(\varepsilon))$. Soit

$$\tilde{x} := \frac{\partial}{\partial \varepsilon} (x(\varepsilon)) \Big|_{\varepsilon=0}, \quad \tilde{a} := \frac{\partial}{\partial \varepsilon} (a(\varepsilon)) \Big|_{\varepsilon=0}, \quad \tilde{\xi} := \frac{\partial}{\partial \varepsilon} (\xi(\varepsilon)) \Big|_{\varepsilon=0} \quad \text{et} \quad \tilde{\alpha} := \frac{\partial}{\partial \varepsilon} (\alpha(\varepsilon)) \Big|_{\varepsilon=0}$$

de sorte que la formule de Taylor, à l'ordre 1 donne :

$$x(\varepsilon) = x + \varepsilon \tilde{x} + o(\varepsilon); \quad a(\varepsilon) = a + \varepsilon \tilde{a} + o(\varepsilon); \quad \xi(\varepsilon) = \xi + \varepsilon \tilde{\xi} + o(\varepsilon) \quad \text{et} \quad \alpha(\varepsilon) = \alpha + \varepsilon \tilde{\alpha} + o(\varepsilon).$$

Par analogie, avec la définition de $\xi_0(\varepsilon)$ et $\alpha_0(\varepsilon)$, on pose donc $\tilde{x}_0 = 0$ et $\tilde{a}_0 = 0$.

L'idée est, en s'inspirant de [15], de trouver une équation différentielle satisfaite par $(\tilde{x}, \tilde{a}, \tilde{\xi}, \tilde{\alpha})$. En effet, on sera alors à même de déterminer $(\tilde{x}_1, \tilde{a}_1, \tilde{\xi}_1, \tilde{\alpha}_1)$ pour toute condi-

tion initiale de la forme $(0, 0, \tilde{\xi}_0, \tilde{\alpha}_0)$. Or,

$$\begin{aligned}\tilde{x}_1 &= \frac{\partial}{\partial \varepsilon} (x_1(\varepsilon)) \Big|_{\varepsilon=0} = \frac{\partial x_1}{\partial \varepsilon} (\xi_0(\varepsilon), \alpha_0(\varepsilon)) \Big|_{\varepsilon=0} \\ &= \frac{\partial x_1}{\partial \xi_0} (\xi_0, \alpha_0) \cdot \frac{\partial}{\partial \varepsilon} (\xi_0(\varepsilon)) \Big|_{\varepsilon=0} + \frac{\partial x_1}{\partial \alpha_0} (\xi_0, \alpha_0) \cdot \frac{\partial}{\partial \varepsilon} (\alpha_0(\varepsilon)) \Big|_{\varepsilon=0} \\ &= \frac{\partial x_1}{\partial \xi_0} (\xi_0, \alpha_0) \cdot \tilde{\xi}_0 + \frac{\partial x_1}{\partial \alpha_0} (\xi_0, \alpha_0) \cdot \tilde{\alpha}_0\end{aligned}$$

et, de même,

$$\tilde{a}_1 = \frac{\partial a_1}{\partial \xi_0} (\xi_0, \alpha_0) \cdot \tilde{\xi}_0 + \frac{\partial a_1}{\partial \alpha_0} (\xi_0, \alpha_0) \cdot \tilde{\alpha}_0.$$

Ceci étant vrai pour toute condition initiale, en résolvant l'équation différentielle pour un nombre suffisant de conditions initiales, on sera donc capable de déterminer les quantités qui nous intéressent, à savoir :

$${}^t \left(\frac{\partial x_1}{\partial \xi_0} (\xi_0, \alpha_0) \right), \quad {}^t \left(\frac{\partial a_1}{\partial \xi_0} (\xi_0, \alpha_0) \right), \quad {}^t \left(\frac{\partial x_1}{\partial \alpha_0} (\xi_0, \alpha_0) \right) \quad \text{et} \quad {}^t \left(\frac{\partial a_1}{\partial \alpha_0} (\xi_0, \alpha_0) \right).$$

b. Établissement de l'équation différentielle

En supposant que \tilde{x} , \tilde{a} , $\tilde{\xi}$ et $\tilde{\alpha}$ soient de régularité suffisante, c'est à dire au moins \mathcal{C}^2 , on a :

$$\dot{\tilde{x}} = \frac{\partial}{\partial \varepsilon} [\dot{x}(\varepsilon)] \Big|_{\varepsilon=0} = \frac{\partial}{\partial \varepsilon} [K_V(x(\varepsilon)) \xi(\varepsilon)] \Big|_{\varepsilon=0} = \frac{\partial}{\partial \varepsilon} [K_V(x(\varepsilon))] \Big|_{\varepsilon=0} \xi + K_V(x) \tilde{\xi}.$$

Or, par linéarité de la multiplication matricielle, pour tout $\varepsilon \in \mathbb{R}$,

$$\frac{\partial}{\partial \varepsilon} [K_V(x(\varepsilon))] \xi = \frac{\partial}{\partial \varepsilon} [K_V(x(\varepsilon)) \xi] = \mathcal{D}_{x(\varepsilon)} [K_V(\cdot) \xi] \cdot \frac{\partial}{\partial \varepsilon} [x(\varepsilon)]$$

d'où, en prenant $\varepsilon = 0$ dans l'expression ci-dessus :

$$\dot{\tilde{x}} = \mathcal{D}_x [K_V(\cdot) \xi] \cdot \tilde{x} + K_V(x) \tilde{\xi}.$$

D'autre part, en procédant de même,

$$\dot{\tilde{a}} = \frac{\partial}{\partial \varepsilon} [K_H(x(\varepsilon))^{-1} \alpha(\varepsilon)] \Big|_{\varepsilon=0} = \mathcal{D}_x [K_H(\cdot)^{-1} \alpha] \cdot \tilde{x} + K_H(x)^{-1} \tilde{\alpha}.$$

Par ailleurs,

$$\begin{aligned}\dot{\xi} &= \frac{\partial}{\partial \varepsilon} \left[-\frac{1}{2} \nabla_x \left[\left\langle \alpha(\varepsilon) \mid K_H(x(\varepsilon))^{-1} \alpha(\varepsilon) \right\rangle + \left\langle \xi(\varepsilon) \mid K_V(x(\varepsilon)) \xi(\varepsilon) \right\rangle \right] \right]_{|\varepsilon=0} \\ &= -\frac{1}{2} \nabla_x \left[\frac{\partial}{\partial \varepsilon} \left[\left\langle \alpha(\varepsilon) \mid K_H(x(\varepsilon))^{-1} \alpha(\varepsilon) \right\rangle \right]_{|\varepsilon=0} \right] - \frac{1}{2} \nabla_x \left[\frac{\partial}{\partial \varepsilon} \left[\left\langle \xi(\varepsilon) \mid K_V(x(\varepsilon)) \xi(\varepsilon) \right\rangle \right]_{|\varepsilon=0} \right]\end{aligned}$$

et en utilisant les calculs précédents, on a

$$\begin{aligned}\frac{\partial}{\partial \varepsilon} \left[\left\langle \xi(\varepsilon) \mid K_V(x(\varepsilon)) \xi(\varepsilon) \right\rangle \right]_{|\varepsilon=0} &= \left\langle \tilde{\xi} \mid K_V(x) \xi \right\rangle + \left\langle \xi \mid \frac{\partial}{\partial \varepsilon} \left[K_V(x(\varepsilon)) \xi(\varepsilon) \right]_{|\varepsilon=0} \right\rangle \\ &= \left\langle \tilde{\xi} \mid K_V(x) \xi \right\rangle + \left\langle \xi \mid \mathcal{D}_x [K_V(\cdot) \xi] \cdot \tilde{x} + K_V(x) \tilde{\xi} \right\rangle \\ &= 2 \left\langle \tilde{\xi} \mid K_V(x) \xi \right\rangle + \mathcal{D}_x \left[\left\langle \xi \mid K_V(\cdot) \xi \right\rangle \right] \cdot \tilde{x}\end{aligned}$$

par linéarité du produit scalaire. Finalement pour tout $\delta x \in \mathbb{R}^{Nd}$,

$$\begin{aligned}\left\langle \nabla_x \left[\frac{\partial}{\partial \varepsilon} \left[\left\langle \xi(\varepsilon) \mid K_V(x(\varepsilon)) \xi(\varepsilon) \right\rangle \right]_{|\varepsilon=0} \right] \mid \delta x \right\rangle \\ &= 2 \left\langle \tilde{\xi} \mid \mathcal{D}_x [K_V(\cdot) \xi] \cdot \delta x \right\rangle + \mathcal{D}_x^2 \left[\left\langle \xi \mid K_V(\cdot) \xi \right\rangle \right] \cdot (\tilde{x}, \delta x) \\ &= \left\langle 2 \left(\mathcal{D}_x [K_V(\cdot) \xi] \right)^* \cdot \tilde{\xi} + \mathcal{D}_x^2 \left[\left\langle \xi \mid K_V(\cdot) \xi \right\rangle \right] \cdot (\tilde{x}, \cdot) \mid \delta x \right\rangle\end{aligned}$$

et donc, en procédant de même pour le second terme,

$$\begin{aligned}\dot{\xi} &= - \left(\mathcal{D}_x [K_V(\cdot) \xi] \right)^* \cdot \tilde{\xi} - \frac{1}{2} \mathcal{D}_x^2 \left[\left\langle \xi \mid K_V(\cdot) \xi \right\rangle \right] \cdot (\tilde{x}, \cdot) \\ &\quad - \left(\mathcal{D}_x [K_H(\cdot)^{-1} \alpha] \right)^* \cdot \tilde{\alpha} - \frac{1}{2} \mathcal{D}_x^2 \left[\left\langle \alpha \mid K_H(\cdot)^{-1} \alpha \right\rangle \right] \cdot (\tilde{x}, \cdot).\end{aligned}$$

Finalement, en confondant les différentielles, éventuellement secondes, et leur écriture matricielle, on a :

$$\left\{ \begin{array}{l} \dot{\tilde{x}} = \mathcal{D}_x [K_V(\cdot) \xi] \tilde{x} + K_V(x) \tilde{\xi} \\ \dot{\tilde{\alpha}} = \mathcal{D}_x [K_H(\cdot)^{-1} \alpha] \tilde{x} + K_H(x)^{-1} \tilde{\alpha} \\ \dot{\tilde{\xi}} = - {}^t \left(\mathcal{D}_x [K_V(\cdot) \xi] \right) \tilde{\xi} - \frac{1}{2} \mathcal{D}_x^2 \left[\left\langle \xi \mid K_V(\cdot) \xi \right\rangle \right] \tilde{x} \\ \quad - {}^t \left(\mathcal{D}_x [K_H(\cdot)^{-1} \alpha] \right) \tilde{\alpha} - \frac{1}{2} \mathcal{D}_x^2 \left[\left\langle \alpha \mid K_H(\cdot)^{-1} \alpha \right\rangle \right] \tilde{x} \\ \dot{\tilde{\alpha}} = 0 \end{array} \right. .$$

Soit, sous forme matricielle :

$$\frac{d}{dt} \begin{pmatrix} \tilde{x} \\ \tilde{a} \\ \tilde{\xi} \\ \tilde{\alpha} \end{pmatrix} = \begin{pmatrix} \mathcal{D}_x[K_V(\cdot)\xi] & 0 & K_V(x) & 0 \\ \mathcal{D}_x[K_H(\cdot)^{-1}\alpha] & 0 & 0 & K_H(x)^{-1} \\ -\mathcal{D}^2 & 0 & -{}^t(\mathcal{D}_x[K_V(\cdot)\xi]) & -{}^t(\mathcal{D}_x[K_H(\cdot)^{-1}\alpha]) \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{a} \\ \tilde{\xi} \\ \tilde{\alpha} \end{pmatrix}$$

où

$$\mathcal{D}^2 := \frac{1}{2} \left(\mathcal{D}_x^2 [\langle \xi \mid K_V(\cdot)\xi \rangle] + \mathcal{D}_x^2 [\langle \alpha \mid K_H(\cdot)^{-1}\alpha \rangle] \right).$$

En fait, par cette méthode, on détermine facilement les quantités sus-citées mais non-transposées. Aussi et pour limiter les calculs nous allons plutôt considérer le problème adjoint au système ci-dessus.

c. Établissement de l'équation adjointe

Soit $\tilde{X} := (\tilde{x}, \tilde{a}, \tilde{\xi}, \tilde{\alpha})$ et

$$M := \begin{pmatrix} \mathcal{D}_x[K_V(\cdot)\xi] & 0 & K_V(x) & 0 \\ \mathcal{D}_x[K_H(\cdot)^{-1}\alpha] & 0 & 0 & K_H(x)^{-1} \\ -\mathcal{D}^2 & 0 & -{}^t(\mathcal{D}_x[K_V(\cdot)\xi]) & -{}^t(\mathcal{D}_x[K_H(\cdot)^{-1}\alpha]) \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

de telle sorte que l'équation précédente s'écrive $\dot{\tilde{X}}_t = M_t \tilde{X}_t$. Le système considéré étant linéaire, il existe une matrice – dite matrice résolvante – que l'on note \mathcal{R}^M et telle que pour tout temps s et t on ait $\tilde{X}_t = \mathcal{R}_{s,t}^M \tilde{X}_s$. En particulier,

$$\frac{d\tilde{X}_t}{dt} = \frac{\partial}{\partial t} (\mathcal{R}_{s,t}^M) \tilde{X}_s = M_t \tilde{X}_t = M_t \mathcal{R}_{s,t}^M \quad \text{et donc} \quad \frac{\partial}{\partial t} (\mathcal{R}_{s,t}^M) = M_t \mathcal{R}_{s,t}^M$$

ce qui est en fait une définition possible pour la matrice résolvante. Soit également les opérateurs de projections suivant

$$\pi_x(x, a, \xi, \alpha) = x \quad , \quad \pi_a(x, a, \xi, \alpha) = a \quad , \quad \pi_\xi(x, a, \xi, \alpha) = \xi \quad , \quad \pi_\alpha(x, a, \xi, \alpha) = \alpha$$

et leur adjoint respectif

$$\pi_x^*(x) = (x, 0, 0, 0) \quad , \quad \pi_a^*(a) = (0, a, 0, 0) \quad , \quad \pi_\xi^*(\xi) = (0, 0, \xi, 0) \quad , \quad \pi_\alpha^*(\alpha) = (0, 0, 0, \alpha)$$

de sorte que

$$\tilde{x}_1 = \pi_x \mathcal{R}_{0,1}^M (\pi_\xi^*(\tilde{\xi}_0) + \pi_\alpha^*(\tilde{\alpha}_0)) = \pi_x \mathcal{R}_{0,1}^M \pi_\xi^* \tilde{\xi}_0 + \pi_x \mathcal{R}_{0,1}^M \pi_\alpha^* \tilde{\alpha}_0$$

et

$$\tilde{a}_1 = \pi_a \mathcal{R}_{0,1}^M \pi_\xi^* \tilde{\xi}_0 + \pi_a \mathcal{R}_{0,1}^M \pi_\alpha^* \tilde{\alpha}_0.$$

Or, d'après ce qui précède et en omettant la dépendance en ξ_0 et α_0 de x_1 et a_1 ,

$$\tilde{x}_1 = \frac{\partial x_1}{\partial \xi_0} \cdot \tilde{\xi}_0 + \frac{\partial x_1}{\partial \alpha_0} \cdot \tilde{\alpha}_0 \quad \text{et} \quad \tilde{a}_1 = \frac{\partial a_1}{\partial \xi_0} \cdot \tilde{\xi}_0 + \frac{\partial a_1}{\partial \alpha_0} \cdot \tilde{\alpha}_0$$

d'où, en identifiant terme à terme,

$$\frac{\partial x_1}{\partial \xi_0} = \pi_x \mathcal{R}_{0,1}^M \pi_\xi^*, \quad \frac{\partial x_1}{\partial \alpha_0} = \pi_x \mathcal{R}_{0,1}^M \pi_\alpha^*, \quad \frac{\partial a_1}{\partial \xi_0} = \pi_a \mathcal{R}_{0,1}^M \pi_\xi^*, \quad \frac{\partial a_1}{\partial \alpha_0} = \pi_a \mathcal{R}_{0,1}^M \pi_\alpha^*$$

autrement dit

$${}^t \left(\frac{\partial x_1}{\partial \xi_0}(\xi_0, \alpha_0) \right) = \pi_\xi {}^t \mathcal{R}_{0,1}^M \pi_x^*, \quad {}^t \left(\frac{\partial x_1}{\partial \alpha_0}(\xi_0, \alpha_0) \right) = \pi_\alpha {}^t \mathcal{R}_{0,1}^M \pi_x^*, \quad \text{etc.}$$

et il suffit donc de déterminer ${}^t \mathcal{R}_{0,1}^M$ pour calculer le gradient de la fonctionnelle d'attache aux données. Par propriété de la résolvante, pour tout t et s ,

$$\begin{aligned} \tilde{X}_s &= (\mathcal{R}_{s,t}^M)^{-1} \tilde{X}_t \quad \text{et} \quad \frac{d\tilde{X}_s}{ds} = -(\mathcal{R}_{s,t}^M)^{-1} \frac{\partial}{\partial s} (\mathcal{R}_{s,t}^M) (\mathcal{R}_{s,t}^M)^{-1} \tilde{X}_t \\ &= -\mathcal{R}_{t,s}^M \frac{\partial}{\partial s} (\mathcal{R}_{s,t}^M) \tilde{X}_s \quad \text{d'une part} \\ &= M_s \tilde{X}_s \quad \text{d'autre part} \end{aligned}$$

et donc, pour tout temps t et s , $-\mathcal{R}_{t,s}^M \frac{\partial}{\partial s} (\mathcal{R}_{s,t}^M) = M_s$ *i.e.* $\frac{\partial}{\partial s} (\mathcal{R}_{s,t}^M) = -\mathcal{R}_{s,t}^M M_s$. D'où, par linéarité de la transposée,

$$\frac{\partial}{\partial s} (\mathcal{R}_{s,t}^{-tM}) = \mathcal{R}_{s,t}^{-tM} {}^t M_s \quad \text{et} \quad \frac{\partial}{\partial s} ({}^t \mathcal{R}_{s,t}^{-tM}) = M_s {}^t \mathcal{R}_{s,t}^{-tM}.$$

Finalement, par unicité de la matrice résolvante, pour tout t et s ,

$${}^t \mathcal{R}_{s,t}^{-tM} = \mathcal{R}_{t,s}^M \quad \text{i.e.} \quad \mathcal{R}_{s,t}^{-tM} = {}^t \mathcal{R}_{t,s}^M.$$

En particulier, ${}^t \mathcal{R}_{0,1}^M = \mathcal{R}_{1,0}^{-tM}$ et donc

$${}^t \left(\frac{\partial x_1}{\partial \xi_0}(\xi_0, \alpha_0) \right) \nabla_{x_1} A(x_1, a_1) = \pi_\xi {}^t \mathcal{R}_{1,0}^{-tM} \pi_x^* \nabla_{x_1} A(x_1, a_1) = \pi_\xi(Y_0)$$

où Y est solution de l'équation (rétrograde) $\dot{Y}_t = -{}^t M_t Y_t$ et $Y_1 = \pi_x^*(\nabla_{x_1} A(x_1, a_1))$. De même, on a :

$${}^t \left(\frac{\partial x_1}{\partial \alpha_0}(\xi_0, \alpha_0) \right) \nabla_{x_1} A(x_1, a_1) = \pi_\alpha(Y_0) \quad ; \quad \begin{cases} \dot{Y}_t &= -{}^t M_t Y_t \\ Y_1 &= \pi_x^*(\nabla_{x_1} A(x_1, a_1)) \end{cases}$$

et en faisant varier la condition finale de l'équation adjointe :

$$\begin{aligned} {}^t\left(\frac{\partial a_1}{\partial \xi_0}(\xi_0, \alpha_0)\right) \nabla_{a_1} A(x_1, a_1) &= \pi_\alpha(Y_0) \\ {}^t\left(\frac{\partial a_1}{\partial \alpha_0}(\xi_0, \alpha_0)\right) \nabla_{a_1} A(x_1, a_1) &= \pi_\alpha(Y_0) \end{aligned} \quad ; \quad \begin{cases} \dot{Y}_t &= -{}^t M_t Y_t \\ Y_1 &= \pi_a^*(\nabla_{a_1} A(x_1, a_1)) \end{cases} .$$

On considère donc dorénavant l'équation $\dot{Y}_t = -{}^t M_t Y_t$ avec pour condition finale

$$Y_1 = \begin{pmatrix} \nabla_{x_1} A(x_1, a_1) \\ \nabla_{a_1} A(x_1, a_1) \\ 0 \\ 0 \end{pmatrix}. \quad \text{Alors,} \quad \text{et} \quad \begin{cases} \nabla_{\xi_0} B(\xi_0, \alpha_0) &= \pi_\xi(Y_0) \quad \text{d'une part,} \\ \nabla_{\alpha_0} B(\xi_0, \alpha_0) &= \pi_\alpha(Y_0) \quad \text{d'autre part.} \end{cases}$$

Autrement dit, on est à même de déterminer le gradient de la fonctionnelle d'attache aux données et ainsi réaliser la descente de gradient annoncée.

2. Pré-requis à l'implémentation numérique

Pour résoudre le problème d'appariement, il faut donc résoudre les deux équations différentielles suivantes :

$$\begin{cases} \dot{x} &= K_V(x)\xi \\ \dot{a} &= K_H(x)^{-1}\alpha \\ \dot{\xi} &= -\frac{1}{2}\nabla_x \left[\langle \alpha \mid K_H(x)^{-1}\alpha \rangle + \langle \xi \mid K_V(x)\xi \rangle \right] \\ \dot{\alpha} &= 0 \end{cases}$$

pour déterminer l'ensemble des courbes géodésiques et $\dot{Y}_t = -{}^t M_t Y_t$ où

$$-{}^t M_t = \begin{pmatrix} -{}^t(\mathcal{D}_x [K_V(\cdot)\xi]) & -{}^t(\mathcal{D}_x [K_H(\cdot)^{-1}\alpha]) & \mathcal{D}^2 & 0 \\ 0 & 0 & 0 & 0 \\ -K_V(x) & 0 & \mathcal{D}_x [K_V(\cdot)\xi] & 0 \\ 0 & -K_H(x)^{-1} & \mathcal{D}_x [K_H(\cdot)^{-1}\alpha] & 0 \end{pmatrix}$$

et

$$\mathcal{D}^2 := \frac{1}{2} \left(\mathcal{D}_x^2 \left[\langle \xi \mid K_V(\cdot)\xi \rangle \right] + \mathcal{D}_x^2 \left[\langle \alpha \mid K_H(\cdot)^{-1}\alpha \rangle \right] \right)$$

pour déterminer celle minimisant le terme d'attache aux données.

Pour l'implémentation numérique, on s'inspire très largement des codes [14] réalisant un appariement de landmarks "classique". Aussi dispose t'on de fonctions permettant de

calculer pour tous vecteurs q et $p, r, \delta q$ dépendants éventuellement de q les expressions suivantes

$$\nabla_q \left\langle p \mid K(q)r \right\rangle \quad ; \quad \mathcal{D}_q K \cdot \delta q^3 \quad \text{et} \quad \mathcal{D}_q^2 \left[\left\langle p \mid K(\cdot)r \right\rangle \right] \cdot (\delta q, \cdot)$$

pour les noyaux reproduisant dits de Gauss et de Cauchy et que l'on va utiliser largement.

En fait, il suffit pour pouvoir réaliser l'implémentation numérique de coder les "pendants" aux fonctions ci-dessus pour l'inverse du noyau K^{-1} . Il vient :

- i. Soit $x \in \mathbb{R}^{Nd}$, $\alpha \in \mathbb{R}^N$ et $\delta x \in \mathbb{R}^{Nd}$. Par linéarité de la multiplication matricielle,

$$\mathcal{D}_x \left[K_H(\cdot)^{-1} \alpha \right] \cdot \delta x = \left[\mathcal{D}_x K_H^{-1} \cdot \delta x \right] \alpha = -K_H(x)^{-1} (\mathcal{D}_x K_H \cdot \delta x) K_H(x)^{-1} \alpha$$

et le résultat s'en suit avec la deuxième fonction ;

- ii. Soit $x \in \mathbb{R}^{Nd}$, $\alpha_1, \alpha_2 \in \mathbb{R}^N$ et $\delta x \in \mathbb{R}^{Nd}$. Par linéarité du produit scalaire et avec le calcul précédent, on a :

$$\begin{aligned} \mathcal{D}_x \left[\left\langle \alpha_1 \mid K_H(\cdot)^{-1} \alpha_2 \right\rangle \right] \cdot \delta x &= - \left\langle \alpha_1 \mid K_H(x)^{-1} (\mathcal{D}_x K_H \cdot \delta x) K_H(x)^{-1} \alpha_2 \right\rangle \\ &= - \left\langle \beta_{x,1} \mid (\mathcal{D}_x K_H \cdot \delta x) \beta_{x,2} \right\rangle \\ &= - \mathcal{D}_x \left[\left\langle \beta_{x,1} \mid K_H(\cdot) \beta_{x,2} \right\rangle \right] \cdot \delta x \end{aligned}$$

où on a posé $\beta_{x,i} = K_H(x)^{-1} \alpha_i$ pour $i = 1, 2$. Par suite,

$$\nabla_x \left[\left\langle \alpha \mid K_H(x)^{-1} \alpha \right\rangle \right] = - \nabla_x \left[\left\langle \beta_x \mid K_H(x) \beta_x \right\rangle \right]$$

et on conclut avec la première des trois fonctions ;

- iii. Soit $x \in \mathbb{R}^{Nd}$, $\delta x \in \mathbb{R}^{Nd}$ et $\alpha \in \mathbb{R}^N$. En utilisant la linéarité du produit scalaire et le caractère symétrique de la matrice $K_H(x)$ pour tout x ,

$$\begin{aligned} \mathcal{D}_x^2 \left[\left\langle \alpha \mid K_H(\cdot)^{-1} \alpha \right\rangle \right] \cdot (\delta x, \cdot) &= \mathcal{D}_x \left[\nabla_x \left[\left\langle \alpha \mid K_H(x)^{-1} \alpha \right\rangle \right] \right] \cdot \delta x \\ &= - \mathcal{D}_x \left[\nabla_x \left[\left\langle \beta_x \mid K_H(x) \beta_x \right\rangle \right] \right] \cdot \delta x \\ &= - \mathcal{D}_x^2 \left[\left\langle \beta_x \mid K_H(\cdot) \beta_x \right\rangle \right] \cdot (\delta x, \cdot) \\ &\quad - 2 \nabla_x \left[\left\langle \mathcal{D}_x \left[K_H(\cdot)^{-1} \alpha \right] \cdot \delta x \mid K_H(x) \beta_x \right\rangle \right] \cdot \end{aligned}$$

3. Alors, pour tout vecteur p on a $\mathcal{D}_q [K(\cdot)p] \cdot \delta q = (\mathcal{D}_q K \cdot \delta q) p$ dès lors que p est indépendant de q .

Modèle de pseudo-métamorphoses pour les mesures

Finalement, il suffit d'appliquer la première des fonctions à $\gamma_x := \mathcal{D}_x [K_H(\cdot)^{-1}\alpha] \cdot \delta x$ et $\beta_x := K_H(x)^{-1}\alpha$ ainsi que la dernière pour conclure.

Enfin, remarquons que pour tout $x, \xi \in \mathbb{R}^{Nd}$, et tout $\tilde{x}, \delta x \in \mathbb{R}^{Nd}$,

$$\left\langle {}^t(\mathcal{D}_x [K_V(\cdot)\xi]) \tilde{x} \mid \delta x \right\rangle = \left\langle \tilde{x} \mid (\mathcal{D}_x [K_V(\cdot)\xi]) \delta x \right\rangle = \mathcal{D}_x \left[\left\langle \tilde{x} \mid K_V(\cdot)\xi \right\rangle \right] \cdot \delta x$$

et donc, pour tout $x, \xi \in \mathbb{R}^{Nd}$ et $\tilde{x} \in \mathbb{R}^{Nd}$,

$${}^t(\mathcal{D}_x [K_V(\cdot)\xi]) \tilde{x} = \nabla_x \left[\left\langle \tilde{x} \mid K_V(\cdot)\xi \right\rangle \right].$$

De même, pour tout $x \in \mathbb{R}^{Nd}$, tout $a \in \mathbb{R}^N$ et $\tilde{a} \in \mathbb{R}^N$,

$${}^t(\mathcal{D}_x [K_H(\cdot)^{-1}\alpha]) \tilde{a} = \nabla_x \left[\left\langle \tilde{a} \mid K_H(\cdot)^{-1}\alpha \right\rangle \right]$$

et l'implémentation numérique est donc entièrement réalisable avec les trois fonctions sus-citées et leur équivalent pour les noyaux inverses.

3. Implémentation numérique

On réalise l'implémentation numérique en MatLab comme annoncé en préambule. Pour l'affichage graphique, on a choisi de faire varier la couleur des trajectoires des différents points selon les valeurs de a . Par soucis de précision, on a préféré une échelle relative à une échelle absolue et on adjoint donc une "barre" au portrait de phase afin de déterminer les variations de a au cours du temps ; même si ces variations sont en fait visibles sur le portrait de phase, la valeur de a modifiant celle de ξ et donc de x ... En particulier, lorsque a s'annule, on peut formellement retirer le point associé de la somme de Dirac, ce qui permet la disparition/apparition de points dans le modèle de pseudo-métamorphoses contrairement au matching de landmarks classique.

Afin de s'assurer que le programme informatique ne donne pas des résultats aberrants vis-à-vis de la théorie, on remarque que si le moment initial α_0 est choisi nul alors le modèle de pseudo-métamorphoses consiste en un appariement (exact), comme décrit au chapitre 1, au paragraphe II.1. En effet, de manière évidente :

Proposition II.3

Soit le système suivant avec pour condition initiale $(x_0, a_0, \xi_0, \alpha_0) = (\bar{y}, \bar{b}, \xi_0, 0)$

$$\begin{cases} \dot{x} &= K_V(x)\xi \\ \dot{a} &= K_H(x)^{-1}\alpha \\ \dot{\xi} &= -\frac{1}{2}\nabla_x \left[\left\langle \alpha \mid K_H(x)^{-1}\alpha \right\rangle + \left\langle \xi \mid K_V(x)\xi \right\rangle \right] \\ \dot{\alpha} &= 0 \end{cases}$$

Alors $a = \bar{b}$ et (x, ξ) est solution du système

$$\begin{cases} \dot{x} &= K_V(x)\xi \\ \dot{\xi} &= -\frac{1}{2}\nabla_x \langle \xi \mid K_V(x)\xi \rangle \end{cases} .$$

Numériquement, on obtient par exemple les résultats suivants si on effectue un shooting par appariement exact (figure II.2a) ou pseudo-métamorphoses avec un moment initial nul (II.2b) ou non (II.2c)

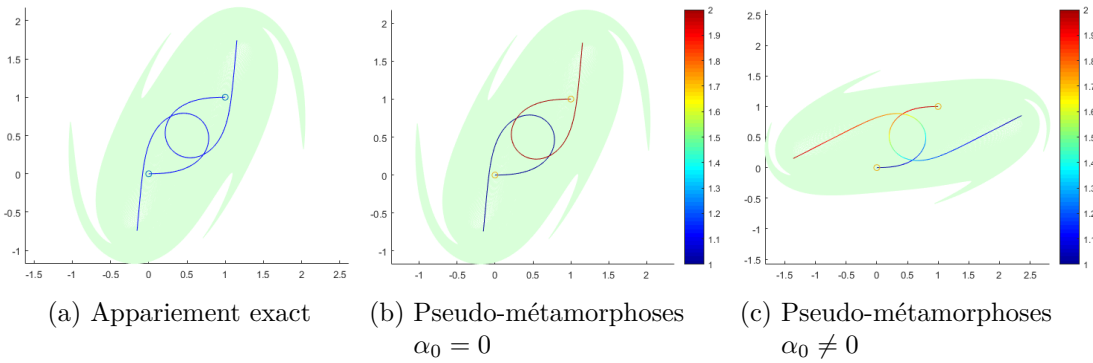


FIGURE II.2. – Appariement exact vs pseudo-métamorphoses

Les résultats ci-dessus sont obtenus pour les conditions initiales suivantes, si définies et non-nulles :

$$x_0 = \begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix} \quad ; \quad a_0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad ; \quad \xi_0 = \begin{pmatrix} 5 & 0 \\ -5 & 0 \end{pmatrix} \quad ; \quad \alpha_0 = \begin{pmatrix} 0.5 \\ -0.5 \end{pmatrix} .$$

Plus précisément, le code MatLab pour les pseudo-métamorphoses est le suivant :

```

1 | % initial points and momenta
2 | x0 = [0,0;1,1] ;
3 | a0 = [1;2] ;
4 | xi0 = 5*[1,0;-1,0] ;
5 | alpha0 = .5*[1;-1] ;
6 | % alpha0 = 0*[1;-1] ;
7 |
8 | d = size(x0,2); % dimension
9 | n = size(x0,1); % number of points
10 | sigma = 1; % kernel width
11 | ker = ScalarKernel(GaussFunction,sigma);
12 |
13 | H = HamiltSysPM(ker,ker,n,d) ;
    
```

```

14 |
15 | figure()
16 | H.PlotShoot(x0, a0, xi0, alpha0, 0, 1)

```

CODE SOURCE II.1 – Shooting par pseudo-métamorphoses

où :

→ `ScalarKernel` implémente la convolution avec les noyaux scalaires dit de Gauss et de Cauchy. Étant donné un poids γ et une constante d'échelle σ , le premier est de la forme

$$\forall x, y \in \mathbb{R}^N, \quad k(x, y) = \gamma \exp\left(-\frac{\|x - y\|^2}{\sigma^2}\right)$$

et le second

$$\forall x, y \in \mathbb{R}^N, \quad k(x, y) = \frac{\gamma \sigma^2}{\sigma^2 - \|x - y\|^2}.$$

→ `HamiltSys` implémente les différentes équations différentielles à résoudre lors de l'appariement de formes : les équations géodésiques d'une part mais également les équations rétrogrades nécessaires à la minimisation de la fonctionnelle d'attache aux données comme expliqué au paragraphe II.1.

Elle intègre également les fonctions d'affichages qui permettent de représenter graphiquement les différentes solutions. En particulier, la fonction `H.PlotShoot` réalise et affiche le shooting par pseudo-métamorphose, à partir de conditions initiales et sur un intervalle de temps donné.

Dans cet exemple on a en particulier choisi un noyau identique – à l'exception de leur dimension respective – pour les espaces H et V mais on peut dissocier ce choix.

4. Expérimentations sur des données tests

Avant d'utiliser notre modèle sur de véritables données médicales, on veut s'assurer de sa validité sur des données tests. Cela devrait également nous permettre de valider nos intuitions quant aux différents rôles joués par les grandeurs engagées.

On a également testé la comptabilité de notre modèle avec celui de métamorphoses développé par Casey L. Richardson et Laurent Younes dans [33]. Ces expériences sont reportées au paragraphe suivant, à la page 260.

a. Validation du Shooting par pseudo-métamorphoses

Du fait du modèle de pseudo-métamorphoses qui demande à considérer $N = m + n$ points, nous allons être amenés à réaliser des shooting faisant intervenir un nombre de points relativement élevé. Aussi, avant de tester l'appariement par matching, on s'assure que le shooting reste stable, même pour un nombre élevé de points.

On commence par regarder l'influence du nombre de points en considérant cette fois ci N points aléatoires dans le carré $[-5, 5] \times [-5, 5]$ de poids initiaux nuls et pour des moments tirés aléatoirement également : pour ξ à coordonnées comprises entre -5 et 5 ; pour α entre -1 et 1 . Autrement dit, pour $N = 10, 20, 50$:

```

1 | x0 = rand(N,2)*(-10)+5 ;
2 | a0 = zeros(N,1) ;
3 | xi0 = rand(N,2)*(-10)+5 ;
4 | alpha0 = rand(N,1)*(-2)+1 ;

```

On obtient, entre autre selon les tirages, les figures suivantes :

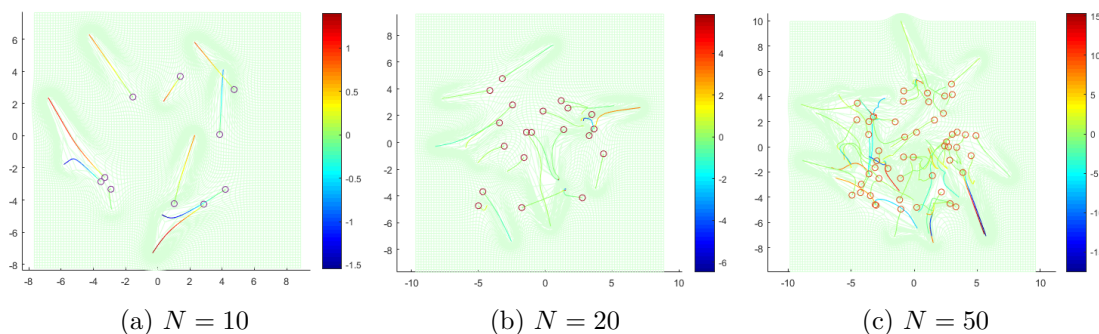


FIGURE II.3. – Influence du nombre de points pour le shooting

La figure II.3 permet de mettre en emphase l'influence des différents points sur la trajectoire empruntée par leur(s) voisin(s). Un premier moyen de s'en convaincre est l'aspect globalement plus rectiligne des trajectoires dans le cas où $N = 10$ (figure II.3a) que lorsque $N = 50$ (figure II.3c), au moins pour des temps petits, *i.e* au voisinage des positions initiales symbolisées par les cercles. En effet, les points étant plus espacés dans la première configuration, ils subissent moins l'influence de leur(s) voisin(s). Un deuxième moyen de s'en convaincre est la présence de trajectoires parallèles sur la figure II.3a. Ceci illustre la propension des points à "voyager" ensemble du fait du modèle de LDDMM qui induit un difféomorphisme transportant le nuage de points dans sa globalité et non point par point.

Pour "mesurer" l'influence des différentes positions initiales les unes sur les autres, on regarde à présent la situation suivante : on reprend les mêmes conditions initiales que pour la figure II.2c même en y adjoignant deux points de poids et moment similaires. A savoir, on considère la situation suivante, pour différentes valeurs de ε :

$$x_0 = \begin{pmatrix} 0 & 0 \\ 1 & -1 \\ \varepsilon & 0 \\ 1 + \varepsilon & -1 \end{pmatrix} ; \quad a_0 = \begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix} ; \quad \xi_0 = \begin{pmatrix} 5 & 0 \\ -5 & 0 \\ -5 & 0 \\ 5 & 0 \end{pmatrix} ; \quad \alpha_0 = \begin{pmatrix} 0.5 \\ -0.5 \\ -0.5 \\ 0.5 \end{pmatrix} .$$

On obtient alors les trajectoires présentées à la figure II.4. Cette figure permet de mettre en évidence d'une part l'influence des différentes positions initiales des points sur les trajectoires de leur(s) voisins(s) comme annoncé mais également le phénomène d'instabilité numérique dû à une accumulation d'erreur arrondis. En effet,

- Sur les figures II.4h et II.4i qui présentent deux configurations où les points sont relativement éloignés les uns des autres, on n'observe pas de perturbation des différentes trajectoires, au moins pour l'intervalle de temps sur lequel on a effectué le shooting, à savoir $[0, 1]$. Sur la figure II.4i, on remarque même une zone de l'espace qui ne subit aucune déformation par le modèle LDDMM entre les deux couples de points.
- Pour $\varepsilon = 2$, figure II.4g, les positions initiales sont suffisamment proches pour que l'influence des autres points soit tangible mais sans pour autant "détruire" complètement l'aspect général des trajectoires non-perturbées.
- Pour ε compris entre 0.25 et 1, *i.e* pour les figures II.4d, II.4e et II.4f, l'influence des points est de plus en plus forte et les trajectoires sont modifiées en "profondeur" par rapport à celles observées aux figures II.4h et II.4i. Ceci s'explique par le fait que le difféomorphisme induit par le modèle doit déplacer l'ensemble du nuage de points et qu'il subit des contraintes "contradictoires" à l'état initial pour les différents moments – et que l'on peut en quelques sorte interpréter comme la vitesse des points – qui sont de directions opposées pour certains.
- A la figure II.4c, les points initiaux sont très proches ($\varepsilon = 0.1$) et les trajectoires sont fortement modifiées par rapport à la situation II.4i. Néanmoins, on observe des couples de points qui "voyagent ensemble" et les valeurs de a_1 et a_2 restent bornées au cours du temps, au moins pour des temps compris entre 0 et 1, aussi le résultat semble cohérent avec la théorie.
- En revanche, on observe aux figures II.4a et II.4b les effets de l'instabilités numérique. En effet, si le modèle est théoriquement valide dans $\mathcal{K}^N(\mathbb{R}^d) = \mathcal{L}^N(\mathbb{R}^d) \times \mathbb{R}^N$, la distance entre les deux paires de points à l'état initial est trop réduite pour que cette condition soit vérifiée numériquement, l'ordinateur n'ayant pas forcément une précision suffisante. Ce phénomène est d'autant plus accentué que l'on est amené à inverser de nombreuses matrices (de noyaux) dans le modèle de pseudo-métamorphoses et qu'un mauvais conditionnement de ces dernières – ce qui va être le cas avec des points excessivement proches, leur déterminant devenant quasiment nul – va induire une accumulation d'erreurs d'arrondis. En pratique, cela se traduit pas une explosion en temps finis, au moins pour les valeurs de a qui se mesurent en 10^4 à la figure II.4a.

En pratique, on constate donc que le shooting par pseudo-métamorphoses est relativement stable mais sujet à explosion en temps finis, notamment du fait des nombreuses inversions de matrices. Cela peut – et va – être un problème si on s'intéresse au matching

de deux configurations. On cherche dorénavant à (in-)valider le matching par métamorphoses en procédant de manière similaire.

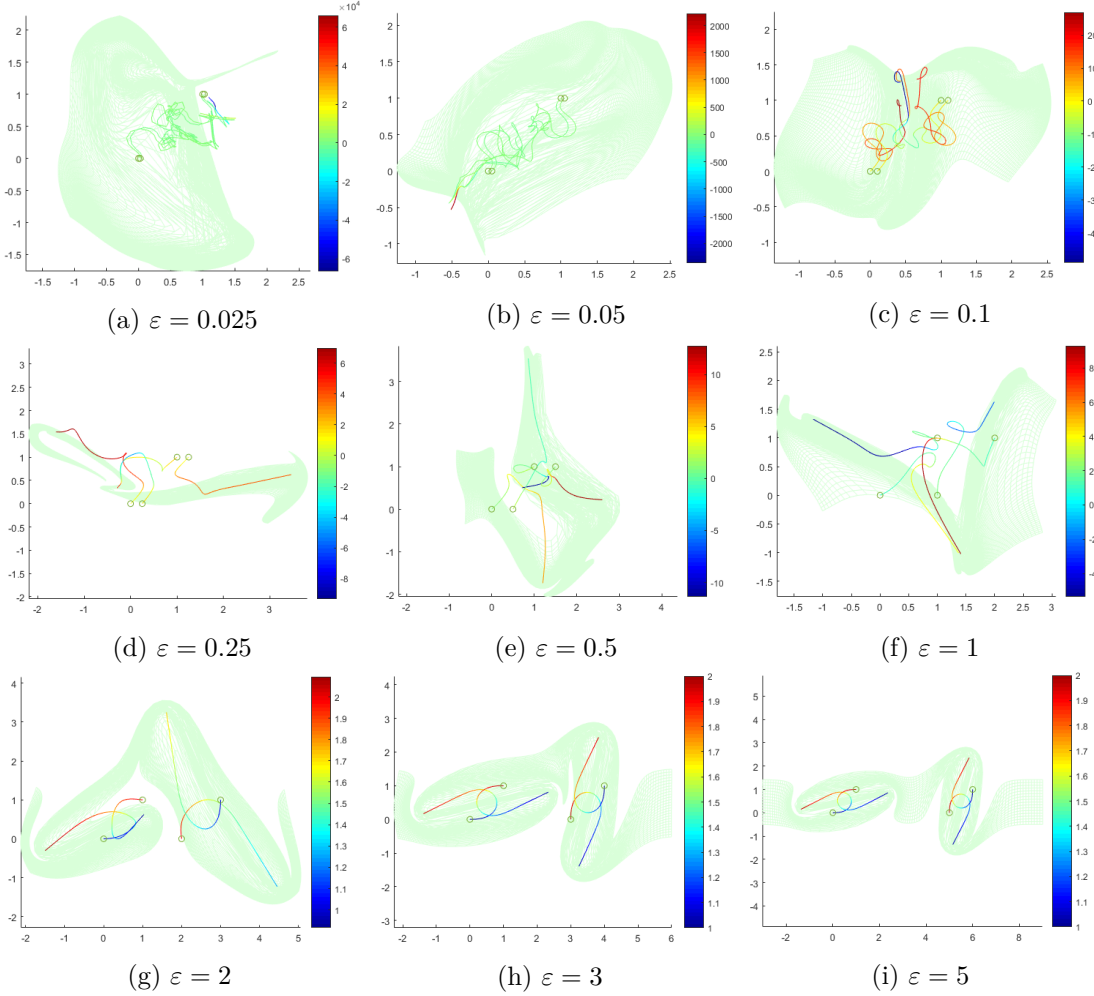


FIGURE II.4. – Influence sur les trajectoires des points voisins et instabilité numérique

b. Validation du Matching par pseudo-métamorphoses

On souhaite à présent tester la fonction réalisant le matching entre deux configurations de points. On commence par essayer d'apparier un point sur deux points de poids identiques. Plus précisément, on cherche à apparier les mesures suivantes :

$$\mu_y := \delta_{(0,0)} \quad \text{et} \quad \mu_z := \frac{1}{2} \left(\delta_{(1, \frac{1}{4})} + \delta_{(1, -\frac{1}{4})} \right)$$

Modèle de pseudo-métamorphoses pour les mesures

et on exécute pour cela le code suivant :

```
1 y = [0,0] ;
2 b = [1] ;
3 z = [1,.25;1,-.25] ;
4 c = [.5;.5] ;
5 m = size(y,1) ;
6 n = size(z,1) ;
7 d = size(y,2) ;
8
9 % choose kernel
10 sigmaV = .5 ; % kernel width
11 sigmaH = .5 ;
12 weightV = 1 ; % kernel weight
13 weightH = 5 ;
14 kerV = ScalarKernel(CauchyFunction,sigmaV,weightV) ; % Cauchy kernel
15 kerH = ScalarKernel(GaussFunction,sigmaH,weightH) ; % Gauss kernel
16
17 % initialize matching
18 nb = m+n ;
19 H = HamiltSysPM(kerV,kerH,nb,d) ;
20 DM = DiffeoMatchPM(H,gammaR) ;
21 OD = OptimDiffeoPM(DM) ;
22
23 % initialize positions and momenta
24 [x0,a0,xi0,alpha0] = DM.Init(y,b,z,c) ;
25
26 % optimize on initial positions and moment
27 % if t = 0, range = 1:m
28 % if t = 1, range = (m+1):(m+n)
29 DM.AddTarget( TargetLandmarksPM(y,b,nb,1:m) ,0 ) ;
30 DM.AddTarget( TargetLandmarksPM(z,c,nb,(m+1):(m+n)) ,1 ) ;
31 [x0,a0,xi0,alpha0] = OD.Optimize(x0,a0,xi0,alpha0) ;
32 % [x0,a0,xi0,alpha0] = OD.Optimize_Energy(x0,a0,xi0,alpha0,m) ;
33
34 % display results
35 figure()
36 clf
37 OD.Plot(x0,a0,xi0,alpha0) ;
```

CODE SOURCE II.2 – Matching par pseudo-métamorphoses

où :

- `OptimDiffeoPM` réalise la minimisation par descente de gradient par une méthode Low-BFGS comme décrite précédemment de la fonctionnelle donnée par `DM`. Plus précisément, la minimisation est réalisée par l'appelle à la fonction `OD.Optimize` à la ligne 31. La fonction `OD.Optimize_Energy` réalise l'optimisation alternée dont il est question à la fin de ce paragraphe.
- `DiffeoMatchPM` initialise la minimisation (avec la convention symétrique, via la

fonction `DM.Init`) et implémente la fonctionnelle d'appariement à minimiser, en l'occurrence le terme d'attache aux données A (cf paragraphe II.1) définie par :

$$A(x_0, a_0, x_1, a_1) = \sum_{i=1}^m |x_0^i - y^i|^2 + \sum_{i=1}^m |a_0^i - b^i|^2 + \sum_{i=1}^n |a_0^{m+i}|^2 + \sum_{i=1}^n |x_1^{m+i} - z^i|^2 + \sum_{i=1}^m |a_1^i|^2 + \sum_{i=1}^n |a_1^{m+i} - c^i|^2$$

et le gradient de cette dernière. Plus précisément, on procède point cible par point cible et le but de la fonction `DM.AddTarget` est d'ajouter des contraintes pour le matching, soit les m premières valeurs du couple (x_0, a_0) et les n dernières de (x_1, a_1) dans le cas des pseudo-métamorphoses.

Mais, en fait le résultat n'est pas aussi probant que ce que l'on aurait pu espérer et on obtient avec l'initialisation "symétrique"

$$x_0 = (y, z) \quad ; \quad a_0 = (b, c) \quad ; \quad \xi_0 = 0_{\mathbb{R}^{ND}} \quad ; \quad \alpha_0 = 0_{\mathbb{R}^N}$$

la figure II.5 où les trajectoires démarrent aux \cdot :

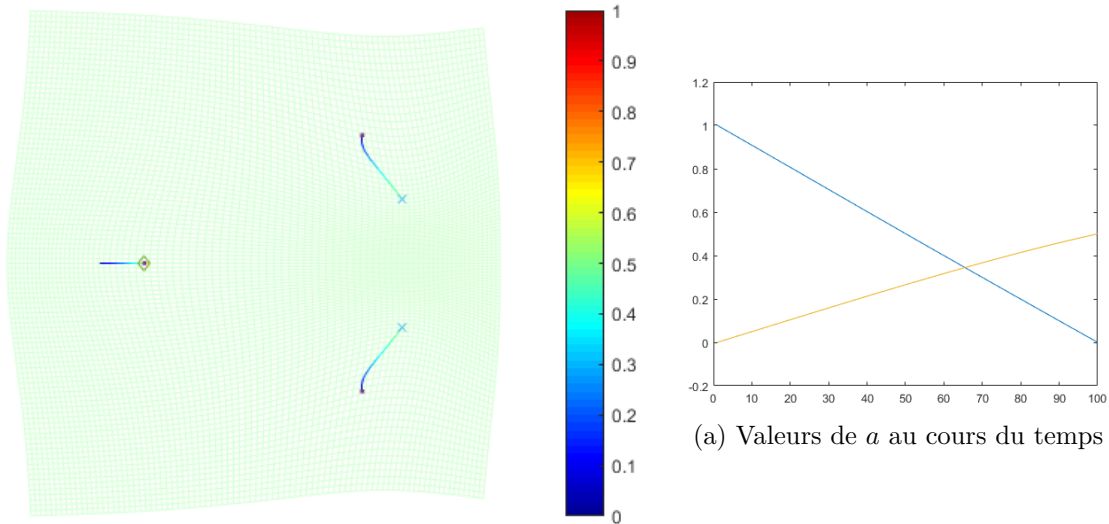


FIGURE II.5. – Matching par pseudo-métamorphoses

ce qui manifestement ne réalise absolument pas le matching entre μ_y et μ_z , du moins au sens où on l'entend. En fait, au vu des contraintes de notre modèle, l'appariement est réalisé. En effet, au temps $t = 0$, on a :

$$x_0 = \begin{pmatrix} 0 & 0 \\ \star & \star \\ \star & \star \end{pmatrix} \quad \text{et} \quad a_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad i.e. \quad \mu_0 = \delta_{(0,0)} = \mu_y$$

tandis qu'au temps $t = 1$:

$$x_1 = \begin{pmatrix} \star & \star \\ 1 & 0.25 \\ 1 & -0.25 \end{pmatrix} \quad \text{et} \quad a_1 = \begin{pmatrix} 0 \\ 0.5 \\ 0.5 \end{pmatrix} \quad \text{i.e.} \quad \mu_1 = \frac{1}{2} \left(\delta_{(1, \frac{1}{4})} + \delta_{(1, -\frac{1}{4})} \right) = \mu_z$$

et donc l'appariement est en fait réalisé et de manière exacte. En fait, le problème tient du fait que l'on a un système avec $2N(d+1)$ degrés de liberté (x et ξ vivent dans \mathbb{R}^{Nd} ; a et α dans \mathbb{R}^N) alors que l'on a $n(d+1) + m(d+1) = N(d+1)$ contraintes (n premières valeurs de x_0 et a_0 ; m dernières de x_1 et a_1). Notre espoir de voir les équations géodésiques contraindre suffisamment les trajectoires possibles pour les différents points n'est donc pas assouvi, on aurait pourtant pû penser que la propension des points à "vouloir voyager ensemble" comme décrit au paragraphe précédent (ce qui permet en fait de minimiser l'énergie nécessaire à la déformation de l'espace) aurait servi à favoriser des trajectoires parallèles. Pour mesurer l'effet de l'initialisation dans le matching, on relance sur la même configuration de points mais avec les initialisations suivantes : on impose toujours $\xi_0 = 0_{\mathbb{R}^{Nd}}$, $\alpha_0 = 0_{\mathbb{R}^N}$ et $a_0 = (b, c) = {}^t(1 \quad 0.5 \quad 0.5)$ – sauf dans le dernier cas – mais on fait varier x_0 comme suit :

$$\begin{array}{ll} \text{a. } x_0 = \begin{pmatrix} 0 & 0 \\ 0 & 0.25 \\ 0 & -0.25 \end{pmatrix}; & \text{d. } x_0 = \begin{pmatrix} 0 & 0 \\ 0 & 0.5 \\ 0 & -0.5 \end{pmatrix}; \\ \text{b. } x_0 = \begin{pmatrix} 0 & 0 \\ -0.05 & 0.05 \\ -0.05 & -0.05 \end{pmatrix}; & \text{e. } x_0 = \begin{pmatrix} 0 & 0 \\ 0 & 5 \\ 0 & -5 \end{pmatrix}; \\ \text{c. } x_0 = \begin{pmatrix} 0 & 0 \\ -0.01 & 0.01 \\ -0.01 & -0.01 \end{pmatrix}; & \text{f. } x_0 = (z, y) \text{ et } a_0 = (c, b). \end{array}$$

On reporte les résultats à la figure II.7, en faisant correspondre le label des figures et ceux de la liste ci-dessus. Les 5 premières initialisations correspondent à des situations artificielles où on a choisi d'initialiser la descente de gradient d'une manière qui nous arrangeait compte tenu de notre problème. Plus précisément :

- a. On cherche à appairer le point $(0, 0)$ et la paire de points $\left\{ (1, \frac{1}{4}); (1, -\frac{1}{4}) \right\}$. Sans contrainte supplémentaire, on a constaté (figure II.5) que le matching ne fonctionnait pas. On aimerait "forcer" les trajectoires des points à être parallèles. On initialise donc x_0 sur la valeur miroir de celle cible. On obtient de bien meilleurs résultats qu'avec l'initialisation symétrique par défaut : la couleur bleue marine correspondant à des poids nuls et donc à des points "inexistants" : on voit assez clairement le point $(0, 0)$ disparaître, tandis que deux points apparaissent de part et d'autre et voient leur poids respectif croître jusqu'à atteindre la valeur seuil $\frac{1}{2}$.

- b. Fort de l'amélioration constatée à la figure II.7a, on poursuit dans la direction suivante : initialiser x_0 au plus près de \bar{y} . On a d'abord voulu initialiser x_0 sur l'entrée $0.05 * (y, z)$ mais le système était alors trop instable numériquement et le matching ne pouvait avoir lieu. On obtient avec l'initialisation retenue le résultat au plus proche de ce qu'on aimerait avoir⁴. Notamment, on a bien disparition du point $(0, 0)$ et apparition de deux autres points comme précédemment.
- c. On poursuit l'idée de la figure II.7b en initialisant le vecteur x_0 comme presque nul. Les points ne se déplacent alors plus du tout et on abandonne l'idée de procéder à une initialisation pour le matching en considérant uniquement des versions translatées de μ_y très proches de cette-dernière, ceci pouvant empêcher le matching, sûrement du fait que l'écart entre les points initiaux ne permet pas de "décider" d'une direction de descente. En effet, l'algorithme L-BFGS s'arrête en seulement 2 itérations et rend comme valeur minimale pour A 2.9896, ce qui montre que la minimisation n'a en fait pas eu lieu.
- d. On repars de l'idée exploitée à la figure II.7a : initialiser x sur la l'axe des ordonnées. On initialise cette fois-ci x de manière plus étalée, *i.e* avec moins d'*a priori* sur le résultat escompté. Le résultat reste plutôt bon : le point $(0, 0)$ qui a l'air de partir dans la direction opposée de celle attendue disparaissant dans les faits assez rapidement...
- e. On continue avec cette idée en initialisant toujours x_0 sur l'axe des ordonnées, mais avec un écart bien plus conséquent. L'idée sous-jacente d'une telle expérience était de se demander si on ne pouvait pas imposer une initialisation non-symétrique mais plus systématique que les précédentes et voir si (ou plutôt comment) une "erreur" dans cette initialisation nuisait au matching. Contrairement aux apparences, le matching n'est pas mauvais (au moins qualitativement) avec cette configuration de départ : en effet, l'algorithme "pousse" le point $(0, 0)$ vers la droite jusqu'à apparition des deux points qui vont chacun converger vers $(1, \pm \frac{1}{4})$ et le point $(0, 0)$ disparaît alors.
- f. Enfin, la figure II.7f correspond à l'initialisation symétrique miroir de celle prise par défaut. Si le matching n'a pas lieu au sens où on l'entend – mais à bien évidemment lieu pour le modèle de peuso-métamorphoses dans $\mathcal{K}^N(\mathbb{R}^d)$ – le résultat obtenu semble meilleur qu'avec l'initialisation "officielle", ce qui est assez contre-intuitif... Par contre, on observe pour la première fois une dissymétrie dans les trajectoires des deux points liés $(1, \pm \frac{1}{4})$, ce qui est contraire à ce qu'on attend *a priori*, les contraintes étant formulées de manière symétrique.

Finalement, même si le matching ne fonctionne pas en l'état, à condition de contraindre (relativement fortement par moment) l'initialisation avant la minimisation, on obtient de meilleurs résultats. On peut donc peut-être envisager de creuser cette piste en vue

4. L'aspect de la grille relevant de problèmes d'affichages...

d'améliorer le modèle en ajoutant des contraintes sur l'initialisation. Cela semble cependant non pleinement satisfaisant : il faut doubler le nombre de contraintes sur le système si on veut qu'il soit entièrement déterminé...

Une autre idée pour améliorer le matching était la suivante : alterner la minimisation de la fonctionnelle d'attache A et de l'énergie de la trajectoire J ; on se rapprocherait alors de l'idée du matching inexact. On a seulement testé si le fait d'introduire quelques étapes de minimisation de l'énergie apportait une amélioration sur les trajectoires. Plus précisément, on a répété trois fois le schéma suivant :

- Réaliser la minimisation de l'attache aux données par l'algorithme L-BFGS ;
- Mettre à jour x_0 , a_0 , ξ_0 et α_0 en conséquence ;
- Réaliser la minimisation de la fonctionnelle d'énergie J ;
- Mettre à jour x_0 en conséquence.

Mais, les résultats n'étant pas probants, on a abandonné cette idée. Déjà, le temps de calcul est d'autant plus allongé – en fait, ce n'est pas complètement vrai : les minimisations suivantes étant la plupart du temps plus rapides que la première – que l'on répète ce motif et l'amélioration n'est pas si nette.

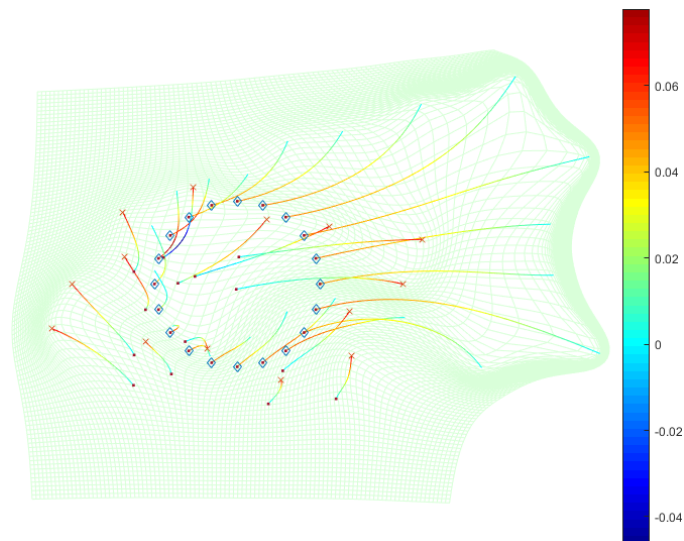


FIGURE II.6. – Appariement d'un cercle sur une courbe

On regarde quand même à la figure II.6 ce qui se passe si on essaie d'apparier un cercle sur une courbe quelconque, en considérant que les poids sont équi-répartis sur les deux courbes. Comme on s'y attend, cela ne fonctionnant déjà pas pour le matching simple de 3 landmarks, on n'apparie absolument pas les deux courbes.

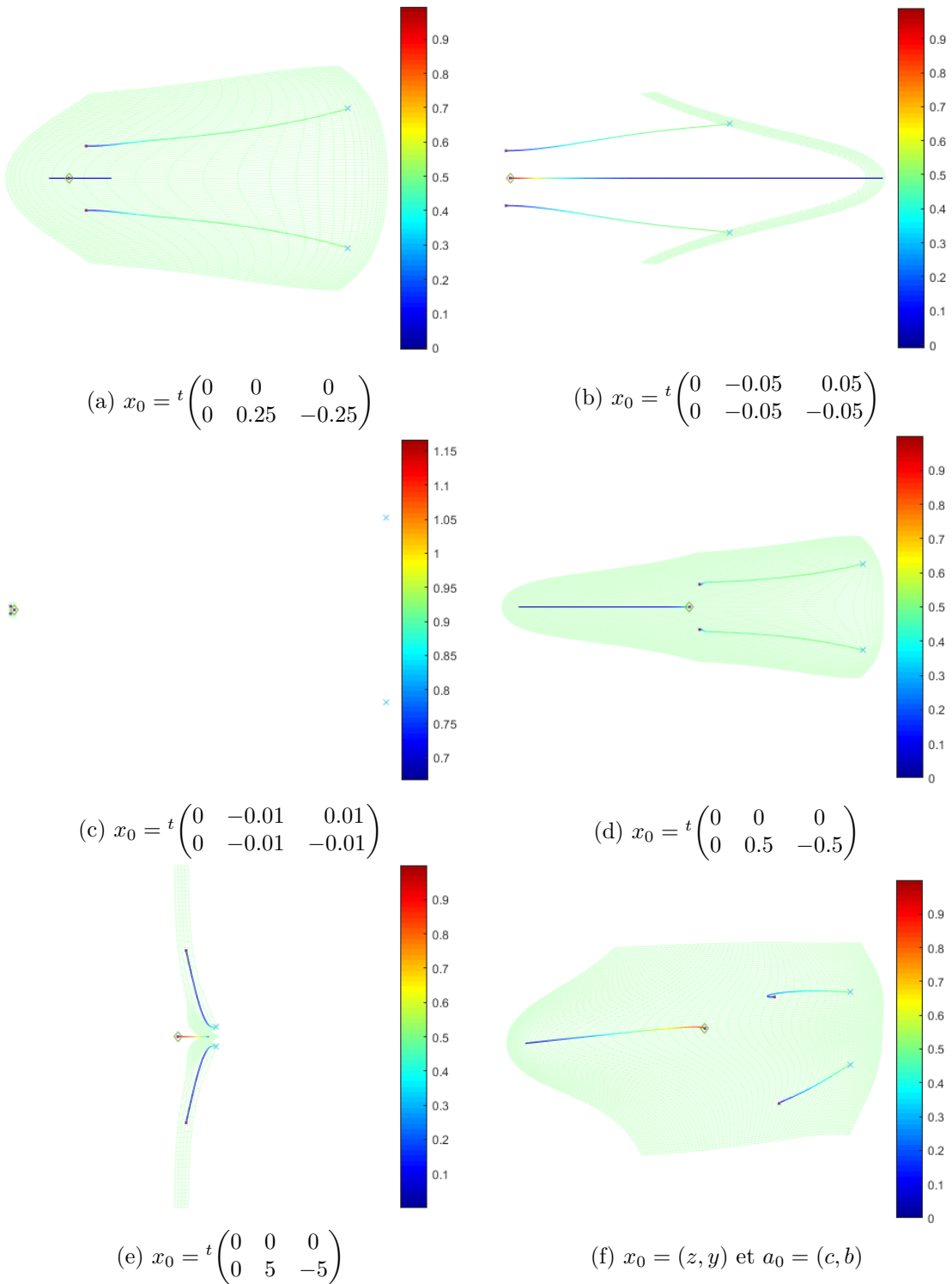


FIGURE II.7. – Influence de l'initialisation sur le matching

III. Métamorphoses vs pseudo-métamorphoses

Enfin, on s'intéresse ici à savoir si les résultats que l'on obtient numériquement par notre modèle de pseudo-métamorphoses sont cohérents avec ceux obtenus par Richardson et Younes dans [33]; ou, du moins, on aurait aimé pouvoir tester la comptabilité de ceux-ci mais, le matching par pseudo-métamorphoses n'étant pas efficient, les tests se sont avérés limités.

Dans l'article sus-cité, on trouve notamment une illustration du matching réalisé par métamorphoses d'un cercle sur un carré que l'on reporte à la figure II.8. On remarque entre autre sur cette figure que les points ne sont pas seulement déplacés par le modèle de métamorphoses mais peuvent également apparaître et disparaître, idée que l'on a voulu reproduire avec notre modèle de pseudo-métamorphoses.

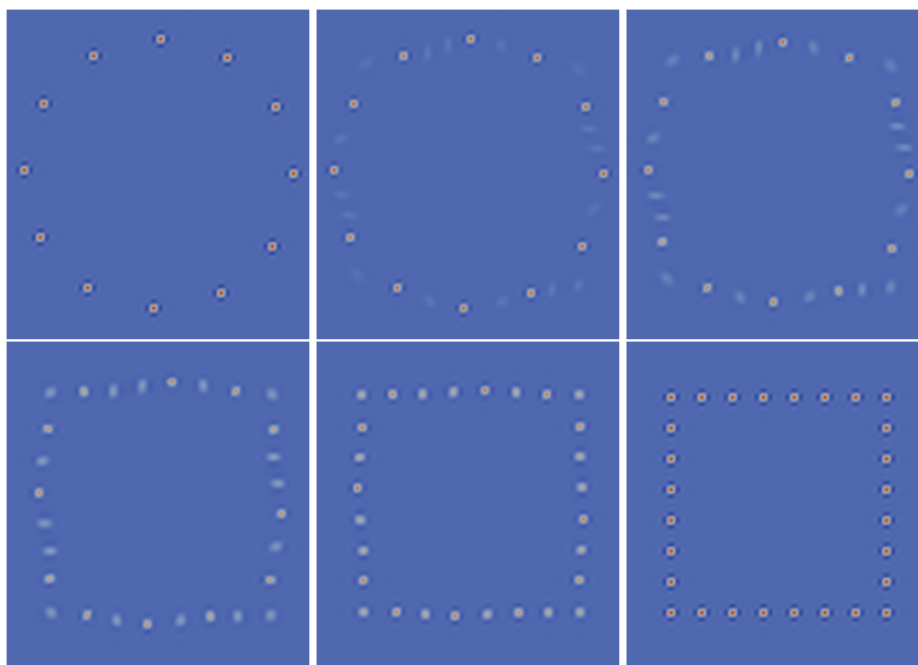


FIGURE II.8. – Appariement d'un cercle sur un carré par métamorphoses.

Crédit : Casey L. Richardson et Laurent Younes

On essaie donc d'apparier un disque sur un carré avec notre modèle. Dès lors que le nombre de points est trop élevés $N = m + n > 27$ par exemple, alors le système explose en temps fini et le matching ne peut avoir lieu. En fait, même pour un faible nombre de points, le matching n'a pas pour autant lieu – au sens où on l'entend – comme en témoigne la figure II.9.

Comme on avait obtenu des résultats moins mauvais en initialisant x_0 et a_0 de manière inverse à la convention naturelle, à savoir

$$x_0 = (y, z) \quad ; \quad a_0 = (b, c) \quad \text{vs} \quad x_0 = (z, y) \quad ; \quad a_0 = (c, b)$$

on a voulu retenter ici pour voir si ce choix ne pouvait pas s'avérer préférable, au moins dans un premier temps et à défaut de "mieux". Ce n'est manifestement pas le cas.

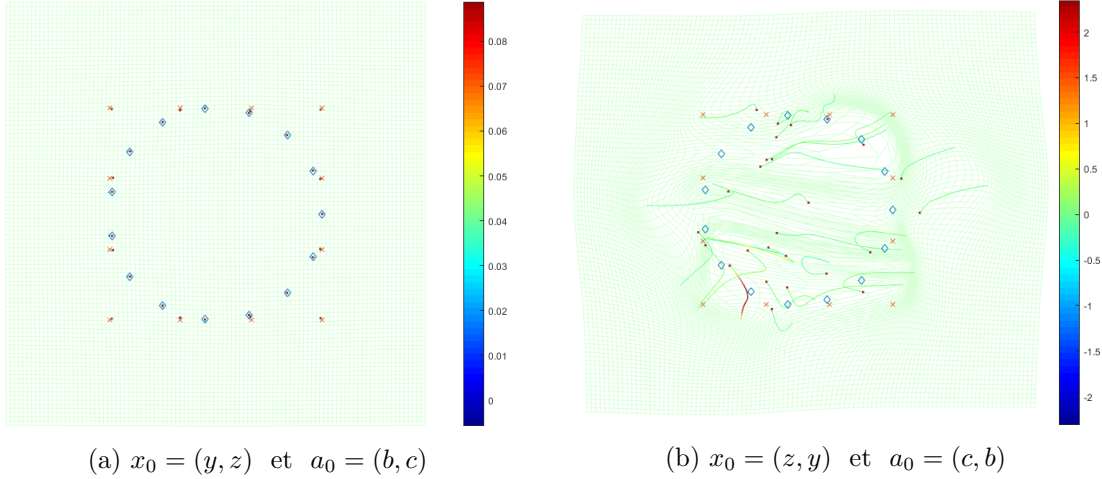


FIGURE II.9. – Appariement d'un cercle sur un carré par pseudo-métamorphoses.

Enfin, on lance une dernière simulation pour illustrer les possibles apparitions/disparitions de points au cours du shooting (et *a fortiori* du matching) selon les valeurs du moment α . En particulier, on démarre le shooting depuis

$$x_0 = \frac{1}{6} \begin{pmatrix} 6 & 6 \\ 6 & 3 \\ 6 & 0 \\ 6 & -3 \\ 6 & -6 \\ 2 & -6 \\ -2 & -6 \\ -6 & -6 \\ -6 & -3 \\ -6 & 0 \\ -6 & 3 \\ -6 & 6 \\ -2 & 6 \\ 2 & 6 \end{pmatrix} \quad ; \quad a_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad ; \quad \xi_0 = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix} \quad ; \quad \alpha_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$$

soit pour x_0 le carré de côté de longueur 2, centré en l'origine et duquel on "retire" l'arrête droite via a_0 . On obtient alors la figure suivante :

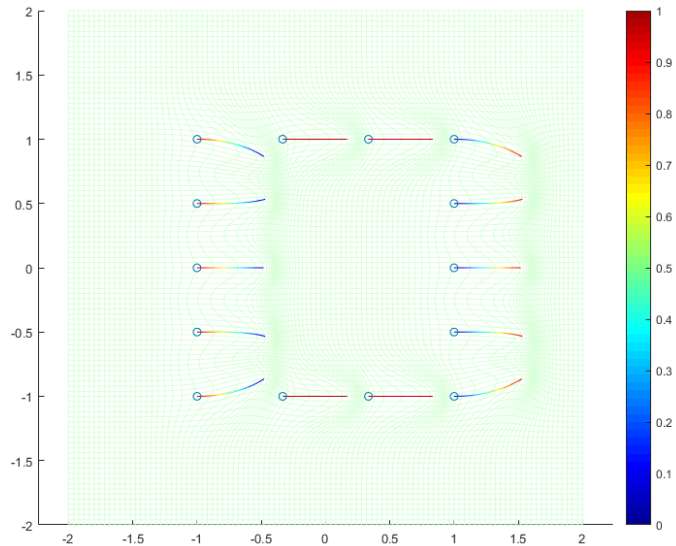


FIGURE II.10. – Un nombre de points qui varie au cours du temps.

En particulier, on voit très clairement que l'arrête droite est "apparue" tandis que l'arrête de gauche a "disparu" : les poids initiaux des points composants l'arrête droite étaient nuls au temps $t = 0$ et valent 1 une fois le shooting effectué ; au contraire, les poids de l'arrête gauche, initialisés sur la valeur 1, ont été annulés par le shooting. Les poids restants sont restés identiques, ce qui est cohérent avec la proposition II.3 qui affirme qu'en cas de nullité du moment initial α_0 , le modèle de pseudo-métamorphoses coïncide avec le matching de landmarks classique.

– Annexe A –

Résultats complémentaires sur les courants

On rappelle ici quelques résultats fondamentaux sur les courants, dont une des utilisations dans le cadre de l'anatomie numérique est "l'encodage de formes" comme décrit au paragraphe I.1.a du chapitre premier.

On se base principalement sur [16] et les chapitres 4 de [30] et [45].

I. Formes extérieures

Soit E un espace vectoriel de dimension n sur \mathbb{R} . On note $\otimes^p E^*$ l'ensemble des formes p -linéaires sur E . C'est un espace vectoriel sur \mathbb{R} .

Définition A.1. Une forme p -linéaire η sur E est dite alternée si pour $i \neq j$ on a

$$\forall x_1, \dots, x_p \in E, \quad \eta(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_{j-1}, x_j, x_{j+1}, \dots, x_p) \\ = -\eta(x_1, \dots, x_{i-1}, x_j, x_{i+1}, \dots, x_{j-1}, x_i, x_{j+1}, \dots, x_p).$$

On note $\Lambda^p E^*$ l'espace des formes p -linéaires alternées sur E , dites aussi formes extérieures.

Exemple : • Par convention, $\Lambda^0 E^* = \mathbb{R}$;

- L'espace $\Lambda^1 E^*$ s'identifie au dual E^* de E , i.e à l'espace des formes linéaires sur E . Il est donc en particulier de dimension n ;
- Le déterminant dans une base est un élément de $\Lambda^n E^*$.

Proposition A.2

Une forme p -linéaire η est alternée si et seulement si pour toute famille liée $(x_1, \dots, x_p) \in E^p$ on a $\eta(x_1, \dots, x_p) = 0$.

Proposition A.3

Soit $\eta \in \Lambda^p E^*$. On a :

$$\forall \sigma \in \mathfrak{S}_p, \quad \forall x_1, \dots, x_p \in E, \quad \eta(x_1, \dots, x_p) = \varepsilon(\sigma) \eta(x_{\sigma(1)}, \dots, x_{\sigma(p)}).$$

où ε est le morphisme signature.

Définition A.4. Soit $\alpha \in \otimes^p E^*$, on appelle partie alternée de α et on note $P(\alpha)$ la forme extérieure vérifiant :

$$\forall x_1, \dots, x_p \in E, \quad P(\alpha)(x_1, \dots, x_p) = \frac{1}{p!} \sum_{\sigma \in \mathfrak{S}_p} \varepsilon(\sigma) \alpha(x_{\sigma(1)}, \dots, x_{\sigma(p)})$$

Résultats complémentaires sur les courants

On vérifie facilement que $P(\alpha)$ est bien dans $\Lambda^p E^*$. On peut maintenant définir un "produit" sur les forme extérieures :

Définition A.5. Soit $\alpha \in \Lambda^p E^*$ et $\beta \in \Lambda^q E^*$. On appelle produit extérieur de α et β et on note $\alpha \wedge \beta$ la $(p+q)$ -forme extérieure qui vérifie : $\forall x_1, \dots, x_{p+q} \in E$,

$$\alpha \wedge \beta(x_1, \dots, x_{p+q}) = \frac{1}{p!q!} \sum_{\sigma \in \mathfrak{S}_{p+q}} \varepsilon(\sigma) \alpha(x_1, \dots, x_p) \beta(x_{p+1}, \dots, x_{p+q}).$$

Proposition A.6

Soit $\alpha \in \Lambda^p E^*$, $\beta \in \Lambda^q E^*$ et $\gamma \in \Lambda^r E^*$. Alors :

$$\alpha \wedge \beta = (-1)^{pq} \beta \wedge \alpha \quad \text{et} \quad \alpha \wedge (\beta \wedge \gamma) = (\alpha \wedge \beta) \wedge \gamma.$$

Le produit extérieur est donc associatif et "presque" commutatif. Il est également distributif. Le produit extérieur permet de doter d'une structure d'algèbre (graduée) l'ensemble suivant que l'on appelle alors *algèbre extérieure* :

$$\Lambda^\bullet E^* := \bigoplus_{p=0}^{+\infty} \Lambda^p E^*.$$

Théorème A.7 (Structure des formes extérieures)

Soit (e_1, \dots, e_p) une base de E et (dx_1, \dots, dx_n) sa base duale. L'espace vectoriel $\Lambda^p E^*$ est engendré par les éléments de la forme

$$dx_{i_1} \wedge \dots \wedge dx_{i_p}, \quad \text{où} \quad 1 \leq i_1 < \dots < i_p \leq n.$$

Par conséquent :

$$\dim_{\mathbb{R}}(\Lambda^p E^*) = \binom{n}{p}.$$

Exemple : Le déterminant dans la base (e_1, \dots, e_n) s'écrit $\det = dx_1 \wedge \dots \wedge dx_n$.

II. Formes différentielles et courants

Définition A.8. On appelle p -forme différentielle sur E , ou forme différentielle de degré p sur E , toute application continue

$$\omega : E \rightarrow \Lambda^p E^*.$$

L'ensemble des p -formes différentielles extérieures se note $\Omega^p(E)$.

Exemple : • $\Omega^0(E)$ est l'ensemble des fonctions continues de E dans \mathbb{R} ;

- Si f est une fonction de classe \mathcal{C}^1 de E dans \mathbb{R} , alors sa différentielle df s'assimile à un élément de $\Omega^1(E)$.

Remarque : D'après le théorème de structure des formes extérieures, un élément $\omega \in \Omega^p(E)$ peut s'écrire sous la forme

$$\omega(y) = \sum_{i_1 < \dots < i_p} g_{i_1, \dots, i_p}(y) dx_{i_1} \wedge \dots \wedge dx_{i_p} .$$

Dans le cas d'une fonction continument différentiable f on retrouve alors pour la 1-forme différentielle df l'écriture

$$df(y) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(y) dx_i .$$

Définition A.9. Notons $\Omega_0^p(E)$ l'ensemble des p -formes différentielles à support compact. On appelle courant au sens de De Rham, ou p -courant tout élément de l'espace dual

$$\left(\Omega_0^p(E) \right)^* .$$

– Annexe B –

Rappels de théorie des distributions

On rappelle ici les éléments de théorie des distributions essentiels à la compréhension des opérateurs pseudo-différentiels, qui font l'objet de l'annexe suivante.

Par la suite, Ω désignera un ouvert de \mathbb{R}^n où $n \in \mathbb{N}$ est fixé. Classiquement et afin de noter les dérivées partielles d'un élément de $\mathcal{C}^k(\Omega)$ de manière compacte, on introduit la notion de *multi-indices* : un multi-indice α est un n -uplet $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ dont on définit le module et la factorielle comme suit :

$$|\alpha| = \alpha_1 + \dots + \alpha_n \quad ; \quad \alpha! = \alpha_1! \dots \alpha_n! .$$

On note alors, pour tout $x = (x_1, \dots, x_n) \in \mathbb{R}^n$,

$$\partial^\alpha := \partial_{x_1}^{\alpha_1} \dots \partial_{x_n}^{\alpha_n} \quad \text{et} \quad x^\alpha := x_1^{\alpha_1} \dots x_n^{\alpha_n}$$

où $\partial_{x_j} = \partial_j = \frac{\partial}{\partial x_j}$ désigne la dérivée (partielle) par rapport à la variable x_j .

I. Distributions sur un ouvert de \mathbb{R}^n **1. Distributions**

Définition B.1. On appelle distribution toute forme linéaire T continue sur l'espace \mathcal{C}_0^∞ des applications \mathcal{C}^∞ à support compact au sens où :

Pour tout compact K , il existe un entier m et une constante C tels que :

$$\forall \phi \in \mathcal{C}_0^\infty(\Omega) \text{ nulle en dehors de } K, \quad \left| \langle T \mid \phi \rangle \right| \leq C \sup_{x \in K} \sup_{|\alpha| \leq m} |\partial^\alpha \phi(x)| .$$

Le plus petit entier m satisfaisant la définition ci-dessus est appelé *ordre* de la distribution : on dit alors que la distribution est *d'ordre fini* si m est fini et *d'ordre infini* dans le cas contraire. L'espace des distributions sur Ω est noté $\mathcal{D}'(\Omega)$ et contient $\mathbb{L}_{loc}^1(\Omega)$ l'espace des fonctions localement intégrables sur Ω , qui sont d'ordre 0, en identifiant fonction et distribution, *i.e*

$$\forall f \in \mathbb{L}_{loc}^1(\Omega), \quad \forall \phi \in \mathcal{C}_0^\infty(\Omega), \quad \langle f \mid \phi \rangle = \int_{\Omega} f(x)\phi(x) dx .$$

Rappels de théorie des distributions

Dans notre cadre d'application, les masses de Dirac jouent un rôle crucial et sont en particulier des distributions :

$$\forall x_0 \in \Omega, \quad \forall \phi \in \mathcal{C}_0^\infty(\Omega), \quad \langle \delta_{x_0} \mid \phi \rangle = \phi(x_0).$$

On généralise aux distributions la notion de *dérivée* en posant, pour tout $j \in \llbracket 1, n \rrbracket$ et toute distribution $T \in \mathcal{D}'(\Omega)$

$$\forall \phi \in \mathcal{C}_0^\infty(\Omega), \quad \langle \partial_{x_j} T \mid \phi \rangle := - \langle T \mid \partial_{x_j} \phi \rangle.$$

Le $-$ provient de la formule d'intégration par partie pour les fonction $\mathcal{C}^1(\Omega)$ – qui sont en particulier des distributions avec l'identification précédente.

Proposition B.2

Pour tout indice j et toute distribution T , $\partial_{x_j} T$ ainsi défini est une distribution. De plus, elle est d'ordre au plus $m + 1$ si T est d'ordre m .

Corollaire B.2.1

Toute distribution est indéfiniment dérivable.

Une autre opération courante est la *multiplication par une fonction* \mathcal{C}^∞ :

Proposition B.3

Pour toute distribution $T \in \mathcal{D}'(\Omega)$ et toute fonction f de classe \mathcal{C}^∞ , on définit une distribution que l'on note fT en posant :

$$\forall \phi \in \mathcal{C}_0^\infty(\Omega), \quad \langle fT \mid \phi \rangle := \langle T \mid f\phi \rangle.$$

Remarque : Si T est d'ordre m alors, il suffit en fait que f soit de classe \mathcal{C}^m pour pouvoir définir fT .

La dérivation étant une opération linéaire, on retrouve des résultats similaires à ceux des fonctions :

$$\forall f \in \mathcal{C}^\infty(\Omega), \quad \forall T \in \mathcal{D}'(\Omega), \quad \partial(fT) = \partial f T + f \partial T.$$

De plus, tout opérateur différentiel $P := \sum_\alpha a_\alpha \partial^\alpha$ se prolonge en une application linéaire de $\mathcal{D}'(\Omega)$ dans lui-même par la formule :

$$\forall T \in \mathcal{D}'(\Omega), \quad \forall \phi \in \mathcal{C}_0^\infty(\Omega), \quad \langle PT \mid \phi \rangle = \langle T \mid {}^t P \phi \rangle$$

où

$$\forall \phi \in \mathcal{C}_0^\infty(\Omega), \quad {}^t P \phi := \sum_\alpha (-1)^{|\alpha|} \partial^\alpha (a_\alpha \phi).$$

Enfin, on introduit une dernière opération sur les distributions, qui sert néanmoins moins fréquemment. Par soucis de lisibilité, dans toute la fin de ce paragraphe on notera T_f la distribution induite par une fonction mesurable f et non uniquement f . Ainsi, on sera à même de distinguer T'_f qui désigne la dérivée de T_f dans l'espace des distributions et $T_{f'}$ qui est la distribution induite par f' .

Proposition B.4

Soit un intervalle I de \mathbb{R} et $f \in \mathbb{L}^1_{loc}(I)$. Alors les solutions de l'équation $T' = f$ dans $\mathcal{D}'(I)$ sont les distributions T_{F+C} où C est une constante et F la fonction *continue* définie par :

$$\text{pour } a \in I \text{ fixé, } \forall x \in I, \quad F(x) = \int_a^x f(t) dt.$$

En particulier, si $u \in \mathcal{H}^p(I)$, l'espace de Sobolev d'ordre p , et en notant u' sa dérivée faible alors $(T_u)' = T_{u'}$.

Corollaire B.4.1

Si I est un intervalle ouvert de \mathbb{R} et $T \in \mathcal{C}_0^\infty(\Omega)$, alors $T' = \mathbf{0}$ si et seulement si T est constante, au sens où elle est associée à une fonction constante.

Enfin, le résultat suivant est souvent très utile en pratique :

Proposition B.5 (Formule des sauts)

Soit $f: \mathbb{R} \rightarrow \mathbb{C}$ une fonction de classe \mathcal{C}^1 par morceaux, discontinues en a_1, \dots, a_n . Alors, si de plus $f \in \mathbb{L}^1_{loc}(\mathbb{R})$,

$$T'_f = T_{f'} + \sum_{k=1}^n [f(a_k^+) - f(a_k^-)] \delta_{a_k}$$

En particulier, la dérivée de la fonction de Heaviside $H := \mathbf{1}_{\mathbb{R}^+}$ vaut δ_0 .

2. Support

Comme restreindre les distributions revient naturellement à restreindre l'espace des fonctions test, on introduit la définition suivante :

Définition B.6. Soit $T \in \mathcal{D}'(\Omega)$ et $\mathcal{U} \subset \Omega$ un ouvert. On dit que T s'annule sur \mathcal{U} si

$$\forall \phi \in \mathcal{C}_0^\infty(\mathcal{U}), \quad \langle T \mid \phi \rangle = 0.$$

On voudrait alors définir, comme pour les fonctions, la notion de support et de support compact, ce qui est rendu possible par la proposition suivante :

Rappels de théorie des distributions

Proposition B.7

Soit $T \in \mathcal{D}'(\Omega)$. Si T s'annule sur chacun des ouverts de la collection $(\mathcal{U}_j)_{j \in \mathcal{J}}$, alors T s'annule sur $\mathcal{U} = \bigcup_{j \in \mathcal{J}} \mathcal{U}_j$.

Il existe donc pour toute distribution un "plus grand" ouvert de Ω – au sens de l'inclusion – sur lequel elle s'annule et on définit donc :

Définition B.8. On appelle support d'une distribution le complémentaire du plus grand ouvert sur lequel elle s'annule. Une distribution est dite à support compact si son support est compact. L'ensemble des distributions à support compact sur Ω sera noté $\mathcal{E}'(\Omega)$.

Pour des raisons techniques liées à la manipulation des supports, le théorème suivant peut être d'une grande utilité.

Théorème B.9 (*Partition de l'unité*)

Soit $(\Omega_j)_{j \in J}$ une famille d'ouverts de Ω telle que $\bigcup_{j \in J} \Omega_j = \Omega$. Il existe une famille $(\phi_j)_{j \in J}$ de fonctions telle que :

- Pour tout j de J , ϕ_j est à support compact inclus dans Ω_j , i.e $\phi_j \in \mathcal{C}_0^\infty(\Omega)$ avec $\text{supp}(\phi_j) \subset \Omega_j$, et valeurs dans $[0, 1]$;
- Pour tout compact K de Ω , l'ensemble $\{j \in J \mid K \cap \text{supp}(\phi_j) \neq \emptyset\}$ est fini ;
- Pour tout x de Ω , $\sum_{j \in J} \phi_j(x) = 1$.

3. Noyaux

Soit Ω_1 et Ω_2 deux ouverts de \mathbb{R}^n et soit $K \in \mathcal{D}'(\Omega_1 \times \Omega_2)$. La relation

$$\forall \phi_1 \in \mathcal{C}_0^\infty(\Omega_1), \quad \forall \phi_2 \in \mathcal{C}_0^\infty(\Omega_2), \quad \left\langle A_K \phi_2 \mid \phi_1 \right\rangle = \left\langle K \mid \phi_1 \otimes \phi_2 \right\rangle \quad (\star)$$

où $\phi_1 \otimes \phi_2(x_1, x_2) := \phi_1(x_1)\phi_2(x_2)$ désigne le produit tensoriel de ϕ_1 et ϕ_2 , définit une application linéaire $A_K : \mathcal{C}_0^\infty(\Omega_2) \rightarrow \mathcal{D}'(\Omega_1)$ continue au sens où :

Pour toute fonction test $\phi_1 \in \mathcal{C}_0^\infty(\Omega_1)$ et tout compact K de Ω_2 , il existe une constante C et un entier m tel que : pour toute fonction test $\phi_2 \in \mathcal{C}_0^\infty(\Omega_2)$ supportée dans K ,

$$\left| \left\langle A_K \phi_2 \mid \phi_1 \right\rangle \right| \leq C \sup_{x \in K} \sup_{|\alpha| \leq m} |\partial^\alpha \phi_2(x)| .$$

Remarque : Lorsque $K \in \mathbb{L}_{loc}^1(\Omega_1 \times \Omega_2)$, la relation (\star) s'écrit plus simplement :

$$A_K \phi_2(x_1) = \int_{\Omega_2} K(x_1, x_2) \phi_2(x_2) dx_2.$$

Définition B.10. On appelle K noyau de l'opérateur A_K .

On peut montrer que K est entièrement déterminé par A_K . Le théorème suivant (dû à Laurent Schwartz) nous assure de l'existence d'un noyau pour tout opérateur dont la régularité est suffisante.

Théorème B.11 (de Schwartz)

Tout opérateur $A: \mathcal{C}_0^\infty(\Omega_2) \rightarrow \mathcal{D}'(\Omega_1)$ continu au sens décrit ci-dessus admet un noyau.

Exemple : • Le noyau de l'opérateur différentiel $P = \sum_{\alpha} a_{\alpha}(x) \partial^{\alpha}$ est donné par :

$$\forall (x_1, x_2) \in \Omega_1 \times \Omega_2, \quad K(x_1, x_2) = \sum_{\alpha} a_{\alpha}(x_1) \partial^{\alpha} \delta_0(x_1 - x_2).$$

• En anticipant sur ce qui suit, le noyau de l'opérateur de convolution par la distribution $T \in \mathcal{D}'(\mathbb{R}^n)$ est la distribution K définie par :

$$\forall (x_1, x_2) \in \Omega_1 \times \Omega_2, \quad K(x_1, x_2) = T(x_1 - x_2).$$

II. Convolution

1. Convolution avec une fonction

Soit $f: \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continue et $\phi \in \mathcal{C}_0^\infty(\Omega)$ une fonction test. La fonction $x \mapsto f(x)\phi(y-x)$ est à support compact, donc intégrable sur \mathbb{R}^n et on peut donc définir :

Définition B.12. Soit $f: \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continue et soit $\phi \in \mathcal{C}_0^\infty(\Omega)$, on appelle produit de convolution de f et ϕ et on note $f * \phi$ la fonction,

$$f * \phi : y \mapsto \int_{\mathbb{R}^n} f(x) \phi(y-x) dx.$$

Proposition B.13

On a $f * \phi = \phi * f$ et pour toute fonction test $\psi \in \mathcal{C}_0^\infty(\Omega)$, $\phi * \psi \in \mathcal{C}_0^\infty(\Omega)$ et

$$(f * \phi) * \psi = f * (\phi * \psi).$$

Rappels de théorie des distributions

La convolution peut être efficacement étendue aux fonctions \mathbb{L}^1 , mais l'extension qui nous intéresse est la suivante :

Définition B.14. Soit $T \in \mathcal{D}'(\Omega)$ et $\phi \in \mathcal{C}_0^\infty(\Omega)$. On appelle produit de convolution de T et ϕ , et on note $T * \phi$, la fonction

$$T * \phi : y \mapsto \left\langle T \mid \phi(y - \cdot) \right\rangle .$$

Proposition B.15

Soit $T \in \mathcal{D}'(\Omega)$ et $\phi \in \mathcal{C}_0^\infty(\Omega)$. La fonction $T * \phi$ est de classe \mathcal{C}^∞ sur Ω et pour tout multi-indice α , on a

$$\partial^\alpha(T * \phi) = (\partial^\alpha T) * \phi = T * (\partial^\alpha \phi) .$$

Le fait que la convolution avec une fonction de $\mathcal{C}_0^\infty(\Omega)$ transforme un objet potentiellement très irrégulier comme une distribution en une fonction de classe \mathcal{C}^∞ justifie l'utilisation de la convolution comme outil "régularisant" :

Théorème B.16

Soit $\phi \in \mathcal{C}_0^\infty(\Omega)$ positive et d'intégrale 1. On pose pour tout $\varepsilon \in \mathbb{R}_+^*$,

$$\phi_\varepsilon(x) := \varepsilon^{-n} \phi\left(\frac{x}{\varepsilon}\right) .$$

Alors, pour toute distribution $T \in \mathcal{D}'(\Omega)$, si l'on pose $T_\varepsilon := T * \phi_\varepsilon$, on a :

$$\forall \psi \in \mathcal{C}_0^\infty(\Omega), \quad \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^n} T_\varepsilon(x) \psi(x) dx = \left\langle T \mid \psi \right\rangle .$$

2. Convolution de deux distributions

Nous allons à présent étendre le produit de convolution au cas de deux distributions, mais, comme nous le constaterons, il y a alors une restriction. Pour plus de lisibilité nous noterons, pour toute fonction f , \check{f} la fonction $x \mapsto f(-x)$. De même, pour une distribution $T \in \mathcal{D}'(\Omega)$, nous noterons \check{T} la distribution telle que

$$\forall \phi \in \mathcal{C}_0^\infty(\Omega), \quad \left\langle \check{T} \mid \phi \right\rangle = \left\langle T \mid \check{\phi} \right\rangle .$$

Définition B.17. Soit $T \in \mathcal{D}'(\Omega)$ et $S \in \mathcal{E}'(\Omega)$, on appelle produit de convolution de T et S et on note $T * S$ la distribution qui vérifie

$$\forall \phi \in \mathcal{C}_0^\infty(\Omega), \quad \left\langle T * S \mid \phi \right\rangle = \left\langle T \mid \check{S} * \phi \right\rangle .$$

Remarque : Du fait que l'on puisse étendre l'ensemble de définition d'une distribution à support compact à $\mathcal{C}^\infty(\Omega)$, écrire $\langle S \mid \check{T} * \phi \rangle$ a un sens pour $S \in \mathcal{E}'(\Omega)$, ce qui permet de définir le produit de convolution $S * T$ de S et de T .

Théorème B.18

Soit $T \in \mathcal{D}'(\Omega)$, $S \in \mathcal{E}'(\Omega)$ et α un multi-indice. On a

$$T * S = S * T$$

ce qui revient à dire que la convolution des distributions est commutative, et

$$\partial^\alpha(T * S) = (\partial^\alpha T) * S = T * (\partial^\alpha S).$$

Théorème B.19

Soit $T \in \mathcal{D}'(\Omega)$ et $R, S \in \mathcal{E}'(\Omega)$. La convolution de R, S et T est associative au sens où

$$R * (S * T) = (R * S) * T.$$

Remarque : On peut donc définir la convolution $R * S * T$, et plus généralement pour $S_1, \dots, S_k \in \mathcal{E}'(\Omega)$ et $T \in \mathcal{D}'(\Omega)$, on peut écrire sans ambiguïté

$$S_1 * \dots * S_k * T.$$

Il faut cependant se méfier de cette associativité, parce que l'hypothèse d'avoir au plus une distribution qui n'est pas à support compact est fondamentale, comme le montre l'exemple suivant :

Exemple : Soit $H = \mathbf{1}_{\mathbb{R}^+}$ la fonction de Heaviside, δ la distribution de Dirac au point 0 et $\mathbf{1}$ la fonction constante égale à 1. On a $H, \mathbf{1} \in \mathcal{D}'(\mathbb{R})$ et $\delta \in \mathcal{E}'(\mathbb{R})$. De plus, δ' , la dérivée de δ est également dans $\mathcal{E}'(\mathbb{R})$ et $H' = \delta$. Cependant,

$$(\mathbf{1} * \delta') * H = (\mathbf{1}' * \delta) * H = \mathbf{0} * H = \mathbf{0}$$

et

$$\mathbf{1} * (\delta' * H) = \mathbf{1} * (\delta * H') = \mathbf{1} * (\delta * \delta) = \mathbf{1} * \delta = \mathbf{1}.$$

3. Support d'un produit de convolution

Proposition B.20

Soit $T \in \mathcal{D}'(\Omega)$ et $S \in \mathcal{E}'(\Omega)$. On a :

$$\text{supp}(T * S) \subset \text{supp}(T) + \text{supp}(S)$$

Rappels de théorie des distributions

La convolution respecte également un autre type de support, le *support singulier*.

Définition B.21. Soit $T \in \mathcal{D}'(\Omega)$, on appelle support singulier de T et on note $\text{Suppsing}(T)$ l'ensemble défini par

$$\text{Suppsing}(T) = \Omega \setminus \left\{ x \in \Omega \mid \exists \mathcal{U} \subset \Omega, \text{ ouvert}, x \in \mathcal{U}, T|_{\mathcal{U}} \in \mathcal{C}^\infty(\mathcal{U}) \right\}.$$

Le support singulier est le complémentaire du plus grand ouvert sur lequel la restriction de la distribution est une fonction indéfiniment différentiable.

Proposition B.22

Soit $T \in \mathcal{D}'(\Omega)$ et $S \in \mathcal{E}'(\Omega)$. Alors :

$$\text{Suppsing}(T * S) \subset \text{Suppsing}(T) + \text{Suppsing}(S).$$

III. Transformation de Fourier

En s'inspirant de ce qu'on a fait pour la dérivation et la convolution, on aimerait pouvoir définir la transformation de Fourier d'une distribution par une formule du type

$$\forall \phi \in \mathcal{C}_0^\infty(\Omega), \quad \langle \mathcal{F}T \mid \phi \rangle = \langle T \mid \mathcal{F}\phi \rangle$$

éventuellement avec un changement de signe. Mais, une telle définition ne peut convenir car, en règle générale, $\mathcal{F}\phi \notin \mathcal{C}_0^\infty(\Omega)$. On va donc construire un espace que l'on note $\mathcal{S}(\Omega)$ et tel que $\phi \in \mathcal{S}(\Omega) \implies \mathcal{F}\phi \in \mathcal{S}(\Omega)$.

Remarque : L'espace $\mathbb{L}^2(\mathbb{R})$ vérifie en particulier cette propriété, la transformée de Fourier étant une isométrie sur $\mathbb{L}^2(\mathbb{R})$. Cependant, $\mathbb{L}^2(\mathbb{R})$ n'est pas stable par passage à la dérivation, d'où la nécessité de se munir d'un espace plus restrictif.

1. Espace de Schwartz

Définition B.23. On note $\mathcal{S}(\mathbb{R}^n)$ l'ensemble des fonctions $f: \mathbb{R}^n \rightarrow \mathbb{C}$ de classe \mathcal{C}^∞ sur \mathbb{R}^n et telles que

$$\forall \alpha, \beta \in \mathbb{N}^n, \quad \sup_{x \in \mathbb{R}^n} |x^\alpha \partial^\beta f| < +\infty.$$

On dit que $\mathcal{S}(\mathbb{R}^n)$ est l'espace de Schwartz des fonctions indéfiniment dérivables à croissance rapide.

On munit l'espace $\mathcal{S}(\mathbb{R}^n)$ des semi-normes ainsi définies lorsque α et β parcourt \mathbb{N}^n :

$$\forall f \in \mathcal{S}(\mathbb{R}^n), \quad N_{\alpha,\beta}(f) := \sup_{x \in \mathbb{R}^n} |x^\alpha \partial^\beta f|.$$

Définition B.24. Soit $f \in \mathcal{S}(\mathbb{R}^n)$. On appelle transformée de Fourier de f l'application \hat{f} définie par

$$\forall \xi \in \mathbb{R}^n, \quad \hat{f}(\xi) := \int_{\Omega} e^{-ix \cdot \xi} f(x) dx$$

où \cdot désigne le produit scalaire usuel dans \mathbb{R}^n et transformation de Fourier l'application (linéaire) suivante :

$$\begin{aligned} \mathcal{F} : \mathcal{S}(\mathbb{R}^n) &\longrightarrow \mathcal{S}(\mathbb{R}^n) \\ f &\longmapsto \hat{f} \end{aligned}.$$

De manière évidente, pour toute fonction f et g de l'espace de Schwartz $\mathcal{S}(\mathbb{R})$ et tout multi-indice $\alpha \in \mathbb{N}^n$, $\partial^\alpha f \in \mathcal{S}(\mathbb{R})$ et $x^\alpha f \in \mathcal{S}(\mathbb{R})$ d'une part et $f \times g \in \mathcal{S}(\mathbb{R})$ et $f \star g \in \mathcal{S}(\mathbb{R})$ d'autre part. De plus, on a les propriétés fondamentales suivantes :

Proposition B.25

L'espace de Schwartz vérifie :

- i. $\mathcal{S}(\mathbb{R}) \subset L^1(\mathbb{R})$ et $f \in \mathcal{S}(\mathbb{R}) \implies \mathcal{F}f \in \mathcal{S}(\mathbb{R})$;
- ii. \mathcal{F} réalise une bijection de $\mathcal{S}(\mathbb{R})$ sur lui-même ;
- iii. Pour tout $f, g \in \mathcal{S}(\mathbb{R})$,

$$\mathcal{F}(f \star g) = (\mathcal{F}f) \times (\mathcal{F}g) \quad \text{et} \quad \mathcal{F}(f \times g) = (\mathcal{F}f) \star (\mathcal{F}g) ;$$

- iv. Pour tout $f \in \mathcal{S}(\mathbb{R})$, $j \in \llbracket 1, n \rrbracket$ et $\xi \in \mathbb{R}^n$,

$$\widehat{\partial_j f}(\xi) = i \xi_j \hat{f}(\xi) \quad \text{et} \quad \widehat{x_j f}(\xi) = i \partial_j \hat{f}(\xi) ;$$

et en notant pour tout $x, y \in \mathbb{R}^n$, $\tau_y f(x) := f(x + y)$

$$\widehat{\tau_y f}(\xi) = e^{iy \cdot \xi} \hat{f}(\xi) \quad \text{et} \quad \widehat{e^{-x \cdot y} f}(\xi) = \tau_y \hat{f}(\xi).$$

Enfin, pour les semi-normes définies ci-dessus, on a le résultat suivant :

Proposition B.26

La transformation de Fourier est continue sur $\mathcal{S}(\mathbb{R})$.

2. Distributions tempérées

Définition B.27. Une forme linéaire $T: \mathcal{C}_0^\infty(\mathbb{R}^n) \rightarrow \mathbb{C}$ est une distribution tempérée s'il existe une constante $C > 0$ et un entier $m \in \mathbb{N}$ tels que :

$$\forall \phi \in \mathcal{C}_0^\infty(\mathbb{R}^n), \quad \left| \langle T \mid \phi \rangle \right| \leq \sup \left\{ N_{\alpha, \beta}(\phi) \mid \alpha, \beta \in \mathbb{N}^n \text{ et } |\alpha|, |\beta| \leq m \right\}.$$

Une telle appellation est notamment justifiée par la proposition qui suit.

Proposition B.28

Toute distribution tempérée est une distribution sur \mathbb{R}^n et est, de plus, d'ordre fini.

Communément, on définit plutôt les distributions tempérées via la proposition B.29 et par analogie avec la notation \mathcal{D}' et \mathcal{E}' , on note $\mathcal{S}'(\mathbb{R}^n)$ l'espace des distributions tempérées. De plus, $\mathcal{E}'(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n)$ et $\mathcal{S}(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n)$ en faisant agir $\mathcal{S}(\mathbb{R}^n)$ sur lui-même par la formule

$$\forall \phi, \psi \in \mathcal{S}(\mathbb{R}^n), \quad \langle \phi \mid \psi \rangle = \int_{\mathbb{R}^n} \phi(x) \psi(x) dx.$$

En fait, on peut même montrer que $\mathcal{S}(\mathbb{R}^n)$ est dense dans $\mathcal{S}'(\mathbb{R}^n)$.

Proposition B.29

Une forme linéaire $T: \mathcal{C}_0^\infty(\mathbb{R}^n) \rightarrow \mathbb{C}$ est une distribution tempérée si et seulement si elle se prolonge en une forme linéaire continue $\tilde{T}: \mathcal{S}(\mathbb{R}^n) \rightarrow \mathbb{C}$.

Corollaire B.29.1

La dérivée de toute distribution tempérée est encore une distribution tempérée.

Une classe d'exemple de distributions tempérées est donnée par les fonctions à croissance lente :

Définition B.30. Une fonction $\phi \in \mathcal{C}^\infty(\mathbb{R}^n)$ est dite à croissance lente si

$$\forall \alpha \in \mathbb{N}^n, \exists m \in \mathbb{N}, \exists C > 0, \quad \forall x \in \mathbb{R}^n, \quad |\partial^\alpha \phi(x)| \leq C (1 + |x|)^m.$$

Exemple : Les fonctions polynomiales sont \mathcal{C}^∞ à croissance lente, de même pour les fonctions $\mathbb{L}^p(\mathbb{R}^n)$ avec $1 \leq p \leq +\infty$.

Proposition B.31

Toute fonction à croissance lente définit une distribution tempérée.

D'après le théorème de Fubini, on a

$$\forall \phi, \psi \in \mathcal{S}(\mathbb{R}^n), \quad \int_{\mathbb{R}^n} \hat{\phi}(\xi) \psi(\xi) \, d\xi = \int_{\mathbb{R}^n} \phi(\xi) \hat{\psi}(\xi) \, d\xi$$

et on définit donc la transformée de Fourier sur $\mathcal{S}'(\mathbb{R}^n)$ en prolongeant, de manière continue, l'application $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathcal{S}'(\mathbb{R}^n)$ à $\mathcal{F} : \mathcal{S}'(\mathbb{R}^n) \rightarrow \mathcal{S}'(\mathbb{R}^n)$.

Définition B.32. Pour toute distribution tempérée T , on définit sa transformée de Fourier $\mathcal{F}T$ comme la distribution tempérée définie par

$$\forall \phi \in \mathcal{S}(\mathbb{R}^n), \quad \langle \mathcal{F}T \mid \phi \rangle := \langle T \mid \mathcal{F}\phi \rangle.$$

On peut vérifier que $\mathcal{F}T$ ainsi définie est bien une distribution tempérée.

Commençons par remarquer que $\hat{\delta} = \mathbf{1}$ et donc, d'après B.25.iv., $\hat{\mathbf{1}} = c\delta$ pour un certain $c \in \mathbb{C}$. D'où, pour toute fonction $\phi \in \mathcal{S}(\mathbb{R}^n)$ de l'espace de Schwartz :

$$\hat{\phi}(0) = \langle \delta \mid \hat{\phi} \rangle = \langle \mathbf{1} \mid \hat{\phi} \rangle = \langle \mathbf{1} \mid \phi \rangle = c\phi(0)$$

en effectuant le calcul dans un cas particulier – typiquement la gaussienne $\phi(x) = e^{-|x|^2/2}$ – on montre que $c = (2\pi)^n$ et avec B.25.iv. il vient que :

$$\forall x \in \mathbb{R}^n, \quad \hat{\phi}(x) = (2\pi)^n \phi(-x).$$

Autrement dit :

$$\forall x \in \mathbb{R}^n, \quad \phi(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{ix \cdot \xi} \hat{\phi}(\xi) \, d\xi := \check{\mathcal{F}}\phi(x)$$

et on a en fait montré la formule d'inversion de Fourier sur $\mathcal{S}(\mathbb{R}^n)$.

Théorème B.33 (Formule d'inversion de Fourier)

La transformée de Fourier \mathcal{F} est un isomorphisme de $\mathcal{S}(\mathbb{R}^n)$ sur $\mathcal{S}(\mathbb{R}^n)$ et également de $\mathcal{S}'(\mathbb{R}^n)$ sur $\mathcal{S}'(\mathbb{R}^n)$. Son inverse est donné par $(2\pi)^{-n} \check{\mathcal{F}}$ où $\check{\mathcal{F}}$ est la composée de \mathcal{F} et de la relation symétrie $T \mapsto \check{T}$ définie sur $\mathcal{S}'(\mathbb{R}^n)$ par :

$$\forall \phi \in \mathcal{S}(\mathbb{R}^n), \quad \langle \check{T} \mid \phi \rangle = \langle T \mid \check{\phi} \rangle.$$

Autrement dit, pour toute distribution tempérée $T \in \mathcal{S}'(\mathbb{R}^n)$:

$$\forall \phi \in \mathcal{S}(\mathbb{R}^n), \quad \langle \check{\mathcal{F}}T \mid \phi \rangle = \langle T \mid \check{\mathcal{F}}\phi \rangle.$$

Rappels de théorie des distributions

Enfin, on conclut en énonçant les résultats suivants à mettre en parallèle avec le point iii. de la proposition B.25 :

i. Pour toute distribution $T \in \mathcal{S}'(\mathbb{R}^n)$ et $R \in \mathcal{E}'(\mathbb{R}^n)$:

$$\mathcal{F}(T \star R) = (\mathcal{F}T) \times (\mathcal{F}R)$$

ii. Pour toute distribution T_1 et $T_2 \in \mathcal{S}'(\mathbb{R}^n)$:

$$\mathcal{F}(T_1 \times T_2) = (2\pi)^{-n} \mathcal{F}T_1 \star \mathcal{F}T_2$$

et le théorème de Plancherel. Pour cela, on munit $\mathbb{L}^2(\mathbb{R}^n)$ de son produit scalaire usuel :

$$\forall f, g \in \mathbb{L}^2(\mathbb{R}^n), \quad (f, g) := \int_{\mathbb{R}^n} f(x) \bar{g}(x) dx.$$

Théorème B.34 (Théorème de Plancherel)

Pour toute fonction $f, g \in \mathbb{L}^2(\mathbb{R}^n)$,

$$(\hat{f}, \hat{g}) = (2\pi)^n (f, g).$$

Autrement dit, \mathcal{F} se prolonge en un automorphisme de $\mathbb{L}^2(\mathbb{R}^n)$ tel que $(2\pi)^{-n/2} \mathcal{F}$ soit un unitaire.

IV. Solutions fondamentales

1. La distribution de Dirac

Nous avons déjà défini et utilisé dans les exemples la distribution de Dirac δ , qui est d'ordre 0 et à support compact, de support et support singulier réduit à $\{0\}$, et définie par la formule

$$\forall \phi \in \mathcal{C}_0^\infty(\mathbb{R}^n), \quad \langle \delta \mid \phi \rangle = \phi(0).$$

Nous allons voir qu'elle est utile pour résoudre des équations aux dérivées partielles. Pour cela, nous avons besoin des propriétés suivantes :

Proposition B.35

Soit $T \in \mathcal{D}'(\mathbb{R}^n)$. On a

$$T \star \delta = \delta \star T = T.$$

Et, comme nous l'avons déjà remarqué :

Proposition B.36

La distribution de Dirac est tempérée et elle vérifie,

$$\hat{\delta} = \mathbf{1} \quad \text{et} \quad \hat{\mathbf{1}} = (2\pi)^n \delta.$$

Démonstration :

La distribution de Dirac est à support compact, donc elle est tempérée. Pour toute fonction $f \in \mathcal{S}(\mathbb{R}^n)$,

$$\langle \hat{\delta} \mid f \rangle = \langle \delta \mid \hat{f} \rangle = \hat{f}(0) = \int_{\mathbb{R}^n} f(x) \, dx = \langle \mathbf{1} \mid f \rangle.$$

Donc $\hat{\delta} = \mathbf{1}$. De plus, pour toute fonction $f \in \mathcal{S}(\mathbb{R}^n)$,

$$\begin{aligned} \langle \hat{\mathbf{1}} \mid f \rangle &= \langle \mathbf{1} \mid \hat{f} \rangle = \int_{\mathbb{R}^n} \hat{f}(x) \, dx \\ &= (2\pi)^n f(0) \quad \text{par la formule de transformée inverse} \\ &= (2\pi)^n \langle \delta \mid f \rangle. \end{aligned}$$

D'où $\hat{\mathbf{1}} = (2\pi)^n \delta$. □

On constate que la distribution la "plus concentrée" qui soit, *i.e* la mesure de Dirac, est envoyée par la transformée de Fourier sur la fonction constante égale à 1, c'est à dire la "plus étalée" que l'on puisse imaginer. Cet exemple traduit un fait plus général : il est impossible de concentrer à la fois la masse d'une distribution et celle de sa transformée de Fourier. C'est cette idée qui conduit en physique quantique au principe d'incertitude de Heisenberg-Gabor.

2. Existence et régularité des solutions

Définition B.37. Soit P un opérateur différentiel à coefficients constants sur \mathbb{R}^n . Une solution fondamentale (aussi appelée solution élémentaire ou fonction de Green) de P est une distribution $E \in \mathcal{D}'(\mathbb{R}^n)$ telle que

$$PE = \delta.$$

Exemple : L'opérateur différentiel

$$P = \frac{\partial}{\partial t} - \Delta,$$

associé à l'équation de la chaleur admet pour solution fondamentale sur $\mathbb{R}_+ \times \mathbb{R}^n$ la fonction

$$E : (t, x) \mapsto H(t) \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{\|x\|^2}{4t}},$$

où H est toujours la fonction de Heaviside.

Rappels de théorie des distributions

Théorème B.38

Soit P un opérateur différentiel à coefficients constants qui admet une solution fondamentale E . Soit $S \in \mathcal{E}'(\mathbb{R}^n)$. La distribution $E * S$ est solution de l'équation

$$PT = S$$

Démonstration :

Posons $T = E * S$. E est solution fondamentale, donc $PE = \delta$. On a alors,

$$PT = P(E * S) = (PE) * S = \delta * S = S$$

d'après les propriétés de la convolution et de la distribution de Dirac. \square

Enfin, on donne la définition de *paramétrix* dont on se sert à l'annexe C et dans la démonstration du théorème de Richardson, [33].

Définition B.39. Soit $P(x)$ un opérateur différentiel à coefficients variables. On appelle paramétrix une distribution E telle que

$$P(x)E(x) = \delta + \omega(x) \quad \text{où} \quad \omega \in \mathcal{C}_0^\infty(\mathbb{R}^n).$$

Proposition B.40

Tout opérateur elliptique de degré 2 à coefficients \mathcal{C}^∞ admet une paramétrix.

Exemple : Soit l'opérateur laplacien $P = \Delta = \partial_1^2 + \dots + \partial_n^2$, qui est bien évidemment elliptique de degré 2. Il suffit pour cela de définir E en posant

$$\forall \xi \in \mathbb{R}^n, \quad \hat{E}(\xi) = -\frac{1 - \chi(\xi)}{|\xi|^2}$$

où χ est une fonction plateau : $\chi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ et $\chi \equiv 1$ au voisinage de $\xi = 0$.
En effet, dans ce cas,

$$\forall \xi \in \mathbb{R}^n, \quad \widehat{PE}(\xi) = -|\xi|^2 \hat{E}(\xi) = 1 - \chi(\xi).$$

Or, on a vu que $\hat{\delta} = \mathbf{1}$ à la proposition B.36.

Le théorème B.38 devient faux dans le cas des paramétrix néanmoins, les paramétrix sont des "solutions approchées" de telles équations : Soit $S \in \mathcal{E}'(\mathbb{R}^n)$ Alors :

$$P(E * S) = (PE) * S = S + \omega * S$$

et la distribution $E * S$ est donc une solution approchée de l'équation $PT = S$ avec une erreur $\omega * S$ qui est de classe \mathcal{C}^∞ .

– Annexe C –

Opérateurs pseudo-différentiels

Dans ce chapitre, on introduit la notion d'opérateurs pseudo-différentiels mais sans entrer dans les détails de la théorie, le but étant "simplement" de comprendre l'article de Richardson et Younes [33] portant sur les métamorphoses de mesures discrètes (cf paragraphe II.2 du chapitre 1) et basé sur cette notion.

Pour aller plus loin, on peut consulter le chapitre 18 du tome III de l'ouvrage de Lars Hörmander [21] sur le sujet ou le chapitre 1 du livre de Patrick Gérard et Serge Alinhac [2] essentiellement éponyme et peut-être plus adapté.

Soit P un opérateur différentiel : $P(x) = \sum_{\alpha} a_{\alpha}(x) \partial_x^{\alpha}$ où, pour tout α , a_{α} vit dans l'espace de Schwartz $\mathcal{S}(\mathbb{R}^n)$. Alors, pour toute distribution u ,

$$\forall x, \xi \in \mathbb{R}^n, \quad \widehat{Pu(x)}(\xi) = p(x, \xi) \hat{u}(\xi) \quad \text{où} \quad p(x, \xi) := \sum_{\alpha} a_{\alpha}(x) \xi^{\alpha}$$

et la formule d'inversion de Fourier donne :

$$\forall x \in \mathbb{R}^n, \quad Pu(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\langle x, \xi \rangle} p(x, \xi) \hat{u}(\xi) d\xi. \quad (\star)$$

Or, toujours d'après la formule d'inversion de Fourier,

$$\forall x \in \mathbb{R}^n, \quad u(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\langle x, \xi \rangle} \hat{u}(\xi) d\xi$$

et l'action de l'opérateur différentiel $P(x)$ sur la distribution u "est" de multiplier la quantité $e^{i\langle x, \xi \rangle} \hat{u}(\xi)$ par $p(x, \xi)$ dans l'intégrale ci-dessus. Plus précisément, le terme fréquentiel $e^{i\langle x, \xi \rangle}$ reste inchangé, tandis que l'amplitude $\hat{u}(\xi)$ est multipliée par $p(x, \xi)$. La quantité $p(x, \xi)$ joue donc un rôle prépondérant dans l'étude des opérateurs différentiels et c'est lui qui va faire porter la dérivation sur u en "modulant" son amplitude : on l'appelle *symbole* de l'opérateur $P(x)$.

I. Symbole d'un opérateur

L'idée des opérateurs pseudo-différentiels est de partir de la formule (\star) mais pour une fonction $p(x, \xi)$ non-nécessairement polynomiale : il suffit pour cela de considérer une fonction $a(x, \xi)$ "raisonnable" auquel correspondra un opérateur que l'on notera $A(x, D)$. Ainsi, on généralise la notion d'opérateurs différentiels en s'autorisant à considérer des dérivées non-entières, l'effet "dérivant" de l'opérateur $A(x, D)$ étant induit par $a(x, \xi)$ qui sera le plus souvent non-polynomial.

1. Définitions et exemples

Définition C.1. Soit $m \in \mathbb{R}$. On note $S^m = S^m(\mathbb{R}^n \times \mathbb{R}^n)$ l'ensemble des fonctions $a \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n)$ telles que :

$$\forall \alpha, \beta \in \mathbb{N}^n, \quad \exists C \in \mathbb{R}, \quad \left| \partial_x^\alpha \partial_\xi^\beta a(x, \xi) \right| \leq C(1 + |\xi|)^{m-|\beta|}.$$

On note de plus $S^{-\infty} = \bigcap_m S^m$ et un élément $a \in S^m$ est appelé symbole d'ordre m .

Par "raisonnable", on entend donc que :

- i. $a(x, \xi)$ doit avoir un comportement de type polynomial en ξ , lorsque $|\xi| \rightarrow +\infty$, et s'améliorant par dérivation en ξ : autrement dit,

$$\forall \beta \in \mathbb{N}^n, \quad \exists C \in \mathbb{R}, \quad \left| \partial_\xi^\beta a(x, \xi) \right| \leq C(1 + |\xi|)^{m-|\beta|};$$

- ii. la variation en x de $a(x, \xi)$ doit rester faible afin de garantir que seule l'amplitude soit modifiée dans l'intégrale (\star). Plus précisément, on demande :

$$\forall \alpha \in \mathbb{N}^n, \quad \exists C \in \mathbb{R}, \quad \left| \partial_x^\alpha a(x, \xi) \right| \leq C(1 + |\xi|)^m.$$

Exemple : • Si $a(x, \xi) = \sum_{|\alpha| \leq m} a_\alpha(x) \xi^\alpha$, avec $a_\alpha \in C^\infty(\mathbb{R}^n)$, bornée ainsi que toutes ses dérivées, alors $a \in S^m$ et on dit que a est un *symbole différentiel*.

- Au contraire, la fonction $a(x, \xi) = e^{i\langle x, \xi \rangle}$ n'est pas un symbole.

De manière évidente, on a :

- i. Si $a \in S^m$ alors, pour tout $\alpha, \beta \in \mathbb{N}^n$, $\partial_x^\alpha \partial_\xi^\beta a \in S^{m-|\beta|}$;
 ii. Si $a \in S^m$ et $b \in S^{m'}$ alors $ab \in S^{m+m'}$;
 iii. Si $a \in S^m$ alors $a \in \mathcal{S}'(\mathbb{R}^n)$.

On munit S^m des semi-normes définies par : pour tout symbole a ,

$$\forall \alpha, \beta \in \mathbb{N}^n, \quad |a|_{\alpha, \beta}^m := \sup \left\{ (1 + |\xi|)^{-(m+|\beta|)} \left| \partial_x^\alpha \partial_\xi^\beta a(x, \xi) \right| \mid x, \xi \in \mathbb{R}^n \right\}$$

ce qui en fait un espace de Fréchet. Pour la topologie induite par ces semi-normes, la dérivation (point i. ci-dessus) et la multiplication (ii.) sont continues : les semi-normes sont, de fait, construites de telle manière à être compatibles avec la structure de S^m .

On dispose enfin d'un lemme d'approximation suivant, utile en pratique. Il permet en effet d'approcher un symbole par une suite de symboles.

Lemme C.2.1. Soit $a \in S^0$. Posons pour tout $\varepsilon > 0$, $a_\varepsilon(x, \xi) := a(x, \varepsilon\xi)$. Alors :

- i. Pour tout $\varepsilon > 0$, a_ε est borné dans S^0 ;
 ii. Pour tout $m > 0$, $a_\varepsilon \rightarrow a_0$ dans S^m lorsque $\varepsilon \rightarrow 0$ au sens où

$$\forall \alpha, \beta \in \mathbb{N}^n, \quad |a_\varepsilon - a_0|_{\alpha, \beta}^m \xrightarrow{\varepsilon \rightarrow 0} 0.$$

2. Sommes asymptotiques et symboles classiques

Soit une suite décroissante $m_j \rightarrow -\infty$ et une famille de symboles $a_j \in S^{m_j}$ pour $j \in \mathbb{N}$. On a envie de pouvoir définir $\sum_j a_j$, dans un sens à définir : on ne peut espérer en effet que la série soit convergente. On introduit donc la notion de *somme asymptotique*.

Définition C.2. Soit $a \in S^m$. On dit que la suite $(a_j)_{j \in \mathbb{N}}$ est une somme asymptotique de a et on écrit $a \sim \sum a_j$ si :

$$\forall k \in \mathbb{N}, \quad a - \sum_{j=0}^{k-1} a_j \in S^{m_k}.$$

Remarque : L'appellation asymptotique tient du fait que la contrainte est d'autant plus forte que $|\xi| \rightarrow +\infty$ puisque la suite $m_j \rightarrow -\infty$.

Proposition C.3

Pour toute telle suite a_j , il existe $a \in S^{m_0}$ tel que $a \sim \sum a_j$.
On peut de plus imposer que $\text{supp}(a) \subset \bigcup_j \text{supp}(a_j)$.

Définition C.4. Un symbole $a \in S^m$ est dit classique si $a \sim \sum_j a_j$ et que les a_j sont des fonctions homogènes de degré $m - j$ pour $|\xi| \geq 1$, i.e

$$\forall \lambda \geq 1, \quad \forall |\xi| \geq 1, \quad a_j(x, \lambda\xi) = \lambda^{m-j} a_j(x, \xi).$$

On note alors $a \in S_{cl}^m$.

Exemple : Les symboles différentiels sont en particuliers classiques.

II. Opérateurs pseudo-différentiels

1. Définition et exemple

Proposition C.5

Soit $a \in S^m$ et $u \in \mathcal{S}(\mathbb{R}^n)$. Alors la formule

$$\forall x \in \mathbb{R}^n, \quad \mathbf{Op}(a)u(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\langle x, \xi \rangle} a(x, \xi) \hat{u}(\xi) d\xi$$

définit une fonction de $\mathcal{S}(\mathbb{R}^n)$ et l'application

$$(a, u) \in S^m \times \mathcal{S}(\mathbb{R}^n) \mapsto \mathbf{Op}(a)u \in \mathcal{S}(\mathbb{R}^n)$$

Opérateurs pseudos-différentiels

est continue. L'application linéaire \mathbf{Op} ainsi définie de S^m dans les opérateurs linéaires de $\mathcal{S}(\mathbb{R}^n)$ est injective et satisfait aux relations :

$$\begin{cases} [Op(a), D_j] &= i \mathbf{Op}(\partial_{x_j} a) \\ [Op(a), x_j] &= -i \mathbf{Op}(\partial_{\xi_j} a) \end{cases}$$

où x_j désigne la multiplication par la fonction $x \mapsto x_j$.

Définition C.6. Pour $a \in S^m$, on appelle opérateur pseudo-différentiel de symbole a l'opérateur $\mathbf{Op}(a)$ et, par analogie avec les opérateurs différentiels, on note le plus souvent $\mathbf{Op}(a) = A(x, D)$.

On note Ψ^m l'espace des opérateurs pseudo-différentiels d'ordre m , *i.e* des opérateurs dont le symbole et d'ordre m et Ψ_{cl}^m l'espace des opérateurs pseudos-différentiels classiques d'ordre m , *i.e* dont le symbole vit dans l'espace S_{cl}^m des symboles classiques d'ordre m .

On s'intéresse pour conclure ce paragraphe à déterminer le noyau (de Schwartz) de tout opérateur pseudo-différentiel. Soit pour cela $a \in S^{-\infty}$. Alors :

$$\begin{aligned} \forall u \in \mathcal{S}(\mathbb{R}^n), \quad A(x, D)u(x) &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\langle x, \xi \rangle} a(x, \xi) \hat{u}(\xi) \, d\xi \\ &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\langle x, \xi \rangle} a(x, \xi) \, d\xi \int_{\mathbb{R}^n} e^{-i\langle y, \xi \rangle} u(y) \, dy \\ &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} u(y) \, dy \int_{\mathbb{R}^n} e^{i\langle x-y, \xi \rangle} a(x, \xi) \, d\xi \end{aligned}$$

et ainsi le noyau de l'opérateur $A(x, D)$ est donné par :

$$\forall x, y \in \mathbb{R}^n, \quad K(x, y) := \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\langle x-y, \xi \rangle} a(x, \xi) \, d\xi.$$

2. Régularité spatiale restreinte

Du fait de notre cadre d'application, on est souvent amené à manipuler des difféomorphismes non-lisses : c'est le cas dès que l'on choisit un noyau K_V qui n'appartient pas à $\Psi^{-\infty}$. On a dans ce cas recourt à des symboles dont la régularité spatiale est moindre.

Soit un espace fonctionnelle X tel que $\mathcal{C}_0^\infty \subset X \subset \mathcal{C}^0$. On note XS^m l'espace des symboles de régularité spatiale restreinte (par X), *i.e* des fonctions définies sur $\mathbb{R}^n \times \mathbb{R}^n$ et telles que :

$$\forall x, \xi \in \mathbb{R}^n, \quad \forall \alpha \in \mathbb{N}^n, \quad \left\| \partial_\xi^\alpha a(x, \xi) \right\|_X \leq C(1 + |\xi|)^{m-|\alpha|}.$$

De même, on introduit l'espace des symboles classiques de régularité spatiale restreinte, celui des opérateurs pseudo-différentiels de régularité spatiale restreinte,...

Enfin, on termine cette annexe en énonçant le résultat suivant, tiré de [33], utile à l'établissement du théorème I.20 concernant les métamorphoses de mesures discrètes :

Proposition C.7

Soit $m \in \mathbb{R}$, $r \in \mathbb{N}$ et $a \in \mathcal{C}_0^r S_{cl}^m$ un opérateur classiques dont la régularité spatiale est "seulement" \mathcal{C}_0^r . Alors, il existe $s \in [-r, r]$ tel que

$$a(x, D): \mathcal{H}^{s+m} \longrightarrow \mathcal{H}^s$$

et

$$a(D, x): \mathcal{H}^s \longrightarrow \mathcal{H}^{s-m}$$

où $a(D, x)$ désigne l'opérateurs défini par

$$\forall u \in \mathcal{S}(\mathbb{R}^n), \quad a(D, x)u = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\langle x-y, \xi \rangle} a(y, \xi) u(y) dy d\xi .$$

Références

- [2] Serge Alinhac and Patrick Gérard. *Opérateurs pseudo-différentiels et théorème de Nash-Moser*. Savoirs actuels. EDP Sciences, 1 edition, 1991.
- [14] Joan Glaunès. Code Matlab pour l'appariement de points, courbes et surfaces. <http://www.mi.parisdescartes.fr/~glaunes/matchine.zip>.
- [15] Joan Glaunès. LDDMM and geodesic equations for discrete matching functionals. Unpublished, 2015.
- [16] Joan Glaunès. *Transport par difféomorphismes de points, de mesures et de courants pour la comparaison de formes et l'anatomie numérique*. PhD thesis, Université Paris 13, 2005.
- [17] Joan Glaunès, Alain Trouvé, and Laurent Younes. Diffeomorphic matching of distributions : a new approach for unlabelled point-sets and sub-manifolds matching. *Computer Vision and Pattern Recognition*, 2004.
- [18] Jean Hare. Mécanique analytique. Notes de cours, L3 FIP, Département de Physique de l'ENS, 2007.
- [19] Darryl D. Holm, Tanya Schmah, and Cristina Stoica. *Geometric Mechanics and Symmetry – From finite to infinite dimensions*. Oxford texts in applied and engineering mathematics. Oxford University Press, 2009.
- [20] Darryl D. Holm, Alain Trouvé, and Laurent Younes. The euler-poincaré theory of metamorphosis. *Quarterly of Applied Mathematics*, 67(4) :661–685, 2008.
- [21] Lars Hörmander. *The analysis of linear partial differential operators III – Pseudo-differential operators*. Classics in Mathematics. Springer-Verlag, 2 edition, 1985.
- [26] Daniel Lie. *Cours d'analyse fonctionnelle*. Références sciences. Ellipses, 2013.
- [30] Pierre Mounoud. Géométrie différentielle. Notes de cours, L3, Université Bordeaux 1, 2012.
- [33] Casey L. Richardson and Laurent Younes. Computing metamorphoses between discrete measures. *Journal of geometric mechanics*, 5(1), 2013.
- [35] Jean-Baptiste Schiratti, Stéphanie Allasonnière, Oliver Colliot, and Stanley Durrleman. Learning spatiotemporal trajectories from manifold-valued longitudinal data. *Neural Information Processing Systems (NIPS)*, 2015.
- [36] Jean-Baptiste Schiratti, Stéphanie Allasonnière, Oliver Colliot, and Stanley Durrleman. A mixed effect model with time reparametrization for longitudinal univariate manifold valued data. *Information Processing in Medical Imaging (IPMI)*, 2015.

-
- [38] Elias M. Stein. *Harmonic Analysis : Real-variable methods, orthogonality and oscillatory integrals*. Princeton University Press, 2 edition, 1995.
- [40] Terence Tao. The Euler-Arnold equation. Online, 2010.
<https://terrytao.wordpress.com/2010/06/07/the-euler-arnold-equation/>.
- [41] Michael E. Taylor. Pseudodifferential operators and nonlinear PDE. Lecture notes.
- [42] D'Arcy Thompson. *On growth and form*. Cambridge University Press, 1917.
- [44] Alain Trounev and Laurent Younes. Metamorphoses through lie group action. *Foundations of computational mathematics*, pages 173–198, 2013.
- [45] Claude Viterbo. Géométrie différentielle. Notes de cours, L3, ENS Ulm, 2013.
- [46] Wikipedia. Computational anatomy ; version of 4 august 2016.
https://en.wikipedia.org/wiki/Computational_anatomy.
- [48] Claude Zuily. *Éléments de distributions et d'équations aux dérivées partielles*. Sciences Sup. Dunod, 2002.

Références

- [1] LAPACK : Linear Algebra PACKage 3.6.0. Symmetric matrix – eigenvalue, Nov 15, 2015. http://www.netlib.org/lapack/explore-html/d3/d88/group__real_s_yeigen.html.
- [2] Serge Alinhac and Patrick Gérard. *Opérateurs pseudo-différentiels et théorème de Nash-Moser*. Savoirs actuels. EDP Sciences, 1 edition, 1991.
- [3] Jonathan E. Atkins, Erik G. Boman, and Bruce Hendrickson. A spectral algorithm for seriation and the consecutive ones problem. *SIAM Journal of Computing*, 28(1) :297–310, 1998.
- [4] Norman T.J Bailey. *The Elements of Stochastic Processes with Applications to the Natural Sciences*. Wiley Classics Library. Wiley-Blackwell, 1990.
- [5] Nicolas Burq and Patrick Gérard. Contrôle optimal des équations aux dérivées partielles. Cours, 2014.
- [6] Nicolas Charon and Alain Trounev. The varifold representation of non-oriented shapes for diffeomorphic registration. *SIAM Journal on Imaging Sciences*, 6(4), 2013.
- [7] The Scipy Community. Scipy.org – numpy.linalg.eigh, Oct 18, 2015. <http://docs.scipy.org/doc/numpy/reference/generated/numpy.linalg.eigh.html>.
- [8] Gaëlle Degrez. Des étudiants de paris-sud obtiennent la médaille d’or à l’iGEM, Dec 10, 2014. <http://www.actu.u-psud.fr/fr/formations/actualites-2014/des-etudiants-de-paris-sud-obtiennent-la-medaille-d-or-a-l-igem.html>.
- [9] Albrecht Dürer. *Vier bücher von menschlicher proportion*. 1528.
- [10] François Dubois, Hervé Lemeur, and Claude Reiss. Mathematical modeling of antigenicity for HIV dynamics. *MathS In Action*, 3(1) :1–35, 2010.
- [11] Walter M. Fitch and Emanuel Margoliash. Construction of phylogenetic trees. *Science*, 155(3760) :279–284, January 1967.
- [12] William Fulton and Joe Harris. *Representation Theory – A first course*. Graduate texts in mathematics. Springer, 1991.
- [13] Sylvestre Gallot, Dominique Hulin, and Jacques Lafontaine. *Riemannian Geometry*. Universitext. Springer, 3 edition, 2004.

- [14] Joan Glaunès. Code Matlab pour l'appariement de points, courbes et surfaces. <http://www.mi.parisdescartes.fr/~glaunes/machine.zip>.
- [15] Joan Glaunès. LDDMM and geodesic equations for discrete matching functionals. Unpublished, 2015.
- [16] Joan Glaunès. *Transport par difféomorphismes de points, de mesures et de courants pour la comparaison de formes et l'anatomie numérique*. PhD thesis, Université Paris 13, 2005.
- [17] Joan Glaunès, Alain Trouvé, and Laurent Younes. Diffeomorphic matching of distributions : a new approach for unlabelled point-sets and sub-manifolds matching. *Computer Vision and Pattern Recognition*, 2004.
- [18] Jean Hare. Mécanique analytique. Notes de cours, L3 FIP, Département de Physique de l'ENS, 2007.
- [19] Darryl D. Holm, Tanya Schmah, and Cristina Stoica. *Geometric Mechanics and Symmetry – From finite to infinite dimensions*. Oxford texts in applied and engineering mathematics. Oxford University Press, 2009.
- [20] Darryl D. Holm, Alain Trouvé, and Laurent Younes. The euler-poincaré theory of metamorphosis. *Quarterly of Applied Mathematics*, 67(4) :661–685, 2008.
- [21] Lars Hörmander. *The analysis of linear partial differential operators III – Pseudo-differential operators*. Classics in Mathematics. Springer-Verlag, 2 edition, 1985.
- [22] iGEM Paris-Saclay. Wiki de l'équipe – this is not a lemon, 2014. http://2014.igem.org/Team:Paris_Saclay.
- [23] Niles Johnson. A visualization of the hopf fibration, 2011. <http://nilesjohnson.net/hopf.html>.
- [24] Jürgen Jost. *Riemannian geometry and geometry analysis*. Universitext. Springer, 3 edition, 2002.
- [25] Jacques Lafontaine. *Introduction aux variétés différentielles*. Grenoble Sciences. EDP Sciences, 2 edition, 2010.
- [26] Daniel Lie. *Cours d'analyse fonctionnelle*. Références sciences. Ellipses, 2013.
- [27] Innar Lilv. Seriation and matrix reordering methods : an historical overview. *Statistical Analysis and Data Mining*, 3(2) :70–91, April 2010.
- [28] Mario Micheli, Peter W. Michor, and David Mumford. Sectional curvature in terms of the co-metric and with applications to the riemannian manifolds of landmarks. *SIAM J. on Imaging Sciences*, 5(1), 2012.
- [29] D.R Miller and H.D Cox. *The Theory of Stochastic Processes*. Methuen & Co, 1965.
- [30] Pierre Mounoud. Géométrie différentielle. Notes de cours, L3, Université Bordeaux 1, 2012.

-
- [31] Pierre-Guy Plamondon. Représentations linéaires des groupes finis. Cours, 2014.
- [32] Eric Renshaw. Modelling biological populations in space and time. 1993.
- [33] Casey L. Richardson and Laurent Younes. Computing metamorphoses between discrete measures. *Journal of geometric mechanics*, 5(1), 2013.
- [34] Ricardo Sa Earp and Eric Toubiana. *Introduction à la géométrie hyperbolique et aux surfaces de Riemann*. Enseignement des mathématiques. Cassini, 2 edition, 2009.
- [35] Jean-Baptiste Schiratti, Stéphanie Allasonnière, Oliver Colliot, and Stanley Durrleman. Learning spatiotemporal trajectories from manifold-valued longitudinal data. *Neural Information Processing Systems (NIPS)*, 2015.
- [36] Jean-Baptiste Schiratti, Stéphanie Allasonnière, Oliver Colliot, and Stanley Durrleman. A mixed effect model with time reparametrization for longitudinal univariate manifold valued data. *Information Processing in Medical Imaging (IPMI)*, 2015.
- [37] Jean-Pierre Serre. *Représentations linéaires des groupes finis*. Méthodes. Hermann, 5 edition, 1998.
- [38] Elias M. Stein. *Harmonic Analysis : Real-variable methods, orthogonality and oscillatory integrals*. Princeton University Press, 2 edition, 1995.
- [39] Eric J. Stewart, Richard Madden, Gregory Paul, and François Taddei. Aging and death in an organism that reproduces by morphologically symmetric division. *PLoS Biology*, 3(2), February 2005.
- [40] Terence Tao. The Euler-Arnold equation. Online, 2010.
<https://terrytao.wordpress.com/2010/06/07/the-euler-arnold-equation/>.
- [41] Michael E. Taylor. Pseudodifferential operators and nonlinear PDE. Lecture notes.
- [42] D’Arcy Thompson. *On growth and form*. Cambridge Université Press, 1917.
- [43] Alain Trouvé. Géométrie et espaces de formes. Cours, 2016.
- [44] Alain Trouvé and Laurent Younes. Metamorphoses through lie group action. *Foundations of computational mathematics*, pages 173–198, 2013.
- [45] Claude Viterbo. Géométrie différentielle. Notes de cours, L3, ENS Ulm, 2013.
- [46] Wikipedia. Computational anatomy ; version of 4 august 2016.
https://en.wikipedia.org/wiki/Computational_anatomy.
- [47] Laurent Younes. *Invariance, déformations et reconnaissance de formes*. Mathématiques et applications 44. Springer, 2003.
- [48] Claude Zuily. *Éléments de distributions et d’équations aux dérivées partielles*. Sciences Sup. Dunod, 2002.