



HAL
open science

Géométrie sous-riemannienne en dimension infinie et applications à l'analyse mathématique des formes

Sylvain Arguillere

► **To cite this version:**

Sylvain Arguillere. Géométrie sous-riemannienne en dimension infinie et applications à l'analyse mathématique des formes. Mathématiques générales [math.GM]. Université Pierre et Marie Curie - Paris VI, 2014. Français. NNT : 2014PA066144 . tel-01126941

HAL Id: tel-01126941

<https://theses.hal.science/tel-01126941>

Submitted on 6 Mar 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université Paris VI - Pierre et Marie Curie

Laboratoire Jacques-Louis Lions

École Doctorale des Sciences mathématiques de Paris Centre

THÈSE DE DOCTORAT

Discipline : Mathématiques et leurs interactions

présentée par

Sylvain Arguillère

**Géométrie sous-riemannienne en dimension infinie
et applications à l'analyse mathématique des
formes**

dirigée par Emmanuel TRÉLAT et Alain TROUVÉ

Rapportée par

M. Ugo BOSCAIN École Polytechnique
M. Andrei AGRACHEV International School for Advanced Studies, Trieste

Soutenue le 10 juillet 2014 devant le jury composé de :

M. Jean-Michel CORON	Université Paris VI	Examineur
M. Pierre PANSU	Université Paris XI	Examineur
M. Ludovic RIFFORD	Université de Nice Sophia Antipolis	Examineur
M. Peter MICHOR	Université de Vienne	Examineur
M. Emmanuel TRÉLAT	Université Paris VI	Directeur de thèse
M. Alain TROUVÉ	ENS de Cachan	Directeur de thèse

Laboratoire Jacques-Louis Lions
Case 187
4, place Jussieu
75 252 Paris cédex 05

École doctorale de sciences
mathématiques de Paris centre Case 290
4 place Jussieu
75 252 Paris cédex 05

À mes deux grand-mères, Monique et Chantal

Remerciements

Tout d'abord, un grand merci à mes directeurs de thèse pour m'avoir donné un sujet aussi passionnant et diversifié, et avoir toujours été disponible pour répondre à mes questions (mathématiques ou autres). Emmanuel Trélat m'a accordé énormément de temps à chacune de mes visites (souvent improvisées). J'ai beaucoup appris de lui sur les plans scientifiques et humains, et je compte m'inspirer de son exemple tout au long de ma carrière. Alain Trouvé a été d'une grande patience et d'une immense gentillesse malgré ses responsabilités, et m'a donné l'occasion de rencontrer de nombreux amis et collègues. Les "shape meeting" annuels ont constitué les moments les plus intéressants (et les plus agréables !) de ma thèse.

Un grand merci à Ugo Boscain et Andrei Agrachev pour avoir été les rapporteurs de ma thèse. En particulier, je les remercie d'avoir rendu leur rapport aussi rapidement.

Merci à tous les membres du jury de me faire l'honneur d'être là : Pierre Pansu, mon tuteur à l'ENS à qui je dois énormément et celui qui m'a présenté à Emmanuel, Ludovic Rifford, qui avait organisé une conférence de géométrie sous-riemannienne à Nice où ma passion pour ce sujet s'est véritablement développée, Peter Michor que j'ai rencontré aux fameux shape meetings, et dont les articles de géométrie et d'analyse de formes m'ont beaucoup inspiré, et Jean-Michel Coron, membre éminent de la communauté du contrôle.

Je remercie ensuite Laurent Younès et toute l'équipe du CIS de Johns, en particulier Hopkins, Michael, Saurabh, Casey, Erika, pour m'avoir invité plusieurs mois dans leur laboratoire.

Je remercie tous les membres du laboratoire Jacques-Louis Lions pour m'avoir accueilli ces trois dernières années, et m'avoir fourni un cadre de travail exceptionnel auprès de grands scientifiques.

Merci à Didier Smets et Albert Cohen pour avoir pris un peu de leur temps pour discuter de mes travaux. Merci également à Laurent Boudin pour son support, et pour l'organisation des cours de LM216, et à Claire David pour le cours de LM121.

Je remercie tous les professeurs de mathématiques et de physique qui m'ont transmis leur passion au cours de mes années de lycée et de prépa : M. Vargas, M. Clary, M. Alfaro et, bien sûr, M. Bayard, du lycée Thiers, et M. Mas et Mme. Baldesi du lycée Sacré-Coeur. Ce sont eux qui m'ont inspiré et permis de développer mes aptitudes en mathématiques, et grâce à qui je suis aujourd'hui, je l'espère, un vrai scientifique.

Merci à mes professeurs de l'ENS pour leurs excellents cours, Pierre Pansu, Frédéric Paulin, Wendelin Werner, Jean Bertoin, Marc Rosso, Olivier Glass et bien d'autres. Merci aussi à mes professeurs de M2, particulièrement à Olivier Biquard, Andreas Höring, et Gennadi Henkin.

Merci à Dmitri Burago de Penn State University pour avoir encadré durant mon mémoire de M2. Il a été le premier à m'introduire à la géométrie sous-riemannienne.

Merci à Davide Barilari d'organiser chaque mois un très bon séminaire de géométrie sous-riemannienne.

Merci à tous les doctorants du LJLL pour leur bienveillance et la bonne ambiance qu'ils contribuent à apporter. Merci plus particulièrement à mes co-bureaux et ex-cobureaux :

Juliette, Malik, Casimir, Guillaume, Jean-Paul, Giacomo, Ludovick et André. C'est grâce à eux et à nos discussions animées que je ne me suis pas ennuyé durant ces trois années. Merci également aux deux Pierre pour l'organisation du GTT, et à Mamadou pour de nombreuses discussions de mathématiques.

Merci aux doctorants, post-doctorants, et ex-doctorants maintenant maîtres de conférence de l'ENS Cachan : Nicolas, Benjamin, Barbara, Irène, Giacomo, Joan, et François-Xavier dont l'amitié a beaucoup compté.

Merci à Artem et Petri pour leur amitié et leur participation au groupe de travail de géométrie sous-riemannienne à l'ENS.

Un immense merci à ma famille, qui m'a toujours encouragé et apporté un support sans faille. Merci à ma mère, mon père et mon frère pour tout ce qu'ils ont fait pour moi. Merci à Pierre pour nos nombreuses conversations de mathématiques. Merci à mes grand-mères pour tout leur amour.

Merci également à ma future belle-famille, particulièrement à Christine, Ludwik, Arthur, Jérémy, Jeanine et André pour m'avoir accepté et accueilli parmi eux comme un membre à part entière.

Merci à Nicolas, Thomas, Xiao, Simon et Louis pour leur amitié à l'ENS et les discussions très amusantes que nous avons eues. Merci aussi pour les vacances au ski !

Un grand merci à mes amis et témoins, Quentin et Olivier, pour le soutien qu'ils m'ont apporté. Je sais que je peux toujours compter sur eux. Merci à tous mes autres amis du collège, du lycée, et de la prépa, trop nombreux pour être tous cités mais que je n'oublie pas.

Et enfin, le plus grand des merci à Séverine, celle qui illumine ma vie. Le plus beau de mes jours restera celui où elle m'a dit oui.

Sylvain ARGUILLÈRE

Géométrie sous-riemannienne en dimension infinie et applications à l'analyse mathématique des formes

Résumé : Cette thèse est dédiée à l'étude de la géométrie sous-riemannienne en dimension infinie, et à ses applications à l'analyse des déformations par difféomorphismes. La première partie du manuscrit est un résumé détaillé des travaux effectués. La seconde compile les articles rédigés pendant ces trois dernières années de travail. On commence par étendre à la dimension infinie le cadre de la géométrie sous-riemannienne classique, en établissant notamment des conditions assurant l'existence d'un flot géodésique. Puis, on applique ces résultats aux structures sous-riemanniennes fortes et invariantes à droite sur le groupe des difféomorphismes d'une variété à géométrie bornée. On définit ensuite rigoureusement la notion d'espaces de formes, jusqu'alors assez vague dans la littérature. Il s'agit de variétés de Banach sur lesquelles un groupe de difféomorphismes a une action satisfaisant certaines propriétés. On construit ensuite diverses structures sous-riemanniennes sur ces espaces de formes grâce à cette action, structures que nous étudierons en détails. Enfin, on ajoute des contraintes aux déformations possibles et on formule les problèmes d'analyse de formes dans un cadre relevant de la théorie du contrôle optimal en dimension infinie. On démontre un principe du maximum de type Pontryagin adapté à ce contexte, permettant d'établir les équations géodésiques contraintes. Des algorithmes pour la recherche de déformations optimales sont ensuite développés et appuyés par des simulations numériques dans le chapitre 7. Ils unifient et étendent des méthodes précédemment établies pour l'analyse de formes dans le domaine de l'image.

Abstract : This manuscript is dedicated to the study of infinite dimensional sub-Riemannian geometry and its applications to shape analysis using diffeomorphic deformations. The first part is a detailed summary of our work, while the second part combines the articles we wrote during the last three years. In the first two chapters, we extend the framework of sub-Riemannian geometry to infinite dimensions, establishing conditions that ensure the existence of a Hamiltonian geodesic flow. In the third chapter, we apply these results to strong right-invariant sub-Riemannian structures on the group of diffeomorphisms of a manifold. We then define rigorously the abstract concept shape spaces. A shape space is a Banach manifold on which the group of diffeomorphisms of a manifold acts in a way that satisfy certain properties. We then define several sub-Riemannian structures on these shape spaces using this action, and study these structures in Chapters 4 and 5. Finally, in Chapters 6 and 7, we add constraints to the possible deformations, and formulate shape analysis problems in an infinite dimensional control theoretic framework. We prove a Pontryagin maximum principle adapted to this context, establishing the constrained geodesic equations. Algorithms for finding optimal deformations are then developed, supported by numerical simulations. These algorithms extend and unify previously established methods in shape analysis.

Table des matières

Introduction	5
I Présentation des travaux	15
1 Géométrie sous-riemannienne en dimension finie	17
1.1 Premières définitions	17
1.2 Un cadre plus général	18
1.2.1 Fibrés tangents relatifs	18
1.2.2 Structure sous-riemannienne sur une variété, systèmes horizontaux . .	19
1.2.3 Longueur et action des courbes horizontales	19
1.3 Orbite et théorème de Chow-Rashevski	20
1.4 Ensemble des courbes horizontales	20
1.4.1 Espace des courbes horizontales	20
1.4.2 Application point final	22
1.5 Formulation hamiltonienne	24
1.5.1 Forme symplectique canonique, gradient symplectique	24
1.5.2 Hamiltonien	25
1.5.3 Réciproque partielle	26
1.6 Hamiltonien réduit et équation hamiltonienne des géodésiques	26
2 Géométrie sous-riemannienne en dimension infinie	29
2.1 Définitions	29
2.1.1 Structure sous-riemannienne sur une variété de Banach	30
2.1.2 Courbes horizontales, action et longueur	30
2.1.3 Distance sous-riemannienne	31
2.1.4 Géodésiques locales	31
2.2 Un équivalent au théorème de Chow-Rashevski	32
2.3 Un exemple de structures sous-riemannienne forte : le produit de groupes de Heisenberg	32
2.4 Application point final en dimension infinie	33
2.5 Absence de principe du maximum en dimension infinie	34
2.6 Géodésiques normales et courbes singulières	35
2.6.1 Forme symplectique faible sur T^*M	36
2.6.2 Hamiltoniens	36
2.7 Espaces cotangents relatifs	37
2.7.1 Fibré cotangent relatif	37
2.7.2 Forme symplectique canonique faible sur τ^*M et gradient symplectique	38
2.7.3 Le cas du hamiltonien	38
2.8 Équation hamiltonienne des géodésiques	38

2.9	Exemple : le groupe des difféomorphismes de \mathbb{R}^d	43
2.9.1	Structures sous-riemanniennes	43
2.9.2	Espace cotangent adapté	44
2.9.3	Équation hamiltonienne des géodésiques	45
3	Géométries invariantes à droite	47
3.1	Groupe des difféomorphismes	47
3.1.1	Variété à géométrie bornée	47
3.1.2	Espace $H_e^s(M, M)$	48
3.1.3	Groupe des difféomorphismes de classe H^s	49
3.1.4	Dérivée logarithmique	49
3.2	Structures sous-riemanniennes invariantes à droite	50
3.2.1	Construction	50
3.2.2	Structures invariantes continues équivalentes à des structures lisses	50
3.2.3	Courbes horizontales, action, longueur et distance	51
3.2.4	Complétude dans le cas fort	51
3.3	Étude de l'orbite de e	52
3.3.1	Contrôlabilité approchée	52
3.3.2	Contrôlabilité exacte	52
3.4	Espaces de Hilbert à noyau reproduisant	53
3.4.1	Définition	53
3.4.2	Construction d'espaces à noyau reproduisant	54
3.5	Hamiltonien, géodésiques normales et anormales	55
3.5.1	Hamiltoniens	55
3.5.2	Hamiltonien réduit	55
3.5.3	Équation hamiltonienne des géodésiques	56
3.5.4	Forme réduite des équations hamiltoniennes sur un groupe de Lie	57
3.5.5	Équation des moments sur $\mathcal{D}^s(M)$	58
3.6	Quelques exemples	58
3.6.1	Noyau diagonal	58
3.6.2	Noyau adapté à une structure sous-riemannienne sur \mathbb{R}^d	59
4	Espaces de formes et analyse des déformations	61
4.1	Définition d'un espace de formes	61
4.2	Structures sous-riemanniennes induites	63
4.2.1	Définition de la structure induite	63
4.2.2	Action d'une trajectoire et relèvement minimal	63
4.2.3	Point de vue dual	64
4.3	Équation hamiltonienne des géodésiques sur un espace de formes	64
4.3.1	Hamiltoniens du système	65
4.3.2	Équations géodésiques	65
4.4	Cas des variétés immergées dans \mathbb{R}^d	66
5	Structures sous-riemanniennes relevées	69
5.1	Applications équivariantes	70
5.2	Structure sous-riemannienne relevée	71
5.3	Équation hamiltonienne des géodésiques	72
5.3.1	Hamiltonien et hamiltonien réduit	72

6	Espaces de formes avec contraintes	75
6.1	Position du problème, existence de solutions	76
6.1.1	Cadre et Notations	76
6.1.2	Première formulation	76
6.1.3	Formulation duale	77
6.2	Un exemple : espaces de multi-formes	78
6.3	Approximation par des espaces de formes de dimension finie	79
6.4	Principe du maximum avec contraintes	80
6.4.1	Énoncé du théorème	80
6.4.2	Équation des géodésiques contraintes sur les espaces de formes	81
7	Algorithmes de minimisation	83
7.1	Minimisation sur l'espace des contrôles	83
7.1.1	Gradient sans contraintes	83
7.1.2	Méthode du lagrangien augmenté pour le cas contraint	84
7.2	Minimisation sur le moment initial par la méthode du tir	86
7.3	Simulations numériques	87
7.3.1	Courbes englobant une aire constante	87
7.3.2	Cas des multiformes	88
7.4	Un algorithme de quasi-Newton	90
7.4.1	Cadre, définitions et énoncé du résultat	91
7.4.2	Application au calcul d'une suite d'inverses	91
8	Conclusion et perspectives	93
II	Compilation des articles	101
9	Right-invariant sub-Riemannian geodesic equations	103
9.1	Sub-Riemannian structure on groups of diffeomorphisms	105
9.1.1	Definition of the manifold $\mathcal{D}^s(M)$	106
9.1.2	Sub-Riemannian structure on $\mathcal{D}^s(M)$	107
9.1.3	The accessibility property on $\mathcal{D}^s(M)$	111
9.1.4	Geodesics on $\mathcal{D}^s(M)$	117
9.1.5	Examples of geodesic equations in $\mathcal{D}(\mathbb{R}^d)$	124
9.2	Shape spaces	127
9.2.1	Definition	127
9.2.2	Sub-Riemannian structure on shape spaces	128
9.2.3	Normal sub-Riemannian geodesic equations on shape spaces	131
9.2.4	Some examples of geodesic equations on classical shape spaces	134
9.3	Lifted Shape Spaces	138
9.3.1	Lifted sub-Riemannian structure	138
9.3.2	The end-point mapping under an additional assumption	141
9.3.3	The lifted Hamiltonian geodesic equations	142
9.3.4	The momentum viewpoint and another form of the lifted geodesic equations	143
9.3.5	Some examples of lifted geodesic equations	145

10	Optimal control and shape deformation	149
10.1	Modelling shape deformation problems with optimal control	152
10.1.1	Preliminaries: deformations and RKHS of vector fields	152
10.1.2	From shape space problems to optimal control	154
10.1.3	Further comments: lifted shape spaces and multishapes	161
10.1.4	Finite dimensional approximation of optimal controls	163
10.1.5	Proof of Theorem 10.1	165
10.2	Constrained geodesic equations in shape spaces	166
10.2.1	First-order optimality conditions: PMP in shape spaces	166
10.2.2	The geodesic equations in a shape space	171
10.3	Algorithmic procedures	173
10.3.1	Problems without constraints	173
10.3.2	Problems with constraints	176
10.4	Numerical examples	179
10.4.1	Matching with constant total volume	179
10.4.2	Multishape matching	180
10.5	Conclusion and open problems	182
11	Symmetric rank-one update	185
11.1	Introduction	185
11.2	Notations and symmetric rank-one algorithm	187
11.3	Main Result	187
11.4	First estimates	188
11.5	Uniform m -span of a sequence and applications	191
11.6	Numerical simulations	192
11.6.1	Approximation of a sequence of matrices	192
11.6.2	Approximation of a sequence of inverses	194
11.7	An application: optimal deformations of constrained shapes	196
11.7.1	Large deformation diffeomorphic metric mapping.	196
11.7.2	Shapes with constraints.	197
11.7.3	Implementation of the symmetric rank-one update.	199
11.7.4	Numerical simulations	199
11.8	Conclusion	201

Introduction

Le but de cette thèse est de définir et d'étudier le concept de variété sous-riemannienne de dimension infinie, afin de l'appliquer à diverses constructions apparaissant dans le domaine de l'image.

Le manuscrit est composé de deux parties. La première, constituée de sept chapitres, fournit un résumé complet des travaux effectués. La seconde regroupe les articles rédigés durant les trois dernières années, avec quelques modifications mineures permettant d'unifier les notations. Dans cette dernière, le lecteur trouvera des détails et des preuves qui n'ont pas été insérés dans la première partie dans un souci de lisibilité.

Commençons par donner une brève vue d'ensemble du domaine de l'analyse mathématique des formes afin de voir où la géométrie sous-riemannienne de dimension infinie intervient.

Analyse mathématique des formes

L'analyse de formes et de déformations d'images a reçu une attention croissante au cours des dernières années. Il s'agit de donner un cadre géométrique au problème afin de permettre, par exemple, l'analyse statistique d'images et d'objets déformables. Ce point de vue a été utilisé avec succès dans de nombreux domaines, et plus particulièrement en anatomie computationnelle. Cette dernière consiste à comparer des images d'organes (généralement des IRM de cerveaux) de différents patients, dans le but de prévoir à l'avance l'apparition de certaines maladies dégénératives [DGM98, GM98, MTY02, MTY06].

La source du problème est que deux formes (par exemple des courbes planes), peuvent représenter des objets similaires pour un observateur humain, mais être très différentes au niveau de la liste de points qu'un ordinateur utilisera pour les représenter. Il faut donc, pour les comparer, trouver des méthodes qui prennent en compte leur géométrie. Plusieurs façons de résoudre ce problème en utilisant des concepts mathématiques rigoureux ont été proposées avec succès.

Pour cela, il faut tout d'abord définir un *espace de formes* \mathcal{S} sur lequel on puisse travailler. Il faut ensuite trouver une manière de comparer deux formes q_0 et q_1 dans \mathcal{S} .

Une des premières solutions a été proposée par Kendall [Ken84]. Il s'agit de définir l'espace des formes discrètes à n points dans \mathbb{R}^d comme l'ensemble des n -uplets de \mathbb{R}^d , pour $n \in \mathbb{N}$, quotienté par l'action des rotations et des translations. On peut alors estimer la différence entre deux formes $q_0 = (x_1, \dots, x_n)$ et $q_1 = (y_1, \dots, y_n)$ en calculant le carré de leur distance à isométrie près :

$$\inf \left\{ \sum_{i=1}^n |T(x_i) - y_i|^2 \mid T : \mathbb{R}^d \rightarrow \mathbb{R}^d \text{ isométrie} \right\}.$$

Plus généralement, on peut considérer une attache aux données g , c'est-à-dire une fonction à valeurs réelles sur $(\mathbb{R}^d)^n$ qui compare la différence entre une forme donnée q et la cible q_1 :

$$\inf \left\{ g(T(q)) \mid T : \mathbb{R}^d \rightarrow \mathbb{R}^d \text{ isométrie} \right\}.$$

On prend généralement $g(q) = d(q, q_1)^2$ où d est une distance qui peut être différente de la distance euclidienne. Ce cadre a été le point de départ de plusieurs méthodes pour l'analyse des déformations [DM98, KBCL99, Sma96]. Toutefois, elles sont très restrictives. On ne peut pas, par exemple, étudier des familles de courbes : il faut d'abord trouver un moyen efficace de les discrétiser. Un problème plus important est que ce modèle ne prend pas en compte toutes les propriétés géométriques des formes discrétisées. Par exemple, deux formes q_0 et q_1 découlant l'une de l'autre par une déformation lisse (par exemple un difféomorphisme de \mathbb{R}^d) devraient être "plus semblables" que si q_1 est un déplacement plus aléatoire de q_0 .

Transformation par difféomorphismes. Une méthode efficace pour tenir compte de ces considérations est d'étudier des transformations induites par des difféomorphismes. En effet, les formes étudiées sont généralement des objets plongés dans une variété M (le plus souvent, $M = \mathbb{R}^d$). Le groupe des difféomorphismes de M agit alors naturellement sur l'espace des formes par $(\varphi, q) \mapsto \varphi \cdot q$. Par exemple, si q est une courbe paramétrée de M , un difféomorphisme φ agit sur q par composition à gauche, transformant q en une nouvelle courbe $\varphi \cdot q = \varphi \circ q$. On estime alors la différence entre deux formes par

$$\inf \{ F(\varphi) + g(\varphi \cdot q_0) \mid \varphi \text{ difféomorphisme de } M \},$$

où $F(\varphi)$ est un coût défini sur le groupe des difféomorphismes, et g une attache aux données estimant une distance entre la forme transformée $\varphi \cdot q_0$ et une forme cible q_1 .

Flots de champs de vecteurs. Vient alors le problème de définir un coût F avec un bon comportement. Pour cela, on se restreint à des difféomorphismes engendrés par des flots

$$\partial_t \varphi^X(t, x) = X(t, \varphi^X(t, x)), \quad (t, x) \in [0, 1] \times M,$$

de champs de vecteurs X dépendant du temps sur M . Un tel flot φ^X est une *déformation* de l'espace ambiant M , et induit une déformation $t \mapsto q(t) = \varphi^X(t) \cdot q_0$ d'une forme initiale q_0 . Un champ de vecteurs X induit donc une *déformation infinitésimale* $X \cdot q$ d'une forme q , qui s'intègre en une déformation de q .

Notons $\Gamma(TM)$ l'espace des champs de vecteurs de M . L'ensemble des déformations infinitésimales autorisées est alors défini comme un ensemble de champs de vecteurs X appartenant à un certain sous-espace pré-hilbertien $(V, \langle \cdot, \cdot \rangle)$ de $\Gamma(TM)$. Les éléments de V sont en général des champs de vecteurs lisses.

La différence entre deux formes q_0 et q_1 est alors mesurée par la minimisation, sur l'ensemble des champs de vecteurs $t \mapsto X(t) \in V$, d'une fonctionnelle de la forme

$$J(\varphi^X) = J(X)^1 = \frac{1}{2} \int_0^1 \langle X(t), X(t) \rangle dt + g(\varphi^X(1) \cdot q_0) = A(X) + g(\varphi^X(1) \cdot q_0). \quad (1)$$

Le premier terme mesure l'action $A(X)$ (ou énergie) de la courbe de difféomorphismes définies par le flot φ^X de X . Celle-ci induit une déformation de l'état initial q_0 . Le second terme est une attache aux données qui mesure la différence entre l'état déformé final $\varphi^X(1) \cdot q_0$ et la forme cible q_1 . Pour le choix de l'espace $(V, \langle \cdot, \cdot \rangle)$, on rencontre deux tendances dans la littérature existante.

Méthode 1 : métriques riemanniennes faibles. On prend pour V l'ensemble des champs de vecteurs de classe C^∞ à support compact. On choisit ensuite une norme pré-hilbertienne (généralement une norme Sobolev) pour $\langle \cdot, \cdot \rangle$. C'est une approche utilisée par exemple dans [BHM11, BHM12, MM07].

1. Remarquons qu'il y a correspondance bijective entre les champs de vecteurs X et les déformations φ^X .

Un tel point de vue engendre naturellement une métrique riemannienne dite "faible" (voir [Cla09]), invariante à droite sur le groupe de Lie (modélisé sur un espace de Fréchet) des difféomorphismes de classe \mathcal{C}^∞ à support compact, de distance associée d_R . L'action d'une déformation $t \in [0, 1] \mapsto \varphi^X(t)$ engendrée par un champ de vecteurs $t \mapsto X(t) \in V$ est alors donnée par

$$A(\varphi^X) = \frac{1}{2} \int_0^1 \langle X(t), X(t) \rangle dt = \frac{1}{2} \int_0^1 \langle \dot{\varphi}^X(t) \circ \varphi^X(t)^{-1}, \dot{\varphi}^X(t) \circ \varphi^X(t)^{-1} \rangle dt.$$

On peut également définir une longueur pour cette courbe, et une distance d_R comme sur les variétés riemanniennes de dimension finie. Minimiser la fonctionnelle J de (1) revient alors à minimiser

$$\frac{1}{2} d_R(\text{Id}_M, \varphi)^2 + g(\varphi \cdot q_0)$$

parmi les difféomorphismes de classe \mathcal{C}^∞ à support compact.

Méthode 2 : espaces de Hilbert à noyau reproduisant. On choisit un espace de Hilbert $(V, \langle \cdot, \cdot \rangle)$ engendré par un *noyau reproduisant* $K(x, y) : T_y^*M \rightarrow T_x M$ [Tro95]. Un tel noyau permet d'effectuer très rapidement certains calculs nécessaires à la minimisation de J . Ce point de vue permet de travailler sur le groupe $\mathcal{D}^s(M)$ des difféomorphismes de classe de Sobolev H^s , qui est une variété de Hilbert et un groupe topologique, plutôt qu'une variété de Fréchet comme dans le cas précédent. Son espace tangent en l'identité est l'espace des champs de vecteurs de classe de Sobolev H^s .

Ce cadre, qui est celui adopté dans cette thèse, offre de nombreux avantages d'un point de vue numérique. En plus de préserver de nombreuses propriétés géométriques des formes étudiées, il a permis le développement d'une famille d'algorithmes efficaces de comparaison de formes : les méthodes LDDMM (*Large Deformation Diffeomorphic Metric Matching*) [BMTY05, DGM98, GTY, GTY06, JM00, MTY02, MTY06, Tro95].

Toutefois, on ne prend généralement pas V égal à tout l'espace des champs de vecteurs de classe de Sobolev H^s , on prend un espace bien plus petit, souvent constitué de champs de vecteurs analytiques. En conséquence, la géométrie invariante à droite engendrée par V n'est pas riemannienne mais *sous-riemannienne*. Les auteurs des méthodes LDDMM n'étaient pas familiers avec la géométrie sous-riemannienne, et ont traité le sujet du point riemannien².

Géométrie sous-riemannienne sur les difféomorphismes. Une structure sous-riemannienne est un triplet (M, Δ, g) , où M est une variété, $\Delta \subset TM$ est un sous-fibré du fibré tangent, dit sous-fibré *horizontal*, et g est une métrique riemannienne sur Δ . On peut alors définir les courbes horizontales, qui sont les courbes tangentes à Δ en tout point, puis une notion de longueur, d'action et de distance sous-riemanniennes de façon analogue au cas riemannien [Mon02, BR96].

Si on revient au cas des flots de champs de vecteurs sur M appartenant à un espace de Hilbert $(V, \langle \cdot, \cdot \rangle)$, V induit sur le groupe des difféomorphismes une structure sous-riemannienne forte et invariante à droite. En effet, il suffit de définir l'espace horizontal en φ par l'ensemble des $X \circ \varphi$ pour $X \in V$. Cette structure se projette alors en une structure sous-riemannienne forte sur les espaces de formes grâce à l'action des difféomorphismes. Cette construction nous

2. Cette confusion a été possible pour deux raisons. Tout d'abord, les auteurs travaillaient généralement sur des espaces de formes de dimension finie, auquel cas la structure sous-riemannienne projetée est, de fait, généralement riemannienne. Ensuite, même en dimension infinie, les méthodes étaient basées sur l'équation hamiltonienne des géodésiques et la théorie du contrôle qui, contrairement aux équations géodésiques classiques, peuvent être appliquées de manière identique en géométrie riemannienne et sous-riemannienne. C'est encore plus vrai lorsque la distribution horizontale est dense dans l'espace tangent, comme c'était généralement le cas.

conduit à étudier la géométrie sous-riemannienne en dimension infinie et ses applications en analyse de formes.

Résumé des travaux

Les objectifs de cette thèse sont les suivants :

1. Définir une notion générale de géométrie sous-riemannienne en dimension infinie adaptée aux espaces de formes. Décrire les équations géodésiques.
2. Étudier les structures sous-riemanniennes fortes invariantes à droites sur le groupe des difféomorphismes d'une variété, et plus particulièrement celles qui peuvent être utilisées dans les méthodes LDDMM.
3. Donner une définition à la fois rigoureuse, générale, et facilement utilisable de la notion d'espace de formes. Étudier la structure sous-riemannienne induite par l'action du groupe des difféomorphismes sur un tel espace.
4. Étendre ce cadre à des problèmes d'analyse de formes avec contraintes dans le cadre du contrôle optimal en dimension infinie, de façon à obtenir de meilleurs modèles des objets étudiés. Développer des algorithmes pour étudier ce nouveau problème.

Le premier point sera étudié dans les chapitres 1 et 2, le second dans le chapitre 3, et le troisième dans les chapitres 4 et 5.

Pour mieux comprendre le quatrième point, supposons qu'on veuille étudier le mouvement d'un cœur lors d'une contraction. Le cadre de déformation par difféomorphismes ne prend pas en compte la mécanique derrière ce mouvement, ce qui permet d'obtenir des algorithmes simples mais pose un problème de modélisation. Pour tenir compte des spécificités du mouvement, on peut par exemple astreindre le cœur à se déformer uniquement le long de ses fibres musculaires. Toutefois, comme ces fibres sont alors elles-mêmes déformées, l'espace des déformations infinitésimales autorisées doit changer en fonction de la position du cœur. On peut modéliser ce problème dans le cadre des LDDMMs en rajoutant des contraintes linéaires sur la vitesse des points du cœur, ce qui requiert un point de vue plus général, développé dans les chapitres 6 et 7.

Chapitres 1 et 2 : Géométrie sous-riemannienne en dimension finie et infinie

On donne tout d'abord une définition générale de structure sous-riemannienne sur une variété M , introduite par Agrachev, Boscaïn et al. dans [ABC⁺10], permettant de prendre en compte des distributions horizontales à rang variable. Il s'agit d'un triplet (\mathcal{H}, ξ, g) , où (\mathcal{H}, g) est un fibré vectoriel riemannien sur M et $\xi : \mathcal{H} \rightarrow TM$ est un morphisme de fibrés vectoriels. On dit que (\mathcal{H}, ξ) est un *fibré tangent relatif* sur M . Les courbes horizontales $t \mapsto q(t)$ sont les courbes telles qu'il existe un *contrôle* $t \mapsto u(t) \in \mathcal{H}_{q(t)}$ tel que $\dot{q}(t) = \xi_{q(t)}u(t)$. Un tel couple $t \mapsto (q(t), u(t))$ est appelé un *système horizontal*. On définit ensuite la longueur et l'action d'un tel système respectivement par

$$L(q, u) = \int \sqrt{g_{q(t)}(u(t), u(t))} dt, \quad A(q, u) = \frac{1}{2} \int g_{q(t)}(u(t), u(t)) dt,$$

et on en déduit une distance sous-riemannienne, ainsi que la notion de géodésique sous-riemannienne, comme dans le cas habituel.

En dimension finie, on sait qu'il existe deux types de géodésiques, données par le principe du maximum de Pontryagin. On définit les hamiltoniens $H^1, H^0 : T^*M \oplus \mathcal{H} \rightarrow \mathbb{R}$ par

$$H^\lambda(q, p, u) = p\xi_q u - \frac{\lambda}{2} g_q(u, u). \quad (2)$$

En $(q, p, u) \in T^*M \oplus \mathcal{H}$ fixé, si $\partial_u H^\lambda(q, p, u) = 0$, ce hamiltonien possèdent un *gradient symplectique partiel* $\nabla^\omega H^\lambda(q, p, u)$ en (q, p) pour la forme symplectique canonique ω de T^*M .

Théorème 0.1. (Pontryagin [PBG64]) Soit $q : [0, 1] \rightarrow M$ une géodésique sous-riemannienne issue d'un système horizontal (q, u) . Alors il existe $(p, \lambda) \neq (0, 0)$, où $p : [0, 1] \rightarrow T^*M$ est de classe de Sobolev H^1 , et $\lambda = 0$ ou 1 tels que

$$\begin{cases} 0 &= \partial_u H^\lambda(q(t), p(t), u(t)), \\ (\dot{q}(t), \dot{p}(t)) &= \nabla^\omega H^\lambda(q(t), p(t), u(t)). \end{cases} \quad (3)$$

Si $\lambda = 0$ la trajectoire q est **singulière** (ou anormale). Elle est dite *géodésique normale* si $\lambda = 1$.

On peut de plus obtenir toutes les géodésiques normales en intégrant le flot du gradient symplectique $\nabla^\omega h$ du *hamiltonien normal* $h : T^*M \rightarrow \mathbb{R}$ défini par

$$h(q, p) = \max_{u \in \mathcal{H}_q} \left(p \xi_q u - \frac{1}{2} g_q(u, u) \right). \quad (4)$$

Grâce à ce dernier, on peut montrer ([Mon02, AS04]) que réciproquement, si $(q(\cdot), p(\cdot))$ satisfait (3), alors, quand $\lambda = 1$, $q(\cdot)$ est bel et bien une géodésique. Lorsque $\lambda = 0$, $q(\cdot)$ est un point singulier de l'application point final E qui à un système horizontal $(q(\cdot), u(\cdot))$ associe son point final $q(1) = (q(\cdot), u(\cdot))$. Cette application est étudiée plus en détails en section 1.4.2

Passons au cas de la dimension infinie. La géométrie riemannienne en dimension infinie est un sujet connu et étudié en analyse sur les variétés [KM97, Omo74, Cla09]. Elle se scinde en deux cas : les géométries *fortes*, dont les topologies induites par la métrique riemannienne et celles induites par la structure de variété coïncident, et qui sont très similaire au cas de la dimension finie, et les géométries *faibles*, où les topologies ne coïncident pas. Le cas faible est plus compliqué, car la connection de Levi-Civita et l'application exponentielle n'existent pas toujours. Toutefois, plusieurs variétés riemanniennes faibles ont été abordées avec succès, avec à la clé de nombreuses applications inspirées de travaux d'Arnold [Arn66], comme l'existence et l'unicité de solutions pour de nombreuses équations aux dérivées partielles conservatives. C'est le cas de l'équation d'Euler des fluides incompressible sur des variétés compactes [EM70], ou de l'équation KdV (voir [KW09] et ses références pour de nombreux autres exemples). En analyse de formes, ces structures ont été particulièrement étudiées par Michor, Mumford et Bauer [BHM11, BHM12, MM07].

La géométrie sous-riemannienne en dimension infinie, en revanche, est encore méconnue. On a bien un résultat de contrôlabilité approché [DS80] dans le cas général, et de contrôlabilité exacte sur le groupe des difféomorphismes d'une variété compacte engendrés par des flots de champs de vecteurs appartenant à un certain $\mathcal{C}^\infty(M)$ -module [AC09]. Un preprint [GMV12] présente également un équivalent à l'équation hamiltonienne des géodésiques dans un cas très restreint.

Même la géométrie sous-riemannienne *forte* de dimension infinie est déjà très différente du cas de la dimension finie. En effet, l'ensemble des points accessibles est le plus souvent un sous-ensemble dense d'intérieur vide, ce qui implique que l'image de la différentielle de l'application point final E est aussi dense d'intérieure vide. En conséquence, il est impossible d'établir un principe du maximum de Pontryagin tel que le théorème 0.1, c'est-à-dire de donner des conditions nécessaires complètes pour qu'une courbe soit une géodésique en toute généralité.

En revanche, on peut se contenter d'une réciproque partielle du principe de Pontryagin qui, elle, reste valide en dimension infinie. Cela permet, à défaut de trouver toutes les géodésiques, d'en produire un grande quantité. On montre en effet le résultat suivant.

Proposition 0.1. *Soit (q, u) un système horizontal dont la trajectoire q est simple. Supposons qu'il existe $(p, \lambda) \neq (0, 0)$, où $p : [0, 1] \rightarrow T^*M$ est de classe H^1 avec $p(t) \in T_{q(t)}^*M$ pour tout t et $\lambda \in \{0, 1\}$ tel que*

$$\begin{cases} 0 &= \partial_u H^\lambda(q(t), p(t), u(t)), \\ (\dot{q}(t), \dot{p}(t)) &= \nabla^\omega H^\lambda(q(t), p(t), u(t)), \end{cases}$$

où les hamiltonien H^λ sont définis comme en dimension finie.

- Si $\lambda = 0$, alors l'image de $dE_{(q,u)}$ est contenue dans un sous-espace fermé strict de $T_{q(1)}M$.
- Si $\lambda = 1$, alors (q, u) est un point critique de l'action A à extrémités fixées. On dit que q est une géodésique normale.

L'étape suivante est d'obtenir un flot géodésique afin de montrer que, comme en dimension finie, les géodésiques normales sont bel et bien des géodésiques au sens usuel. Cela demande de résoudre les équations hamiltoniennes sur T^*M . Le problème est qu'à priori, cela n'est possible que si la métrique g est forte sur \mathcal{H} car dans le cas contraire,

$$h(q, p) = \max_{u \in \mathcal{H}_q} \left(p \xi_q u - \frac{1}{2} g_q(u, u) \right) = +\infty$$

si p n'est pas de la forme $g_q(u, \cdot)$. Il faut donc restreindre h à un sous-fibré dense du fibré cotangent τ^*M , appelé *fibré cotangent relatif adapté*, sur lequel h est bien défini. Remarquons que la restriction de la forme symplectique canonique à τ^*M est encore une forme symplectique. Dans le cas riemannien, $\mathcal{H} = TM$, et on peut simplement prendre $\tau^*M = g(TM, \cdot)$.

Pour pouvoir ensuite définir et résoudre l'équation hamiltonienne des géodésiques, il faut ensuite que h admette un gradient symplectique suffisamment régulier sur τ^*M .³

Le théorème principal du chapitre est alors le suivant.

Théorème 0.2. *Si le gradient symplectique de h est \mathcal{C}^2 , alors la projection $t \mapsto q(t) \in M$ de courbes intégrales $t \mapsto (q(t), p(t)) \in \tau^*M$ du champ de gradients symplectiques de h est une géodésique locale⁴.*

Chapitre 3 : structures sous-riemanniennes invariantes à droite sur le groupe des difféomorphismes

Ce chapitre est un résumé de la première partie de [AT], reproduit en partie I dans la section 9.1.

On commence par détailler la structure du groupe $\mathcal{D}^s(M)$, $s > \dim(M)/2 + 1$, des difféomorphismes de classe de Sobolev H^s d'une variété M à géométrie bornée. Il s'agit à la fois d'une variété de Hilbert lisse et d'un groupe topologique. Ce n'est toutefois pas un groupe de Lie car la composition à gauche n'est pas différentiable. La composition à droite, elle, est de classe \mathcal{C}^∞ .

Soit alors $(V, \langle \cdot, \cdot \rangle)$ un sous-espace de Hilbert de l'ensemble des champs de vecteurs sur M avec inclusion continue dans l'espace des champs de classe H^{s+k} , pour $k \in \mathbb{N}$. On peut alors définir sur $\mathcal{D}^s(M)$ une structure sous-riemannienne (\mathcal{H}, ξ, g) invariante à droite par

$$\mathcal{H}_\varphi = V, \quad \xi_\varphi(X) = X \circ \varphi, \quad g(X, X) = \langle X, X \rangle,$$

3. Dans le cas riemannien, cette hypothèse équivaut à l'existence d'une connection de Levi-Civita et d'une application exponentielle [Cla09].

4. Une géodésique locale est une courbe $t \mapsto q(t)$ qui minimise l'action en temps court et à extrémités fixées sur l'ensemble des courbes horizontales *qui restent au voisinage de q* . On sait que c'est le mieux qu'on puisse espérer en général : il existe des structures riemanniennes sur lesquelles certaines géodésiques locales ne sont pas de véritables géodésiques [Cla09].

où $\varphi \in \mathcal{D}^s(M)$, $X \in V$. On en déduit une action, une longueur, et une distance d_{SR} sous-riemannienne sur $\mathcal{D}^s(M)$. On montre alors le résultat suivant.

Théorème 0.3. *Tout couple de points $(\varphi, \psi) \in \mathcal{D}^s(M)^2$ tels que $d_{SR}(\varphi, \psi) < \infty$ peuvent être joints par une géodésique minimisante. De plus, l'espace métrique $(\mathcal{D}^s(M), d_{SR})$ est complet.*

Le problème suivant que l'on étudie est celui de l'accessibilité. En s'inspirant d'un résultat d'Agrachev et Caponigro [AC09], on montre que dans un cas très particulier, on a un résultat de contrôlabilité *exacte*. C'est un théorème remarquable car, comme on l'a mentionné plus haut, c'est très rare en dimension infinie.

Théorème 0.4. *Soit M une variété compacte munie d'une structure sous-riemannienne lisse $(\mathcal{H}^0, \xi^0, g^0)$ complètement non-intégrable⁵. Supposons que V soit l'ensemble des champs de vecteurs sur M tangents à $\xi^0(\mathcal{H}^0)$ et de classe H^s .*

Alors la topologie induite par la distance sous-riemannienne sur $\mathcal{D}^s(M)$ coïncide avec sa topologie intrinsèque de variété. En conséquence, tous $\varphi, \psi \in \mathcal{D}^s(M)$ appartenant à la même composante connexe peuvent être joints par une courbe horizontale.

Le reste du chapitre est dédié à l'étude des équations géodésiques.

Comme V a une injection continue dans l'espace des champs de vecteurs de classe \mathcal{C}^0 , il possède un *noyau reproduisant*

$$(x, y) \mapsto K(x, y) \in \mathcal{L}(T_y^*M, T_xM),$$

où $\mathcal{L}(E, F)$ est l'espace des opérateurs linéaires continus entre deux espaces de Banach E et F , tel que

$$pX(x) = \langle K(\cdot, x)p, X \rangle, \quad (x, p) \in T^*M, \quad X \in V.$$

On en déduit en particulier une forme relativement simple de l'équation hamiltonienne des géodésiques sur une courbe $t \mapsto (\varphi(t), P(t)) \in T^*\mathcal{D}^s(M)$:

$$\begin{cases} \partial_t \varphi(t, x) = \int_M K(\varphi(t, x), \varphi(t, y)) P(t, y) dy \\ \partial_t P(t, x) = -P(t, x) \int_M \partial_1 K(\varphi(t, x), \varphi(t, y)) P(t, y) dy. \end{cases}$$

Ces équations ont précédemment été obtenues dans [TY11, You10], dans les cas $M = \mathbb{R}^d$ en supposant implicitement que la structure sous-jacente était riemannienne. Comme l'équation hamiltonienne des géodésique a la même formulation dans les deux cas, et comme les espaces V utilisés étaient denses dans les champs de vecteurs de classe H^s , les résultats étaient tout de même valides. Nous obtenons ici l'équation des géodésiques dans un cadre général.

Chapitres 4 et 5 : structures sous-riemanniennes sur les espaces de formes

Ces chapitres résument respectivement la deuxième et la troisième partie de [AT], reproduit en partie I dans les sections 9.2 et 9.3.

La notion d'espace de formes est assez vague dans la littérature. Il peut s'agir d'un espace de points [JM00, You10], de courbes [MM07, You10], ou de sous-variétés immergées [BHM11, BHM12]. Tous ces espaces ont un point commun : le groupe des difféomorphismes a une action de groupe naturelle sur ces espaces. Cette action est de plus compatible avec la structure particulière de $\mathcal{D}^s(M)$. Nous nous servons de cette observation pour définir un espace de formes abstrait dans une variété M à géométrie bornée.

5. C'est-à-dire que tout vecteur tangent à M est une combinaison linéaire de crochets de champs de vecteurs horizontaux pour cette structure.

Définition 0.1. Soit \mathcal{S} une variété de Banach métrisable modélée sur un espace de Banach B , muni d'une distance compatible D . On suppose que $\mathcal{D}^s(M)$, où $s > \dim(M)/2 + \ell$ et $\ell \in \mathbb{N}^*$, a une action continue sur \mathcal{S} via

$$(\varphi, q) \mapsto \varphi \cdot q, \quad (\varphi, q) \in \mathcal{D}^s(M) \times \mathcal{S}.$$

On dit que \mathcal{S} est un **espace de formes d'ordre ℓ** si les trois conditions suivantes sont satisfaites :

1. L'action est semi-lipschitzienne.
2. Pour tout q dans \mathcal{S} fixé, l'application $R_q : \varphi \mapsto \varphi \circ q$ est lisse. Sa dérivée en Id_M est notée ξ_q .
3. Pour tout k dans \mathbb{N} , les applications

$$\begin{aligned} R : \mathcal{S} \times \mathcal{D}^{s+k}(M) &\longrightarrow \mathcal{S}, & \xi : \mathcal{S} \times \Gamma^{s+k}(TM) &\longrightarrow T\mathcal{S}, \\ (q, \varphi) &\longmapsto \varphi \cdot q & (q, X) &\longmapsto \xi_q X \end{aligned} \quad (5)$$

sont de classe C^k . On dit que $q \in \mathcal{S}$ est à support compact si $\varphi \circ q$ ne dépend que de la restriction de φ à un compact fixé.

Cette définition permet l'étude de nouveaux types d'espaces de formes, qui peuvent être très utiles pour la modélisation d'objets. Par exemple, le fibré tangent $T\mathcal{S}$ d'un espace de formes \mathcal{S} d'ordre ℓ a une structure naturelle d'espace de formes d'ordre $\ell+1$ par la différentielle de l'action. On peut se servir de cela pour étudier le mouvement de fibres sur une forme, par exemple les fibres musculaires d'un coeur.

Soit alors $(V, \langle \cdot, \cdot \rangle)$ un sous-espace de Hilbert de $\Gamma(TM)$ avec inclusion continue dans l'espace des champs de vecteurs de classe H^{s+k} , pour $k \in \mathbb{N}^*$. On déduit de l'action de $\mathcal{D}^s(M)$ sur \mathcal{S} une structure sous-riemannienne forte sur \mathcal{S} et invariante par cette action, définie par le triplet

$$(\mathcal{S} \times V, \xi, \langle \cdot, \cdot \rangle).$$

Les courbes horizontales pour cette structure sont les solutions de

$$\dot{q}(t) = \xi_{q(t)} X(t), \quad X(t) \in V,$$

et sont exactement de la forme

$$q(t) = \varphi^X(t) \cdot q(0),$$

où φ^X est le flot de X . On note $d^{\mathcal{S}}$ la distance sous-riemannienne correspondante.

Théorème 0.5. Soient $q_0 \in \mathcal{S}$ à support compact, et \mathcal{O}_{q_0} l'orbite de q_0 , c'est-à-dire l'ensemble des états $q \in \mathcal{S}$ tels que $d^{\mathcal{S}}(q, q_0) < \infty$. Alors, pour tous $q, q' \in \mathcal{O}_{q_0}$, il existe une géodésique minimisante reliant q à q' . De plus, $(\mathcal{O}_{q_0}, d^{\mathcal{S}})$ est un ensemble métrique complet.

On donne ensuite l'équation hamiltonienne des géodésiques pour cette structure. On mentionne également un exemple très utilisé d'espace de formes : l'espace des plongements continus d'une même sous-variété. Ces résultats unifient dans un cadre bien plus général des résultats antérieurs [TY11, You10].

Dans le chapitre 5, on définit une autre sorte de structure sous-riemannienne sur un espace de formes : les structures relevées par des submersions équivariantes. Il s'agit d'un analogue aux structures sous-riemanniennes définies par une connection de Ehresmann sur un fibré sur une variété riemannienne [Mon02].

Il s'agit là d'un cadre nouveau pour l'analyse des déformations, qui admet de nombreux cas applicatifs inédits en permettant la modélisation de formes "d'ordre supérieur". Un exemple de submersion équivariante est donnée par la projection canonique $\pi : T\mathcal{S} \rightarrow \mathcal{S}$. La structure relevée à $T\mathcal{S}$ par π , par exemple, peut être utilisée pour étudier comment la déformation d'un objet déforme des fibres sur cet objet. On peut donc utiliser ce cadre pour modéliser la contraction d'un coeur.

Chapitres 6 et 7 : espaces de formes contraintes et algorithmes de minimisation

Ces chapitres résument [ATTY13], reproduit en partie I au chapitre 10. On travaille dans le cadre simplifié suivant : $M = \mathbb{R}^d$, \mathcal{S} est un espace de formes d'ordre $\ell \in \mathbb{N}^*$ dans \mathbb{R}^d et un ouvert d'un espace de Banach B .

Rappelons qu'en analyse des déformations, on veut minimiser une fonctionnelle

$$J(X) = \frac{1}{2} \int_0^1 \langle X(t), X(t) \rangle dt + g(\varphi^X(1) \cdot q_0) = \frac{1}{2} \int_0^1 \langle X(t), X(t) \rangle dt + g(q(1)),$$

où $X(t) \in V$ un espace de Hilbert de champs de vecteurs, $\dot{q}(t) = \xi_{q(t)} X(t)$ et $q(0) = q_0 \in \mathcal{S}$. De nombreuses techniques ont été développées pour minimiser J dans diverses exemples d'espaces de formes. Toutefois, l'idée d'ajouter sur les déformations possibles des contraintes qui varient avec la forme est nouvelle. Le seul papier la mentionnant est [You12], et il s'agit de contraintes très particulières, essentiellement utilisées en dimension finie.

On souhaite donc étendre notre cadre à des espaces de formes avec contraintes. Autrement dit, minimiser J sur les champs de vecteurs $t \mapsto X(t)$ tels que

$$C_{q(t)} X(t) = 0$$

pour presque tout t , où $C_q \in \mathcal{L}(B, Y)$. On voit tout d'abord que sous des hypothèses très faibles, un minimiseur contraint de J existe toujours.

Théorème 0.6. *Si V a une inclusion continue dans $H^{s+1}(\mathbb{R}^d, \mathbb{R}^d)$ avec $s > d/2 + \ell$, et si $q \mapsto C_q$ est continue, g est semicontinue inférieurement et bornée inférieurement, et q_0 est à support compact, alors J possède au moins un minimiseur contraint.*

On prouve également que dans la plupart des cas, on peut approcher le minimiseur contraint de J par une suite de minimiseurs contraints sur des espaces de formes dimension finie.

Le résultat principal de ce chapitre est une version adaptée à la dimension infinie du principe du maximum de Pontryagin avec contraintes. Pour cela, notons $K^V : V^* \rightarrow V$ l'isométrie induite par $\langle \cdot, \cdot \rangle$, et $K_q = \xi_q K^V \xi_q^* : T_q^* M \rightarrow T_q M$. L'opérateur K_q se calcule facilement à l'aide du noyau reproduisant de V . On définit le hamiltonien du problème $h : T^* \mathcal{S} \rightarrow \mathbb{R}$ par

$$h(q, p) = \inf_{X \in \ker C_q} \left\{ p \xi_q X - \frac{1}{2} \langle X, X \rangle \right\}.$$

Théorème 0.7. *Supposons que pour tout $q \in \mathcal{S}$, C_q soit surjective. Soit $X \in L^2(0, 1; V)$ un minimiseur de J le long des contraintes. Alors il existe $p \in H^1(0, 1; B^*)$ et $\lambda \in L^2(0, 1; Y^*)$ tels que*

$$p(1) + dg(q(1)) = 0$$

et

$$\begin{cases} \dot{q}(t) = \partial_p h(q(t), p(t)) &= K_{q(t)} p(t) - \xi_{q(t)} K^V C_{q(t)}^* \lambda(t), \\ \dot{p}(t) = -\partial_q h(q(t), p(t)) &= -\partial_q (K_{q(t)} X(t))^* p(t) + \partial_q (C_{q(t)} X(t))^* \lambda(t) \end{cases}$$

presque partout sur $[0, 1]$, où le **multiplicateur de Lagrange** λ est donné par

$$C_{q(t)} K^V C_{q(t)}^* \lambda(t) = C_{q(t)} K^V \xi_{q(t)}^* p(t),$$

et

$$X(t) = K^V (\xi_{q(t)}^* p(t) - C_{q(t)}^* \lambda(t)).$$

Dans le chapitre 7, on développe et on compare deux types d'algorithmes pour la minimisation de J , à la fois dans le cas avec contraintes et le cas sans contraintes. La première méthode consiste à minimiser directement J sur l'espace total des contrôles $L^2(0, 1; V)$. La deuxième, appelée méthode du tir, utilise le principe du maximum de Pontryagin pour se restreindre aux géodésiques partant de q_0 , paramétrées par $T_{q_0}^* \mathcal{S}$.

Le premier algorithme requiert le calcul du gradient de J sur l'espace $L^2(0, 1; V)$. On obtient

$$\nabla J(X)(t) = X(t) - K^V \xi_{q(t)}^* p(t), \quad X \in L^2(0, 1; V),$$

où $\dot{q}(t) = \xi_{q(t)} X(t)$, et $p \in H^1(0, 1; B^*)$ tel que $p(1) + dg(q(1)) = 0$ et $\dot{p}(t) = -\frac{1}{2} \partial_q(p(t) K_{q(t)} p(t))$ presque partout. Un algorithme du gradient permet alors d'obtenir un minimiseur sans contraintes.

Il s'agit d'une formule connue, utilisée avec succès sur de nombreux exemples dans les méthodes LDDMM [TY11, You10]. L'avantage de notre version est qu'elle est valide pour tous les espaces de formes, et même pour des problèmes de contrôle optimal plus généraux.

Si il y a des contraintes, c'est plus compliqué, et comme les formes contraintes n'ont jamais été étudiées, nous avons appliqué la *méthode du lagrangien augmenté* [IK08]. Plus précisément, il s'agit de minimiser successivement des fonctionnelles de la forme

$$J_A(X, \lambda_r, \mu_r) = \frac{1}{2} \int_0^1 \left(\langle X(t), X(t) \rangle - \lambda_r(t) C_{q(t)} X(t) + \frac{1}{2\mu_r} \|C_{q(t)} X(t)\|_Y^2 \right) dt + g(q(1)),$$

où $\lambda_r \in L^2(0, 1; Y)$ et $\mu_r > 0$. On obtient, par exemple en utilisant un algorithme du gradient, un contrôle optimal X_r sans contraintes et une trajectoire q_r associée qui minimisent cette fonctionnelle. On met ensuite à jour $\lambda_{r+1} = \lambda_r - \frac{1}{\mu_r} C_{q_r} X_r$ et $\mu_{r+1} < \mu_r$. Alors, sous des hypothèses appropriées, la suite X_r converge vers un contrôle optimal contraint X , et la suite λ_r vers le multiplicateur de Lagrange associé à ces contraintes.

Le second algorithme développé, la méthode du tir, utilise notre version du principe du maximum de Pontryagin [PBG64, AS04, Tré08], selon lequel un minimiseur contraint de J satisfait

$$\exists p \in H^1(0, 1; B^*), \quad (\dot{q}(t), \dot{p}(t)) = \nabla^\omega h(q(t), p(t)) = (\partial_p h(q(t), p(t)), -\partial_q h(q(t), p(t))).$$

Cette équation différentielle a une unique solution maximale à condition initiale $p_0 \in T_{q_0}^* \mathcal{S} = B^*$ fixée. Si une telle courbe est définie sur $[0, 1]$ entier, alors il existe un contrôle $X \in L^2(0, 1; V)$ de trajectoire associée q , tel que

$$J(X) = \frac{1}{2} h(q_0, p_0) + g(q(1)) = \tilde{J}(p_0).$$

Minimiser J revient donc à minimiser \tilde{J} sur B^* . Il suffit, là encore, de calculer le gradient de \tilde{J} , que l'on peut ensuite minimiser. Cela s'effectue par la résolution d'une équation adjointe.

Le calcul du gradient de \tilde{J} est plus lent que celui des fonctionnelles J_A , mais permet une satisfaction totale des contraintes, plutôt qu'une satisfaction approchée. Des simulations numériques de ces algorithmes sont ensuite fournies, sur des exemples ayant des applications en anatomie computationnelle.

L'analyse des déformations avec contraintes promet de très nombreuses applications. La plus importante semble être la modélisation de multi-formes : l'étude simultanée d'objets indépendants, chacun étant déformé par un flot de champ de vecteurs distinct, mais possédant des interactions les uns avec les autres. Par exemple, pour étudier simultanément l'hypothalamus et l'hippocampe de différents cerveaux humains, il est préférable de les modéliser comme deux objets distincts. Toutefois, ces objets sont immergés dans un milieu continu (le cerveau lui-même), ce qui les empêche de se toucher. On peut modéliser cette configuration par des contraintes sur des multi-formes, voir section 6.2.

Première partie

Présentation des travaux

Chapitre 1

Géométrie sous-riemannienne en dimension finie

On commence par rappeler quelques résultats de géométrie sous-riemannienne en dimension finie. Cela permettra de mieux comprendre les difficultés qui se posent en dimension infinie.

Une variété sous-riemannienne est une variété lisse sur laquelle on astreint les vitesses des courbes à appartenir à un sous-fibré du fibré tangent, appelé le fibré *horizontal*. De telles courbes sont dites **horizontales**. Ce sous-fibré est également équipé d'une métrique, qui permet de définir les notions usuelles de longueur, d'action, de distance et de géodésique comme sur une variété riemannienne.

Après avoir rappelé la définition usuelle d'une structure sous-riemannienne, on en donnera une version plus générale, tirée de [ABC⁺10] que nous utiliserons dans la suite du manuscrit. Elle permet d'étudier des cas où la dimension des fibres horizontales peut varier de points en points, et est particulièrement adaptée à une généralisation en dimension infinie. On évoquera ensuite le théorème de Chow-Rashevski, qui donne une condition suffisante pour que tout couple de points d'une variété sous-riemannienne puissent être joints par une courbe horizontale.

Nous définirons ensuite l'application *point final* E , qui à une courbe horizontale associe son point final. Cela nous permettra d'énoncer le *principe du maximum de Pontryagin* (ou PMP). Celui-ci établit qu'une géodésique qui n'est pas une courbe singulière pour E satisfait l'équation hamiltonienne des géodésiques. On formulera ce principe à l'aide des gradients symplectiques du hamiltonien de la structure.

Nous finirons en donnant une réciproque partielle au principe du maximum : les courbes satisfaisant l'équation hamiltonienne des géodésiques sont effectivement des géodésiques.

1.1 Premières définitions

Définition 1.1. *Une variété sous-riemannienne est un triplet (M, Δ, g) , avec M une variété lisse de dimension finie, $\Delta \subset TM$ un sous-fibré du fibré tangent, et g une métrique riemannienne sur Δ . [Mon02, BR96]*

Remark 1.1. Lorsque $\Delta = TM$, (M, g) est une variété riemannienne.

On appelle Δ la *distribution horizontale*, et on note k la dimension de ses fibres. Un objet tel qu'un champ de vecteur ou une courbe absolument continue est dit *horizontal* s'il est tangent à Δ presque partout.

On procède ensuite comme sur une variété riemannienne. On définit la longueur $L(q)$ d'une courbe horizontale $q : I \rightarrow M$ par

$$L(q) = \int_I \sqrt{g_{q(t)}(\dot{q}(t), \dot{q}(t))} dt.$$

On définit ensuite la distance sous-riemannienne $d(q_0, q_1)$, $q_0, q_1 \in M$:

$$d(q_0, q_1) = \inf L(q), \tag{1.1}$$

où l'infimum est pris sur l'ensemble des courbes horizontales entre q_0 et q_1 .

Définition 1.2. Une courbe horizontale $q : [a, b] \rightarrow M$ telle que $L(q) = d(q(a), q(b))$ est appelée une **géodésique minimisante**.

Une **géodésique** est une courbe $q : I \rightarrow M$, I intervalle, qui est une géodésique minimisante sur des sous-arcs suffisamment petits : pour deux temps $t_0 < t_1$ assez proches, $d(q(t_0), q(t_1)) = L(q|_{[t_0, t_1]})$, où $q|_{[t_0, t_1]}$ est la restriction de q au sous-intervalle $[t_0, t_1]$.

On peut également définir l'action d'une telle courbe horizontale par

$$A(q) = \frac{1}{2} \int_I \|\dot{q}(t)\|_{q(t)}^2 dt. \tag{1.2}$$

L'inégalité de Cauchy-Schwarz implique que, pour une courbe paramétrée telle que $t \mapsto \|\dot{q}(t)\|_{q(t)}$ soit constante, minimiser l'action à extrémités fixées revient à minimiser la longueur. En conséquence, pour trouver les géodésiques sous-riemanniennes, il suffit de trouver les courbes qui minimisent de l'action. Celle-ci a généralement de meilleures propriétés que la longueur.

Ces définitions ne comprennent toutefois pas certaines structures plus générales que nous souhaitons étudier dans la suite de cette thèse. Par exemple, le cas où la dimension de la fibre Δ_q peut varier est ici exclus.

Un exemple classique est celui du plan de Grushin, dont la structure est dite *presque riemannienne* sur \mathbb{R}^2 , avec $\Delta_{(x,y)}$ l'espace engendré par les vecteurs $\frac{\partial}{\partial x}$ et $x \frac{\partial}{\partial y}$.

1.2 Un cadre plus général

Par la suite, si $E \rightarrow M$ est un fibré vectoriel sur M , on notera $\Gamma(E)$ l'espace des sections $e : M \rightarrow E$ de ce fibré.

1.2.1 Fibrés tangents relatifs

Pour étudier des cas où la dimension de Δ est variable, Agrachev, Boscain et al. ont introduit dans [ABC⁺10] une définition plus générale, reposant sur la notion de *fibré tangent relatif*.

Définition 1.3. Un **fibré tangent relatif** sur une variété M est la donnée d'un couple (\mathcal{H}, ξ) , où $\pi : \mathcal{H} \rightarrow M$ est un fibré vectoriel lisse sur M , et $\xi : \mathcal{H} \rightarrow TM$ est un morphisme de fibrés lisses au-dessus de M . L'image $\Delta = \xi(E) \subset TM$ de ce morphisme est appelée la **distribution horizontale**.

Bien qu'Agachev, Boscain et al. n'utilisent pas le terme de "fibré tangent relatif", il s'agit d'une notion bien connue dans l'étude des algèbroïdes de Banach-Lie (voir par exemple [Mac05]). Les définitions de champ de vecteurs horizontal et de courbe horizontale sont adaptées à ce cadre comme suit.

Définition 1.4. Soit (\mathcal{H}, ξ) un fibré tangent relatif sur une variété M . Un champ de vecteur $X \in \Gamma(TM)$ est dit **horizontal** s'il existe une section $u \in \Gamma(\mathcal{H})$ telle que $X = \xi u$.

Une courbe **horizontale** est une courbe absolument continue $q : I \rightarrow M$ au-dessus de laquelle il existe une application mesurable $(q, u) : I \rightarrow \mathcal{H}$ telle que

$$\dot{q}(t) = \xi_{q(t)}u(t), \quad p.p. t \in I.$$

On appelle le couple (q, u) un **système horizontal de trajectoire** q , et de **contrôle** u .

1.2.2 Structure sous-riemannienne sur une variété, systèmes horizontaux

Définition 1.5. Une structure sous-riemannienne sur une variété M est un triplet (\mathcal{H}, ξ, g) , où (\mathcal{H}, ξ) est un espace tangent relatif sur M , et g est une métrique riemannienne sur le fibré vectoriel \mathcal{H} .

On définit alors la *longueur* d'un système horizontal $(q, u) : I \rightarrow \mathcal{H}$ par

$$L(q, u) = \int_I \sqrt{g_{q(t)}(u(t), u(t))} dt,$$

et son *action* par

$$A(q, u) = \frac{1}{2} \int_I g_{q(t)}(u(t), u(t)) dt.$$

On peut ensuite définir, comme dans la section 1.1 qui précède, la *distance sous-riemannienne* d entre deux points de M comme l'infimum des longueurs de systèmes horizontaux dont la trajectoire connecte ces deux points. De plus, minimiser la longueur revient ici encore à minimiser l'action.

1.2.3 Longueur et action des courbes horizontales

Si ξ n'est pas injective, une même courbe horizontale q peut être la trajectoire de plusieurs systèmes horizontaux, car plusieurs contrôles u peuvent engendrer la même trajectoire.

Pour définir directement l'action et la longueur d'une courbe horizontale q , il faut prendre

$$A^M(q) = \inf A(q, u), \quad L^M(q) = \inf L(q, u),$$

où l'infimum est pris sur l'ensemble des systèmes horizontaux de trajectoire q .

On peut calculer A^M et L^M d'une façon un peu plus directe. Notons $\Delta_q = \xi_q(\mathcal{H}_q) \subset T_qM$, pour q un point de M . Alors ξ_q induit une norme euclidienne sur Δ_q par

$$\|w\|_q^2 = \inf_{u \in \mathcal{H}_q, \xi_q u = w} g_q(u, u).$$

On obtient alors, pour toute courbe horizontale $q : I \rightarrow M$,

$$A^M(q) = \frac{1}{2} \int_I \|\dot{q}(t)\|_{q(t)}^2 dt, \quad L^M(q) = \int_I \|\dot{q}(t)\|_{q(t)} dt.$$

Remark 1.2. En utilisant cette définition, on peut montrer facilement qu'il existe un unique contrôle u_q tel que

$$A^M(q) = A(q, u_q).$$

Il est caractérisé pour presque tout t par les conditions

$$\forall u \in \mathcal{H}_{q(t)}, \quad g_{q(t)}(u_q(t), u) = 0.$$

La distance sous-riemannienne entre deux points de M est donc aussi égale à l'infimum des longueurs des courbes horizontales. On définit alors les géodésiques et les géodésiques minimisantes comme dans la section 1.1.

Toutefois, comme on le verra plus loin, l'espace des systèmes horizontaux (q, u) a une meilleure structure que celui des courbes horizontales. On travaillera donc sur les systèmes (q, u) plutôt que sur les courbes q .

1.3 Orbite et théorème de Chow-Rashevski

Dans toute la suite, M désigne une variété lisse, et (\mathcal{H}, ξ, g) une structure sous-riemannienne lisse sur M . On notera également $\Delta = \xi(\mathcal{H}) \subset TM$ la distribution horizontale, et k la dimension des fibres de \mathcal{H} .

Il n'est pas acquis que deux points puissent toujours être joints par une courbe horizontale.

Définition 1.6. *Soit q un point de M . L'orbite de q , noté \mathcal{O}_q , est l'ensemble des points q' qui peuvent être joints à q par une courbe horizontale.*

Notons que $q' \in \mathcal{O}_q$ si et seulement si $\mathcal{O}_q = \mathcal{O}_{q'}$ si et seulement si $d(q, q') < +\infty$.

Désignons également par Δ à la fois l'image $\xi(\mathcal{H}) \subset TM$ de ξ et les champs de vecteurs de classe \mathcal{C}^∞ et horizontaux. Comme ces deux objets se déduisent naturellement l'un de l'autre, il n'y a pas de confusion. Le crochet de Lie de champs de vecteurs permet alors de définir la sous-algèbre de Lie $\mathcal{L} \subset \Gamma(TM)$ de champs de vecteurs engendrée par Δ .

En tout point q de M , on en déduit un sous-espace $\mathcal{L}_q \subset T_qM$ défini par

$$\mathcal{L}_q = \{X(q), X \in \mathcal{L}\}.$$

On a alors le théorème suivant, dû à Sussmann.

Théorème 1.1. Théorème de l'orbite de Sussmann [Sus73] *Pour tout q dans M , \mathcal{O}_q est une sous-variété lisse immergée de M , et $\mathcal{L}_q \subset T_q\mathcal{O}_q$. Si \mathcal{H} et ξ sont analytiques, on a même $\mathcal{L}_q = T_q\mathcal{O}_q$.*

On en déduit le théorème de Chow-Rashevski.

Théorème 1.2. Théorème de Chow-Rashevski [Mon02, BR96, AS04, Tré08] *Si, pour tout q dans M , $\mathcal{L}_q = T_qM$, alors tout couple de points de M peuvent être joints par une courbe horizontale.*

1.4 Structure de l'ensemble des courbes horizontales, application point final et adjoint

Pour trouver une géodésique reliant deux points q_0 et q_1 de M , il faut minimiser la fonctionnelle d'action A sur l'ensemble $\Omega_{q_0, q_1}^{\mathcal{H}}$ des systèmes horizontaux à extrémités fixés q_0 et q_1 . Il convient donc de chercher les points critiques de A sur cet ensemble. Pour une variété riemannienne, cela conduit aux équations géodésiques habituelles. Pour une variété sous-riemannienne, un autre cas peut se présenter.

1.4.1 Espace des courbes horizontales

Commençons par définir l'ensemble $H^1 \times L^2(0, 1; \mathcal{H})$ comme l'ensemble des applications mesurables $(q, u) : [0, 1] \rightarrow \mathcal{H}$ de classe L^2 (dans une trivialisatoin) et dont la projection $q : [0, 1] \rightarrow M$ est de classe de Sobolev H^1 . Dans ce cas, si (q, u) est un système horizontal, alors

$$\int_0^1 g_{q(t)}(u(t), u(t)) dt < +\infty.$$

Cela découle du fait que l'image de q est compacte. Rappelons que $H^1 \times L^2(0, 1; \mathcal{H})$ est une variété lisse modelée sur un espace de Hilbert (voir par exemple [EM70, Mon02]). C'est même un fibré vectoriel d'espaces de Hilbert au-dessus de la variété de Hilbert $H^1(0, 1; M)$.

Notons qu'une courbe horizontale $(q, u) : [0, 1] \rightarrow \mathcal{H}$ d'action $A(q, u)$ finie appartient à $H^1 \times L^2(0, 1; \mathcal{H})$. On a alors le théorème suivant, dû à Bismut [Bis84].

Théorème 1.3. *L'ensemble $\Omega_{q_0}^{\mathcal{H}}$ des courbes horizontales $(q, u) : [0, 1] \rightarrow \mathcal{H}$ d'action finie, et de point initial $q(0) = q_0$, est une sous-variété de $H^1 \times L^2(0, 1; \mathcal{H})$.*

Preuve. Nous nous contentons de prouver le résultat plus faible suivant : si (\bar{q}, \bar{u}) est un système horizontal avec \bar{q} simple, alors on peut trouver une paramétrisation lisse de $\Omega_{q_0}^{\mathcal{H}}$ au voisinage de (\bar{q}, \bar{u}) par un ouvert de $L^2(0, 1; H)$ où H est un espace vectoriel isomorphe à une fibre de \mathcal{H} . Pour la preuve complète incluant le cas où \bar{q} a des intersections, voir l'appendice D de [Mon02].

On fixe $(\bar{q}, \bar{u}) \in \Omega_{q_0}^{\mathcal{H}}$, et on suppose que la trajectoire \bar{q} est simple. On peut donc considérer U un voisinage de $\text{im}(\bar{q})$ que l'on peut identifier, grâce à une carte, à un ouvert de \mathbb{R}^d , où d est la dimension de M . Un tel ouvert sera appelé un *système de coordonnées* au voisinage de $\text{im}(\bar{q})$, et l'identification d'un tel ouvert avec un ouvert de \mathbb{R}^d sera sous-entendue.

On peut alors définir une trivialisaton $\mathcal{H}|_U \simeq U \times H$ de \mathcal{H} au-dessus de U , où H est un espace vectoriel isomorphe à une fibre de \mathcal{H} . Dans cette trivialisaton, on peut identifier ξ à une application $U \mapsto \mathcal{L}(H, \mathbb{R}^d)$. On peut également identifier \bar{u} à un élément de $L^2(0, 1; H)$.

Dans ces coordonnées, on a toujours

$$\dot{\bar{q}}(t) = \xi_{\bar{q}(t)} \bar{u}(t), \quad \text{p.p. } t \in [0, 1].$$

Les théorèmes usuels de Cauchy-Lipshitz à paramètres dans le cas d'un champ de vecteur intégrable en temps et lisse en espace (voir par exemple [Mon02], appendice D, ou [Tré08]) impliquent alors l'existence d'un voisinage \mathcal{U} de \bar{u} dans $L^2(0, 1; H)$ tel que le problème de Cauchy

$$\dot{q}(t) = \xi_{q(t)} u(t), \quad q(0) = q_0, \quad u \in \mathcal{U}$$

ait une unique solution $t \mapsto q(t, u)$, appartenant à $H^1(0, 1; U)$. On appelle $u \mapsto q(\cdot, u)$ l'*application trajectoire*. De plus, l'application $\mathcal{U} \rightarrow H^1 \times L^2(0, 1; \mathcal{H})$ définie par

$$u \in \mathcal{U} \mapsto (q(\cdot, u), u)$$

est lisse, injective et d'image localement fermée.

Dans le système de coordonnées U , la dérivée de l'application trajectoire $u \mapsto q(\cdot, u)$ en \bar{u} dans la direction $\delta u \in L^2(0, 1; H)$, notée

$$\delta q = \partial_u(q(\cdot, \bar{u})) \delta u \in H^1(0, 1; \mathbb{R}^d),$$

est donnée par la solution du problème de Cauchy linéaire sur $[0, 1]$ suivant :

$$\dot{\delta q}(t) = \partial_q(\xi_{\bar{q}(t)} \bar{u}(t)) \delta q(t) + \xi_{\bar{q}(t)} \delta u(t), \quad \delta q(0) = 0. \quad (1.3)$$

Notons qu'à $u \in H$ fixé, $q \mapsto \partial_q(\xi_q u)$ est une application à valeur dans l'espace $\mathcal{L}(\mathbb{R}^d, \mathbb{R}^d)$ des endomorphismes linéaires de \mathbb{R}^d .

La méthode de variation de la constante montre que la solution de cette équation est donnée par

$$\delta q(t) = R(t) \int_0^t R(s)^{-1} \xi_{q(s)} \delta u(s) ds, \quad (1.4)$$

où R est la résolvante du système homogène, c'est-à-dire la courbe sur $[0, 1]$ à valeurs dans $M_d(\mathbb{R})$ solution du problème de Cauchy

$$\dot{R}(t) = \partial_q(\xi_{q(t)} u(t)) R(t), \quad R(0) = I_d,$$

avec I_d la matrice identité.

On obtient finalement à la fois une paramétrisation locale et une carte d'un voisinage de \bar{q} dans $\Omega_{q_0}^{\mathcal{H}}$. On a prouvé que $\Omega_{q_0}^{\mathcal{H}}$ est localement une sous-variété immergée. On admettra que $\Omega_{q_0}^{\mathcal{H}}$ est bien une véritable sous-variété au sens global (pas d'auto-intersections par exemple). \square

Remark 1.3. Cette preuve reste vraie telle quelle dans le cas de structures sous-riemanniennes sur une variété M de dimension infinie, modélisée sur un espace de Banach B . Il suffit de remplacer chaque occurrence de \mathbb{R}^d par B .

L'esprit de la preuve reste le même lorsque q n'est pas une courbe simple, mais certaines difficultés techniques doivent être surmontées, voir [Mon02]. Toutefois, la méthode pour surmonter ces difficultés ne se généralise pas à la dimension infinie.

Cela dit, les géodésiques minimisantes sont toujours simples, et les courbes horizontales avec auto-intersection ne sont donc pas intéressantes quand il s'agit de calculer la distance sous-riemannienne entre deux points. On travaillera donc uniquement au voisinage de courbes horizontales simples à partir de maintenant.

Pour un système U de coordonnées au voisinage de $\text{im}(q)$, avec $(q, u) \in \Omega_{q_0}^{\mathcal{H}}$ fixé, on appelle *espace de contrôle* induit par U une carte

$$\mathcal{U} \subset L^2(0, 1; H)$$

de $\Omega_{q_0}^{\mathcal{H}}$ au voisinage de (q, u) obtenue comme dans la preuve ci-dessus.

Remark 1.4. Le choix d'un repère (e_1, \dots, e_k) de $\mathcal{H}|_U$, orthonormé par rapport à la métrique sous-riemannienne g , permet d'identifier \mathcal{U} à un ouvert de $L^2(0, 1; \mathbb{R}^k)$, sur lequel

$$A(q(\cdot), u) = \frac{1}{2} \sum_{i=1}^k \int_0^1 (u^i(t))^2 dt = \frac{1}{2} \sum_{i=1}^k \int_0^1 u(t)^T u(t) dt, \quad u \in \mathcal{U} \subset L^2(0, 1; \mathbb{R}^k).$$

Dans la dernière égalité, on a considéré $u(t) \in \mathbb{R}^k$ comme une matrice colonne de taille d de transposée $u(t)^T$.

1.4.2 Application point final

Définissons maintenant l'application *point final*, notée E , qui à un système horizontal $(q, u) \in \Omega_{q_0}^{\mathcal{H}}$ associe le point final de sa trajectoire $q(1)$. Cette application est lisse.

Pour q_1 élément de M , on note $\Omega_{q_0, q_1}^{\mathcal{H}} = E^{-1}(\{q_1\})$ l'ensemble des systèmes horizontaux dont les trajectoires joignent q_0 et q_1 . Les géodésiques minimisantes de q_0 à q_1 sont alors exactement les minimiseurs de A sur $\Omega_{q_0, q_1}^{\mathcal{H}}$. Il est donc logique de chercher les points critiques de la restriction de A à $\Omega_{q_0, q_1}^{\mathcal{H}}$.

Toutefois, contrairement au cas riemannien, E n'est pas forcément une submersion de $\Omega_{q_0}^{\mathcal{H}}$ sur M en tout point. Cela signifie que $\Omega_{q_0, q_1}^{\mathcal{H}}$ n'est pas forcément une variété lisse, et que l'on ne peut pas appliquer le théorème des extrema liés tel quel pour obtenir tous les points critiques possibles : on doit rajouter la possibilité de points critiques, ou *anormaux*.

Lemme 1.1. Soient \mathcal{M} une variété de Banach, $F : \mathcal{M} \rightarrow \mathbb{R}$ et $G : \mathcal{M} \rightarrow M$ deux applications de classe \mathcal{C}^1 . Soient $q_1 \in M$ et u un élément de \mathcal{M} qui minimise F sur l'ensemble $G^{-1}(\{q_1\})$. Alors il existe $p_1 \in T_{q_1}^* M$ tel qu'au moins l'une des deux possibilités suivantes soit satisfaite :

1. $dF(u) = dG(u)^* p_1$. On dit que p_1 est le multiplicateur de Lagrange du problème en u .
2. $dG(u)^* p_1 = 0$ avec $p_1 \neq 0$. Dans ce cas, u est un point critique de G , et $G^{-1}(\{q_1\})$ n'est peut-être pas une sous-variété régulière au voisinage de u .

En effet, si on n'est pas dans le cas 2, $dG(u)$ est surjective et on peut appliquer le théorème des extrema liés.

Nous nous intéressons bien entendu au cas $\mathcal{M} = \Omega_{q_0}^{\mathcal{H}}$, $F = A$, et $G = E$. Pour utiliser le lemme 1.1, il nous faut donc calculer l'adjoint de dE . Pour cela, on a le résultat crucial suivant. De nombreux résultats de cette thèse sont conséquences de variantes en dimension infinie de la proposition suivante.

Proposition 1.1. Soit $(q, u) \in \Omega_{q_0}^{\mathcal{H}}$ un système horizontal dont la trajectoire est simple. Soit $p_1 \in T_{q(1)}^*M$. On définit $P \in T_{(q,u)}^*\Omega_{q_0}^{\mathcal{H}}$ par

$$P(\delta q, \delta u) = (dE_{q,u})^*p_1(\delta q, \delta u) = p_1\delta q(1), \quad (\delta q, \delta u) \in T_{(q,u)}\Omega_{q_0}^{\mathcal{H}}.$$

Soit \mathcal{U} un espace de contrôles induit par un système de coordonnées U au voisinage de $im(q)$, identifié à un ouvert de \mathbb{R}^d .

Alors, dans ces coordonnées, on a

$$(dE_{q,u})^*p_1(\delta q, \delta u) = \int_0^1 p(t)\xi_{q(t)}\delta u(t)dt,$$

où $(q, p) \in H^1 \times H^1(0, 1; T^*M)$ et p est l'unique solution sur $[0, 1]$ du problème de Cauchy linéaire adjoint

$$\dot{p}(t) = \partial_q(\xi_{q(t)}u(t))^*p(t) = -p(t)\partial_q(\xi_{q(t)}u(t)), \quad p(1) = p_1, \quad (1.5)$$

où on a identifié $p(t)$ à une matrice ligne de taille d .

Remark 1.5. L'équation (1.5) dépend de la trivialisaton $\mathcal{H}|_U \simeq U \times H$ choisie et du système de coordonnées U . Ce lemme n'a donc de sens qu'en coordonnées.

Preuve. La preuve est une application directe de (1.4). En effet, si $(\delta q, \delta u) \in T_{(q,u)}\Omega_{q_0}^{\mathcal{H}}$, alors

$$\dot{\delta q}(t) = \partial_q(\xi_{q(t)}u(t))\delta q(t) + \xi_{q(t)}\delta u(t), \quad \text{p.p. } t \in [0, 1], \quad \delta q(0) = 0.$$

De plus, $dE_{(q,u)}(\delta q, \delta u) = \delta q(1)$.

En conséquence,

$$(\partial_u q(1, u))^*P(\delta u) = \int_0^1 p_1 R(1)R(t)^{-1}\xi_{q(t)}\delta u(t)dt,$$

où $R : [0, 1] \rightarrow M_d(\mathbb{R})$ est la solution du problème de Cauchy linéaire

$$\dot{R}(t) = \partial_q(\xi_{q(t)}u(t))R(t), \quad R(0) = I_d.$$

Or p est solution de (1.5) si et seulement si $p(t) = p_1 R(1)R(t)^{-1}$, ce qui implique notre proposition. \square

Soit donc $\mathcal{U} \subset L^2(0, 1; \mathbb{R}^k)$ un espace de contrôle induit par un système de coordonnées U au voisinage de $Im(q)$ et un repère orthonormé (e_1, \dots, e_k) de $\mathcal{H}|_U$. Soient aussi $q_0, q_1 \in M$ et $(q, u) \in \Omega_{q_0, q_1}^{\mathcal{H}}$. Dans ce cas, on a

$$A(q, u) = A(u) = \frac{1}{2} \sum_{i=1}^k \int_0^1 (u^i(t))^2 dt.$$

L'action A est donc lisse, de dérivée directionnelle

$$dA(u)\delta u = \sum_{i=1}^k \int_0^1 u^i(t)\delta u^i(t)dt = \int_0^1 u(t)^T \delta u(t)dt.$$

En appliquant le lemme 1.1 et la proposition 1.1, on obtient une condition nécessaire pour que la trajectoire q de (q, u) soit une géodésique minimisante.

Lemme 1.2. Soit (q, u) un système horizontal dont la trajectoire q est simple et minimise l'action A sur $\Omega_{q_0, q_1}^{\mathcal{H}}$. Soit $\mathcal{U} \subset L^2(0, 1; \mathbb{R}^k)$ un espace de contrôles au voisinage de (q, u) induit par un système de coordonnées U au voisinage de $\text{im}(q)$ et un repère orthonormé de $\mathcal{H}|_U$. Alors il existe $p : [0, 1] \rightarrow T^*M$ au-dessus de q et de classe H^1 , tel que pour presque tout t dans $[0, 1]$,

$$\dot{p}(t) = -p(t)\partial_q(\xi_{q(t)}u(t)),$$

et tel qu'au moins l'une des deux conditions suivantes soit satisfaite.

1. Pour presque tout t ,

$$u(t)^T = \xi_{q(t)}^* p(t).$$

En particulier, $u(t) = (u^1(t), \dots, u^k(t)) \in \mathbb{R}^k$ est entièrement déterminé par $p(t)$, avec $u^i(t) = p(t)\xi_{q(t)}e_i(q(t))$.

2. $p(0) \neq 0$ (ce qui implique $p(t) \neq 0$ pour tout t), et $p(t)\xi_{q(t)} = 0$. Autrement dit, $\xi_{q(t)}(\mathcal{H}_{q(t)}) = \Delta_{q(t)} \subset \ker p(t)$ pour tout t . Le système (q, u) est alors un point critique de E . On dit que q est une trajectoire **singulière**.

Remark 1.6. Bien que ce ne soit pas le sujet principal de cette thèse, les courbes anormales sont des objets très étudiés en géométrie sous-riemannienne depuis que Montgomery a montré [BR96] l'existence de géodésiques singulières minimisantes qui ne satisfont pas l'équation hamiltonienne des géodésiques. Citons par exemple [CJT06], où il est montré que, génériquement, si la dimension de la distribution horizontale est au moins égale à 3, aucune courbe singulière n'est une géodésique. En revanche, si la dimension de cette distribution est égale à 2, alors génériquement, l'ensemble des géodésiques singulières ne satisfaisant pas l'équation hamiltonienne est un sous-ensemble de dimension $d - 4$ [LS95].

Un problème dans cette formulation est qu'elle dépend du choix de l'espace de contrôle. Une formulation hamiltonienne, décrite dans la section 1.5, permet de donner une autre forme à ces conditions, à la fois indépendante du repère et plus compacte.

1.5 Formulation hamiltonienne

1.5.1 Forme symplectique canonique, gradient symplectique

Définition 1.7. Une forme symplectique sur une variété N est une 2-forme différentielle fermée ω telle que

$$\forall (q, w) \in TN, \exists w' \in TN, \quad \omega(w, w') \neq 0.$$

L'espace cotangent T^*M d'une variété M est équipé d'une forme symplectique canonique ω [Arn89]. Toute fonction $f : T^*M \rightarrow \mathbb{R}$ différentiable en (q, p) possède alors un *gradient symplectique* en (q, p) , noté $\nabla^\omega f(q, p) \in T_{(q,p)}T^*M$ et défini implicitement par

$$\omega_{(q,p)}(\nabla^\omega f(q, p), \cdot) = df(q, p).$$

Soit U est un système de coordonnées dans lequel on note $q = (x^1, \dots, x^d)$. Alors U induit un système de coordonnées canoniques

$$(q, p) = (x^1, \dots, x^d, p_1, \dots, p_d) \in U \times (\mathbb{R}^d)^*$$

sur $T^*M|_U$. Dans ces coordonnées, on a

$$\omega(q, p) = dx^1 \wedge dp_1 + \dots + dx^d \wedge dp_d.$$

On en déduit

$$\nabla^\omega f(q, p) = (\partial_p f(q, p), -\partial_q f(q, p)) \in \mathbb{R}^d \times (\mathbb{R}^d)^*.$$

Remarquons que $\partial_p f$ est un élément de $(\mathbb{R}^d)^{**}$, qui s'identifie canoniquement à \mathbb{R}^d . Écrire que $\partial_p f(q, p) \in \mathbb{R}^d$ a donc bien un sens.

Gradient symplectique partiel. Considérons maintenant un cas légèrement plus compliqué. Soit $\mathcal{E} \rightarrow M$ un fibré vectoriel sur M , et notons $T^*M \oplus_M \mathcal{E}$ le fibré sur M dont la fibre en $q \in M$ est la somme directe de T_q^*M et de \mathcal{E}_q . Soit $f : T^*M \oplus_M \mathcal{E} \rightarrow \mathbb{R}$ différentiable en un point (q_0, p_0, e_0) , et telle $\partial_e f(q_0, p_0, e_0) = 0$ (rappelons qu'une dérivée partielle le long d'une fibre peut bien être définie de façon invariante).

On peut alors définir un *gradient symplectique partiel* de f par rapport aux variables (q, p) en (q_0, p_0, e_0) . En effet, considérons une section lisse $(q, p) \mapsto e(q, p)$ du fibré

$$T^*M \oplus_M \mathcal{E} \rightarrow T^*M$$

au voisinage de (q_0, p_0) telle que $e(q_0, p_0) = e_0$. Alors l'application $\tilde{f}(q, p) = f(q, p, e(q, p))$ est différentiable en (q_0, p_0) , et possède donc un gradient symplectique. Celui-ci est donné en coordonnées canoniques par

$$\nabla^\omega \tilde{f} = (\partial_p f + \partial_e f \partial_p e, -\partial_q f - \partial_e f \partial_q e).$$

Comme $\partial_e f(q_0, p_0, e_0) = 0$,

$$\nabla^\omega \tilde{f}(q_0, p_0) = (\partial_p f(q_0, p_0, e_0), -\partial_q f(q_0, p_0, e_0)),$$

qui ne dépend pas de la section e . On note $\nabla^\omega f(q_0, p_0, e_0) \in T_{(q_0, p_0)} T^*M$ ce vecteur, et on l'appelle le *gradient symplectique partiel* de f en (q_0, p_0, e_0) .

1.5.2 Hamiltonien

Définissons maintenant les hamiltoniens $H^\lambda : T^*M \oplus_M \mathcal{H} \rightarrow \mathbb{R}$, avec λ égal à 0 ou 1, par

$$H^\lambda(q, p, u) = p\xi_q u - \frac{\lambda}{2} g_q(u, u).$$

Soit alors un système de coordonnées U sur M identifié à un ouvert de \mathbb{R}^d , et e_1, \dots, e_k un repère orthonormé de $\mathcal{H}|_U$. Dans ces coordonnées, on a

$$H^\lambda(q, p, u) = \sum_{i=1}^k u^i p \xi_q e_i(q) - \frac{\lambda}{2} u^T u.$$

De plus, dans des coordonnées canoniques sur $T^*M|_U$, on peut écrire

$$\begin{cases} \partial_q H^\lambda(q, p, u) = p \partial_q (\xi_q u) & \in (\mathbb{R}^d)^* \\ \partial_p H^\lambda(q, p, u) = \xi_q u & \in T_q M \\ \partial_u H^\lambda(q, p, u) = \xi_q^* p - u^T & \in \mathcal{H}_q^*. \end{cases}$$

Remark 1.7. En particulier, $\partial_q H^0 = \partial_q H^1$ et $\partial_p H^0 = \partial_p H^1$.

On en déduit la reformulation suivante du lemme 1.2, qui est un cas particulier du fameux principe du maximum de Pontryagin [PBG64].

Théorème 1.4. Soit (q, u) un système horizontal minimisant l'action A sur $\Omega_{q_0, q_1}^{\mathcal{H}}$. Alors, il existe $(p, \lambda) \neq (0, 0)$, où $p : [0, 1] \rightarrow T^*M$ de classe H^1 et $\lambda = 0$ ou 1 tels que

$$\begin{cases} 0 & = \partial_u H^\lambda(q(t), p(t), u(t)), \\ (\dot{q}(t), \dot{p}(t)) & = \nabla^\omega H^\lambda(q(t), p(t), u(t)). \end{cases}$$

Si $\lambda = 0$, alors $p(t) \neq 0$ et $\Delta_{q(t)} = \xi_{q(t)}(\mathcal{H}_{q(t)}) \subset \ker p(t)$ pour tout t et (q, u) est un point singulier de E : la trajectoire q est *singulière* (ou *anormale*).

Si $\lambda = 1$, alors u est entièrement déterminé par p , de telle sorte que $g_{q(t)}(u(t), \cdot) = \xi_{q(t)}^* p(t)$. On dit que la courbe (q, u) est une *géodésique normale*.

Cette formulation, elle, est indépendante du système de coordonnées et du repère.

1.5.3 Réciproque partielle

Le principe du maximum ne se généralise pas à la dimension infinie, car, comme on l'évoquera dans la section 1.6, le lemme 1.1 n'est pas vrai en dimension infinie. En revanche, ce lemme admet la réciproque partielle qui, elle, reste vrai en dimension infinie. Commençons par une définition.

Définition 1.8. Soit \mathcal{M} une variété de Banach, $D \subset \mathcal{M}$, et $F : \mathcal{M} \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 . On dit que $u \in \mathcal{M}$ est un **point critique** de F sur D si, pour toute courbe $s \mapsto u_s \in D$ dérivable en 0 et telle que $u_0 = u$, on a

$$\frac{d}{ds}F(u_s)|_{s=0} = 0.$$

Le résultat suivant est trivial.

Lemme 1.3. Soient $\mathcal{M}, \mathcal{M}'$ deux variétés de Banach, $F : \mathcal{M} \rightarrow \mathbb{R}$, $G : \mathcal{M} \rightarrow \mathcal{M}'$ deux applications de classe \mathcal{C}^1 , q_1 un élément de \mathcal{M}' et u un élément de \mathcal{M} .

1. S'il existe $p_1 \neq 0 \in T_{q_1}^* \mathcal{M}'$ tel que $dG(u)^* p_1 = 0$, u est un point singulier de G , au sens où $dG(u)$ n'est pas surjective, et $G^{-1}(\{q_1\})$ n'est donc peut-être pas une variété au voisinage de u .
2. S'il existe $p_1 \in T_{q_1}^* \mathcal{M}'$ tel que

$$dF(u) = dG(u)^* p_1,$$

alors u est un point critique de F sur $G^{-1}(\{q_1\})$. On dit que p_1 est le multiplicateur de Lagrange du problème en u .

Appliqué à notre cadre, on obtient la proposition suivante.

Proposition 1.2. Soit $(q, u) : [0, 1] \rightarrow \mathcal{H}$ un système horizontal. Supposons qu'il existe $(p, \lambda) \neq (0, 0)$, où $p : [0, 1] \rightarrow T^*M$, de classe H^1 avec $p(t) \in T_{q(t)}^* M$, et $\lambda = 0$ ou 1 , tels que

$$\begin{cases} 0 &= \partial_u H^\lambda(q(t), p(t), u(t)), \\ (\dot{q}(t), \dot{p}(t)) &= \nabla^\omega H^\lambda(q(t), p(t), u(t)), \end{cases}$$

Alors, si $\lambda = 0$, $p(t) \neq 0$ pour tout t et (q, u) est un point singulier de E .

Si $\lambda = 1$, alors (q, u) est un point critique de A sur $\Omega_{q(0), q(1)}^{\mathcal{H}}$.

Les géodésiques normales sont donc des points critiques de l'action à extrémités fixées.

1.6 Hamiltonien réduit et équation hamiltonienne des géodésiques

Dans cette section, on va voir que les géodésiques normales minimisent l'action à extrémités fixées, c'est-à-dire que ce sont bien des géodésiques. De plus, en utilisant le hamiltonien normal (ou, plus simplement, le hamiltonien), on va voir pouvoir donner toutes les géodésiques normales en résolvant une équation différentielle sur T^*M .

Soit $(q, p, u) \in T^*M \oplus_M \mathcal{H}$. Alors l'équation $\partial_u H^1(q, p, u) = \xi_q^* p - g_q(u, \cdot) = 0$ admet une unique solution $u(q, p)$, lisse en (q, p) . On obtient $g_q(u(q, p), u(q, p)) = p \xi_q u(q, p)$, ce qui nous permet de définir le *hamiltonien normal* $h : T^*M \rightarrow \mathbb{R}$ par la formule

$$h(q, p) = H^1(q, p, u(q, p)) = \frac{1}{2} p \xi_q u(q, p).$$

Dans ce cas, $\nabla^\omega h(q, p) = \nabla^\omega H^1(q, p, u(q, p))$, qui est bien défini car $\partial_u H^1(q, p, u(q, p)) = 0$.

Remark 1.8. Comme H^1 est strictement concave par rapport à u à $(q, p) \in T^*M$ fixé, on a

$$h(q, p) = \max_{u \in \mathcal{H}_q} H^1(q, p, u) = \max_{u \in \mathcal{H}_q} (p\xi_q u - \frac{1}{2}g_q(u, u)).$$

Théorème 1.5. Géodésiques normales.

Soit $(q_0, p_0) \in T^*M$. Il existe une unique courbe maximale $(q, p) : I \ni 0 \rightarrow T^*M$ de classe H^1_{loc} , telle que

$$(\dot{q}(t), \dot{p}(t)) = \nabla^\omega h(q(t), p(t)) \quad p.p. \ t \in [0, 1].$$

Alors q est la trajectoire du système horizontal $(q, u(q, p))$, et est une géodésique : elle minimise localement la longueur à extrémités fixes. De plus,

$$t \mapsto g_{q(t)}(u(q(t), p(t)), u(q(t), p(t))) = \|\dot{q}(t)\|_{q(t)}^2$$

est constante.

On dit que (q, p) satisfait l'équation hamiltonnienne des géodésiques.

La preuve est assez longue, et on ne la détaillera pas car nous allons prouver une version plus générale en dimension infinie. Pour des preuves en dimension finie, voir le chapitre 1 de [Mon02] pour le cas où ξ est injective, et le chapitre 17 de [AS04] pour le cas général.

Chapitre 2

Géométrie sous-riemannienne en dimension infinie

Dans ce chapitre, nous décrivons nos travaux les plus récents. Comme le papier correspondant est encore en cours de rédaction, nous fournissons les preuves complètes des résultats nouveaux énoncés.

Le but principal est d'étendre la géométrie sous-riemannienne à la dimension infinie. Pour des applications à venir, nous voulions considérer un cas très général. Par exemple, il était important d'avoir la possibilité d'étudier une distribution Δ de sous-espaces tangents qui ne soit pas fermée dans le fibré tangent.

Pour cela, nous commençons par définir les structures sous-riemanniennes par le biais de fibrés tangents relatifs, comme dans la section 1.2. Puis, après avoir mentionné une généralisation déjà connue du théorème de Chow-Rashevski, nous donnerons des exemples simples qui montreront quelques-unes des difficultés qui peuvent apparaître en dimension infinie.

On s'attaquera ensuite à la recherche des géodésiques. Les structures sous-riemanniennes faibles posent particulièrement problème, car l'équation $\partial_u H^1 = 0$ n'a pas toujours de solution, et le hamiltonien normal n'est donc pas défini sur tout l'espace cotangent de notre variété. Il nous faudra introduire des sous-fibrés vectoriels denses dans celui-ci, appelés fibrés cotangents relatifs, sur lesquels $\partial_u H^1 = 0$ a toujours des solutions. On devra se restreindre à de tels fibrés pour résoudre l'équation hamiltonienne des géodésiques.

On prouvera ensuite notre résultat principal : les courbes satisfaisant cette équation hamiltonienne sont des géodésiques locales. On appliquera ces résultats à un exemple de géométrie sous-riemannienne faible sur le groupe des difféomorphismes de \mathbb{R}^d de classe de Sobolev H^s , $s > 1 + d/2$.

Quelques exemples de structures sous-riemanniennes fortes sont traités plus en détails dans les prochaines parties et dans les preprints [ATTY13, AT].

2.1 Définitions

En dimension infinie, alors que la géométrie riemannienne est très étudiée [EM70, KM97, Cla09], avec par exemple des applications en mécanique des fluides [EM70, MP10], le cas sous-riemannien est un domaine très peu exploré. On a tout de même quelques résultats de contrôlabilité dans le groupe des difféomorphismes par Agrachev et Caponigro [AC09]. Une pré-publication de Grong, Markina et Vasil'ev, disponible sur arxiv [GMV12], calcule certaines géodésiques dans le cas particulier d'une distribution régulière Δ de sous-espaces tangents fermés, admettant un fibré supplémentaire lisse et fermé, et dont la métrique est la restriction d'une métrique riemannienne faible admettant un flot géodésique.

2.1.1 Structure sous-riemannienne sur une variété de Banach

Soit donc M une variété de Banach. On définit une structure sous-riemannienne sur M de façon analogue à la dimension finie (définition 1.5).

Définition 2.1. Une structure sous-riemannienne sur une variété de Banach M est un triplet (\mathcal{H}, ξ, g) , où \mathcal{H} est un fibré vectoriel lisse de Banach sur M , ξ est un morphisme lisse de fibrés vectoriels de \mathcal{H} sur M (donc (\mathcal{H}, ξ) est un espace tangent relatif sur M), et g est une métrique riemannienne lisse sur \mathcal{H} , c'est-à-dire que g définit sur chaque fibre \mathcal{H}_q une forme bilinéaire symétrique définie positive.

Une première différence avec la dimension finie est la dichotomie entre le cas des structures dites *fortes*, et celui des structures *faibles*.

Définition 2.2. Une structure sous-riemannienne (\mathcal{H}, ξ, g) sur une variété de Banach M est dite **forte** si la topologie intrinsèque des fibres \mathcal{H}_q coïncide avec la topologie définie par g_q pour tout $q \in M$. Dans ce cas, (h_q, g_q) est un espace de Hilbert. Elle est dite **faible** dans le cas contraire.

Remark 2.1. C'est une dichotomie bien connu dans les variétés riemanniennes de dimension infinie [KM97, Cla09]. Comme nous le verrons plus tard, un des principaux obstacles pour obtenir des géodésiques est que, pour une métrique faible, l'équation

$$g_q(u, \cdot) = \xi_q^* p \quad (2.1)$$

n'a pas toujours de solution, puisque l'opérateur $u \mapsto g_q(u, \cdot)$ n'est pas forcément surjectif de \mathcal{H}_q sur \mathcal{H}_q^* .

Dans toute la suite, on fixe M une variété de Banach lisse équipée d'une structure sous-riemannienne (\mathcal{H}, ξ, g) . On notera également $\Delta = \xi(\mathcal{H}) \subset TM$.

2.1.2 Courbes horizontales, action et longueur

Pour faire plus simple, on va se restreindre aux courbes horizontales de classe de Sobolev H^1 . Une courbe $q \in H^1(I; M)$, I intervalle, est horizontale s'il existe $(q, u) \in H^1 \times L^2(I; \mathcal{H})$ au-dessus de q tel

$$\dot{q}(t) = \xi_{q(t)} u(t), \quad \text{p.p. } t \in I.$$

Remark 2.2. Lorsqu'on note $(q, u) \in H^1 \times L^2(I; \mathcal{H})$, cela signifie que u est de carré intégrable pour la topologie de Banach intrinsèque sur \mathcal{H} , et non au sens de la métrique g .

On appelle un tel couple $t \in I \mapsto (q(t), u(t)) \in \mathcal{H}$ un système horizontal, q sa trajectoire et u son contrôle. On définit également la longueur

$$L(q, u) = \int_I \sqrt{g_{q(t)}(u(t), u(t))} dt,$$

et l'action

$$A(q, u) = \frac{1}{2} \int_I g_{q(t)}(u(t), u(t)) dt$$

d'un tel système, ainsi que l'action

$$A^M(q) = \inf_{u, (q,u) \text{ horizontal}} A(q, u),$$

et la longueur

$$L^M(q) = \inf_{u, (q,u) \text{ horizontal}} L(q, u)$$

de toute courbe horizontale, comme en section 1.2.

Remark 2.3. Il n'existe pas forcément de contrôle u tel que $A^M(q) = A(q, u)$, contrairement à ce qu'il se passe en dimension finie. Toutefois, si un tel u existe, il est unique, déterminé pour presque tout t par

$$\forall v \in \ker \xi_{q(t)}, \quad g_{q(t)}(u(t), v) = 0.$$

Pour être assuré de l'existence d'un tel u , il faut que $\ker \xi$ admette un fibré supplémentaire orthogonal $(\ker \xi)^\perp$. C'est par exemple vrai pour une structure forte, ou si ξ est injective.

2.1.3 Distance sous-riemannienne

On peut de même définir la distance sous-riemannienne $d(q_0, q_1)$ comme l'infimum des longueurs des systèmes horizontaux dont la trajectoire relie q_0 et q_1 . Toutefois, d n'est à priori qu'une semi-distance lorsque la structure est faible : on peut éventuellement avoir $d(q_0, q_1) = 0$ et $q_0 \neq q_1$.

Proposition 2.1. *Pour une structure forte, la topologie induite par d est plus fine que la topologie intrinsèque de M . En conséquence, c'est une vraie distance :*

$$\forall q_0 \neq q_1 \in M, \quad d(q_0, q_1) > 0.$$

Preuve. Soient $q_0 \in M$ et U un système de coordonnées au voisinage de q_0 , identifié à un ouvert d'un espace de Banach $(B, \|\cdot\|_B)$ dans lequel q_0 est identifié à 0. Il faut montrer qu'il existe $C > 0$ tel que pour tout $\varepsilon > 0$ assez petit, et pour tout $q_1 \in U$ tel que $\|q_1\|_B = \varepsilon$, $d(0, q_1) \geq C\varepsilon$.

On trivialisait $\mathcal{H}|_U \simeq U \times H$ au-dessus de U . Comme ξ est continue, il existe alors $\varepsilon_0, C > 0$ tels que

$$\forall (q, u) \in U \times H, \quad \|q\| \leq \varepsilon_0 \quad \Rightarrow \quad C\|\xi_q u\|_B \leq \sqrt{g_q(u, u)} \quad (2.2)$$

Soit q_1 tel que $\|q_1\|_B = \varepsilon < \varepsilon_0$. On va montrer que tout système horizontal dont la trajectoire relie q_0 et q_1 a une longueur supérieure à $C\varepsilon$.

Soit $I = [a, b]$ un intervalle de \mathbb{R} , et $(q, u) : I \rightarrow \mathcal{H}$ un système horizontal tel que $q(a) = q_0$ et $q(b) = q_1$. Comme q est continue, il existe au moins un $t \in [a, b]$ tel que $\|q(t)\|_B = \varepsilon$. Soit t_ε le premier d'entre eux. Pour tout $t \in [a, t_\varepsilon]$, on obtient $\|q(t_\varepsilon)\|_B \leq \varepsilon$. On peut donc appliquer (2.2) et obtenir

$$\begin{aligned} C\varepsilon = C\|q(t_\varepsilon)\|_B &\leq \int_a^{t_\varepsilon} C\|\xi_{q(t)}u(t)dt\|_B dt \\ &\leq \int_a^{t_\varepsilon} \sqrt{g_{q(t)}(u(t), u(t))} dt \\ &\leq L(q|_{[a, t_\varepsilon]}, u|_{[a, t_\varepsilon]}) \\ &\leq L(q, u). \end{aligned}$$

□

2.1.4 Géodésiques locales

On définit un nouveau type de géodésiques n'ayant d'intérêt qu'en dimension infinie et pour des structures faibles.

Définition 2.3. *Une **géodésique locale** est une courbe horizontale $q : I \rightarrow M$ telle que, pour $t_0 < t_1$ assez proches, il existe un voisinage U de $q([t_0, t_1])$ tel que $q|_{[t_0, t_1]}$ minimise la longueur parmi toutes les courbes horizontales **à valeurs dans** U reliant q_0 et q_1 .*

Dans le cas d'une structure faible, une géodésique locale n'est pas forcément une géodésique au sens habituel, même dans le cas riemannien [Cla09]. En revanche, si la structure est forte, on voit facilement que les deux notions coïncident en utilisant la même méthode que dans la preuve de la proposition 2.1.

2.2 Un équivalent au théorème de Chow-Rashevski

Dans toute la suite, on fixe M une variété de Banach lisse équipée d'une structure sous-riemannienne (\mathcal{H}, ξ, g) . On notera également $\Delta = \xi(\mathcal{H}) \subset TM$.

Soit $q_0 \in M$. L'orbite \mathcal{O}_{q_0} de q_0 est l'ensemble des points qui peuvent être joints à q_0 par une courbe horizontale. Autrement dit, $\mathcal{O}_{q_0} = \{q \in M, d(q_0, q) < +\infty\}$. Bien entendu, $q_1 \in \mathcal{O}_{q_0}$ si et seulement si $\mathcal{O}_{q_0} = \mathcal{O}_{q_1}$.

Soit \mathcal{L} l'algèbre de Lie des champs de vecteurs sur M engendrée par les champs de vecteurs tangents à Δ en tout point. On définit pour $q \in M$

$$\mathcal{L}_q = \{X(q), X \in \mathcal{L}\} \subset T_qM.$$

Pour un sous-ensemble quelconque D de M , on définit T_qD l'ensemble des vitesses initial possibles de courbes \mathcal{C}^1 sur M partant de q et qui restent dans D . On vérifie facilement, par la même méthode qu'en dimension finie, que $\mathcal{L}_q \subset T_q\mathcal{O}_q$.

Toutefois, comme \mathcal{O}_q n'est à priori pas une sous-variété immergée, on ne peut pas en déduire que $\mathcal{L}_q = T_qM$ pour tout q dans M implique $\mathcal{O}_q = M$. Ce théorème serait de toute façon peu utile : comme \mathcal{L} est défini de façon algébrique par récurrence, il est très rare que \mathcal{L}_q soit fermé dans T_qM . Par contre, il a été montré le résultat de contrôlabilité approché suivant.

Proposition 2.2. [DS80] *Si \mathcal{L}_q est dense dans T_qM pour tout q , alors toutes les orbites sont denses dans M .*

2.3 Un exemple de structures sous-riemannienne forte : le produit de groupes de Heisenberg

On rappelle [Mon02, CDPT07, BR96] que le groupe de Heisenberg \mathbb{H} est égal à \mathbb{R}^3 muni de la structure sous-riemannienne $(\mathbb{R}^3 \times \mathbb{R}^2, \xi, g)$, où

$$g_q(u, v; u, v) = u^2 + v^2, \quad (q, u, v) \in \mathbb{R}^3 \times \mathbb{R}^2,$$

et

$$\xi_{(x,y,z)}(u, v) = u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + \frac{1}{2}(vx - uy) \frac{\partial}{\partial z}.$$

On note $A^{\mathbb{H}}$ l'action associée sur les courbes horizontales, et $d^{\mathbb{H}}$ la distance sous-riemannienne induite. On a

$$\frac{1}{C}(x^2 + y^2 + |z|) \leq d^{\mathbb{H}}(0, (x, y, z))^2 \leq C(x^2 + y^2 + |z|)$$

pour un certain $C > 0$ (voir [Mon02, BR96]).

On considère maintenant $M = \ell^2(\mathbb{N}, \mathbb{R}^3)$ l'espace des suites de triplets $q = (x_n, y_n, z_n)_{n \in \mathbb{N}}$ réels de carré sommable. C'est un espace de Hilbert pour la norme ℓ^2 .

On définit sur M la structure sous-riemannienne (\mathcal{H}, ξ, g) suivante : $\mathcal{H} = M \times H$, où $H = \ell^2(\mathbb{N}, \mathbb{R}^2)$ est l'espace des suites $u = (u_n, v_n)_{n \in \mathbb{N}}$ de couples réels de carré sommables,

$$g_q(u, u) = \sum_{n=0}^{\infty} u_n^2 + v_n^2,$$

et

$$\xi_q u = \sum_{i=0}^{\infty} u_n \frac{\partial}{\partial x_n} + v_n \frac{\partial}{\partial y_n} + \frac{1}{2}(v_n x_n - u_n y_n) \frac{\partial}{\partial z_n}.$$

Les champs de vecteurs horizontaux sont des $C^\infty(M)$ -combinaisons (de carré sommable en chaque point) des champs X_0, Y_0, \dots , où

$$X_n = \frac{\partial}{\partial x_n} - \frac{1}{2}y_n \frac{\partial}{\partial z_n}, \quad Y_n = \frac{\partial}{\partial y_n} + \frac{1}{2}x_n \frac{\partial}{\partial z_n}, \quad n \in \mathbb{N}.$$

De plus, $Z_n = [X_n, Y_n] = \frac{\partial}{\partial z_n}$, donc $\mathcal{L}_q \ni X_0(q), Y_0(q), Z_0(q), \dots$, est dense dans $T_q M$. Donc toutes les orbites sont denses.

Toutefois, si $q = (x_n, y_n, z_n)_{n \in \mathbb{N}} : [0, 1] \rightarrow M$ est une courbe horizontale, alors $q_n = (x_n, y_n, z_n) : [0, 1] \rightarrow \mathbb{R}^3$ est une courbe horizontale pour la structure sous-riemannienne usuelle du groupe de Heisenberg \mathbb{H} sur \mathbb{R}^3 . De plus, on a

$$A^M(q) = \sum_{n=0}^{\infty} A^{\mathbb{H}}(x_n, y_n, z_n).$$

On en déduit aisément que, pour tout $q = (x_n, y_n, z_n)_{n \in \mathbb{N}} \in M$,

$$\frac{1}{C} \left(\sum_{n=0}^{\infty} x_n^2 + y_n^2 + |z_n| \right) \leq d(0, q)^2 \leq C \left(\sum_{n=0}^{\infty} x_n^2 + y_n^2 + |z_n| \right).$$

Cela implique que $\mathcal{O}_0 = \ell^2(\mathbb{N}, \mathbb{R}^2) \times \ell^1(\mathbb{N}, \mathbb{R})$, qui est un sous-espace stricte de $\ell^2(\mathbb{N}, \mathbb{R}^3)$. De plus, la topologie induite par d sur \mathcal{O}_0 coïncide avec celle de $\ell^2(\mathbb{N}, \mathbb{R}^2) \times \ell^1(\mathbb{N}, \mathbb{R})$, qui est plus fine que celle induite par M .

Remark 2.4. La structure (\mathcal{H}, ξ, g) se restreint en une structure sous-riemannienne lisse sur $M' = \ell^2(\mathbb{N}, \mathbb{R}^2) \times \ell^1(\mathbb{N}, \mathbb{R})$. En se restreignant à M' , on grossit l'espace cotangent, puisque

$$(\ell^2(\mathbb{N}, \mathbb{R}^2) \times \ell^1(\mathbb{N}, \mathbb{R}))^* = \ell^2(\mathbb{N}, \mathbb{R}^2) \times \ell^\infty(\mathbb{N}, \mathbb{R}) \supset \ell^2(\mathbb{N}, \mathbb{R}^3)^*.$$

On trouvera donc "plus" de géodésiques normale en se restreignant à M' .

Remark 2.5. Si on procède de la même manière dans $M = \ell^2(\mathbb{N}, \mathbb{R}^4)$, avec une structure de produit infini de groupes de Engel (voir [Mon02]), variété sous-riemannienne \mathbb{E} de dimension 4 sur laquelle la distance $d^{\mathbb{E}}$ vérifie

$$\frac{1}{C'} (x^2 + y^2 + |z| + |w|^{2/3}) \leq d^{\mathbb{E}}(0, (x, y, z)) \leq C' (x^2 + y^2 + |z| + |w|^{2/3}),$$

on trouve que, sur M ,

$$\frac{1}{C} \left(\sum_{n=0}^{\infty} x_n^2 + y_n^2 + |z_n| + |w_n|^{2/3} \right) \leq d(0, q)^2 \leq C \left(\sum_{n=0}^{\infty} x_n^2 + y_n^2 + |z_n| + |w_n|^{2/3} \right).$$

On obtient donc $\mathcal{O}_0 = \ell^2(\mathbb{N}, \mathbb{R}^2) \times \ell^1(\mathbb{N}, \mathbb{R}) \times \ell^{2/3}(\mathbb{N}, \mathbb{R})$, et la topologie coïncide avec celle induite par $d^{\mathbb{E}}$. Dans ce cas, l'orbite n'est donc même pas un espace localement convexe.

2.4 Application point final en dimension infinie

Dans toute la suite, on fixe M une variété de Banach lisse équipée d'une structure sous-riemannienne (\mathcal{H}, ξ, g) . On notera également $\Delta = \xi(\mathcal{H}) \subset TM$.

Rappelons que $\Omega_{q_0}^{\mathcal{H}}$ est l'ensemble des systèmes horizontaux sur $[0, 1]$ partant de q_0 . On a vu qu'en dimension finie, $\Omega_{q_0}^{\mathcal{H}}$ est une sous-variété de $H^1 \times L^2(0, 1; \mathcal{H})$, elle-même dotée d'une structure de variété de Hilbert. Toutefois, ce n'est plus forcément le cas en dimension infinie.

L'espace $H^1 \times L^2(0, 1; \mathcal{H})$ n'est pas toujours muni d'une structure naturelle de variété. En effet, pour construire une telle structure, il faut que M admette une "addition locale" : un difféomorphisme lisse d'un voisinage de la section nulle dans TM sur un voisinage de la diagonale dans $M \times M$ qui envoie la section nulle sur la diagonale. C'est le cas, par exemple, lorsque M admet une métrique riemannienne possédant une connexion de Levi-Civita lisse. Voir [KM97] pour plus de détails. Pour la même raison, $\Omega_{q_0}^{\mathcal{H}}$ n'est pas forcément muni d'une structure de variété.

En revanche, dans le cas où la trajectoire q de $(q, u) \in \Omega_{q_0}^{\mathcal{H}}$ est *simple*, on peut reprendre mot à mot la construction décrite dans la preuve de la proposition 1.3, remplaçant simplement chaque occurrence de \mathbb{R}^d par l'espace de Banach B sur lequel est modelé M . Cela permet de définir un espace de contrôle \mathcal{U} qui donne une paramétrisation lisse de $\Omega_{q_0}^{\mathcal{H}}$ au voisinage de (q, u) , induit par un système de coordonnées U au voisinage de $\text{im}(q)$. Dans cette construction, \mathcal{U} est identifié à un ouvert de $L^2(0, 1; H)$, où H est un espace de Banach isomorphe aux fibres de \mathcal{H} .

De plus, on obtient une application trajectoire lisse $\mathcal{U} \rightarrow H^1(0, 1; U)$, qui au contrôle $u \in \mathcal{U}$ associe la trajectoire $q(\cdot, u)$. La dérivée en u dans la direction $\delta u \in L^2(0, 1; H)$ de cette application, notée δq , est donnée par l'unique solution sur $[0, 1]$ du problème de Cauchy linéaire

$$\dot{\delta q}(t) = \partial_q(\xi_{q(t)}u(t))\delta q(t) + \xi_{q(t)}\delta u(t), \quad \delta q(0) = 0.$$

Dans ce cas, dire que la restriction de E à \mathcal{U} est lisse a bel et bien un sens, et on peut l'utiliser pour déterminer les courbes minimisant l'action à extrémités fixées et dont la trajectoire reste dans U . On obtient ainsi les mêmes formules pour l'application adjointe de sa différentielle.

Proposition 2.3. *Soit $(q, u) \in \Omega_{q_0}^{\mathcal{H}}$ avec q est simple. On considère un certain $p_1 \in T_{q(1)}^*M$. On définit $P \in T_{(q,u)}^*\Omega_{q_0}^{\mathcal{H}}$ par*

$$P(\delta q, \delta u) = (dE_{q,u})^*p_1(\delta q, \delta u) = p_1\delta q(1).$$

Soit un espace de contrôle \mathcal{U} induit par un système de coordonnées U au voisinage de $\text{im}(q)$ correspondant.

Alors on a

$$(dE_{(q,u)})^*p_1(\delta q, \delta u) = \int_0^1 p(t)\xi_{q(t)}\delta u(t)dt,$$

*où $(q, p) \in H^1 \times H^1(0, 1; T^*M)$ et p est l'unique solution sur $[0, 1]$ du problème de Cauchy linéaire adjoint*

$$\dot{p}(t) = \partial_q(\xi_{q(t)}u(t))^*p(t), \quad p(1) = p_1. \quad (2.3)$$

La preuve est identique à celle du lemme 1.1.

2.5 Absence de principe du maximum en dimension infinie

En dimension infinie, le lemme 1.1 donnant des conditions nécessaires de minimisation contrainte est fautive : il faut rajouter une troisième possibilité. En effet, soient \mathcal{M} et M deux variétés modelées sur des espaces de Banach. Soit $G : \mathcal{M} \rightarrow M$ et $F : \mathcal{M} \rightarrow \mathbb{R}$, deux applications de classe \mathcal{C}^1 . Soit $q_1 \in G(\mathcal{M}) \subset M$, et $\bar{u} \in G^{-1}(\{q_1\})$, minimiseur de F sur $G^{-1}(\{q_1\})$. Dans ce cas, l'application $Q : u \in \mathcal{M} \mapsto (G(u), F(u)) \in M \times \mathbb{R}$ est également de classe \mathcal{C}^1 , et sa différentielle en \bar{u} n'est pas surjective (sinon, comme Q serait alors une submersion en \bar{u} , on pourrait trouver $(q_1, a) \in Q(\mathcal{M})$ avec $a < F(\bar{u})$). Il y a alors deux possibilités : soit $\text{im}(dQ(\bar{u}))$ est dense dans $T_{q_1}M \times \mathbb{R}$, soit elle est incluse dans un sous-espace strict fermé de $T_{q_1}M \times \mathbb{R}$ (on dit qu'il est de *codimension positive*). Le premier cas est équivalent à dire que $\ker(dQ(\bar{u})^*) = \{0\}$, est n'est possible que si M est de dimension infinie.

Dans le second cas, il existe $(p_1, \lambda) \neq (0, 0) \in T_{q_1}^*M \times \mathbb{R}$ tel que $\lambda dF(\bar{u}) = dG(\bar{u})^*p_1$, et on peut normaliser (p_1, λ) λ de telle sorte que $\lambda \in \{0, 1\}$.

Lemme 2.1. *Soient \mathcal{M} une variété de Banach, et $F : \mathcal{M} \rightarrow \mathbb{R}$ et $G : \mathcal{M} \rightarrow M$ deux applications de classe \mathcal{C}^1 . Soient q_1 un élément de M et u un élément de \mathcal{M} qui minimise F sur l'ensemble contraint $G^{-1}(\{q_1\})$. Alors une et une seule des deux possibilités suivantes est satisfaite :*

1. *L'image de $(dG(u), dF(u))$ est incluse dans un sous-espace fermé strict de $T_{q_1}M \times \mathbb{R}$. Il $(p_1, \lambda) \neq (0, 0)$ tel que*

$$dG(u)^*p_1 = \lambda dF(u).$$

Si $\lambda = 1$, on dit que u est un minimiseur normal, et p_1 est appelé le multiplicateur de Lagrange associé à u . Si $\lambda = 0$, $\text{im}(dG(u))$ de codimension positive : on dit que u est un minimiseur singulier.

2. *L'image de $(dG(u), dF(u))$ est un sous-espace strict dense de $T_{q_1}M \times \mathbb{R}$. On dit que u est un minimiseur **élusif** de F sur $G^{-1}(\{q_1\})$.*

Le deuxième cas n'apparaît qu'en dimension infinie. Il s'agit d'une difficulté connue en contrôle optimal de dimension infinie, qui explique l'absence de principe du maximum de Pontryagin. Voir par exemple [LY95].

Il est intéressant de remarquer que même dans les exemples (non-triviaux) les plus simples, comme dans le cas d'un produit infini de groupes de Heisenberg (section 2.3), l'image de la différentielle de l'application point final est *toujours* un sous-espace dense strict. En conséquence, toutes les courbes horizontales qui ne sont pas des géodésiques normales satisfont la condition 2. Cette condition ne donne donc aucune information particulière sur un minimiseur qui est dans le cas 2.

Il est donc inutile d'espérer utiliser ce résultat pour obtenir un principe du maximum comme le théorème 1.4 pour donner de bonnes conditions nécessaires pour qu'une courbe soit une géodésique.

Remark 2.6. Ce problème est lié à un défaut topologique de M : la topologie finale induite par E sur \mathcal{O}_{q_0} est souvent plus fine que la topologie héritée de M . T^*M n'est donc pas assez "gros".

Une manière de rajouter de rendre normaux ou singuliers certaines géodésiques élusives est de considérer une sous-variété de Banach $M' \hookrightarrow M$ dense, avec $q_0 \in M'$ sur laquelle $\Delta_q \subset T_qM' \subset T_qM$. Dans ce cas, $\iota_{q_0} \subset M'$ et $(\mathcal{H}_{|M'}, \xi_{|\mathcal{H}_{M'}}, i, g_{|\mathcal{H}_{M'}})$ est une structure sous-riemannienne sur M' . Comme l'inclusion $T_qM' \hookrightarrow T_qM$ est dense et continue, la topologie de T_qM' est plus fine que celle de T_qM , et il est donc "plus difficile" d'être un sous-espace dense de $T_qM' \times \mathbb{R}$, ce qui nous permet d'obtenir plus de minimiseurs normaux et singuliers avec la condition 1. Dans le cas du produit infini du groupe de Heisenberg, \mathcal{O}_0 lui-même est une sous-variété immergée dense, avec en plus une topologie qui coïncide avec celle induite par la distance sous-riemannienne.

2.6 Hamiltoniens, géodésiques normales et courbes singulières

Malgré l'absence d'un principe du maximum de Pontryagin, on peut donner des conditions *suffisantes* pour qu'une courbe horizontale soit singulière ou soit un point critique de l'action à extrémités fixées. En effet, rappelons que la réciproque partielle du lemme 1.1, donnée dans le lemme 1.3, reste valable en dimension infinie. Grâce à la proposition 2.3, on en déduit immédiatement le résultat suivant, comme en dimension finie.

Cela va nous permettre de généraliser la proposition 1.2 à la dimension infinie.

2.6.1 Forme symplectique faible sur T^*M

Sur une variété de Banach, il y a deux catégories de formes symplectiques : les formes symplectiques fortes et les formes symplectiques faibles.

Définition 2.4. Une 2-forme ω sur une variété de Banach N est dite **symplectique faible** si elle est fermée et si l'opérateur $v \mapsto \omega_q(v, \cdot)$ est injectif de T_qN sur T_q^*N pour point tout $q \in N$. Elle est dite **symplectique forte** lorsque cet opérateur est également surjectif [KM97].

Un vecteur v est le gradient symplectique d'une fonction f en un point q si on a $df(q) = \omega_q(v, \cdot)$. Si ω n'est pas forte, contrairement à la dimension finie, une fonction différentiable en q n'admet pas forcément de gradient symplectique.

L'espace cotangent T^*M à une variété de Banach M modelé sur un espace de Banach B possède une forme symplectique faible canonique ω . En coordonnées canoniques $(q, p) \in B \times B^*$, on a

$$\omega_{(q,p)}(\delta q, \delta p; \delta q', \delta p') = \delta p(\delta q') - \delta p'(\delta q).$$

Remark 2.7. En particulier, la forme symplectique canonique ω sur T^*M est forte si et seulement si B est reflexif.

Ainsi, une fonction $f : T^*M \rightarrow \mathbb{R}$ différentiable en $(q_0, p_0) \in T^*M$ admet un gradient symplectique $\nabla^\omega f(q_0, p_0)$ si et seulement si $\partial_p f(q, p)$ appartient à T_qM à travers l'identification $T_qM = j(T_qM)$, où $j : T_qM \rightarrow T_q^{**}M$ est l'injection canonique de T_qM dans son bidual. Dans ce cas, dans un système de coordonnées canoniques, on obtient

$$\nabla^\omega f(q_0, p_0) = (\partial_p f(q_0, p_0), -\partial_q f(q_0, p_0)).$$

Soit maintenant \mathcal{E} un fibré vectoriel sur M et $f : T^*M \oplus_M \mathcal{E}$ est différentiable en (q_0, p_0, e_0) . Si

$$\partial_e f(q_0, p_0, e_0) = 0$$

et

$$\partial_p f(q_0, p_0, e_0) \in T_{(q_0, p_0)}M,$$

alors, comme en dimension finie, f possède un gradient symplectique partiel $\nabla^\omega f(q_0, p_0, e_0) \in T_{q_0, p_0}T^*M$, donné dans un système de coordonnées canoniques par

$$\nabla^\omega f(q_0, p_0, e_0) = (\partial_p f(q_0, p_0, e_0), -\partial_q f(q_0, p_0, e_0)) \in B \times B^*.$$

2.6.2 Hamiltoniens

Soit $\lambda \in \{0, 1\}$. On définit les hamiltoniens $H^\lambda : T^*M \oplus_M \mathcal{H} \rightarrow \mathbb{R}$ par

$$H^\lambda(q, p, u) = H^1(q, p, u) = p\xi_q u - \frac{\lambda}{2}g_q(u, u).$$

Alors H^λ est lisse, et

$$\partial_p H^\lambda(q, p, u) = \xi_q u \in T_qM.$$

Donc $\partial_p H^\lambda(q, p, u)$ s'identifie toujours à un élément de T_qM à travers l'injection canonique $T_qM \hookrightarrow T_q^{**}M$. En particulier, dès que $\partial_u H(q, p, u) = 0$, H possède un gradient symplectique partiel en (q, p, u) .

En utilisant rigoureusement le même raisonnement que dans le cas de dimension finie, c'est-à-dire en combinant les lemmes 1.3, 2.3, et les valeurs des dérivées partielles de H^λ , on obtient le résultat suivant.

Proposition 2.4. *Soit (q, u) un système horizontal dont la trajectoire q est simple, et \mathcal{U} un espace de contrôles au voisinage de (q, u) . Supposons qu'il existe $(p, \lambda) \neq (0, 0)$, où $p : [0, 1] \rightarrow T^*M$ est de classe H^1 avec $p(t) \in T_{q(t)}^*M$ pour tout t et $\lambda \in \{0, 1\}$ tel que*

$$\begin{cases} 0 &= \partial_u H^\lambda(q(t), p(t), u(t)), \\ (\dot{q}(t), \dot{p}(t)) &= \nabla^\omega H^\lambda(q(t), p(t), u(t)), \end{cases}$$

- Si $\lambda = 0$, alors, pour tout $t \in [0, 1]$, $p(t) \neq 0$, $\xi_{q(t)}^* p(t) = 0$ et l'image de $dE_{(q,u)}$ est contenue dans un sous-espace fermé strict de $T_{q(1)}^*M$. On dit que q est une courbe singulière.
- Si $\lambda = 1$, alors (q, u) est un point critique de A sur $\Omega_{q(0), q(1)}^{\mathcal{H}} \cap \mathcal{U}$. On dit que q est une géodésique normale.

2.7 Espaces cotangents relatifs

L'étape suivante est de montrer que les géodésiques normales sont des géodésiques locales (voir définition 2.3). Le problème est que cela requiert l'utilisation d'un hamiltonien normal, qui lui-même nécessite de résoudre en u l'équation

$$\partial_u H^1(q, p, u) = \xi_q^* p - g_q(u, \cdot) = 0$$

pour tout $(q, p) \in T^*M$. Or, si la structure est faible, cette équation peut ne pas avoir de solution. Pour remédier à ce problème, on est obligé de restreindre H^1 à un sous-fibré dense de T^*M , appelé fibré cotangent relatif adapté, sur lequel cette équation peut se résoudre.

2.7.1 Fibré cotangent relatif

Définition 2.5. *On appelle **fibré cotangent relatif** un fibré vectoriel d'espaces de Banach τ^*M sur M qui est aussi un sous-fibré dense de T^*M avec inclusion continue.*

*On note encore H^1 la restriction du hamiltonien à τ^*M , qui est toujours lisse. On dit que τ^*M est **adapté** à la structure (\mathcal{H}, ξ, g) si l'équation en u*

$$\partial_u H^1(q, p, u) = \xi_q^* p - g_q(u, \cdot) = 0.$$

*a une solution pour tous $(q, p) \in \tau^*M$. Dans ce cas, cette solution est unique, et on la note $u(q, p)$. Cela nous permet de définir le hamiltonien normal $h : \tau^*M \rightarrow \mathbb{R}$ correspondant par*

$$h(q, p) = H^1(q, p, u(q, p)) = \frac{1}{2} p \xi_q u(q, p) = \frac{1}{2} g_q(u(q, p), u(q, p)).$$

Remark 2.8. Supposons que \mathcal{H} soit un sous-fibré fermé de TM et ξ l'inclusion, et que g soit la restriction d'une métrique riemannienne faible \bar{g} sur TM . Alors, si \mathcal{H} admet un fibré supplémentaire orthogonal fermé, $\tau^*M = TM$ avec l'injection $TM \hookrightarrow T^*M$ définie par $v \mapsto \bar{g}(v, \cdot)$ est un fibré cotangent adapté.

Dans ce cas, $u(q, \bar{g}(v, \cdot)) = u(q, v)$ est la projection orthogonale de v sur \mathcal{H}_q . C'est ce qui est utilisé de façon implicite dans [GMV12], et de façon explicite dans [MM07] pour certaines structures riemanniennes sur des espaces de sous-variétés.

Remark 2.9. Bien entendu, pour une structure forte, T^*M lui-même est un fibré cotangent relatif adapté.

2.7.2 Forme symplectique canonique faible sur τ^*M et gradient symplectique

Fixons un fibré cotangent relatif τ^*M sur M . Alors la forme symplectique canonique ω sur T^*M se restreint à une forme symplectique sur τ^*M .

Lemme 2.2. *La restriction de la forme symplectique canonique ω de T^*M à τ^*M , c'est-à-dire son tiré en arrière par l'inclusion $\tau^*M \hookrightarrow T^*M$, est une forme symplectique faible, encore notée ω .*

Preuve. La restriction de ω à τ^*M est fermée car ω est fermée. De plus, pour tout $(q, p) \in \tau^*M$, $T_{(q,p)}\tau^*M$ est dense dans $T_{q,p}T^*M$ par définition d'un fibré cotangent relatif. Donc si $\omega_{(q,p)}(\delta q, \delta p; \cdot, \cdot)$ s'annule sur $T_{q,p}\tau^*M$, elle s'annule sur $T_{q,p}T^*M$ et donc $(\delta q, \delta p) = 0$. \square

Une fonction $f : \tau^*M \rightarrow \mathbb{R}$ différentiable en (q, p) possède alors un gradient symplectique si et seulement si

1. la dérivée partielle $\partial_p f(q, p)$ s'identifie à un vecteur tangent de $T_q M$ à travers l'injection canonique de $T_q M$ dans $T_q^* M \subset (\tau^* M)^*$.
2. Dans un système de coordonnées canonique $U \subset B \times B^*$ autour de (q, p) dans T^*M , la dérivée partielle $\partial_q f(q, p)$, appartient au sous-espace dense $\tau_q^* M$ de $T_q^* M$. Si $\tau^* M = T^* M$, comme dans le cas d'une structure forte, cette condition est toujours vérifiée.

De même, une fonction $f : \tau^* M \oplus_M \mathcal{H} \rightarrow \mathbb{R}$ admet un gradient symplectique partiel en (q, p, u) si et seulement si les deux conditions précédentes sont satisfaites et si $\partial_u f(q, p, u) = 0$.

2.7.3 Le cas du hamiltonien

On a déjà vu que $\partial_p H^1(q, p, u)$ peut toujours s'identifier au vecteur $\xi_q u$. Ainsi, H^1 possède un gradient symplectique partiel sur $\tau^* M$ en un point (q, p, u) tel que $\partial_u H(q, p, u) = 0$ si et seulement si

$$\partial_q H^1(q, p, u) = \frac{1}{2} \partial_q g_q(u, u) - \partial_q (\xi_q u)^* p \in \tau_q^* M.$$

Dans ce cas, le hamiltonien réduit h admet également un gradient symplectique, donné par

$$\nabla^\omega h(q, p) = \nabla^\omega H^1(q, p, u(q, p)) = \left(\xi_q u(q, p), \frac{1}{2} \partial_q g_q(u(q, p), u(q, p)) - \partial_q (\xi_q u(q, p))^* p \right).$$

2.8 Équation hamiltonienne des géodésiques normales sur un espace cotangent relatif adapté

On a maintenant tous les outils nécessaires à la preuve de notre résultat principale sur les géodésiques sous-riemanniennes en dimension infinie. Commençons par la proposition suivante, conséquence im-médiate de la proposition 2.4.

Proposition 2.5. Géodésiques normales *Soit τ^*M un espace cotangent relatif adapté à (\mathcal{H}, ξ, g) sur M . On suppose que l'application $(q, p) \mapsto u(q, p)$ est de classe \mathcal{C}^1 , et que le hamiltonien réduit $h : \tau^*M \rightarrow \mathbb{R}$ possède un gradient symplectique $\nabla^\omega h$ en tout point, et que ce champ de gradient est de classe \mathcal{C}^1 .*

*Soit $(q_0, p_0) \in \tau^*M$. Alors il existe une unique courbe maximale $(\bar{q}, \bar{p}) : I \ni 0 \rightarrow T^*M$ de classe H_{loc}^1 , avec I un intervalle, telle que $(q(0), p(0)) = (q_0, p_0)$ et (q, p) satisfasse l'équation hamiltonienne des géodésiques*

$$(\dot{\bar{q}}(t), \dot{\bar{p}}(t)) = \nabla^\omega h(\bar{q}(t), \bar{p}(t)).$$

Alors \bar{q} est une géodésique normale, et donc point critique de l'action à extrémités fixées. De plus, $t \mapsto \|u(\bar{q}(t), \bar{p}(t))\|_{\bar{q}(t)}^2$ est constante.

On dit que (\bar{q}, \bar{p}) satisfait l'équation hamiltonnienne des géodésiques sur τ^*M .

Preuve. Le fait que \bar{q} soit géodésique normale dans le cas où le champ de gradients est de classe C^1 est évident. De plus,

$$t \mapsto \|u(\bar{q}(t), \bar{p}(t))\|_{\bar{q}(t)}^2 = 2h(\bar{q}(t), \bar{p}(t))$$

est constante égale à $2h(q_0, p_0)$ car une fonction est constante le long des trajectoires de son gradient symplectique (voir l'appendice de [Mon02]). \square

Lorsque le gradient symplectique de h est plus régulier, on peut montrer que les géodésiques normales sont bien des géodésiques locales. C'est le résultat principal de cette partie.

Théorème 2.1. Géodésiques normales. *Supposons que le champ de gradients symplectiques $\nabla^\omega h$ soit de classe C^2 sur τ^*M . Soit $(q_0, p_0) \in \tau^*M$, tel que $\text{im}(\xi_{q_0})$ ne soit pas incluse dans $\ker p_0$, et soit $(q, p) : I \ni 0 \rightarrow \tau^*M$ la solution maximale de l'équation hamiltonnienne des géodésiques telle que $(q(0), p(0)) = (q_0, p_0)$.*

Alors \bar{q} est une géodésique locale. Si la structure est forte, c'est donc une géodésique.

Remark 2.10. Ce théorème implique le théorème 1.5 sur les géodésiques normales en dimension finie.

Remark 2.11. L'espace cotangent relatif adapté n'est pas forcément unique. Selon celui que l'on choisit, on peut obtenir plus ou moins de géodésiques normales.

Remark 2.12. Dans le cas d'une structure forte, T^*M est toujours adapté, et le hamiltonien réduit admet toujours un gradient symplectique lisse. On a donc une unique géodésique normale de point initial q_0 et de covecteur initial p_0 pour tout $(q_0, p_0) \in T^*M$.

Preuve. La preuve s'inspire à la fois du cas riemannien classique et du cas sous-riemannien de dimension finie. On commence par choisir un voisinage de coordonnées U de q_0 dans M , qu'on identifie à l'espace de Banach B sur lequel M est modelé.

Il reste donc à montrer le point principal, c'est-à-dire que si $\nabla^\omega h$ est de classe C^2 , alors \bar{q} est une géodésique locale. Autrement dit, que pour $0 < t_1$ assez petit, il existe un voisinage \bar{U} de $\bar{q}([0, t_1])$ tel que pour tout système horizontal $(q, u) : [0, 1] \rightarrow \mathcal{H}_{|\bar{U}}$ avec $q(0) = q_0$ et $q(1) = \bar{q}(t_1)$, on ait

$$L(\bar{q}, \bar{u}) \leq L(q, u).$$

Soit U un système de coordonnées au voisinage de q_0 , que l'on identifiera à un ouvert d'un espace de Banach B . On va travailler sur une trivialisations $\tau^*M \oplus_M \mathcal{H}|_U \simeq B \times G \times H$, où G est isomorphe à $\tau_{q_0}^*M$ et H à \mathcal{H}_{q_0} .

On va montrer qu'on peut trouver une calibration de la trajectoire $t \mapsto \bar{q}(t)$ au voisinage de q_0 . C'est-à-dire une 1-forme fermée θ telle qu'il existe une constante $c > 0$ pour tout t assez proche de 0,

$$\theta_{\bar{q}(t)}(\dot{\bar{q}}(t)) = c\sqrt{h(\bar{q}(t), \bar{p}(t))} = c\sqrt{g_{\bar{q}(t)}(u(\bar{q}(t), \bar{p}(t)), u(\bar{q}(t), \bar{p}(t)))}$$

et telle que pour tout $(q, u) \in \mathcal{H}$ avec q au voisinage de q_0 ,

$$\theta_q(\xi_q u) \leq c\sqrt{g_q(u, u)}.$$

Une fois θ trouvée, quitte à réduire le voisinage, on peut supposer que θ est exacte. Soit $(q, u) \in \Omega_{q_0}^{\mathcal{H}}$ tel que q reste au voisinage de q_0 , $q_0 = q(0)$ et $\bar{q}(t_1) = q(1)$, avec $t_1 > 0$ assez petit. Alors

$$L(q, u) = \int_0^1 \sqrt{g_{q(t)}(u(t), u(t))} dt \geq \frac{1}{c} \int_0^1 \theta_{q(t)}(\underbrace{\xi_{q(t)} u(t)}_{=\dot{q}(t)}) dt = \frac{1}{c} \int_q \theta.$$

Puisque θ est exacte, on a

$$\int_q \theta = \int_{\bar{q}|_{[0, t_1]}} \theta.$$

On obtient donc

$$L(q, u) \geq \frac{1}{c} \int_{\bar{q}|_{[0, t_1]}} \theta = \frac{1}{c} \int_0^{t_1} \theta(\dot{\bar{q}}(t)) dt = \int_0^{t_1} \sqrt{g_{\bar{q}(t)}(\bar{u}(t), \bar{u}(t))} dt = L(\bar{q}|_{[0, t_1]}, \bar{u}|_{[0, t_1]}),$$

où on a noté $\bar{u}(t) = u(\bar{q}(t), \bar{p}(t))$: (\bar{q}, \bar{u}) est bien une géodésique locale au voisinage de q_0 .

Il reste à construire la forme θ . On va utiliser le flot Φ du champ de gradient symplectique de h . En effet, comme $\nabla^\omega h$ est, par hypothèse, un champ de vecteurs \mathcal{C}^2 sur τ^*M , il possède un flot local, soit une application

$$\Phi = (\Phi_M, \Phi_\tau) :]-\varepsilon, \varepsilon[\times U_0 \times V_0 \rightarrow U_1 \times V_1,$$

de classe \mathcal{C}^2 , avec U_0 et U_1 des voisinages de q_0 dans B , V_0 et V_1 des voisinages de p_0 dans G , et $\varepsilon > 0$. À t fixé $(q, p) \mapsto \Phi(t, q, p)$ est un difféomorphisme sur son image, avec $\Phi(0, q, p) = (q, p)$.

D'un autre côté, p_0 s'identifie à un élément de B^* . On note $U'_0 = \ker p_0 \cap U_0$. On peut alors définir l'application $\varphi :]-\varepsilon, \varepsilon[\times U_0 \rightarrow U_1$ par

$$\varphi(t, q) = \Phi_M \left(t, q, \sqrt{\frac{h(q, p_0)}{h(q_0, p_0)}} p_0 \right).$$

Pour faire plus court, on notera $n(q) = \sqrt{\frac{h(q_0, p_0)}{h(q, p_0)}}$. Ce nombre est un facteur de normalisation :

$$\forall q \in U'_0, \quad h(q, n(q)p_0) = h(q_0, p_0).$$

Notons que un tel q , la courbe $t \mapsto \varphi(t, q)$ est la projection sur M de la géodésique normale de condition initiale $(q, n(q)p_0)$ dans nos coordonnées locales.

On prouve alors le lemme suivant.

Lemme 2.3. *Quitte à réduire U'_0 et ε , φ est un difféomorphisme de $]-\varepsilon, \varepsilon[\times U'_0$ sur $U_2 = \varphi(]-\varepsilon, \varepsilon[\times U'_0)$*

Preuve. En effet, pour tout δq dans $\ker p_0$, on obtient

$$\partial_q \varphi(0, q_0) \delta q = \partial_q (\Phi_M(0, q_0, p_0)) \delta q = \delta q.$$

Mais $\partial_t \varphi(0, q_0) = \dot{\bar{q}}(0) = \xi_{q_0} u(q_0, p_0)$. Comme $\text{im}(\xi_{q_0}) \not\subseteq \ker p_0$, $u(q_0, p_0)$ est non nul. Donc

$$h(q_0, p_0) = \frac{1}{2} p_0 \xi_{q_0} u(q_0, p_0) \neq 0,$$

d'où le fait que $\partial_t \varphi(0, q_0)$ n'appartienne pas à $\ker p_0$. Donc $d\varphi(0, q_0)$ est un isomorphisme linéaire. Le théorème d'inversion local implique termine la preuve. \square

On écrira, pour q dans U_2 , $\varphi^{-1}(q) = (t(q), q_0(q))$. Cela nous permet alors de définir sur U_2 la 1-forme

$$\theta_q = \Phi_\tau(t_q, q_0(q), n(q_0(q))p_0) \in \tau_q^*M \subset T_q^*M.$$

Autrement dit, θ_q est le moment en t_q de la géodésique hamiltonienne partant de $q_0(q)$ et de moment initial $n(q_0(q))p_0$. Cette remarque a plusieurs conséquences :

- On a $\theta_{\bar{q}(t)} = \bar{p}(t)$ pour tout t , d'où

$$\theta_{\bar{q}(t)}\dot{q}(t) = \bar{p}(t)\xi_{\bar{q}(t)}\bar{u}(t) = g(\bar{u}(t), \bar{u}(t)) = \sqrt{2h(q_0, p_0)}\sqrt{g_{\bar{q}(t)}(\bar{u}(t), \bar{u}(t))}.$$

- Pour tout (q, u) dans $U_2 \times H$,

$$\theta_q\xi_q u = g_q(u(q, \theta_q), u) \leq \sqrt{g_q(u(q, \theta_q), u(q, \theta_q))}\sqrt{g_q(u, u)}. \quad (2.4)$$

De plus, comme le hamiltonien est constant le long des trajectoires de son gradient, on obtient

$$g_q(u(q, \theta_q), u(q, \theta_q)) = 2h(q, \theta_q) = 2h(q_0, n(q_0(q))p_0) = 2h(q_0, p_0).$$

En injectant cette remarque dans (2.4), on obtient

$$\theta_q\xi_q u \leq \sqrt{2h(q_0, p_0)}\sqrt{g_q(u, u)}.$$

On en déduit que si θ est fermée, c'est bien une calibration de \bar{q} sur un voisinage de q_0 .

Il ne nous reste donc plus qu'à montrer que θ est effectivement fermée sur U_2 . Comme φ est un difféomorphisme sur U_2 , il suffit de montrer que $\varphi^*\theta$ est fermée.

Maintenant, fixons (t, q) dans $] - \varepsilon, \varepsilon[\times U'_0$. Pour tout $(\delta t, \delta q) \in \mathbb{R} \times \ker p_0$, on a

$$\begin{aligned} \varphi^*\theta_{t,q}(\delta t, \delta q) &= \theta_{\varphi(t,q)}(\partial_t\varphi(t, q)\delta t + \partial_q\varphi(t, q)\delta q) \\ &= \Phi_\tau(t, q, n(q)p_0)(\partial_t\varphi(t, q)\delta t + \partial_q\varphi(t, q)\delta q) \\ &= p(t)(\partial_t\varphi(t, q)\delta t + \partial_q\varphi(t, q)\delta q), \end{aligned}$$

où $p(t)$ est telle que la courbe $t \mapsto (q(t), p(t))$ soit la solution de l'équation géodésique de condition initiale $(q, n(q)p_0)$ sur $] - \varepsilon, \varepsilon[$ (on a alors $\varphi(t, q) = q(t)$ et $p(t) = \theta_{q(t)}$).

Mais comme $\varphi(t, q) = q(t)$, on obtient $\partial_t\varphi(t, q) = \dot{q}(t)$, d'où

$$p(t)(\partial_t\varphi(t, q)\delta t) = p(t)\xi_{q(t)}(u(q(t), p(t)), u(q(t), p(t))) = 2h(q(t), p(t)) = 2h(q_0, p_0).$$

Il reste à calculer $p(t)(\partial_q\varphi(t, q)\delta q)$.

Lemme 2.4. *On a $p(t)(\partial_q\varphi(t, q)\delta q) = 0$.*

Preuve. Commençons par calculer $\partial_q\varphi(t, q)\delta q$. On pose, pour un réel s assez proche de 0, et r dans $[0, t]$, $q_s(r) = \varphi(r, q + s\delta q)$. Remarquons que $q(r) = q_0(r)$ pour tout r .

L'application $(s, t) \mapsto q_s(t)$ est de classe \mathcal{C}^2 , et chaque courbe $r \mapsto q_s(r)$ est une la trajectoire géodésique normale et vérifie

$$\dot{q}_s(r) = \xi_{q_s(r)}u_r(s),$$

où $u_r(s) = u(q_s(r), \theta_{q_s(r)})$ est de classe \mathcal{C}^1 . On notera $u_0(r) = u(r)$. Enfin, toutes ces courbes ont la même action

$$A(q_s, u_s) = h(q_0, p_0)t.$$

On a également $\partial_q\varphi(t, q)\delta q = \partial_s(q_s(t))_{s=0}$.

Pour plus de simplicité, on va noter

$$\partial_s(q_s(r))_{s=0} = \delta q(r), \quad r \in [0, t],$$

et

$$\partial_s(u_s(r))_{s=0} = \delta u(r), \quad r \in [0, t].$$

Comme les dérivées partielles commutent, on obtient

$$\dot{\delta}q(r) = \partial_q(\xi_{q(r)}u(r))\delta q(r) + \xi_{q(r)}\delta u(r). \quad (2.5)$$

D'un autre côté, $\partial_s A(q_s, u_s)_{s=0} = 0$, donc

$$0 = \int_0^t \frac{1}{2} \partial_q(g_{q(r)}(u(r), u(r)))\delta q(r) + g_{q(r)}(u(r), \delta u(r))dr. \quad (2.6)$$

Or, $r \mapsto p(r)$ vérifie

$$\begin{aligned} \dot{p}(r)\delta q(r) &= -\partial_q h(q(r), p(r))\delta q(r) \\ &= -\partial_q H(q(r), p(r), u(r))\delta q(r) \\ &= \frac{1}{2} \partial_q(g_{q(r)}(u(r), u(r)))\delta q(r) - p(r)\partial_q(\xi_{q(r)}u(r))\delta q(r). \end{aligned}$$

En injectant cette expression dans (2.6), et en utilisant successivement l'égalité $p(t)\xi_{q(t)} = g_{q(t)}(u(t), \cdot)$ puis (2.5) on obtient

$$\begin{aligned} 0 &= \int_0^t \dot{p}(r)\delta q(r) + p(r)\partial_q(\xi_{q(r)}u(r))\delta q(r) + g_{q(r)}(u(r), \delta u(r))dr \\ &= \int_0^t \dot{p}(r)\delta q(r) + p(r) \left(\partial_q(\xi_{q(r)}u(r))\delta q(r) + \xi_{q(r)}\delta u(r) \right) dr \\ &= \int_0^t \dot{p}(r)\delta q(r) + p(r)\dot{\delta}q(r)dr \\ &= \int_0^1 (p\dot{\delta}q)(r)dr = p(t)\delta q(t) - p(0)\delta q(0). \end{aligned}$$

Autrement dit, la quantité $t \mapsto p(t)\delta q(t)$ est conservée le long de la trajectoire. Comme $\delta q(0) = \delta q \in \ker p_0$ et $p(0) = n(q(0))p_0$, on a $p(0)\delta q(0) = 0$, d'où

$$p(t)\delta q(t) = p(t)\partial_q \varphi(t, q)\delta q = 0.$$

□

Finalement,

$$\varphi^* \theta_{t,q}(\delta t, \delta q) = 2h(q_0, p_0)\delta t,$$

d'où

$$\varphi^* \theta_{t,q} = 2h(q_0, p_0)dt.$$

La 1-forme $\varphi^* \theta$ est bien fermée, et θ est donc une calibration de \bar{q} sur U_2 .

□

2.9 Exemple d'application sur le groupe des difféomorphismes de \mathbb{R}^d

Les exemples qui suivent sont une variante sous-riemannienne de certaines métriques étudiées par Ebin et Marsden dans [EM70]. Dans la suite, on identifiera les éléments v d'un espace vectoriel \mathbb{R}^n à des vecteurs colonnes de taille n , et on notera v^T leur transposée.

Soit d un entier strictement positif. On note $e = \text{Id}_{\mathbb{R}^d}$. On considère, pour s un réel strictement supérieur à $d/2 + 1$, l'espace $\mathcal{D}^s(\mathbb{R}^d)$ des difféomorphismes φ de classe \mathcal{C}^1 tels que $\varphi - e$ appartienne à l'espace de Sobolev $H^s(\mathbb{R}^d, \mathbb{R}^d)$. Il s'agit d'un ouvert de l'espace affine $H_e(\mathbb{R}^d, \mathbb{R}^d) = e + H^s(\mathbb{R}^d, \mathbb{R}^d)$, et donc d'une variété de Hilbert. La structure des groupes de difféomorphismes de classe H^s d'une variété sera plus précisément définie dans la section 3.

2.9.1 Structures sous-riemanniennes

Considérons une structure sous-riemannienne $(\mathcal{H}^0, \xi^0, g^0)$ sur \mathbb{R}^d , avec $\mathcal{H}^0 = \mathbb{R}^d \times \mathbb{R}^k$, $\xi^0 : \mathbb{R}^d \rightarrow \mathcal{L}(\mathbb{R}^k, \mathbb{R}^d)$ de classe \mathcal{C}^∞ dont les dérivées sont bornées à tout ordre, et $g_x^0(u, u) = u^T u$.

Alors, pour tout entier $r \in \{0, \dots, s\}$, la structure (\mathcal{H}, ξ, g_r) suivante définit une structure sous-riemannienne de classe \mathcal{C}^∞ sur $\mathcal{D}^s(\mathbb{R}^d)$:

- \mathcal{H} est le fibré trivial $\mathcal{D}^s(\mathbb{R}^d) \times H^s(\mathbb{R}^d, \mathbb{R}^k)$.
- Pour $(\varphi, u) \in \mathcal{H}$, $\xi_\varphi u \in T_\varphi \mathcal{D}^s(\mathbb{R}^d) = H^s(\mathbb{R}^d, \mathbb{R}^d)$ est donné par

$$(\xi_\varphi u)(x) = \xi_{\varphi(x)}^0 u(x).$$

Remarquons que comme ξ^0 est lisse et bornée à tout ordre, ξ est bien de classe \mathcal{C}^∞ , et l'application $(\xi^0 \circ \varphi)u$ appartient bien à $H^s(\mathbb{R}^d, \mathbb{R}^d) = T_\varphi(\mathcal{D}^s(\mathbb{R}^d))$ (voir [EM70, ES96, Sch04]).

- La valeur de $g_{r,\varphi}(u, u) = g_r(u, u)$ ne dépend pas de φ , et est donnée par le carré de la norme de Sobolev H^r de u :

$$g_r(u, u) = \int_{\mathbb{R}^d} L_r u(x)^T u(x) dx = \sum_{i=1}^k \int_{\mathbb{R}^d} L_r u^i(x) u^i(x) dx, \quad (2.7)$$

où $L_r : H^s(\mathbb{R}^d) \rightarrow H^{s-2r}(\mathbb{R}^d)$ est l'opérateur $1 + (-\Delta)^r$ si $r \geq 1$ avec Δ le laplacien euclidien usuel, et $L_0 = 1$. L_r est un opérateur symétrique et elliptique d'ordre $2r$, donc un isomorphisme linéaire et continu [BBI01].

Remark 2.13. Si $s - 2r$ est négatif, (2.7) reste valide, mais est à comprendre au sens des distributions.

Remark 2.14. Une courbe $t \mapsto \varphi(t)$ de difféomorphismes est horizontale si et seulement si pour tout $x \in \mathbb{R}^d$, la courbe $t \mapsto \varphi(t, x)$ est horizontale sur \mathbb{R}^d pour $(\mathcal{H}^0, \xi^0, g^0)$. En effet, notant u un contrôle pour φ et $x(t) = \varphi(t, x)$, on a

$$\dot{x}(t) = \dot{\varphi}(t, x) = (\xi_{\varphi(t)} u(t))(x) = \xi_{\varphi(t,x)}^0 u(t, x) = \xi_{x(t)}^0 u(t, x),$$

qui est bien de la forme $\dot{x}(t) = \xi_{x(t)}^0 v(t)$, pour $v(t)$ dans \mathbb{R}^k . De plus, l'action A_0 de cette courbe pour la métrique g_0 est donnée par

$$\begin{aligned} A_0(\varphi, u) &= \frac{1}{2} \sum_{i=1}^k \int_0^1 \left(\int_{\mathbb{R}^d} (u^i(t, x))^2 dx \right) dt \\ &= \int_{\mathbb{R}^d} \underbrace{\left(\frac{1}{2} \sum_{i=1}^k \int_0^1 (u^i(t, x))^2 dt \right)}_{A^0(\varphi(\cdot, x), u(\cdot, x))} dx \\ &\geq \frac{1}{2} \int_{\mathbb{R}^d} d^0(\varphi(0, x), \varphi(1, x))^2 dx, \end{aligned}$$

où A^0 et d^0 sont respectivement l'action et la distance sous-riemannienne sur \mathbb{R}^d . En particulier, la distance sous-riemannienne correspondant à la métrique g_0 sur $\mathcal{D}^s(\mathbb{R}^d)$ est une vraie distance. De plus, si, pour tout x , $t \mapsto \varphi(t, x)$ est une géodésique minimisante sur \mathbb{R}^d , alors φ en est une sur $\mathcal{D}^s(\mathbb{R}^d)$. Il est aussi intéressant de remarquer que le problème est très similaire à celui de transport optimal de masse sur la variété sous-riemannienne \mathbb{R}^d avec la structure $(\mathcal{H}^0, \xi^0, g^0)$ [FR10].

Remark 2.15. Il découle de la remarque précédente que toutes les distances d_r correspondant aux métriques g_r sont de vraies distances, pas simplement des semi-distances.

Par la suite, on va noter X_i le champ de vecteur $x \mapsto \xi_x^0 e_i$, où (e_1, \dots, e_k) est la base canonique de \mathbb{R}^k . On peut alors écrire en coordonnées

$$(\xi_\varphi u)(x) = \sum_{i=1}^k u_i(x) X_i(\varphi(x)).$$

Chaque X_i est de classe \mathcal{C}^∞ , et est borné à tout ordre.

2.9.2 Espace cotangent adapté

On a

$$T^*\mathcal{D}^s(\mathbb{R}^d) = \mathcal{D}^s(\mathbb{R}^d) \times H^{-s}(\mathbb{R}^d, (\mathbb{R}^d)^*),$$

où

$$H^{-s}(\mathbb{R}^d, (\mathbb{R}^d)^*) = H^s(\mathbb{R}^d, \mathbb{R}^d)^*.$$

Un élément p de l'espace $H^{-s}(\mathbb{R}^d, (\mathbb{R}^d)^*)$ peut être vu comme un champ de 1-formes $p = (p_1, \dots, p_d)$, identifié à un champ de matrices lignes, dont les coefficients p_i sont des distributions appartenant à $H^{-s}(\mathbb{R})$. On dit que p est un *co-courant* d'ordre s .

Pour un champ de vecteurs $X = (X^1, \dots, X^d)$ dans $H^s(\mathbb{R}^d, \mathbb{R}^d)$, on a alors les notations

$$p(X) = p X = (p|X)_{H^{-s}, H^s} = \int_{\mathbb{R}^d} p(x) X(x) dx = \sum_{i=1}^d \int_{\mathbb{R}^d} p_i(x) X^i(x) dx = \sum_{i=1}^d (p_i|X^i)_{H^{-s}, H^s}.$$

Le hamiltonien du problème $H : \mathcal{D}^s(\mathbb{R}^d) \times H^{-s}(\mathbb{R}^d, (\mathbb{R}^d)^*) \times H^s(\mathbb{R}^d, \mathbb{R}^k) \rightarrow \mathbb{R}$ est donc

$$H(\varphi, p, u) = p \xi_\varphi u - \frac{1}{2} g_r(u, u) = \sum_{i=1}^k \int_{\mathbb{R}^d} p(x) X_i(\varphi(x)) u^i(x) - \frac{1}{2} L_r u^i(x) u^i(x) dx.$$

On en déduit que $\partial_u H(\varphi, p, u) = 0$ si et seulement si, pour tout i dans $\{1, \dots, k\}$,

$$\underbrace{L_r u^i}_{\in H^{s-2r}(\mathbb{R}^d, \mathbb{R})} = p X_i \circ \varphi = \underbrace{p_1 X_i^1 \circ \varphi + \dots + p_d X_i^d \circ \varphi}_{\in H^{-s}(\mathbb{R}^d, \mathbb{R})}.$$

Remarquons que comme les X_i sont lisses de dérivées bornées à tout ordre, les produits des $X_i^j \circ \varphi$ par les distributions p_j d'ordre s sont bien définis.

Cette équation n'a de solution que si chaque $p_1 X_i^1 + \dots + p_d X_i^d$ est en fait une fonction de classe de Sobolev au moins H^{s-2r} . C'est le cas, par exemple, si

$$p \in H^{s-2r}(\mathbb{R}^d, (\mathbb{R}^d)^*).$$

Cela nous pousse à prendre l'espace cotangent relatif

$$\tau^*\mathcal{D}^s(\mathbb{R}^d) = \mathcal{D}^s(\mathbb{R}^d) \times H^{s-2r}(\mathbb{R}^d, (\mathbb{R}^d)^*).$$

Comme $\tau^*\mathcal{D}^s(\mathbb{R}^d)$ est dense dans $T^*\mathcal{D}^s(\mathbb{R}^d)$, la discussion qui précède montre que $\tau^*\mathcal{D}^s(\mathbb{R}^d)$ est adapté à la structure (\mathcal{H}, ξ, g_r) .

Dans ce cas, si $(\varphi, p) \in \mathcal{D}^s(\mathbb{R}^d) \times H^{s-2r}(\mathbb{R}^d, \mathbb{R}^d)$, on a $u(\varphi, p) = \sum_{i=1}^k u^i(\varphi, p)e_i$, où

$$u^i(\varphi, p) = L_r^{-1}(pX_i \circ \varphi).$$

Remarquons que $(\varphi, p) \mapsto u(\varphi, p)$ est de classe \mathcal{C}^∞ , et on obtient le hamiltonien réduit

$$h(\varphi, p) = \sum_{i=1}^k \int_{\mathbb{R}^d} p(x)X_i(\varphi(x))L_r^{-1}(pX_i \circ \varphi)(x)dx.$$

Remark 2.16. Si $r > d/2$, alors L_r^{-1} est obtenu par convolution avec un noyau $G : \mathbb{R}^d \rightarrow \mathbb{R}$ [Eva10], d'où

$$u^i(\varphi, p)(x) = \int_{\mathbb{R}^d} G(x-y)p(y)X_i(\varphi(y))dy.$$

2.9.3 Équation hamiltonienne des géodésiques

Il nous faut maintenant calculer

$$\partial_\varphi h(\varphi, p) = \partial_\varphi H(\varphi, p, u(\varphi, p)) \in H^{-s}(\mathbb{R}^d, (\mathbb{R}^d)^*),$$

et vérifier qu'elle appartient à $H^{s-2r}(\mathbb{R}^d, (\mathbb{R}^d)^*)$.

On a, pour tout $\delta\varphi \in H^s(\mathbb{R}^d, \mathbb{R}^d)$, et tout $(\varphi, p, u) \in \tau^*\mathcal{D}^s(\mathbb{R}^d) \times H^s(\mathbb{R}^d, \mathbb{R}^k)$,

$$\partial_\varphi H(\varphi, p, u)\delta\varphi = \partial_\varphi(\xi_\varphi u)^*p\delta\varphi = \sum_{i=1}^k \int_{\mathbb{R}^d} p(x)^T dX_i(\varphi(x))\delta\varphi(x)u^i(x)dx.$$

On peut donc écrire

$$\partial_\varphi H(\varphi, p, u) = \sum_{i=1}^k u^i p dX_i \circ \varphi.$$

Or, chaque $u_i dX_i \circ \varphi$ est un champ de matrices à coefficients dans H^s , et p appartient à H^{s-2r} . On a bien que $\partial_\varphi H(\varphi, p, u) \in H^{s-2r}$. De plus, comme les matrices dX_i sont de classe \mathcal{C}^∞ et bornées à tout ordre et, on voit que $(\varphi, p, u) \mapsto \partial_\varphi H(\varphi, p, u)$ est de classe \mathcal{C}^∞ .

On obtient finalement le gradient symplectique de h :

$$\left\{ \begin{array}{l} \partial_p h(\varphi, p)(x) = \xi_\varphi(u(\varphi, p))(x) = \sum_{i=1}^k L_r^{-1}(p^T X_i \circ \varphi)(x)X_i(\varphi(x)) \in H^s(\mathbb{R}^d, \mathbb{R}^d), \\ -\partial_\varphi h(\varphi, p)(x) = -\sum_{i=1}^k L_r^{-1}(p^T X_i \circ \varphi)(x)(dX_i(\varphi(x)))^T p(x) \in H^{s-2r}(\mathbb{R}^d, \mathbb{R}^d), \end{array} \right.$$

de classe \mathcal{C}^∞ sur $\tau^*\mathcal{D}^s(\mathbb{R}^d)$. On en déduit, en appliquant le théorème 2.1, les équations des géodésiques normales.

Proposition 2.6. Soit $(\varphi_0, p_0) \in \mathcal{D}^s(\mathbb{R}^d) \times H^{s-2r}(\mathbb{R}^d, \mathbb{R}^d)$, et considérons $(\varphi, p) : I \rightarrow \mathcal{D}^s(\mathbb{R}^d) \times H^{s-2r}(\mathbb{R}^d, \mathbb{R}^d)$ une solution maximale de l'équation différentielle

$$\forall x \in \mathbb{R}^d, t \in I, \quad \left\{ \begin{array}{l} \partial_t \varphi(t, x) = \sum_{i=1}^k L_r^{-1}(p(t)X_i \circ \varphi(t))(x)X_i(\varphi(t, x)), \\ \partial_t p(t, x) = -\sum_{i=1}^k L_r^{-1}(p(t)X_i \circ \varphi(t))(x)p(t, x)dX_i(\varphi(t, x)). \end{array} \right.$$

Alors φ est une géodésique locale pour (\mathcal{H}, ξ, g_r) sur $\mathcal{D}^s(\mathbb{R}^d)$.

Remark 2.17. Si $r = 0$, on remarque que (φ, p) satisfait l'équation des géodésiques si et seulement si, pour tout x dans \mathbb{R}^d , $t \mapsto (\varphi(t, x), p(t, x)) \in \mathbb{R}^d \times (\mathbb{R}^d)^*$ satisfait l'équation hamiltonienne des géodésiques sur \mathbb{R}^d pour $(\mathcal{H}^0, \xi^0, g^0)$.

Chapitre 3

Géométries sous-riemanniennes invariantes à droite sur les groupes de difféomorphismes

Dans cette section, nous décrivons les résultats de la première partie de [AT], reproduite en section 9.1 dans laquelle seront données les différentes preuves, ainsi que quelques détails supplémentaires. Il s'agit d'étudier les structures sous-riemanniennes fortes et invariantes à droite sur un groupe de difféomorphismes d'une variété M à géométrie dite bornée [ES96, Sch04].

On commence par rappeler brièvement la structure du groupe $\mathcal{D}^s(M)$ des difféomorphismes H^s sur une telle variété M de dimension d , pour s un entier plus grand que $d/2+1$. On définit ensuite les structures sous-riemanniennes fortes invariantes à droite sur un tel groupe.

On montre ensuite que sur de telles structures, deux difféomorphismes d'une même orbite peuvent toujours être joints par une géodésique minimisante. De plus, la distance sous-riemannienne induite est toujours complète.

On finit par l'étude des géodésiques et des courbes singulières. Une des difficultés qui apparaîtra est que ces structures ne sont à priori pas lisses. Cela vient du fait que le groupe des difféomorphismes, bien que munit d'une structure naturelle de variété de Hilbert, n'est pas un groupe de Lie pour cette structure. Plus précisément, on va voir que pour que la structure invariante à droite engendrée par un espace de Hilbert V de champs de vecteurs soit de classe \mathcal{C}^k , il faut supposer que possède une inclusion continue de V dans l'espace $\Gamma^{s+k}(TM)$ des champs de vecteurs sur M de classe de Sobolev H^{s+k} .

On pourra alors appliquer les théorèmes du chapitre 2, et donner l'équation hamiltonienne des géodésiques.

3.1 Groupe des difféomorphismes

3.1.1 Variété à géométrie bornée

Définition 3.1. [ES96, Sch04] Une variété lisse riemannienne (M, g) de dimension finie d est dite à géométrie bornée lorsque son rayon d'injectivité $\text{inj}(M)$ est non nul et que l'une des deux conditions équivalentes suivantes est satisfaite.

1. Le tenseur de courbure et toutes ses dérivées covariantes sont uniformément bornées.
2. Pour tout $i \in \mathbb{N}$, il existe une constante $C_i > 0$ telle que dans toute carte exponentielle de rayon $\text{inj}(M)/2$, les dérivées d'ordre i des coefficients de g sont uniformément bornées par C_i .

Une telle structure riemannienne permet de généraliser à des variétés non compactes des résultats classiques dans les variétés compacts [ES96, Sch04].

Pour le reste du chapitre, on fixe une variété (M, g) de dimension finie d à géométrie bornée.

3.1.2 Espace $H_e^s(M, M)$

Pour tout entier s et tout fibré vectoriel riemannien (\mathcal{E}, g') sur M de façon invariante, on peut définir l'espace de Banach $\Gamma^{s,p}(\mathcal{E})$ des sections de classe de Sobolev $W^{s,p}$, $p \in [1, +\infty]$. Il suffit de considérer la connection ∇ induite sur \mathcal{E} par g et g' . Alors $\Gamma^{s,p}(\mathcal{E})$ est la complétion de l'espace des sections de classe C^∞ pour la norme

$$\|f\|_{s,p}^p = \sum_{r=0}^s (g'_x(\nabla^r f(x), \nabla^r f(x)))^{p/2} d\text{vol}(x) < \infty,$$

où on a noté $d\text{vol}$ la forme volume sur M . Alors

$$(\Gamma^{s,p}(\mathcal{E}), \|\cdot\|_{s,p})$$

est un espace de Banach, et même de Hilbert si $p = 2$. Voir [ES96, Sch04] pour plus de détails. Dans la suite, on notera simplement $\Gamma^{s,2}(\mathcal{E}) = \Gamma^s(\mathcal{E})$.

Fixons maintenant $s \in \mathbb{N}$ et notons $\exp : TM \rightarrow M$ l'application exponentielle riemannienne. On définit $e = \text{Id}_M$ et

$$H_e^s(M, M) = \{\exp \circ X, X \in \Gamma^s(TM)\}.$$

Exemple 3.1. Si $M = \mathbb{R}^d$, alors $H_e^s(\mathbb{R}^d, \mathbb{R}^d) = e + H^s(\mathbb{R}^d, \mathbb{R}^d)$, où $H^s(\mathbb{R}^d, \mathbb{R}^d)$ est l'espace des champs de vecteurs de classe de Sobolev H^s sur \mathbb{R}^d .

On peut voir [ES96, Sch04] que si $s > d/2 + 1$, $H_e^s(M, M)$ est muni d'une structure de variété de Hilbert. De plus, les injections de Sobolev habituelles

$$H_e^s(M, M) \rightarrow C^\ell(M, M), \quad s > d/2 + \ell,$$

sont définies et continues.

Pour le voir, on note pour $f \in H_e^s(M, M)$

$$T_f H_e^s(M, M) = \{X : M \rightarrow TM \mid X \circ \varphi^{-1} \in \Gamma^s(TM)\} = \Gamma^s(f^*TM).$$

L'application $X \mapsto \exp \circ X$, définie sur $T_f H_e^s(M, M)$ définit une carte lisse de $H_e^s(M, M)$ d'un voisinage de la section nulle sur un voisinage de f , et il est facile de voir que les changements de cartes sont de classe C^∞ . On peut alors définir la distance complète

$$d_{H_e^s(M, M)}(f, g) = \inf\{\|X\|_{s,2}, X \in \Gamma^s(f^*TM), g = \exp \circ X\}.$$

Remarquons que $T_e H_e^s(M, M) = \Gamma^s(TM)$, et que son dual $\Gamma^s(TM)^*$, noté $\Gamma^{-s}(T^*M)$, s'identifie à l'ensemble des 1-formes P sur M à coefficients dans $H^{-s}(M)$, aussi appelés *co-courants* d'ordre s sur M . On peut alors écrire, pour tout X dans $\Gamma^s(TM)$,

$$P(X) = \int_M P(x)X(x)d\text{vol}(x),$$

qui se calcule, par exemple, grâce à des partitions de l'unité.

Remark 3.1. L'application $(f, X) \mapsto (f, \exp \circ X)$, définie sur $TH_e^s(M, M)$, définit une *addition locale lisse* sur $H_e^s(M, M)$. Cela permet de munir des espaces telles que $H^1(0, 1; H^s(M, M))$ d'une structure de variété, voir [KM97].

3.1.3 Groupe des difféomorphismes de classe H^s

À partir de maintenant, on fixe $s > d/2+1$. On note $\text{Diff}(M)$ l'ensemble des difféomorphismes de classe \mathcal{C}^1 de M . Alors

$$\mathcal{D}^s(M) = \text{Diff}(M) \cap H_e^s(M, M)$$

est un ouvert de $H_e^s(M, M)$, appelé *groupe des difféomorphismes de classe H^s* sur M .

Remark 3.2. Il y a d'autres applications considérées comme des difféomorphismes de classe H^s sur M dans la littérature [ES96, Sch04], mais nous ne nous intéresserons qu'à ceux qui font partie de $\mathcal{D}^s(M)$. Il s'agit de la composante connexe de l'identité dans l'espace total de ces difféomorphismes.

Proposition 3.1. [ES96, Omo74, Sch04] *Le groupe $\mathcal{D}^s(M)$ est un groupe topologique pour la composition $(\varphi, \psi) \mapsto \varphi \circ \psi$. De plus, la multiplication à droite $R_\psi \varphi \mapsto \varphi \circ \psi$ est de classe \mathcal{C}^∞ , de dérivée en $e = \text{Id}_M$ donnée par $dR_\psi(e)X = X \circ \psi$, $X \in \Gamma^s(TM)$.*

Enfin, pour tout entier k , l'application

$$\begin{aligned} \mathcal{D}^s(M) \times \Gamma^{s+k}(TM) &\rightarrow T\mathcal{D}^s(M). \\ (\varphi, X) &\mapsto X \circ \varphi \end{aligned}$$

est de classe \mathcal{C}^k .

Remark 3.3. Il est important de signaler que $\mathcal{D}^s(M)$ n'est PAS un groupe de Lie. Les applications de composition et d'inversion sont continues, mais pas différentiables.

La suite décroissante $(\mathcal{D}^s(M))_{s > d/2+1}$ définit en revanche sur son intersection $\mathcal{D}^\infty(M)$ une vraie structure de groupe de Lie de Fréchet, qui a la particularité d'être une structure de groupe de Lie Limite Inverse d'espaces de Hilbert, ou LIH [Omo74, Sch04].

On définit alors l'ensemble des champs de vecteurs invariants à droite sur $\mathcal{D}^s(M)$ comme l'espace des champs de vecteurs $X \in \Gamma(T\mathcal{D}^s(M))$ tels que,

$$\forall \varphi \in \mathcal{D}^s(M), \quad X(\varphi) = X(e) \circ \varphi.$$

Comme pour les groupes de Lie, cet espace s'identifie naturellement à $T_e\mathcal{D}^s(M) = \Gamma^s(TM)$.

3.1.4 Dérivée logarithmique

Soit $\varphi : I \ni 0 \rightarrow \mathcal{D}^s(M)$ une courbe de classe de Sobolev H_{loc}^1 sur $\mathcal{D}^s(M)$. Alors pour tout t dans I on a

$$\varphi(t) = \varphi^X(t) \circ \varphi(0),$$

où φ^X est le *flot* du champ de vecteurs X sur M dépendant du temps et donné par

$$X(t) = \dot{\varphi}(t) \circ \varphi^{-1}(t) \in \Gamma^s(TM), \quad t \in I,$$

c'est-à-dire que φ^X satisfait $\varphi^X(0) = e$ et

$$\dot{\varphi}^X(t) = X(t) \circ \varphi^X(t).$$

De plus, un tel X est unique. On appelle $X : I \rightarrow \Gamma^s(TM)$ la *dérivée logarithmique* de φ .

3.2 Structures sous-riemanniennes invariantes à droite

3.2.1 Construction

À partir de maintenant, on fixe (M, g) une variété de dimension d à géométrie bornée, et on note $e = \text{Id}_M$. On pose également s_0 le plus petit entier tel que $s_0 > d/2$, et $s \geq s_0 + 1$ un autre entier. Une structure sous-riemannienne $(\mathcal{H}, \xi, \bar{g})$ sur $\mathcal{D}^s(M)$ est *invariante à droite* si :

- Le fibré \mathcal{H} est de la forme $\mathcal{H} = \mathcal{D}^s(M) \times V$, où $(V, \|\cdot\|_V)$ est un espace de Banach de champs de vecteurs sur M avec une inclusion continue $V \subset \Gamma^s(TM)$. De plus, $\xi_e X = X$ pour tout $X \in V$.
- À X dans V fixé, le champ de vecteurs $\varphi \mapsto \xi_\varphi X$ est invariant à droite :

$$\xi_\varphi X = (\xi_e X) \circ \varphi = X \circ \varphi.$$

- La métrique \bar{g} est indépendante de φ :

$$\bar{g}_\varphi(X, Y) = \langle X, Y \rangle,$$

où $\langle \cdot, \cdot \rangle$ est une application bilinéaire symétrique définie positive sur V .

Une telle structure est entièrement déterminée par l'espace de champs de vecteurs V et le produit scalaire $\langle \cdot, \cdot \rangle$. Réciproquement, pour tout espace de Banach de champs de vecteurs V muni d'un produit scalaire $\langle \cdot, \cdot \rangle$, il existe une unique structure sous-riemannienne invariante à droite, appelée *structure sous-riemannienne induite* (ou engendrée) par $(V, \langle \cdot, \cdot \rangle)$.

Il faut toutefois faire attention : cette structure n'est pas forcément lisse ! Plus précisément, la proposition 3.1 montre que ξ n'est de classe \mathcal{C}^k que si $V \subset \Gamma^{s+k}(TM)$ avec inclusion continue.

3.2.2 Structures invariantes continues équivalentes à des structures lisses

On dit que deux structures sous-riemanniennes (\mathcal{H}, ξ, g) et (\mathcal{H}', ξ', g') sont équivalentes lorsqu'il existe un isomorphisme de fibrés vectoriels $f : \mathcal{H} \mapsto \mathcal{H}'$ tels que $\xi' = \xi \circ f$, et $g = f^* g'$. Dans ce cas, les courbes horizontales et la distance sous-riemannienne sont identiques pour les deux structures.

Il existe des structures lisses sur $\mathcal{D}^s(M)$ qui sont équivalentes à certaines structures invariantes à droites non lisses. C'est par exemple le cas lorsque $V = \Gamma^s(TM)$ (on est donc dans le cas riemannien), et que

$$\langle X, Y \rangle = \int_M g_x(L_r(X)(x), Y(x)) dx,$$

où $L_r : \Gamma^s(TM) \rightarrow \Gamma^{s-2r}(TM)$ est un opérateur différentiel linéaire, symétrique, elliptique, à coefficients lisses et uniformément bornés à tout ordre sur M . La structure induite est équivalente à $(TM, \text{Id}, \bar{g}')$, avec

$$\bar{g}'_\varphi(X, Y) = \langle X \circ \varphi^{-1}, Y \circ \varphi^{-1} \rangle = \int_M g_x(L_r(X(\varphi^{-1}(x)), Y(\varphi^{-1}(x))) dx,$$

pour $X, Y \in \Gamma^s(T_\varphi M)$.

Bien que l'inversion $\varphi \mapsto \varphi^{-1}$ soit seulement continue, un changement de variable $x = \varphi(y)$ permet de voir que \bar{g}' est bien lisse. Voir [BV13, MP10] pour plus de détails.

C'est par ce changement de structure qu'Ebin et Marsden ont montré dans [EM70] l'existence et l'unicité des solutions de l'équation d'Euler sur les fluides incompressibles sur des variétés compactes. Dans ce cas, la structure invariante à droite est donnée par $V = \ker(\text{div}) \cap \Gamma^s(TM)$ et $L_r = 1$.

Remark 3.4. Les structures sous-riemanniennes définies en section 2.9 ne sont *pas* équivalentes à des structures invariantes à droite.

3.2.3 Courbes horizontales, action, longueur et distance

Considérons une structure invariante à droite sur $\mathcal{D}^s(M)$ induite par $(V, \langle \cdot, \cdot \rangle)$. On note k le plus grand entier tel que V soit continuellement inclus dans $\Gamma^{s+k}(TM)$. Pour cette structure, une courbe $t \mapsto \varphi$ de classe de Sobolev H^1 est horizontale s'il existe $t \mapsto X(t) \in V$ de carré intégrable tel que

$$\dot{\varphi}(t) = X(t) \circ \varphi(t), \quad \text{p.p. } t.$$

En conséquence, φ est horizontale si et seulement si sa dérivée logarithmique $X = \dot{\varphi} \circ \varphi^{-1}$ est dans V et de carré intégrable. De plus, $\varphi \circ \psi$ est horizontale pour tout ψ fixé de $\mathcal{D}^s(M)$.

On dira parfois que le flot d'un champ de vecteurs $X \in L^2(0, 1; V)$ est une *déformation* de M .

Les théorèmes généraux sur les équations différentielles montrent que, réciproquement, pour tout $X \in L^2(0, 1; V)$, on a

$$t \mapsto \sup_{x \in M} (g_x(X(t, x), X(t, x)))$$

intégrable, et donc X est complet sur $[0, 1]$: X admet bien un unique flot φ^X défini sur tout $[0, 1] \times M$. On en déduit que dans le cas particulier des structures invariantes à droite, $\Omega_{\varphi_0}^{\mathcal{H}}$ a une structure de variété lisse avec seule carte globale $L^2(0, 1; V) \rightarrow \Omega_{\varphi_0}^{\mathcal{H}}$ (voir lemme 9.1).

L'action et la longueur d'une courbe horizontale $\varphi : [0, 1] \rightarrow \mathcal{D}^s(M)$ prennent les formes simplifiées suivantes

$$\begin{aligned} A^{\mathcal{D}^s(M)}(\varphi) &= A(\varphi, X) = A(X) = \frac{1}{2} \int_0^1 \langle X(t), X(t) \rangle dt, \\ L^{\mathcal{D}^s(M)}(\varphi) &= L(\varphi, X) = L(X) = \int_0^1 \sqrt{\langle X(t), X(t) \rangle} dt. \end{aligned}$$

On déduit aisément des observations qui précèdent que la distance sous-riemannienne d_{SR} est invariante à droite : pour tous φ, ψ dans $\mathcal{D}^s(M)$,

$$d_{SR}(\varphi, \psi) = d_{SR}(e, \varphi \circ \psi^{-1}).$$

L'orbite \mathcal{O}_e de l'identité est donc un sous-groupe de $\mathcal{D}^s(M)$.

3.2.4 Complétude dans le cas fort

À partir de maintenant, on considère uniquement des structures fortes. Cela signifie que $(V, \langle \cdot, \cdot \rangle)$ est un espace de Hilbert, et donc $L^2(0, 1; V)$ aussi.

Théorème 3.1. *Dans le cas d'une structure forte, la distance sous-riemannienne est une véritable distance. De plus, tout couple de points $(\varphi, \psi) \in \mathcal{D}^s(M)^2$ tels que $d_{SR}(\varphi, \psi)$ soit finie peuvent être joints par une géodésique minimisante. Enfin, l'espace métrique $(\mathcal{D}^s(M), d_{SR})$ est complet.*

La preuve, donnée en section 9.1.2, repose sur le lemme suivant, prouvé en section 10.1.5. C'est une conséquence simple du théorème d'Arzela-Ascoli. On commence par définir $d_{H_e^s(U, M)}(\varphi|_U, \psi|_U)$, pour $U \subset M$, $\varphi, \psi \in \mathcal{D}^s(M)$ et $s > d/2 + 1$ par

$$\inf \left\{ \|X\|_{\Gamma^s(\varphi^* T\varphi(U))} \mid X \in \varphi^* \Gamma^s(T\varphi(U)), \psi|_U = \exp \circ X \right\}.$$

Lemme 3.1. *Soit $(X_n)_{n \in \mathbb{N}}$ une suite bornée de l'espace de Hilbert $L^2(0, 1; V)$, et $(\varphi_n)_{n \in \mathbb{N}}$ la suite de flots associée. Alors il existe une sous-suite $(X_{n_j})_{j \in \mathbb{N}}$ qui converge vers $\bar{X} \in L^2(0, 1; V)$ de flot $\bar{\varphi}$ tel que X_{n_j} converge faiblement vers \bar{X} et, pour tout domaine borné U de M ,*

$$\sup_{t \in [0, 1], x \in C} d_{H_e^{s-1}(U, M)}(\varphi_{n_j}|_U(t), \bar{\varphi}|_U(t)) \xrightarrow{j \rightarrow +\infty} 0.$$

3.3 Étude de l'orbite de e

Ici, $(V, \langle \cdot, \cdot \rangle)$ est un espace de Hilbert de champs de vecteurs avec inclusion continue dans $\Gamma^s(TM)$, et on considère la structure sous-riemannienne invariante à droite associée sur $\mathcal{D}^s(M)$, et on note d_{SR} la distance sous-riemannienne associée.

Le but de cette section est de donner des conditions pour obtenir un théorème de Chow-Rashevski, c'est-à-dire des conditions pour que l'orbite \mathcal{O}_e de e soit dense dans (voire égal à) la composante connexe $\mathcal{D}_0^s(M)$ de e dans $\mathcal{D}^s(M)$.

3.3.1 Contrôlabilité approchée

Comme on l'a vu en section 2.2, il existe quelques résultats qui assurent la densité de \mathcal{O}_e dans $\mathcal{D}_0^s(M)$.

Proposition 3.2. *Si V est dense dans $\Gamma^s(TM)$, alors \mathcal{O}_e est dense dans $\mathcal{D}_0^s(M)$.*

Un résultat plus général peut être déduit du résultat de contrôlabilité d'Agrachev et Caponigro [AC09] lorsque M est compact.

Proposition 3.3. (Agrachev-Caponigro) *Supposons que M soit compact et soient X_1, \dots, X_r des champs de vecteurs de classe \mathcal{C}^∞ sur M tels que tout couple de points de M peuvent être joints par une courbe horizontale pour la structure sous-riemannienne sur M induite par les X_i .*

Enfin, supposons que

$$\left\{ \sum_{i=1}^r u_i X_i \mid u_1, \dots, u_r \in \mathcal{C}^\infty(M) \right\} \subset V.$$

Alors \mathcal{O}_e contient la composante connexe de e de l'ensemble $\mathcal{D}^\infty(M)$ des difféomorphismes de M de classe \mathcal{C}^∞ . En particulier, \mathcal{O}_e est dense dans $\mathcal{D}_0^s(M)$.

Le résultat principal de [AC09] est en fait plus fort. Il implique que si M est compact et si tout couple de points $x, y \in M$ peuvent être reliés par une courbe horizontale sur M pour la structure sous-riemannienne induite par les X_i , alors

$$\exists m \in \mathbb{N}, \forall \varphi \in \mathcal{D}_0^\infty(M), \exists u_1, \dots, u_m \in \mathcal{C}^\infty(M), i_1, \dots, i_m \in \{1, \dots, r\}$$

$$\varphi = \varphi^{u_1 X_{i_1}}(1) \circ \dots \circ \varphi^{u_m X_{i_m}}(1).$$

3.3.2 Contrôlabilité exacte

Comme on l'a vu au chapitre 2, il n'est en général pas possible d'obtenir des résultats de contrôlabilité exacte en géométrie sous-riemannienne. Toutefois, le théorème d'Agrachev et Caponigro [AC09] est remarquable car il fait exactement cela : sous certaines conditions naturelles, tout difféomorphisme lisse peut être connecté à e par une courbe horizontale.

Nous allons généraliser ce théorème au cas des difféomorphismes H^s . Mieux encore, nous allons voir que si V est l'ensemble des champs de vecteurs horizontaux de classe H^s sur M compact pour une structure sous-riemannienne lisse sur M qui satisfait les conditions de non-intégrabilité usuelles, alors la topologie induite par d_{SR} sur $\mathcal{D}^s(M)$ coïncide avec sa topologie intrinsèque.

Soit M une variété compacte de dimension d , et soit (\mathcal{H}, ξ, g) une structure sous-riemannienne lisse sur M . On suppose que l'algèbre de Lie \mathcal{L} engendrée par les champs de vecteurs horizontaux sur M pour cette structure satisfait $\mathcal{L}_x = T_x M$ pour tout x .

Soit alors e_1, \dots, e_r des sections C^∞ engendrant \mathcal{H} entier, et notons $X_1 = \xi e_1, \dots, X_r = \xi e_r$. On peut alors trouver $m \in \mathbb{N}$ et des familles I_1, \dots, I_m , avec $I_l = \{i_l^1, \dots, i_l^{j_l}\}$, où $j_l \leq j_{l+1}$, telles que pour tout $x \in M$,

$$T_x M = \text{vect}(X_{I_l}(x))_{l \in \{1, \dots, m\}},$$

où X_{I_l} est le crochet de Lie itéré des $X_{i_l^1}, \dots, X_{i_l^{j_l}}$.

Theorem 3.1. *On conserve le cadre décrit ci-dessus, et on suppose que V est l'ensemble des champs de vecteurs de classe \mathcal{H}^s et horizontaux pour la structure (\mathcal{H}, ξ, g) sur M , et que $\langle \cdot, \cdot \rangle$ induit une norme de Hilbert équivalente à la norme H^s .*

Alors il existe une application $\phi : H^s(M, \mathbb{R}^m) \rightarrow \mathcal{D}^s(M)$ de classe C^1 avec $\phi(0) = e$ et une constante $C > 0$ telle que

$$d_{SR}(e, \phi(u_1, \dots, u_m)) \leq C \sum_{i=1}^m \|u_i\|_s^{1/j^i}.$$

En particulier, $\mathcal{O}_e = \mathcal{D}_0^s(M)$ et la topologie induite sur $\mathcal{D}^s(M)$ par d_{SR} coïncide avec sa topologie intrinsèque

La preuve, donnée au chapitre 9. Il s'agit d'une variante du cas de la dimension finie.

3.4 Espaces de Hilbert à noyau reproduisant

3.4.1 Définition

Un espace de Hilbert à noyau reproduisant de champs de vecteurs de M [You10] est un espace de Hilbert $(V, \langle \cdot, \cdot \rangle)$, avec $V \subset \Gamma(TM)$ et une inclusion continue de V dans l'espace des champs de vecteurs continus sur M . On note alors $K^V : V^* \rightarrow V$ la réciproque de l'isométrie $V \rightarrow V^*$ donnée par $X \mapsto \langle X, \cdot \rangle$ [BBI01].

Dans un tel cas, pour tout $(x, p) \in T^*M$, la forme linéaire $p \otimes \delta_x$ définie par

$$p \otimes \delta_x(Y) = pY(x), \quad Y \in V$$

est continue sur V , et appartient donc à V^* . Il existe alors un unique champ de vecteur $X = K^V p \otimes \delta_x$ de V tel que, sur V ,

$$p \otimes \delta_x = \langle X, \cdot \rangle.$$

De plus, $p \mapsto K^V p \otimes \delta_x$ est linéaire pour tout x . En notant $K(y, x)p = K^V p \otimes \delta_x(y)$, on obtient une section K de $TM \otimes TM$, appelée le *noyau reproduisant* de V , où

$$TM \otimes TM = \{(x, y, A) \in M \times M \times T_x M \otimes T_y M = \mathcal{L}(T_y^* M, T_x M)\}.$$

On a noté $\mathcal{L}(A, B)$ l'espace des applications linéaires continues entre deux espaces de Banach A et B . De plus, $y \mapsto K(y, x)p$ appartient à V et

$$p(Y(x)) = \langle K(\cdot, x)p, Y \rangle, \quad (x, p) \in T^*M, Y \in V.$$

Remark 3.5. Comme $K(x, y) \in T_x M \otimes T_y M$, on peut également considérer que $K(x, y)$ appartient à $\mathcal{L}(T_x^* M, T_y M)$ plutôt que $\mathcal{L}(T_y^* M, T_x M)$. On appelle cette application la *transposée* de $K(x, y)$, et on la note $K(x, y)^T$. Elle vérifie, pour tout p_1 dans $T_x^* M$ et p_2 dans $T_y^* M$,

$$\underbrace{p_1 K(x, y) p_2}_{\in T_x M} = \underbrace{p_2 K(x, y)^T p_1}_{\in T_y M}.$$

Alors on voit facilement que $K(x, y)^T = K(y, x)$. En effet,

$$p_1 K(x, y) p_2 = \langle K(\cdot, x) p_1, K(\cdot, y) p_2 \rangle = p_2 K(y, x) p_1.$$

On dit que K est une section *symétrique* de $TM \otimes TM$.

De plus, si V s'injecte continuellement dans $\Gamma^s(TM)$, on vérifie facilement (voir section 10.1.1) que pour tout co-courant P d'ordre s , l'élément $K^V P \in V$ tel que $P = \langle K^V P, \cdot \rangle$ est donné par la convolution

$$K^V P(x) = \int_M \underbrace{K(x, y) P(y)}_{\in T_x M} dy, \quad x \in M.$$

En particulier, pour tous co-courants P, P' de classe H^{-s} ,

$$\langle K^V P, K^V P' \rangle = P(K^V P') = P'(K^V P) = \int \int_{M \times M} P(x) K(x, y) P'(y) dy dx.$$

En conséquence, K est au moins de classe H^s .

Remark 3.6. On en déduit que pour toute 1-forme lisse à support compact P sur M , on a

$$\int \int_{M \times M} P(x) K(x, y) P(y) dy dx \geq 0,$$

avec égalité si et seulement si le champ de vecteurs $K_V P$ donné par

$$K^V P(x) = \int_M K(x, y) P(y) dy,$$

est nul pour tout x dans M . On dit que K est une section *définie positive*.

L'espace V est entièrement déterminé par son noyau reproduisant.

3.4.2 Construction d'espaces à noyau reproduisant

Réciproquement, soit K une section symétrique définie positive de $TM \otimes TM$ de classe H^{2s} pour s un entier plus grand que $d/2 + 1$. Alors il existe un unique espace de Hilbert $(V, \langle \cdot, \cdot \rangle)$ de champs de vecteurs de classe H^s admettant K comme noyau reproduisant [You10, Aro50].

En effet, V est donné par la complétion de l'espace préhilbertien $(V_0, \langle \cdot, \cdot \rangle)$, où V_0 est l'ensemble des champs de vecteurs de la forme

$$K^V P(\cdot) = \int_M K(\cdot, y) P(y) dy,$$

où P est une 1-forme lisse à support compact, et

$$\langle K^V P, K^V P' \rangle = \int \int_{M \times M} P(x) K(x, y) P'(y) dy dx.$$

Cette construction permet de définir directement un noyau K , et de travailler uniquement avec celui-ci sans jamais expliciter V .

C'est une méthode très utilisée en analyse des déformations. Un tel noyau permet en effet de résoudre facilement et explicitement les problèmes d'interpolation de splines sur V (voir [TY05] par exemple).¹

1. On rappelle qu'un problème d'interpolation de splines sur V est celui de trouver, pour $(x_1, v_1), \dots, (x_n, v_n)$ dans TM fixés, le champ de vecteurs X de norme minimale dans V tel que pour tout i dans $\{1, \dots, n\}$, $X(x_i) = v_i$.

3.5 Hamiltonien, géodésiques normales et anormales

On reprend les notations de la section 3.2, et note k le plus grand entier tel que V ait une inclusion continue dans $\Gamma^{s+k}(TM)$. On supposera que $k \geq 1$.

3.5.1 Hamiltoniens

Les hamiltoniens $H^\lambda : T^*\mathcal{D}^s(M) \times V \rightarrow \mathbb{R}$, $\lambda \in \{0, 1\}$, de notre système sont donnés par

$$H^\lambda(\varphi, P, X) = P(X \circ \varphi) - \frac{\lambda}{2} \langle X, X \rangle = dR_\varphi^* P(X) - \frac{\lambda}{2} \langle X, X \rangle.$$

Les H^λ sont de classe \mathcal{C}^k . On peut trouver les géodésiques normales et les courbes singulières en appliquant la proposition 2.4. On rappelle que, dans notre cas, $\Omega_{\varphi_0}^{\mathcal{H}}$ est une variété lisse que l'on peut identifier à l'espace de contrôles $L^2(0, 1; V)$ par $X \mapsto \varphi^X \circ \varphi_0$, et que E désigne l'application point final.

Proposition 3.4. *Soit $\varphi : [0, 1] \rightarrow \mathcal{D}^s(M)$ une courbe horizontale de contrôle associé $X \in L^2(0, 1; V)$. Supposons qu'il existe $(P, \lambda) \neq (0, 0)$, où $(\varphi, P) : [0, 1] \rightarrow T^*\mathcal{D}^s(M)$ est de classe H^1 et $\lambda = 0$ ou 1 tel que*

$$\begin{cases} 0 & = \partial_X H^\lambda(\varphi(t), P(t), X(t)), \\ (\dot{\varphi}(t), \dot{P}(t)) & = \nabla^\omega H^\lambda(\varphi(t), P(t), X(t)), \end{cases} \quad (3.1)$$

- Si $\lambda = 0$, alors, pour tous $t \in [0, 1]$, $P(t) \neq 0$ et $V \circ \varphi \subset \ker P(t)$. L'image de $dE_{(q,u)}$ est incluse dans un sous-espace fermé strict de $T_{\varphi(1)}\mathcal{D}^s(M) : \varphi$ est une courbe singulière.
- Si $\lambda = 1$, (φ, X) est un point critique de A à extrémités fixées : φ est une géodésique normale.

Étudions plus en détails les conséquences de ces équations. Dans un système de coordonnées canoniques, on a

$$\partial_P H^\lambda(\varphi, P)(x) = X(\varphi(x)), \quad x \in M,$$

et

$$-\partial_\varphi H^\lambda(\varphi, P, X)(x) = -\partial_\varphi (X(\varphi(x)))^* P(x) = -P(x) dX(\varphi(x)), \quad x \in M.$$

Remark 3.7. Cette dernière équation est un abus de notation : elle signifie que pour tout $\delta\varphi$ dans $T_\varphi\mathcal{D}^s(M)$,

$$-\partial_\varphi H^\lambda(\varphi, P, X)\delta\varphi = - \int_M P(x) \underbrace{dX(\varphi(x))\delta\varphi(x)}_{\in T_{\varphi(x)}M} dx.$$

En conséquence, dans la proposition 3.4, le support de $t \mapsto P(t)$ reste constant.

3.5.2 Hamiltonien réduit

Dans le cas $\lambda = 1$,

$$\partial_X H^1(\varphi, P, X) = 0$$

si et seulement si

$$X = X(\varphi, P) = K^V dR_\varphi^* P.$$

Cela permet de calculer le hamiltonien normal

$$h(\varphi, P) = H^1(\varphi, P, X(\varphi, P)) = \frac{1}{2} P(K^V dR_\varphi^* P) \circ \varphi.$$

On définira, pour plus de simplicité, l'opérateur $K_\varphi^V : T_\varphi^* \mathcal{D}^s(M) \rightarrow T_\varphi \mathcal{D}^s(M)$ par $K_\varphi^V = dR_\varphi K^V dR_\varphi^*$, de sorte que

$$h(\varphi, P) = \frac{1}{2} P K_\varphi^V P.$$

Remark 3.8. Grâce au noyau, on peut écrire

$$X(\varphi, P)(x) = \int_M K(x, \varphi(y)) P(y) dy, \quad x \in M,$$

et

$$h(\varphi, P) = \frac{1}{2} \int \int_{M \times M} P(x) K(\varphi(x), \varphi(y)) P(y) dy dx.$$

3.5.3 Équation hamiltonienne des géodésiques

Rappelons qu'on a supposé que V a une inclusion continue dans $\Gamma^{s+k}(TM)$, $k \in \mathbb{N}^*$. Dans ce cas, les H^λ et h sont de classe \mathcal{C}^k . De plus, h possède un champ de gradients symplectiques de classe \mathcal{C}^{k-1} .

Ce champ de gradient est donné dans un système de coordonnées canoniques par

$$\nabla^\omega h(\varphi, P) = (\partial_P H(\varphi, P, X), -\partial_\varphi H(\varphi, P, X)) = (X \circ \varphi, -P dX \circ \varphi),$$

où $X = K^V dR_\varphi^* P$.

Le gradient symplectique est facile à calculer si on connaît le noyau reproduisant K de V . En effet, dans un système de coordonnées canoniques, et en utilisant la symétrie de K , on obtient pour tout x dans M ,

$$\partial_P h(\varphi, P)(x) = K^V dR_\varphi^* P = \int_M K(\varphi(x), \varphi(y)) P(y) dy$$

et

$$-\partial_\varphi h(\varphi, P)(x) = -P(x) \int_M \partial_1 K(\varphi(x), \varphi(y)) P(y) dy,$$

où $\partial_1 K(\varphi(x), \varphi(y))$ désigne la dérivée de K au point $(\varphi(x), \varphi(y))$ par rapport à la première variable.

Remark 3.9. Cette équation est aussi un abus de notation : elle signifie que pour tout $\delta\varphi$ dans $T_\varphi \mathcal{D}^s(M)$,

$$-\partial_\varphi h(\varphi, P) \delta\varphi = - \int \int_{M \times M} P(x) (\partial_1 K(\varphi(x), \varphi(y)) \delta\varphi(x)) P(y) dy dx.$$

On peut maintenant appliquer le théorème 2.1. Comme la structure sous-riemannienne considérée est forte, on peut prendre $\tau^* M = T^* M$.

Théorème 3.2. *Équation hamiltonienne des géodésiques pour des structures invariantes à droite.*

Supposons que $(V, \langle \cdot, \cdot \rangle)$ possède une inclusion continue dans $\Gamma^{s+k}(TM)$, avec k un entier supérieur ou égal à 2. Soit $(\varphi_0, P_0) \in T^* \mathcal{D}^s(M)$, tel que $dR_\varphi(V)$ ne soit pas inclus dans $\ker P_0$. Il existe une unique courbe maximale $(\varphi, P) : I \ni 0 \rightarrow T^* \mathcal{D}^s(M)$ de classe H_{loc}^1 , avec I un intervalle, telle que $(\varphi(0), P(0)) = (\varphi_0, P_0)$ et pour presque tout t dans I ,

$$(\dot{\varphi}(t), \dot{P}(t)) = \nabla^\omega h(\varphi(t), P(t)).$$

Dans un système de coordonnées canoniques, cette équation s'écrit, pour tout (t, x) dans $I \times M$,

$$\begin{cases} \partial_t \varphi(t, x) = \int_M K(\varphi(t, x), \varphi(t, y)) P(t, y) dy \\ \partial_t P(t, x) = -P(t, x) \int_M \partial_1 K(\varphi(t, x), \varphi(t, y)) P(t, y) dy. \end{cases} \quad (3.2)$$

Alors φ est une géodésique normale, et donc cette courbe est un point critique de l'action à extrémités fixées pour la structure sous-riemannienne invariante à droite induite par V sur $\mathcal{D}^s(M)$.

De plus, si $k \geq 3$, φ est une géodésique pour cette structure, c'est-à-dire qu'elle minimise la longueur entre $\varphi(t_0)$ et $\varphi(t_1)$ pour t_0 et t_1 assez proches l'un de l'autre.

Enfin, $t \mapsto \langle K_{\varphi(t)}^V P(t), K_{\varphi(t)}^V P(t) \rangle$ est constante.

Remark 3.10. Une conséquence particulièrement intéressante de la formule (3.2) est que le support de $P(t)$ ne dépend pas de t : le flot géodésique hamiltonien conserve le support de la variable duale.

De même, si on a $P_0 \in \Gamma^{s'}(T^*M)$, avec $s' \in \{-s, \dots, s\}$, alors $P(t) \in \Gamma^{s'}(T^*M)$ pour tout t . De même, si P_0 est une 1-forme à coefficients dans les mesures de Radon, $P(t)$ en est également une pour tout t . Autrement dit, le flot géodésique hamiltonien conserve la régularité de la variable duale.

3.5.4 Forme réduite des équations hamiltoniennes sur un groupe de Lie

Les résultats décrits dans cette section sont prouvés dans [MR99]. Soit G un groupe de Lie de dimension finie, d'élément neutre e , et d'algèbre de Lie des champs de vecteurs invariants à droite $(\mathfrak{g}, [\cdot, \cdot])$. On définit l'application moment $\mu : T^*G \mapsto \mathfrak{g}^*$ par $\mu(g, p) = dR_g^* p$, où R_g est la multiplication à droite par g . Considérons un hamiltonien lisse sur $\mathfrak{h} : T^*G \rightarrow \mathbb{R}$ invariant à droite, c'est-à-dire tel que

$$\mathfrak{h}(g, p) = \mathfrak{h}(e, \mu(g, p)) =: \mathfrak{h}_e(\mu(g, p)).$$

La fonction $\mathfrak{h}_e : \mathfrak{g}^* \rightarrow \mathbb{R}$ est la restriction de \mathfrak{h} à $\mathfrak{g}^* \simeq T_e^*G$. Soit une courbe intégrale $t \mapsto (q(t), p(t))$ de $\nabla^\omega \mathfrak{h}$, et notons $\mu(t) = \mu(q(t), p(t))$. Alors $\mu(t)$ satisfait

$$\dot{\mu}(t) = -ad_{dh_e(\mu(t))}^* \mu(t). \quad (3.3)$$

Ici, pour v dans \mathfrak{g} , $ad_v : \mathfrak{g} \rightarrow \mathfrak{g}$ est l'application $ad_v(w) = [v, w]$, et

$$dh_e(\nu) \in \mathfrak{g}^{**} = \mathfrak{g}, \quad \nu \in \mathfrak{g}^*.$$

L'équation différentielle 3.3 est appelée équation hamiltonienne réduite, ou *équation des moments*.

Réciproquement, soit $g_0 \in G$ un point fixé, et $t \mapsto \mu(t)$ une courbe absolument continue sur \mathfrak{g}^* satisfaisant (3.3). Définissons alors une courbe $t \mapsto (g(t), p(t))$ sur T^*G de classe H_{loc}^1 par

$$g(0) = g_0, \quad \dot{g}(t) = dR_{g(t)} dh_e(\mu(t)), \quad p(t) = dR_{g(t)}^{-1} \mu(t), \quad \text{p.p. } t.$$

Dans ce cas, cette courbe est une courbe intégrale de $\nabla^\omega \mathfrak{h}$, la seule telle que $g(0) = g_0$ et $\mu(t) = \mu(g(t), p(t))$ pour tout t .

3.5.5 Équation des moments sur $\mathcal{D}^s(M)$

Les mêmes résultats s'appliquent dans le cas du hamiltonien normal d'une structure sous-rieman-nienne invariante à droite sur $\mathcal{D}^s(M)$. Toutefois, comme $\mathcal{D}^s(M)$ n'est pas un groupe de Lie, l'équation des moment perd de sa régularité. On suppose que la structure est induite par un espace de Hilbert de champs de vecteurs sur M avec inclusion continue dans $\Gamma^{s+1}(TM)$. On définit sur $T^*\mathcal{D}^s(M)$ l'application moment $\mu(\varphi, P) = dR_\varphi^*P$. Cette application est continue, mais pas forcément différentiable.

Soit alors $t \mapsto (\varphi(t), P(t))$ satisfaisant l'équation des géodésiques, et définissons $\mu(t) = \mu(\varphi(t), P(t))$. On note également $X(t) = K^V dR_{\varphi(t)}^*P(t) = \dot{\varphi}(t) \circ \varphi(t)^{-1}$. Soit enfin Y dans $\Gamma^{s+1}(TM)$. Alors on a

$$\begin{aligned} \frac{d}{dt} (\mu(t)Y) &= \dot{P}(t)(Y \circ \varphi(t)) + P(t)(dY \circ \varphi(t)(\dot{\varphi}(t))) \\ &= -P(t)(dX(t) \circ \varphi(t)(Y(t) \circ \varphi(t)) + P(t)(dY \circ \varphi(t)(X(t) \circ \varphi(t))) \\ &= P(t)([X(t), Y(t)] \circ \varphi(t)) \\ &= \mu(t)([X(t), Y(t)]). \end{aligned}$$

Or, le crochet de Lie des champs de vecteurs s'identifie à l'opposé des crochets de Lie induits sur $\Gamma^{s+1}(TM)$ par la structure de groupe $\mathcal{D}^s(M)$ (voir par exemple [Omo74] ou [KM97]). On obtient donc

$$\dot{\mu}(t) = -ad_{X(t)}^* \mu(t).$$

Enfin, $h(e, \nu) = \frac{1}{2}\nu K^V \nu$ pour tout $\nu \in \Gamma^{-s}(T^*M)$, donc

$$\partial_p h(e, \mu(t)) = K^V \mu(t) = K^V dR_{\varphi(t)}^*P(t) = X(t).$$

Finalement, on a bien l'équation des moments

$$\dot{\mu}(t) = -ad_{K^V \mu(t)}^* \mu(t).$$

Remarquons que $\dot{\mu}(t)$ n'est définie que sur $\Gamma^{s+1}(TM)$, et n'appartient donc pas forcément à $\Gamma^{-s}(T^*M) = T_e^*\mathcal{D}^s(M)$: il y a une perte de régularité. En conséquence, pour s'assurer que l'équation des moments a une solution, il faut passer par la transformation inverse qui envoie $t \mapsto \mu(t)$ sur une courbe $t \mapsto (\varphi(t), P(t))$ qui satisfait alors l'équation hamiltonienne des géodésiques décrites dans le théorème 3.2.

3.6 Quelques exemples

Dans cette section, on donne quelques exemples d'espaces de champs de vecteurs à noyaux reproduisants sur $M = \mathbb{R}^d$, ainsi que les équations géodésiques correspondantes sur $\mathcal{D}^s(\mathbb{R}^d)$, $s > d/2 + 1$. Un élément de $(\mathbb{R}^d)^*$ est identifié à une matrice ligne. Tout $x \in \mathbb{R}^d$ est identifié à une matrice colonne de transposée x^T . On note $|x| = \sqrt{x^T x}$ la norme euclidienne canonique de x .

3.6.1 Noyau diagonal

On considère une fonction γ telle que $\gamma \in \mathcal{C}^\infty(\mathbb{R}) \cap H^r(\mathbb{R})$ pour tout entier r . On définit le noyau $K : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathcal{L}((\mathbb{R}^d)^*, \mathbb{R}^d)$ par

$$K(x, y)p = \gamma(|x - y|^2)p^T, \quad x, y \in \mathbb{R}^d, p \in (\mathbb{R}^d)^*.$$

On suppose que K est un noyau symétrique défini positif, c'est-à-dire que pour tout $P \in \mathcal{C}^1(\mathbb{R}^d, (\mathbb{R}^d)^*)$ à support compact,

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} \int_{\mathbb{R}^d \times \mathbb{R}^d} P(x)K(x, y)P(y)dydx = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \gamma(|x - y|^2)P(x)P(y)^T dydx \geq 0,$$

avec égalité si et seulement si

$$\forall x \in \mathbb{R}^d, \quad \int_{\mathbb{R}^d} K(x, y)P(y)dy = \int_{\mathbb{R}^d} \gamma(|x - y|^2)P(y)^T dy = 0.$$

Dans ce cas, K est le noyau reproduisant d'un unique espace de Hilbert de champs de vecteurs $(V, \langle \cdot, \cdot \rangle)$ sur \mathbb{R}^d , avec inclusion continue dans tous les $H^r(\mathbb{R}^d, \mathbb{R}^d)$, $r \geq 0$. Par exemple, $\gamma(t) = \exp(-t)$ convient.

Soit $s > d/2 + 1$ un entier. On équipe $\mathcal{D}^s(\mathbb{R}^d)$ de la structure sous-riemannienne invariante à droite associée. Cette structure est forte, et de classe \mathcal{C}^∞ .

L'équation des géodésiques est alors donnée par

$$\begin{aligned} \partial_t \varphi(t, x) &= \int_{\mathbb{R}^d} \gamma(|\varphi(t, x) - \varphi(t, y)|^2)P(t, y)^T dy, \\ \partial_t P(t, x) &= -2P(t, x) \underbrace{\int_{\mathbb{R}^d} \gamma'(|\varphi(t, x) - \varphi(t, y)|^2)P(t, y)^T}_{\in \mathbb{R}} \underbrace{(\varphi(t, x) - \varphi(t, y))}_{\in M_{d,d}(\mathbb{R})} dy, \end{aligned}$$

pour tout x et t .

Par exemple, prenons $P(0) = p(0) \otimes \delta_{x_0}$, $p(0) \in (\mathbb{R}^d)^*$ et $x_0 \in \mathbb{R}^d$ fixés. On sait que le flot géodésique préserve le support et l'ordre de régularité, donc $P(t)$ est forcément de la forme $p(t) \otimes \delta_{x_0}$ pour tout t . On obtient $\dot{p}(0) = 0$ et donc, pour tout x et tout t ,

$$\partial_t \varphi(t, x) = \gamma(|\varphi(t, x) - \varphi(t, x_0)|^2)p(0).$$

3.6.2 Noyau adapté à une structure sous-riemannienne sur \mathbb{R}^d

On considère $\gamma : \mathbb{R}_+ \rightarrow \mathbb{R}$ comme précédemment. Soient X_1, \dots, X_k des champs de vecteurs lisses et bornés à tout ordre. On définit le noyau $K : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathcal{L}((\mathbb{R}^d)^*, \mathbb{R}^d)$ par

$$K(x, y)p = \sum_{i=1}^k \gamma(|x - y|^2)pX_i(y)X_i(x).$$

Supposons qu'il s'agisse d'un noyau symétrique défini positif. C'est le cas si $\gamma(t) = \exp(-t)$ (voir section 9.1.5 pour une preuve). L'équation hamiltonienne des géodésiques est alors donnée pour tous x et t par

$$\begin{aligned} \partial_t \varphi(t, x) &= \sum_{i=1}^k \left(\int_{\mathbb{R}^d} \gamma(|\varphi(t, x) - \varphi(t, y)|^2)P(t, y)X_i(\varphi(t, y))dy \right) X_i(\varphi(t, x)), \\ \partial_t P(t, x) &= -2P(t, x) \sum_{i=1}^k \int_{\mathbb{R}^d} \gamma'(|\varphi(t, x) - \varphi(t, y)|^2)P(t, y)X_i(\varphi(t, y))(\varphi(t, x) - \varphi(t, y))^T X_i(\varphi(t, x))dy \\ &\quad - P(t, x) \sum_{i=1}^k \left(\int_{\mathbb{R}^d} \gamma(|\varphi(t, x) - \varphi(t, y)|^2)P(t, y)X_i(\varphi(t, y))dy \right) dX_i(\varphi(t, x)). \end{aligned}$$

Dans le cas où $P(0) = p(0) \otimes \delta_{x_0}$, on obtient pour tous x et t

$$\begin{aligned} \partial_t \varphi(t, x) &= \sum_{i=1}^k \gamma(|\varphi(t, x) - \varphi(t, x_0)|^2)p(t)X_i(\varphi(t, x_0))X_i(\varphi(t, x)), \\ \dot{p}(t) &= -p(t) \sum_{i=1}^k \gamma(0)p(t)X_i(\varphi(t, x_0))dX_i(t, x_0). \end{aligned}$$

Remark 3.11. On voit que si (φ, P) satisfait une telle équation des géodésiques, alors chaque courbe $t \mapsto \varphi(t, x)$ est horizontale sur \mathbb{R}^d pour la structure sous-riemannienne induite par les champs de vecteurs X_i . Si de plus $P(0) = p_0 \otimes \delta_{x_0}$, en supposant que $\gamma(0) = 1$, la courbe $t \mapsto (\varphi(t, x_0), p_0)$ vérifie l'équation hamiltonienne des géodésiques pour cette structure.

Exemple 3.2. Plaçons-nous dans le cas du plan de Grushin. On a $d = 2$, $X_1 = \frac{\partial}{\partial x}$ et $X_2 = \frac{\partial}{\partial y}$. On prend γ tel que $\gamma(0) = 1$. Pour

$$P(0) = p(0) \otimes \delta_{(x_0, y_0)} = (p_1(0), p_2(0)) \otimes \delta_{(x_0, y_0)},$$

on notera

$$P(t) = p(t) \otimes \delta_{(x_0, y_0)} = (p_1(t), p_2(t)) \otimes \delta_{(x_0, y_0)}.$$

Posons alors $(x_0(t), y_0(t)) = \varphi(t, (x_0, y_0))$. On obtient

$$\begin{cases} \dot{x}_0(t) = p_1(t), \\ \dot{y}_0(t) = p_2(t)x_0(t), \\ \dot{p}_1(t) = -x_0(t)p_2^2(t), \\ \dot{p}_2(t) = 0, \end{cases}$$

d'où $p_2(t) = p_2(0) = p_2$, et

$$\begin{cases} x_0(t) = x_0 + \frac{p_1}{p_2} \sin(p_2 t), \\ y_0(t) = y_0 - \frac{p_1}{p_2} (\cos(p_2 t) - 1), \\ p_1(t) = p_1 \cos(p_2 t). \end{cases}$$

Finalement, pour toute courbe $(x(t), y(t)) = \varphi(t, (x, y))$, on a

$$\begin{cases} \dot{x}(t) = \gamma \left(\left(x(t) - x_0 - \frac{p_1}{p_2} \sin(p_2 t) \right)^2 + \left(y(t) - y_0 + \frac{p_1}{p_2} (\cos(p_2 t) - 1) \right)^2 \right) p_1 \cos(p_2 t), \\ \dot{y}(t) = \gamma \left(\left(x(t) - x_0 + \frac{p_1}{p_2} \sin(p_2 t) \right)^2 + \left(y(t) - y_0 - \frac{p_1}{p_2} (\cos(p_2 t) - 1) \right)^2 \right) \left(x_0 + \frac{p_1}{p_2} \sin(p_2 t) \right) p_2. \end{cases}$$

Chapitre 4

Espaces de formes et analyse des déformations

Ce chapitre est un résumé de la deuxième partie de l'article [AT], reproduit en section 9.2.

La littérature ne donne pas de sens précis au terme d'espace de formes plongées dans une variété M . De nombreux exemples d'espaces de formes ont été donnés [You10, TY05], tel que l'espace des courbes fermées continues sur M [MM07], ou encore l'ensemble $\text{Emb}^1(S, M)$ des plongements d'une variété compacte S dans M [BHM11, BHM12], sur lesquels les difféomorphismes agissent par composition à gauche. Toutefois, ces définitions nous privent de possibilités plus générales, telle que des variétés munies de fibres pour modéliser du tissu musculaire.

Il ressort cependant des exemples qu'un espace de formes dans une variété lisse à géométrie bornée M est, moralement, une variété de Banach \mathcal{S} sur lequel le groupe difféomorphismes de M a une action compatible avec sa structure de groupe.

Une notion générale et abstraite d'espace de formes n'a jamais été définie. Un des buts de cette thèse est de donner un cadre rigoureux et général à cette notion. Dans cette section, on commencera en donnant la définition d'un espace de formes \mathcal{S} dans M englobant tous les exemples connus, comme un espace muni sur lequel $\mathcal{D}^s(M)$ agit d'une manière particulière, compatible avec sa structure de groupe de Lie LIH [Omo74].

On continuera sur l'étude de la structure sous-riemannienne sur \mathcal{S} induite par une structure sous-riemannienne forte invariante à droite sur $\mathcal{D}^s(M)$ et par son action sur \mathcal{S} . On montrera ensuite des résultats de complétude pour la distance, et que deux points à distance finie peuvent être reliés par une géodésique minimisante. On donnera alors, comme dans les cas précédents, des conditions pour qu'une courbe horizontale soit une géodésique normale ou une courbe singulière, puis on donnera l'équation hamiltonienne des géodésiques.

Dans toute la suite, on fixe M une variété lisse de dimension d à géométrie bornée. On notera également s_0 le plus petit entier strictement plus grand que $d/2$.

4.1 Définition d'un espace de formes

Un espace de formes \mathcal{S} dans M est une variété lisse modelée sur un espace de Banach sur lequel un groupe de difféomorphismes $\mathcal{D}^s(M)$ a une action continue et compatible avec la structure de groupe décrite dans la proposition 3.1.

Définition 4.1. *Soit \mathcal{S} une variété de Banach métrisable modelée sur un espace de Banach B , muni d'une distance compatible D . On suppose que $\mathcal{D}^s(M)$, où $s = s_0 + \ell$ et $\ell \in \mathbb{N}^*$, agit*

continûment sur \mathcal{S} via l'action

$$(\varphi, q) \mapsto \varphi \cdot q, \quad (\varphi, q) \in \mathcal{D}^s(M) \times \mathcal{S}.$$

Alors \mathcal{S} est un **espace de formes d'ordre ℓ** si les trois conditions suivantes sont satisfaites :

1. L'action est semi-lipschitzienne : pour tout $q \in \mathcal{S}$ fixé, il existe $\gamma > 0$ tel que

$$\forall \varphi_1, \varphi_2 \in \mathcal{D}^s(M), \quad D(\varphi_1 \cdot q, \varphi_2 \cdot q) \leq \gamma d_{H_e^s(M, M)}(\varphi_1, \varphi_2). \quad (4.1)$$

2. Pour tout q dans \mathcal{S} fixé, l'application $R_q : \varphi \mapsto \varphi \cdot q$ est de classe \mathcal{C}^∞ . Sa différentielle en e est notée $\xi_q : \Gamma^s(TM) \rightarrow T_q\mathcal{S}$, et est appelée **l'action infinitésimale** de $\Gamma^s(TM)$ en q .

3. Pour tout k dans \mathbb{N} , les applications

$$\begin{aligned} R : \mathcal{S} \times \mathcal{D}^{s+k}(M) &\longrightarrow \mathcal{S}, & \xi : \mathcal{S} \times \Gamma^{s+k}(TM) &\longrightarrow T\mathcal{S}, \\ (q, \varphi) &\longmapsto \varphi \cdot q & (q, X) &\longmapsto \xi_q(X) \end{aligned} \quad (4.2)$$

sont de classe \mathcal{C}^k .

En particulier, l'action de $\mathcal{D}^s(M)$ sur lui-même par composition en fait un espace de formes d'ordre ℓ .

Les éléments de \mathcal{S} sont généralement appelés des *états*.

Définition 4.2. Un état q d'un espace de formes d'ordre ℓ est dit à **support compact** s'il existe un sous-ensemble compact U de M tel que (4.1) dans la condition 1 de la définition 4.1 puisse être remplacée par

$$\forall \varphi_1, \varphi_2 \in \mathcal{D}^s(M), \quad D(\varphi_1 \cdot q, \varphi_2 \cdot q) \leq \gamma d_{H_e^s(U, M)}(\varphi_1, \varphi_2).$$

Exemple 4.1. Soit S est une variété compacte. Alors l'ensemble $\mathcal{S} = \text{Emb}^1(S, M)$ des plongements \mathcal{C}^1 de S dans M est un espace de formes d'ordre 1 dont tous les états sont à support compact pour la composition à gauche par des éléments de $\mathcal{D}^s(M)$.

Un cas particulier, le plus étudié dans les applications [You10] puisqu'étant de dimension finie, est celui où S est de dimension 0. Alors $S = \{s_1, \dots, s_n\}$ est un ensemble fini, et $\mathcal{S} = \mathcal{C}^1(S, M)$ est diffeomorphe à l'espace de formes $Lmk_n(M)$ des n landmarks sur M

$$Lmk_n(M) = \{(x_1, \dots, x_n) \in M^n \mid i \neq j \Rightarrow x_i \neq x_j\},$$

sur lequel un diffeomorphisme φ agit diagonalement :

$$\varphi \cdot (x_1, \dots, x_n) = (\varphi(x_1), \dots, \varphi(x_n)) \in Lmk_n(M).$$

Dans ces exemples, l'action infinitésimale $\xi_q X$ de $X \in \Gamma^s(TM)$ est le champ de vecteurs le long de q donné par la composition à gauche $X \circ q : S \rightarrow TM$.

Remark 4.1. Soit $t \mapsto X(t) \in \Gamma^s(TM)$ un champ de vecteurs dépendant du temps, et $q_0 \in \mathcal{S}$. Pour $t \mapsto \varphi^X(t)$ le flot de X , on définit la courbe $q : t \mapsto q(t) = \varphi^X(t) \cdot q_0$. Alors on a

$$\dot{q}(t) = \xi_{q(t)} X(t) \quad \text{p.p. } t.$$

Réciproquement, si $t \mapsto q(t) \in \mathcal{S}$ est une courbe telle qu'il existe $t \mapsto X(t) \in \Gamma^{s+1}(TM)$ avec $\dot{q}(t) = \xi_{q(t)} X(t)$ presque partout, alors $q(t) = \varphi^X(t) \cdot q(0)$.

4.2 Structures sous-riemanniennes induites

Soit \mathcal{S} un espace de formes d'ordre ℓ dans M , $s = s_0 + \ell$ comme défini dans la section 4.1 précédente.

4.2.1 Définition de la structure induite

Soit $(V, \langle \cdot, \cdot \rangle)$ un espace de Hilbert à noyau reproduisant de champs de vecteurs sur M avec inclusion continue dans $\Gamma^{s+1}(TM)$. Alors $(\mathcal{S} \times V, \xi, \langle \cdot, \cdot \rangle)$ est une structure sous-riemannienne forte sur \mathcal{S} : c'est la *structure sous-riemannienne induite* par V sur \mathcal{S} . On en déduit une notion de courbe horizontale q , d'actions $A(q, X)$ et $A^{\mathcal{S}}(q)$, de longueurs $L(q, X)$ et $L^{\mathcal{S}}(q)$, et de distance sous-riemannienne $d^{\mathcal{S}}$. Une courbe $q : I \ni 0 \rightarrow \mathcal{S}$ est horizontale pour cette structure si et seulement si il existe un contrôle $X : I \rightarrow V$ tel que $\dot{q}(t) = \xi_{q(t)}X(t)$ presque partout. En notant φ^X le flot de X , on obtient $q(t) = \varphi^X(t) \cdot q(0)$.

On appellera parfois une telle courbe horizontale une *déformation* de l'état initial $q(0)$. Comme φ^X est horizontale sur $\mathcal{D}^s(M)$ pour la structure invariante à droite engendrée par V , on en déduit qu'une courbe $t \ni I \mapsto q(t)$ est horizontale si et seulement si il existe $q_0 \in \mathcal{S}$ fixé et une courbe de difféomorphismes horizontaux telle que

$$q(t) = \varphi(t) \cdot q_0, \quad \text{p.p. } t.$$

De plus,

$$A(q, X) = A(\varphi, X) = \int_I \langle X(t), X(t) \rangle dt.$$

Une telle courbe φ est appelée un *relevé horizontal* à $\mathcal{D}^s(M)$ de la trajectoire q .

4.2.2 Action d'une trajectoire et relèvement minimal

Une même trajectoire q peut correspondre à de nombreux systèmes (q, X) , et donc admettre plusieurs relèvements horizontaux (φ, X) . Rappelons que l'action d'une courbe horizontale q (indépendamment de tout contrôle associé), est donnée par

$$A^{\mathcal{S}}(q) = \inf_{X \in L^2(0,1;V), \dot{q}=\xi_q X} A(q, X).$$

On en déduit que pour tout relevé horizontal φ de q associé au contrôle X ,

$$A^{\mathcal{S}}(q) \leq A(\varphi, X) = A^{\mathcal{D}^s(M)}(\varphi).$$

Il convient alors de se demander s'il existe, pour une trajectoire q fixée, un *relèvement minimal*, c'est-à-dire un relèvement d'action minimale. Pour cela, notons $\Delta = \xi(\mathcal{S} \times V)$. Alors ξ_q munit son image $\Delta_q = \xi_q(V)$ d'un produit de Hilbert $\langle \cdot, \cdot \rangle_q$ défini par la formule

$$\langle w, w \rangle_q = \inf_{X \in V, w=\xi_q X} \langle X, X \rangle.$$

De plus, si $\xi_q X = w$, on a une décomposition unique $X = X_0 + X_1$, avec $\xi_q X_0 = 0$ et $X_1 \in (\ker \xi_q)^\perp$. On a alors

$$\langle w, w \rangle_q = \langle X_1, X_1 \rangle,$$

et on en déduit que la restriction de ξ_q à $(\ker \xi_q)^\perp$ est une isométrie sur son image Δ_q (voir sections 9.2.2 et 10.1.2 pour plus de détails).

On notera $\xi_q^{-1} : \Delta_q \rightarrow (\ker \xi_q)^\perp$ l'inverse de cette image. Il s'agit en fait du *pseudoinverse* de ξ_q , non de son inverse, mais on s'autorisera cet abus de notation. Le produit Hilbertien sur Δ_q permet de définir calculer l'action $A^{\mathcal{S}}(q)$ d'une courbe horizontale $q : I \rightarrow \mathcal{S}$. On a en effet

$$A^{\mathcal{S}}(q) = \frac{1}{2} \int_I \langle \dot{q}(t), \dot{q}(t) \rangle_{q(t)} dt.$$

Proposition 4.1. *Soit $q : [0, 1] \rightarrow \mathcal{S}$ une courbe horizontale sur \mathcal{S} . Alors q a un unique relèvement minimal $t \mapsto \bar{\varphi}(t)$ tel que $\bar{\varphi}(0) = e$, qui vérifie*

$$A^{\mathcal{D}^s(M)}(\bar{\varphi}) = A^{\mathcal{S}}(q).$$

De plus, le contrôle associé $\bar{X}(t) = \dot{\bar{\varphi}} \cdot \varphi^{-1}$ appartient à $(\ker \xi_{q(t)})^\perp$ pour presque tout t , ce qui équivaut à dire que

$$\bar{X}(t) = \xi_{q(t)}^{-1} \dot{q}(t) \quad p.p. \ t \in [0, 1].$$

Voir section 9.2.2 pour plus de détails.

En d'autres termes, la déformation d'un état q se relève à une unique déformation φ minimale de l'espace M tout entier. En utilisant cette propriété, on peut montrer la complétude de la métrique sous-riemannienne $d^{\mathcal{S}}$ dans le cas d'états à supports compacts.

Théorème 4.1. *Soient q_0 dans \mathcal{S} à support compact, et \mathcal{O}_{q_0} l'orbite de q_0 , c'est-à-dire l'ensemble des états à distance finie de q_0 . Alors, pour tous $q, q' \in \mathcal{S}$, il existe une géodésique minimisante reliant q à q' . De plus, $(\mathcal{O}_{q_0}, d^{\mathcal{S}})$ est un ensemble métrique complet.*

Voir section 9.2.2 pour la preuve.

4.2.3 Point de vue dual

Trouver le relèvement minimal d'une courbe horizontale donnée est difficile, car l'opérateur inverse ξ_q^{-1} est très difficile à calculer en pratique. Il est plus simple de construire directement un champ de vecteurs dans $(\ker \xi_q)^\perp$, qui est automatiquement le relèvement minimal de la courbe horizontale correspondante par la proposition 4.1. On travaille pour cela sur l'espace cotangent $T^*\mathcal{S}$.

Lemme 4.1. *Soit $(q, p) \in T^*\mathcal{S}$. Alors $K^V \xi_q^* p$ appartient à $(\ker \xi_q)^\perp$, et*

$$\langle K^V \xi_q p, K^V \xi_q p \rangle = p \xi_q K^V \xi_q^* p.$$

On notera $K_q = \xi_q K^V \xi_q^ : T_q^*\mathcal{S} \rightarrow T_q\mathcal{S}$.*

C'est un opérateur symétrique (i.e., $p_1 K_q p_2 = p_2 K_q p_1$), semi-défini positif (i.e., $p K_q p \geq 0$). L'opérateur K_q définit un morphisme de fibrés vectoriel $T^\mathcal{S} \rightarrow T\mathcal{S}$ de classe \mathcal{C}^k , où k est le plus grand entier tel que V ait une inclusion continue dans $\Gamma^{s+k}(TM)$. Voir les sections 9.2.2 et 10.1.2 pour plus de détails.*

Remark 4.2. Il peut donc être utile, lors de la recherche de géodésiques, de considérer simplement des contrôles de la forme $X(t) = K^V \xi_{q(t)}^* p(t)$. On n'obtient pas toutes les courbes horizontales de cette façon (sauf si Δ_q est fermé dans $T_q M$), mais les courbes q obtenues satisfont alors

$$A^{\mathcal{S}}(q) = \frac{1}{2} \int p(t) K_{q(t)} p(t) dt.$$

De plus, le relèvement minimal de q est le flot de X . On verra en particulier dans la section suivante que toutes les géodésiques normales sont de cette forme.

4.3 Équation hamiltonienne des géodésiques sur un espace de formes

On se place dans le cadre suivant : on prend \mathcal{S} un espace de formes d'ordre ℓ dans M , on pose $s = s_0 + \ell$, et on considère un espace $(V, \langle \cdot, \cdot \rangle)$ de Hilbert à noyau reproduisant de champs de vecteurs sur M avec injection continue dans $\Gamma^{s+k}(TM)$, $k \in \mathbb{N}^*$. On considère la structure sous-riemannienne induite par V sur \mathcal{S} .

4.3.1 Hamiltoniens du système

Les hamiltoniens $H_{\mathcal{S}}^{\lambda} : T^*\mathcal{S} \times V \rightarrow \mathbb{R}$, $\lambda \in \{0, 1\}$, du système sont donnés par

$$H_{\mathcal{S}}^{\lambda}(q, p, X) = p\xi_q X - \frac{\lambda}{2} \langle X, X \rangle.$$

les $H_{\mathcal{S}}^{\lambda}$ sont de classe \mathcal{C}^k . On peut donc appliquer la proposition 2.4.

Proposition 4.2. *Soit $q : [0, 1] \rightarrow \mathcal{D}^s(M)$ une courbe horizontale, de contrôle associé $X \in L^2(0, 1; V)$. Supposons qu'il existe $(p, \lambda) \neq (0, 0)$, où $p : [0, 1] \rightarrow T^*\mathcal{S}$ est de classe H^1 et $\lambda = 0$ ou 1 tels que*

$$\begin{cases} 0 & = & \partial_X H^{\lambda}(q(t), p(t), X(t)) & = & \xi_{q(t)}^* p(t) - \lambda \langle X(t), \cdot \rangle, \\ (\dot{q}(t), \dot{p}(t)) & = & \nabla^{\omega} H^{\lambda}(q(t), p(t), X(t)) & = & (\xi_{q(t)} X(t), -\partial_q(\xi_{q(t)} X(t))^* p(t). \end{cases} \quad (4.3)$$

- Si $\lambda = 0$, alors, pour tout $t \in [0, 1]$, $p(t) \neq 0$, et $\Delta_{q(t)} \subset \ker p(t)$: q est une courbe singulière.
- Si $\lambda = 1$, (q, X) est un point critique de A à extrémités fixées : q est une géodésique normale.

On a donc $\partial_X H^{\lambda}(q, p, X) = 0$ si et seulement si $\xi_q^* p = \langle X, \cdot \rangle$, autrement dit si $X = K^V \xi_q^* p$.

Remark 4.3. En particulier, les géodésiques normales sont bien de la forme décrite en remarque 4.2.

Le hamiltonien normal $h^{\mathcal{S}} : T^*M \rightarrow \mathbb{R}$ vaut alors

$$h(q, p) = \frac{1}{2} p \xi_q K^V \xi_q^* p = \frac{1}{2} p K_q p.$$

On en déduit facilement son gradient symplectique, donné dans un système de coordonnées canoniques par

$$\nabla^{\omega} h(q, p) = (K_q p, -\frac{1}{2} \partial_q(p K_q p)).$$

4.3.2 Équations géodésiques

En appliquant alors le théorème 2.1, et on obtient directement le

Théorème 4.2. Géodésiques normales sur les espaces de formes.

Soit $(V, \langle \cdot, \cdot \rangle)$ un espace de Hilbert à noyau reproduisant K de champs de vecteurs sur une variété M lisse à géométrie bornée. On suppose que V a une inclusion continue dans $\Gamma^{s+k}(TM)$, avec k un entier supérieur ou égal à 2 et $s = s_0 + \ell$, $\ell \in \mathbb{N}^$. Soit \mathcal{S} un espace de formes dans M d'ordre ℓ , et $(\mathcal{S} \times V, \xi, \langle \cdot, \cdot \rangle)$ la structure sous-riemannienne engendrée par V sur \mathcal{S} .*

Soit $(q_0, p_0) \in T^\mathcal{S}$, tel que $\Delta_{q_0} = \xi_{q_0}(V)$ ne soit pas inclus dans $\ker p_0$. Il existe une unique courbe maximale $(q, p) : I \ni 0 \rightarrow T^*\mathcal{S}$ de classe H_{loc}^1 , avec I un intervalle, telle que $(q(0), p(0)) = (q_0, p_0)$ et pour presque tout t dans I ,*

$$(\dot{q}(t), \dot{p}(t)) = \nabla^{\omega} h(q(t), p(t)).$$

Dans un système de coordonnées canoniques, cette équation s'écrit

$$\begin{cases} \dot{q}(t) = K_{q(t)} p(t) \\ \dot{p}(t) = -\frac{1}{2} \partial_q(p(t) K_{q(t)} p(t)). \end{cases} \quad (4.4)$$

Alors q est la trajectoire du système $(q, X) = (q, K^V \xi_q^* p)$, et ce système est un point critique de l'action à extrémités fixées. Le relèvement minimal de q est le flot de $t \mapsto K^V \xi_{q(t)}^* p(t)$. De plus, si k est au moins égal à 3, q est une géodésique pour cette structure, c'est-à-dire qu'elle minimise la longueur entre $q(t_0)$ et $q(t_1)$ pour t_0 et t_1 assez proches l'un de l'autre.

Enfin,

$$t \mapsto \left\langle K^V \xi_{q(t)}^* p(t), K^V \xi_{q(t)}^* p(t)(t) \right\rangle = p(t) K_{q(t)} p(t)$$

est constante, égale à $2h(q_0, p_0)$.

Grâce à l'équation des moments sur $\mathcal{D}^s(M)$, on obtient aussi le résultat suivant.

Proposition 4.3. *Si $t \mapsto (q(t), p(t))$ satisfait l'équations géodésique sur T^*S , alors*

$$t \mapsto (\varphi^X(t), \xi_{q(t)}^* p(t)) \in T^* \mathcal{D}^s(M),$$

où $X(t) = K^V \xi_{q(t)}^* p(t)$, satisfait l'équations géodésique sur $T^*D^s(M)$, où φ est le relèvement minimal de q .

Autrement dit, le relèvement minimal d'une géodésique normale est une géodésique normale.

4.4 Cas des variétés immergées dans \mathbb{R}^d

On prend $M = \mathbb{R}^d$. On peut alors identifier le noyau K d'un espace de Hilbert à noyau reproduisant de champs de vecteurs sur \mathbb{R}^d à une application $K : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathcal{L}((\mathbb{R}^d)^*, \mathbb{R}^d)$. On considère S une variété lisse compacte et $\mathcal{S} = \mathcal{C}^0(S, \mathbb{R}^d)$. \mathcal{S} est un espace de formes d'ordre 1 pour l'action donnée par la composition à gauche. On a $T^* \mathcal{S} = \mathcal{S} \times \mathcal{C}^0(S, \mathbb{R}^d)^*$, et

$$\mathcal{C}^0(S, \mathbb{R}^d)^* = \mathcal{M}(S) \otimes (\mathbb{R}^d)^*,$$

où $\mathcal{M}(S)$ désigne l'ensemble des mesures de Radon sur S .

En utilisant les propriétés de K , on voit que pour $q : S \rightarrow \mathbb{R}^d$ continue, $x \in \mathbb{R}^d$, et $p \in \mathcal{M}(S) \otimes (\mathbb{R}^d)^*$,

$$K^V \xi_q^* p(x) = \int_S K(x, q(s)) dp(s),$$

d'où

$$K_q p(s) = \int_S K(q(s), q(s')) dp(s')$$

et

$$p K_q^{\mathcal{S}} p = \int \int_{S \times S} dp(s) K(q(s), q(s')) dp(s')$$

En utilisant le fait que $K(x, y)^T = K(y, x)$ pour tous $x, y \in \mathbb{R}^d$, on obtient la forme suivante pour l'équation normale des géodésiques :

$$\begin{cases} \partial_t q(t, s) = \int_S K(q(t, s), q(t, s')) dp(t, s') \\ \partial_t p(t, s) = -p(t, s) \int_S \partial_1 K(q(t, s), q(t, s')) dp(t, s'). \end{cases} \quad (4.5)$$

Là encore, le support et la régularité de p sont conservés. Dans le cas d'un espace de landmarks $\mathcal{S} \simeq Lmk_n(\mathbb{R}^d)$, avec

$$Lmk_n(\mathbb{R}^d) = \{(x_1, \dots, x_n) \in (\mathbb{R}^d)^n \mid i \neq j \Rightarrow x_i \neq x_j\},$$

on a $q(t) = (x_1(t), \dots, x_n(t))$, $p(t) = (p_1(t), \dots, p_n(t)) \in T_{q(t)}^* \mathcal{S} = ((\mathbb{R}^d)^*)^n$, et on obtient

$$\begin{cases} \dot{x}_i(t) = \sum_{j=1}^n K(x_i(t), x_j(t)) p_j(t), & i = 1, \dots, n, \\ \dot{p}_i(t) = -p_i(t) \sum_{j=1}^n \partial_1 K(x_i(t), x_j(t)) p_j(t), & i = 1, \dots, n. \end{cases} \quad (4.6)$$

Chapitre 5

Structures sous-riemanniennes relevées sur un espace de formes

Dans cette section, on décrit un autre type de structure sous-riemannienne, appelée structure *relevée*. Il s'agit d'une variante de la construction permettant de doter une fibration sur une variété riemannienne d'une structure sous-riemannienne grâce à une connection d'Ehresmann (voir [Mon02], chapitre 11). On peut utiliser de telles structures pour étudier comment la déformation d'un état q_0 appartenant à un certain espace de formes \mathcal{S} peut engendrer une déformation d'un état \hat{q}_0 plus "gros".

Par exemple, soit $\hat{\mathcal{S}} = \mathcal{S} \times \mathcal{S}'$ un produit d'espaces de formes d'ordre ℓ dans une variété à géométrie bornée M . Alors $\hat{\mathcal{S}}$ est un espace de formes d'ordre ℓ pour l'action diagonale

$$\varphi \cdot (q, q') = (\varphi \cdot q, \varphi \cdot q'), \quad (q, q') \in \hat{\mathcal{S}} = \mathcal{S} \times \mathcal{S}'.$$

Soit maintenant un espace de Hilbert de champs de vecteurs à noyau reproduisant V avec inclusion continue dans $\Gamma^{s+1}(TM)$, où $s > \ell + d/2$ avec $d = \dim(M)$. On munit $\hat{\mathcal{S}}$, \mathcal{S} et \mathcal{S}' des structures sous-riemanniennes induites.

Soit $\hat{q}_0 = (q_0, q'_0) \in \hat{\mathcal{S}}$. Considérons une déformation $q : [0, 1] \rightarrow \mathcal{S}$ du premier terme du couple (q_0, q'_0) . Alors q a un unique relèvement minimal $\varphi : [0, 1] \rightarrow \mathcal{D}^s(M)$, de contrôle $X(t) = \xi_{q(t)}^{-1} \dot{q}(t)$. Ce relèvement induit à son tour une déformation $\hat{q} : [0, 1] \rightarrow \hat{\mathcal{S}}$ de l'état total par

$$\hat{q}(t) = \varphi(t) \cdot \hat{q}_0 = (\varphi(t) \cdot q_0, \varphi(t) \cdot q'_0),$$

et on a

$$\dot{\hat{q}}(t) = (\dot{q}(t), \underbrace{\xi_{q(t)}^{-1} \dot{q}(t)}_{\in (\ker \xi_{q(t)})^\perp}).$$

C'est la seule courbe horizontale partant de \hat{q}_0 sur $\hat{\mathcal{S}}$ dont la projection sur \mathcal{S} soit égale à q et dont l'action $A^{\hat{\mathcal{S}}}(\hat{q})$ soit égale à celle de q : c'est le *relèvement minimal* de q partant de \hat{q}_0 à travers la projection canonique $\pi : \mathcal{S} \times \mathcal{S}' \rightarrow \mathcal{S}$. La structure relevée sur $\hat{\mathcal{S}}$ à travers π sera alors une structure pour laquelle les courbes horizontales seront exactement les relèvements minimaux de courbes horizontales sur \mathcal{S} .

On se placera dans le cadre plus général où on a une submersion équivariante $\pi : \hat{\mathcal{S}} \rightarrow \mathcal{S}$ entre deux espaces de formes. Après avoir étudié quelques propriétés d'une telle application, on définira la structure sous-riemannienne relevée induite par π sur $\hat{\mathcal{S}}$, puis on donnera une équation hamiltonienne des géodésiques sous certaines hypothèses.

Toutes les preuves sont détaillées dans la troisième partie de [AT], reproduit en section 9.3 avec plusieurs exemples explicites d'équations géodésiques.

5.1 Applications équivariantes

Soient $\hat{\mathcal{S}}$ et \mathcal{S} deux espaces de formes d'ordre $\ell \in \mathbb{N}^*$ dans M une variété à géométrie bornée, et soit $s > d/2 + \ell$. On suppose qu'on a une submersion lisse $\pi : \hat{\mathcal{S}} \rightarrow \mathcal{S}$ qui est *équivariante* sous l'action de $D^s(M)$:

$$\forall(\varphi, \hat{q}) \in \mathcal{D}^s(M) \times \hat{\mathcal{S}}, \quad \pi(\varphi \cdot \hat{q}) = \varphi \cdot \pi(\hat{q}).$$

Soit $(V, \langle \cdot, \cdot \rangle)$ un espace de Hilbert à noyau reproduisant de champs de vecteurs sur M avec inclusion continue dans $\Gamma^{s+k}(TM)$, $k \in \mathbb{N}^*$. On équipe \mathcal{S} et $\hat{\mathcal{S}}$ des structures sous-riemanniennes induites par V .

Définition 5.1. *Un relèvement horizontal au travers de π d'une courbe horizontale $q : [0, 1] \rightarrow \mathcal{S}$ est une courbe horizontale $\hat{q} : [0, 1] \rightarrow \hat{\mathcal{S}}$ telle que $q = \pi \circ \hat{q}$.*

Si $(q, X) : I \ni 0 \rightarrow \mathcal{S} \times V$ est un système horizontal tel que $q(0) \in \pi(\hat{\mathcal{S}})$, alors pour tout $\hat{q}_0 \in \pi^{-1}(\{q(0)\})$, $\varphi^X \cdot \hat{q}_0$ est un relèvement horizontal de q . Réciproquement, une courbe horizontale est un relèvement de la courbe horizontale $\pi \circ \hat{q} : I \rightarrow \mathcal{S}$, et on a

$$A^{\mathcal{S}}(\pi(\hat{q})) \leq A^{\hat{\mathcal{S}}}(\hat{q}),$$

car $d\pi_{\hat{q}}\xi_{\hat{q}}X = \xi_{\pi(\hat{q})}X$ (voir section 9.3.1).

Lemme 5.1. *Soit $\hat{q} \in \hat{\mathcal{S}}$ et $q = \pi(\hat{q})$. Notons $(\ker d\pi_{\hat{q}} \cap \Delta_{\hat{q}})^\perp$ l'orthogonal de $\ker d\pi_{\hat{q}}$ dans $\Delta_{\hat{q}} = \xi_{\hat{q}}(V)$ pour le produit scalaire $\langle \cdot, \cdot \rangle_{\hat{q}}$ induit par $\xi_{\hat{q}}$.*

Alors la restriction

$$d\pi_q : (\ker d\pi_{\hat{q}} \cap \Delta_{\hat{q}})^\perp \rightarrow \Delta_q = \xi_q(V),$$

est une isométrie, d'inverse

$$d\pi_{\hat{q}}^{-1} = \xi_{\hat{q}}\xi_q^{-1}.$$

En particulier,

$$\Pi_{\hat{q}} = \xi_{\hat{q}}\xi_q^{-1}d\pi_{\hat{q}}$$

est la projection orthogonal de $\Delta_{\hat{q}}$ sur $(\ker d\pi_{\hat{q}} \cap \Delta_{\hat{q}})^\perp$.

Voir section 9.3.1 pour la preuve.

Proposition 5.1. *Soit $\hat{q}_0 \in \hat{\mathcal{S}}$, $q_0 = \pi(\hat{q}_0)$, et $q : I \ni 0 \rightarrow \mathcal{S}$ une courbe horizontale avec $q(0) = q_0$. Cette courbe admet un unique relèvement horizontal $\hat{q} : I \rightarrow \hat{\mathcal{S}}$ tel que $\hat{q}(0) = \hat{q}_0$ et*

$$A^{\hat{\mathcal{S}}}(\hat{q}) = A^{\mathcal{S}}(q).$$

*Ce relèvement, appelé **relèvement minimal** de q partant de \hat{q}_0 , est déterminé par*

$$\dot{\hat{q}}(t) = \xi_{\hat{q}(t)}\xi_{q(t)}^{-1}\dot{q}(t), \quad p.p. t \in I. \quad (5.1)$$

Pour l'existence, on voit que $t \mapsto X(t) = \xi_{q(t)}^{-1}\dot{q}(t)$ appartient bien à $L^2(0, 1; V)$, et l'unique solution de (5.1) est donc donnée par $\hat{q} = \varphi^X \cdot \hat{q}_0$. Le reste du résultat est conséquence immédiate du lemme 5.1.

Exemple 5.1. On prend $\mathcal{S} = \text{Lmk}_n(M)$, $\hat{\mathcal{S}} = TS$, et $\pi : TS \rightarrow \mathcal{S}$ la projection canonique. Si $q = (x_1, \dots, x_n) \in \mathcal{S}$, alors $w \in T_q\mathcal{S}$ s'écrit

$$w = (w_1, \dots, w_n) \in T_{x_1}M \times \dots \times T_{x_n}M.$$

Alors \mathcal{S} et $\hat{\mathcal{S}}$ sont des espaces de formes d'ordre 1 pour les actions respectives

$$\varphi \cdot (x_1, \dots, x_n) = (\varphi(x_1), \dots, \varphi(x_n)),$$

et

$$\varphi \cdot (x_1, \dots, x_n; w_1, \dots, w_n) = (\varphi(x_1), \dots, \varphi(x_n); d\varphi(x_1)w_1, \dots, d\varphi(x_n)w_n),$$

pour $(x_1, \dots, x_n) \in M^n$ et $w_i \in T_{x_i}M$ pour tout $i = 1, \dots, n$. L'application π est équivariante : on a bien

$$\varphi \cdot (x_1, \dots, x_n) = \pi(\varphi \cdot (x_1, \dots, x_n; w_1, \dots, w_n)).$$

5.2 Structure sous-riemannienne relevée

La structure sous-riemannienne relevée induite sur $\hat{\mathcal{S}}$ par $(V, \langle \cdot, \cdot \rangle)$ et π est une structure pour laquelle les courbes horizontales sur $\hat{\mathcal{S}}$ sont exactement les relèvements minimaux de courbes horizontales sur \mathcal{S} . Elle est définie par $(\hat{\mathcal{S}} \times V, \zeta, \langle \cdot, \cdot \rangle)$, avec

$$\zeta_{\hat{q}} = \Pi_{\hat{q}} \xi_{\hat{q}},$$

où

$$\Pi_{\hat{q}} = \xi_{\hat{q}} \xi_{\pi_{\hat{q}}}^{-1} d\pi_{\hat{q}}$$

est la projection orthogonale sur $(\ker d\pi_{\hat{q}} \cap \Delta_{\hat{q}})^\perp$ comme introduit dans le lemme 5.1. L'action et la longueur des courbes coïncident avec celles de la structure sous-riemannienne non relevée.

Remark 5.1. L'application $q \mapsto \zeta_q$ n'a aucune raison d'être régulière. Même en dimension finie, il suffit que le rang de $\xi_{\pi(\hat{q})}$ varie pour que ζ ne soit pas continue en q .

Proposition 5.2. *Le relèvement minimal d'une géodésique minimisante est une géodésique minimisante pour la structure sous-riemannienne relevée (voir section 5).*

Comme Π_q est difficile à calculer directement, on utilise généralement un point de vue dual pour trouver des relèvements minimaux.

Lemme 5.2. *Soient $\hat{q} \in \hat{\mathcal{S}}$, $q = \pi(\hat{q})$, et $u \in T_q^* \mathcal{S}$. Alors*

$$\zeta_{\hat{q}} K^V \xi_q^* u = \xi_{\hat{q}} K^V \xi_q^* u.$$

En particulier, un relèvement minimal à $\hat{\mathcal{S}}$ d'une courbe $t \mapsto q(t)$ sur \mathcal{S} telle que $\dot{q}(t) = K_{q(t)} u(t)$ presque partout satisfait

$$\dot{\hat{q}}(t) = \xi_{\hat{q}(t)} K^V \xi_{q(t)}^* u(t), \quad p.p. t.$$

Cette courbe est donc horizontale pour la structure relevée.

Voir section 9.3.4 pour la preuve.

Exemple 5.2. On suppose que $M = \mathbb{R}^d$. Soient $\mathcal{S} = Lmk_n(\mathbb{R}^d)$, $\hat{\mathcal{S}} = T\mathcal{S}$, et $\pi : T\mathcal{S} \rightarrow \mathcal{S}$. Soit

$$\hat{q}_0 = (q_0, w_0) = (x_{1,0}, \dots, x_{n,0}; w_{1,0}, \dots, w_{n,0}) \in T\mathcal{S}.$$

En utilisant un point de vue dual pour une courbe horizontales $t \mapsto q(t) = (x_1(t), \dots, x_n(t))$ partant de q_0 sur \mathcal{S} , on obtient pour tout temps t des vitesses de la forme

$$\dot{x}_i(t) = \sum_{j=1}^n K(x_i(t), x_j(t)) u_j(t), \quad i \in \{1, \dots, n\}$$

et

$$u(t) = (u_1(t), \dots, u_n(t)) \in T_{q(t)}^* \mathcal{S} = T_{x_1}^* M \times \dots \times T_{x_n}^* M.$$

De plus,

$$K^V \xi_{q(t)}^* u(t)(x) = \sum_{j=1}^n K(x, x_j(t)) u_j(t).$$

Donc, l'unique relèvement minimal $t \mapsto \hat{q}(t) = (q(t), w(t))$ à \hat{S} de la courbe q partant de \hat{q}_0 est donné par la solution de l'équation différentielle

$$\begin{aligned} \dot{x}_i(t) &= \sum_{j=1}^n K(x_i(t), x_j(t)) u_j(t) \\ \dot{w}_i(t) &= \sum_{j=1}^n \partial_1(K(x_i(t), x_j(t)) u_j(t)) w_i. \end{aligned}$$

5.3 Équation hamiltonienne des géodésiques

5.3.1 Hamiltonien et hamiltonien réduit

On va se restreindre à la recherche de géodésiques normales. Considérons le hamiltonien $H_\pi^1 : T^* \hat{S} \times V \rightarrow \mathbb{R}$ défini par

$$H_\pi^1(\hat{q}, \hat{p}, X) = \hat{p} \zeta_{\hat{q}} X - \frac{1}{2} \langle X, X \rangle.$$

Alors

$$0 = \partial_X H_\pi(\hat{q}, \hat{p}, X) = \zeta_{\hat{q}}^* \hat{p} - \langle X, \cdot \rangle$$

équivalent à

$$X = K^V \zeta_{\hat{q}}^* \hat{p} = K^V \xi_{\hat{q}}^* \Pi_{\hat{q}}^* \hat{p}.$$

Le hamiltonien réduit est donné par

$$h_\pi(\hat{q}, \hat{p}) = \frac{1}{2} \hat{p} \zeta_{\hat{q}} K^V \zeta_{\hat{q}}^* \hat{p}.$$

À première vue, ce hamiltonien (et ses dérivées éventuelles) est extrêmement difficile à calculer. Dans [AT], on lui a donné une forme un peu plus simple sous certaines conditions, qui nous a permis de prouver la version suivante de l'équation hamiltonienne des géodésiques. Elle s'appuie sur le lemme 5.2.

Proposition 5.3. *Supposons que ζ soit lisse, et que l'équation linéaire en $u \in T_{\pi(\hat{q})}^* \mathcal{S}$ suivante :*

$$K_{\pi(\hat{q})} u = \xi_{\pi(\hat{q})} K^V \xi_{\hat{q}}^* \hat{p},$$

ait au moins une solution u pour tout $(\hat{q}, \hat{p}) \in T^ \hat{S}$. Alors, pour un tel u ,*

$$h_\pi(\hat{q}, \hat{p}) = \frac{1}{2} u K_{\pi(\hat{q})} u.$$

Soit maintenant

$$(\hat{q}, \hat{p}, u) \in H^1 \times H^1 \times L^2(0, 1; T^* \hat{S} \oplus_{\hat{S}} \pi^* T^* \mathcal{S})$$

tel que, pour presque tout t ,

$$\left\{ \begin{array}{l} K_{\pi(\hat{q}(t))}u(t) = \xi_{\pi(\hat{q}(t))}K^V \xi_{\hat{q}(t)}^* \hat{p}(t) \\ \dot{\hat{q}}(t) = \xi_{\hat{q}(t)}K^V \xi_{\pi(\hat{q}(t))}^* u(t), \\ \dot{\hat{p}}(t) = -(\hat{p}(t) - \frac{1}{2}d\pi_{\hat{q}(t)}^* u(t))(\partial_{\hat{q}}K_{\hat{q}(t)}d\pi_{\hat{q}(t)}^*)u(t) \\ \quad - (\hat{p}(t) - d\pi^* p_{\hat{q}(t)}u(t))K_{\hat{q}(t)}(\partial_{\hat{q}}d\pi_{\hat{q}(t)}^*)^* u(t). \end{array} \right.$$

Alors \hat{q} est une géodésique pour la structure relevée par π sur $\hat{\mathcal{S}}$.

Voir section 9.3.4 pour la preuve. Dans la section 9.3.5, on utilise cette forme de l'équation hamiltonienne pour donner quelques exemples de géodésiques dans des espaces de formes relevés de dimension finie.

Remark 5.2. On peut également modéliser les structures relevées en rajoutant des contraintes de la forme $C_{\hat{q}}X = 0$ plutôt qu'en définissant une nouvelle structure sous-riemannienne. On va étudier ce type de problème dans la section suivante, et on trouvera des équations très similaires à celles de la proposition 5.3.

Chapitre 6

Espaces de formes avec contraintes

Les résultats de ce chapitre sont tirés des deux premières parties de [ATTY13], reproduit dans les sections 10.1 et 10.2.

La structure d'espace de formes peut être utilisée pour faire des études statistiques sur une famille d'objets déformables. Par exemple, en anatomie computationnelle, on les utilise pour comparer les cerveaux de différents patients les uns avec les autres. L'avantage d'utiliser les espaces de formes pour modéliser les déformations d'un objet physique est que les algorithmes de minimisation sont relativement simples à implémenter. Beaucoup plus, par exemple, que des modèles se basant sur une analyse des lois qui régissent la dynamique de l'objet à modéliser. En contrepartie, on perd en qualité de modélisation. Les déformations par difféomorphismes préservent tout de même des aspects géométriques locaux et globaux des objets étudiés (nombre d'intersections, régularité) qui rendent souvent ces pertes acceptables.

L'ajout de contraintes sur les déformations possibles permettrait d'améliorer la modélisation. Par exemple, si on veut étudier plusieurs objets indépendants, plongés dans un milieu continu, de manière simultanée, cela n'a pas de sens de les déformer tous à l'aide d'un même difféomorphisme : il faut utiliser un difféomorphisme différent pour chaque objet, et un pour le milieu. Toutefois, si on veut conserver les propriétés du système, il faut ajouter des contraintes forçant les frontières des objets à coïncider avec celles du milieu ambiant.

Pour résoudre un tel problème, nous allons utiliser le langage de la théorie du contrôle optimal [AS04, Tré08] plutôt que celui de la géométrie. De plus, on travaillera dans \mathbb{R}^d plutôt que dans une variété à géométrie bornée quelconque, et sur un espace de formes \mathcal{S} qui est un ouvert d'un espace de Banach B . C'est dans ce contexte qu'est écrit [ATTY13].

On va commencer par poser rigoureusement le problème de contrôle optimal contraint. On conserve les notations des deux chapitres précédents. On cherche à comparer deux états q_0, q_1 d'un même espace de formes. Si \mathcal{S} est de dimension infinie, il est très rare d'avoir $q_1 \in \mathcal{O}_{q_0}$. Donc, plutôt que de chercher une géodésique horizontale reliant q_0 à q_1 , on va minimiser une fonctionnelle de la forme

$$J(X) = A(\varphi^X, X) + g(q(1)) = \int_0^1 \langle X(t), X(t) \rangle dt + g(\varphi^X(1) \cdot q_0).$$

Dans la formule ci-dessus, on a pris $X \in L^2(0, 1; V)$ avec V un espace de champs de vecteurs à noyau reproduisant, φ^X le flot de X , et $t \mapsto q(t) = \varphi^X(t) \cdot q_0$ la courbe horizontale correspondante sur \mathcal{S} partant de q_0 . La fonction $g : \mathcal{S} \rightarrow \mathbb{R}$ est une fonction d'attache aux données mesurant la différence entre $\varphi^X(1) \cdot q_0$ et un état cible q_1 , voir [GTU, VG05] pour des exemples de telles fonctions.

Pour ajouter des contraintes, on définit également $C : \mathcal{S} \rightarrow \mathcal{L}(V, Y)$, avec Y espace de Banach, et on minimisera J sur le sous-ensemble des X tels que, pour presque tout $t \in [0, 1]$,

$$C_{q(t)}X(t) = 0.$$

Une fois le cadre correctement posé, on voit que sous des hypothèses très faibles, ce problème contraint admet bien une solution. On verra également que, dans la plupart des cas, les solutions du problème de minimisation avec contraintes sur un espace de formes de dimension infini peuvent être approchées par une suite de solutions de problèmes sur des espaces de formes de dimension finie.

On donne ensuite une variante en dimension infinie du principe du maximum de Pontryagin qui donne des conditions de minimisation au premier ordre sur les contrôles. Un des avantages de minimiser une fonctionnelle $J(X) = A(X) + g(\varphi^X(1) \cdot q_0)$, plutôt que de minimiser l'action à extrémités fixées est que tous les minimiseurs satisfont un équivalent contraint de l'équation hamiltonienne des géodésiques. On n'aura pas à s'inquiéter des minimiseurs singuliers ou élusifs par exemple (voir section 2.5).

6.1 Position du problème, existence de solutions

6.1.1 Cadre et Notations

Dans tout ce chapitre, on prendra d un entier strictement positif et \mathcal{S} un ouvert d'un espace de Banach B qui est également un espace de formes dans \mathbb{R}^d d'ordre $\ell \in \mathbb{N}^*$ pour une action

$$(\varphi, q) \mapsto \varphi \cdot q, \quad (\varphi, q) \in \mathcal{D}^s(\mathbb{R}^d) \times \mathcal{S},$$

dont l'action infinitésimale sera notée

$$(q, X) \mapsto \xi_q X, \quad (q, X) \in \mathcal{S} \times \Gamma^s(T\mathbb{R}^d) = \mathcal{S} \times H^s(\mathbb{R}^d, \mathbb{R}^d),$$

avec s le plus petit entier strictement supérieur à $d/2 + \ell$. On fixe également un état initial q_0 dans \mathcal{S} . On définit ensuite $V \subset \mathcal{C}^0(\mathbb{R}^d, \mathbb{R}^d)$ un espace de Hilbert de champs de vecteurs sur \mathbb{R}^d à noyau reproduisant avec inclusion continue dans $H^s(\mathbb{R}^d, \mathbb{R}^d)$. On notera $K = \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathcal{L}((\mathbb{R}^d)^*, \mathbb{R}^d)$ son noyau.

On prendra des contraintes $C : \mathcal{S} \rightarrow \mathcal{L}(V, Y)$, où $\mathcal{L}(V, Y)$ est l'ensemble des opérateurs linéaires continus de V dans Y , avec Y un espace de Banach. Enfin, on fixera une attache aux données $g : \mathcal{S} \rightarrow \mathbb{R}$.

6.1.2 Première formulation

Notons $H_{q_0}^1(0, 1; \mathcal{S}) = \{q \in H^1(0, 1; \mathcal{S}), q(0) = q_0\}$.

Problème 1. *Minimiser la fonctionnelle*

$$J_1(q, X) = \frac{1}{2} \int_0^1 \langle X(t), X(t) \rangle dt + g(q(1))$$

sur l'ensemble des couples

$$(q, X) \in H_{q_0}^1(0, 1; \mathcal{S}) \times L^2(0, 1; V),$$

tels que

$$\dot{q}(t) = \xi_{q(t)} X(t), \quad C_{q(t)} X(t) = 0, \quad p.p. t \in [0, 1].$$

Il s'agit d'un problème de contrôle optimal avec contraintes. Un cas particulièrement intéressant est celui de *contraintes cinétiques*, c'est-à-dire de contraintes de la forme

$$C'(t)\dot{q}(t) = C'_{q(t)} \xi_{q(t)} X(t) = 0.$$

Par exemple, des contraintes d'état pures $C(q(t)) = 0$, pour $C : \mathcal{S} \rightarrow Y$ une application différentiable, sont équivalentes aux contraintes cinétiques $dC(q(t))\xi_{q(t)} X(t) = 0$, et ce cas est donc compris dans le problème 1.

Remark 6.1. On peut équiper \mathcal{S} d'une structure sous-riemannienne pour laquelle les courbes horizontales q sont exactement celles pour lesquelles il existe un contrôle X tel que (q, X) satisfasse les contraintes. Il suffit de remplacer ξ_q par $\xi_q \Pi_q$, $q \in \mathcal{S}$, où Π_q est la projection orthogonale sur $\ker C_q$. Dans le cas où $\ker C$ est un sous-fibré de $\mathcal{S} \times V$, on peut tout simplement considérer la structure $(\ker C, \xi, \langle \cdot, \cdot \rangle)$.

Théorème 6.1. *Si V a une inclusion continue dans $H^{s+1}(\mathbb{R}^d, \mathbb{R}^d)$, si $q \mapsto C_q$ est continue, si g est semicontinue inférieurement et bornée inférieurement, et si q_0 est à support compact, alors J_1 possède au moins un minimiseur.*

6.1.3 Formulation duale

Comme on définit généralement V par son noyau, on connaît rarement explicitement tous ses éléments. Il peut être préférable de considérer une formulation duale en utilisant les résultats de la section 4.2.3. Rappelons que pour $(q, p) \in \mathcal{S} \times B$, on a

$$\langle K^V \xi_q^* p, K^V \xi_q^* p \rangle = p K_q p,$$

où $K^V : V^* \rightarrow V$ est l'opérateur inverse de l'isométrie de Riesz

$$X \mapsto \langle X, \cdot \rangle$$

et

$$K_q = \xi_q K^V \xi_q^* : B^* \rightarrow B.$$

Cela nous pousse à considérer le problème suivant.

Problème 2. *Minimiser la fonctionnelle*

$$J_2(q, u) = J_1(q, K^V \xi_q^* u) = \frac{1}{2} \int_0^1 u(t) K_{q(t)} u(t) dt + g(q(1))$$

sur l'ensemble des couples

$$(q, u) \in H_{q_0}^1(0, 1; \mathcal{S}) \times L^2(0, 1; B^*),$$

tels que

$$\dot{q}(t) = K_{q(t)} u(t), \quad C_{q(t)} K^V \xi_{q(t)}^* u(t) = 0, \quad \text{p.p. } t \in [0, 1]. \quad (6.1)$$

Dans le cas de contraintes cinétiques $C_q = C'_q \xi_q$, (6.1) devient

$$\dot{q}(t) = K_{q(t)} u(t), \quad C'_{q(t)} K_{q(t)}^* u(t) = 0, \quad \text{p.p. } t \in [0, 1].$$

Bien sûr, on se prive de certaines possibilités, il faut donc se demander dans quels cas les problèmes 1 et 2 sont équivalents.

Proposition 6.1. *Supposons que pour tout $q \in \mathcal{S}$, $\ker \xi_q \subset \ker C_q$, et que $\Delta_q = \xi_q(V)$ soit fermé. Alors les problèmes 1 et 2 sont équivalents au sens où $\inf J_1 = \inf J_2$. De plus, si (\bar{q}, \bar{u}) minimise J_2 , alors $(\bar{q}, K^V \xi_{\bar{q}}^* \bar{u})$ minimise J_1 , et, réciproquement, si (\bar{q}, \bar{X}) minimise J_1 il existe \bar{u} tel que $\bar{X} = K^V \xi_{\bar{q}}^* \bar{u}$.*

Voir fin de la section section 10.1.2 pour la preuve. Ces conditions sont toujours vérifiées si \mathcal{S} est de dimension finie et si les contraintes sont cinétiques. Cela dit, en dimension infinie, Δ_q n'est presque jamais fermée. Toutefois, le principe du maximum nous permettra de donner un autre cas où il y a équivalence.

Proposition 6.2. *Supposons que V ait une inclusion continue dans $H^{s+1}(\mathbb{R}^d, \mathbb{R}^d)$, que les contraintes sont cinétiques, de classe \mathcal{C}^1 , et que $C_q(V) = Y$ pour tout q dans \mathcal{S} . Alors pour tout point critique (q, X) de J_1 , il existe $u \in L^2(0, 1; B^*)$ tel que $X = K^V \xi_q^* u$. En conséquence, les problèmes 1 et 2 sont équivalents.*

6.2 Un exemple : espaces de multi-formes

Les problèmes d'analyses de formes demandent parfois d'étudier simultanément plusieurs objets avec des propriétés très distinctes les uns des autres. Dans ce cas, il est plus logique que chaque forme se déplace grâce à son propre espace de champs de vecteurs.

On considère $\mathcal{S}_1, \dots, \mathcal{S}_k$ des espaces de formes dans \mathbb{R}^d , ouverts d'espaces de Banach B_1, \dots, B_k . On considère également V_1, \dots, V_k des espaces de Hilbert de champs de vecteurs sur \mathbb{R}^d à noyau reproduisant avec inclusions continues dans $H^s(\mathbb{R}^d, \mathbb{R}^d)$, où $s > d/2 + 1$. Pour $i \in \{1, \dots, k\}$, on note $\xi_{i,q_i} X_i$ l'action infinitésimale de $X_i \in V_i$ sur $q_i \in \mathcal{S}_i$. On cherche à minimiser

$$\frac{1}{2} \sum_{i=1}^k \int_0^1 \langle X_i(t), X_i(t) \rangle dt + g(q_1(1), \dots, q_k(1))$$

sur tous les $X = (X_1, \dots, X_k) \in L^2(0, 1; V_1 \times \dots \times V_k)$, où $q_i(1) = \varphi^{X_i}(1) \cdot q_{i,0}$, avec $q_{i,0} \in \mathcal{S}_i$ fixé, $i \in \{1, \dots, k\}$. Du point de vue dual, on peut à la place minimiser

$$\frac{1}{2} \sum_{i=1}^k \int_0^1 u_i(t) K_{i,q_i} u_i(t) dt + g(q_1(1), \dots, q_k(1))$$

sur tous les $u = (u_1, \dots, u_k) \in L^2(0, 1; B_1^* \times \dots \times B_k^*)$, où $q_i(0) = q_{i,0}$ et

$$\dot{q}_i(t) = K_{i,q_i(t)} u_i(t)$$

presque partout sur $[0, 1]$, avec $K_{i,q_i} = \xi_{i,q_i} K^{V_i} \xi_{i,q_i}^*$. Bien entendu, si on n'ajoute pas de contraintes, il n'y a presque aucune différence avec le fait d'étudier séparément chaque espace de formes \mathcal{S}_i : la seule interaction possible vient de l'attache aux données g .

Prenons l'exemple de deux objets immergés dans un milieu continu, représentés par des ouverts U_1 et U_2 de \mathbb{R}^d à bords lisses. Chaque objet est déformé par un espace de Hilbert de champs de vecteurs à noyau reproduisant, respectivement noté V_1 et V_2 . Pour modéliser un tel cas de figure, on considère alors deux états $q_{1,0} = \text{Id}_{\partial U_1}$ et $q_{2,0} = \text{Id}_{\partial U_2}$ dans leurs espaces de formes respectifs $\mathcal{S}_1 = \mathcal{C}^1(\partial U_1, \mathbb{R}^d)$ et $\mathcal{S}_2 = \mathcal{C}^1(\partial U_2, \mathbb{R}^d)$, où \mathcal{S}_1 et \mathcal{S}_2 sont des hypersurfaces de \mathbb{R}^d . Les états $q_{1,0}$ et $q_{2,0}$ représentent les bords des ouverts respectifs U_1 et U_2 de \mathbb{R}^d .

Le milieu continu dans lequel ces objets sont immergés est alors repéré par le complémentaire $U_3 = (U_1 \cup U_2)^c$. Il est déformé par des flots de champs de vecteurs d'un troisième espace de Hilbert de champs de vecteurs V_3 . On considère alors un troisième état $q_{3,0}$ qui représente le bord de ce milieu continu. On peut écrire $q_3 = (q_3^1, q_3^2) \in \mathcal{S}_3 = \mathcal{S}_1 \times \mathcal{S}_2$.

Une déformation de ce système

$$t \mapsto (q_1(t), q_2(t), q_3^1(t), q_3^2(t)) = (\varphi^{X^1}(t) \cdot q_1, \varphi^{X^2}(t) \cdot q_2, \varphi^{X^3}(t) \cdot q_3), \quad X_i \in L^2(0, 1; V_i), \quad i = 1, 2, 3,$$

représente la façon dont chaque bord ∂U_i est déformé par le flot φ^{X_i} .

Pour que cette déformation conserve la configuration voulue, il faut imposer que le bord du milieu se déplace avec les bords des deux objets, ce qui revient à imposer pour tout temps les contraintes suivantes :

$$q_1(t) = q_3^1(t), \quad q_2(t) = q_3^2(t).$$

On parle de contraintes *collées*. On peut utiliser ce modèle pour étudier des déformations d'objets indépendants qui ne doivent pas se rencontrer, puisque la transformation

$$x \mapsto \varphi^{X^i}(t, x) \text{ lorsque } x \in \bar{U}_i$$

est alors un homéomorphisme de \mathbb{R}^d pour tout t .

Si on veut que la satisfaction des contraintes ne dépendent pas de la paramétrisation q_i du bord de U_i tout au long de la déformation, il faut laisser les bords glisser les uns sur les autres. Autrement dit, on doit imposer

$$\exists f_i \in \text{Diff}(\partial U_i), \quad q_i(t) \cdot f = q_3^i(t), \quad i = 1, 2.$$

On parle alors de contraintes *glissées*.

6.3 Approximation par des espaces de formes de dimension finie

On suppose que V a une inclusion continue dans $H^{s+1}(\mathbb{R}^d, \mathbb{R}^d)$, et que q_0 est à support compact. Soient alors $(Y^n)_{n \in \mathbb{N}}$ une suite d'espaces de Banach, $C^n : \mathcal{S} \rightarrow \mathcal{L}(V, Y^n)$ une suite de contraintes, et (g^n) une suite de fonctions $\mathcal{S} \rightarrow \mathbb{R}$ semicontinues inférieurement et possédant une borne inférieure indépendante de n . Pour tout entier n , on considère le problème P^n suivant : minimiser

$$J_1^n(q, X) = \frac{1}{2} \int_0^1 \langle X(t), X(t) \rangle dt + g^n(q(1))$$

sur l'ensemble des couples $(q, X) \in H_{q_0}^1(0, 1; \mathcal{S}) \times L^2(0, 1; V)$ tels que

$$\dot{q}(t) = \xi_{q(t)} X(t), \quad C_{q(t)}^n X(t) = 0, \quad \text{p.p. } t \in [0, 1].$$

Par le théorème 6.1, P^n admet au moins une solution (q^n, X^n) . Alors, sous certaines conditions, on peut extraire de la suite (q^n, X^n) une sous-suite qui converge vers une solution du problème 1.

Proposition 6.3. *Supposons que V ait une inclusion continue dans $H^{s+1}(\mathbb{R}^d, \mathbb{R}^d)$, et que pour tout $q \in \mathcal{S}$, la suite $(\ker C_q^n)_{n \in \mathbb{N}}$ soit décroissante pour l'inclusion et satisfasse*

$$\bigcap_{n \in \mathbb{N}} \ker C_q^n = \ker C_q.$$

On suppose également que g^n converge vers g sur tout compact de \mathcal{S} , et que q_0 est à support compact. Alors la suite $(X^n)_{n \in \mathbb{N}}$ est bornée dans $L^2(0, 1; V)$, et tout point d'adhérence \bar{X} de cette suite pour la topologie faible sur $L^2(0, 1; V)$, $(\bar{q}, \bar{X}) = (\varphi^{\bar{X}} \cdot q_0, \bar{X})$ satisfait les contraintes et minimise J_1 .

Plus précisément, pour toute sous-suite $(X^{n_j})_{j \in \mathbb{N}}$ faiblement convergente vers un certain $\bar{X} \in L^2(0, 1; V)$, la suite de courbes $(q^{n_j})_{j \in \mathbb{N}}$ converge uniformément vers la courbe $\bar{q} = \varphi^{\bar{X}} \cdot q_0$, $C_{\bar{q}(t)} \bar{X}(t) = 0$ presque partout, et

$$J_1^{n_j}(q^{n_j}, X^{n_j}) \xrightarrow{j \rightarrow +\infty} J_1(\bar{q}, \bar{X}) = \inf J_1.$$

On peut utiliser ce résultat pour approcher une solution du problème 1 sur un espace de formes de dimension infinie \mathcal{S} par une suite de solutions au problème 1 sur des espaces de formes de dimension finie. En effet, soit (\mathcal{S}^n) une suite d'espaces de formes d'ordre ℓ de dimension finie, et supposons qu'il existe des applications équivariantes $\pi^n : \mathcal{S} \rightarrow \mathcal{S}^n$ et $\pi_m^n : \mathcal{S}^m \rightarrow \mathcal{S}^n$ telles que $\pi_m^n \pi^m = \pi^n$ pour tous $m \geq n \in \mathbb{N}$. On note $q_0^n = \pi^n(q_0)$. On définit alors des contraintes $\tilde{C}^n : \mathcal{S}^n \rightarrow Y^n$ et des attaches aux données $\tilde{g}^n : \mathcal{S}^n \rightarrow \mathbb{R}$ sur les espaces de formes de dimension finie \mathcal{S}^n .

Soit (q^n, X^n) une solution du problème 1 sur \mathcal{S}^n avec contraintes \tilde{C}^n et attache aux données \tilde{g}^n . Alors $(\varphi^{X^n} \cdot q_0, X^n)$ est une solution du problème P^n sur \mathcal{S} ci-dessus avec contraintes $C_q^n = \tilde{C}_{\pi^n(q)}^n$ et attache aux données $g^n = \tilde{g}^n \circ \pi^n$. On trouve donc des solutions aux problèmes P^n en résolvant le problème 1 sur des espaces de formes de dimension finie.

Exemple 6.1. Soient S^1 le cercle unité, et $\mathcal{S} = \mathcal{C}^1(S^1, \mathbb{R}^d)$ l'ensemble des courbes fermées de \mathbb{R}^d . C'est un espace de formes d'ordre 1 sur lequel $\mathcal{D}^s(\mathbb{R}^d)$ agit par composition. On prend l'attache aux données

$$g(q) = \int_{S^1} (q(s) - q_1(s))^T (q(s) - q_1(s)) ds,$$

pour q_1 une courbe cible. Soit $F \in \mathcal{C}^1(\mathbb{R}^d, \mathcal{L}(\mathbb{R}^d, \mathbb{R}^k))$. On définit des contraintes $C_q \in \mathcal{L}(V, \mathcal{C}^1(S^1, \mathbb{R}^k))$ par

$$C_q X(s) = F(q(s))X(q(s)), \quad s \in S^1.$$

Soit alors $(s_n)_{n \in \mathbb{N}^*}$ une suite dense sur S^1 .

On peut alors définir $\mathcal{S}^n = (\mathbb{R}^d)^n$,

$$\pi^n(q) = (q(s_1), \dots, q(s_n)), \quad \pi_m^n(x_1, \dots, x_m) = (x_1, \dots, x_n)$$

avec $m \geq n \in \mathbb{N}$, et $(x_1, \dots, x_m) \in \mathcal{S}^m$, ainsi que

$$\tilde{C}_{(x_1, \dots, x_n)}^n X = (F(x_1)X(x_1), \dots, F(x_n)X(x_n)), \quad (x_1, \dots, x_n) \in (\mathbb{R}^d)^n$$

et

$$\tilde{g}^n(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n (x_i - q_1(s_i))^T (x_i - q_1(s_i)).$$

Alors, pour $g^n = \tilde{g}^n \cdot \pi^n$ et $C_q^n = \tilde{C}_{\pi^n(q)}^n$, on a bien que $(\ker C^n)_{n \in \mathbb{N}}$ est croissante, que

$$\bigcap_{n \in \mathbb{N}} \ker C_q^n = \ker C_q,$$

et que g^n converge vers g uniformément sur tout compact.

6.4 Principe du maximum avec contraintes

6.4.1 Énoncé du théorème

On se place dans le cadre suivant, qui englobe celui des espaces de formes avec contraintes. On considère V un espace de Hilbert, \mathcal{S} un ouvert d'un espace de Banach B , et Y un espace de Banach.

Soient $\xi : \mathcal{S} \times V \rightarrow B$ et $C : \mathcal{S} \times V \rightarrow Y$, deux applications de classe \mathcal{C}^1 linéaires par rapport à la seconde variable. On définit des fonctions $L : \mathcal{S} \times V \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 et $g : \mathcal{S} \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 . On suppose que L satisfait, pour des fonctions continues $\gamma_0 : \mathbb{R} \rightarrow \mathbb{R}$ valant 0 en 0 et $\gamma_1 : B \rightarrow \mathbb{R}$ fixées,

$$\|dL(q', X') - dL(q, X)\|_{B^* \times V^*} \leq \gamma_0(\|q' - q\|_B) + \gamma_1(q' - q)\|X' - X\|_V, \quad (6.2)$$

pour tous $(q, q', X, X') \in \mathcal{S}^2 \times V^2$, où on note $\|\cdot\|_E$ la norme d'un espace de Banach E . Fixons q_0 dans M .

Problème 3. *Minimiser*

$$J(q, X) = \int_0^1 L(q(t), X(t)) dt + g(q(1))$$

sur tous les couples $(q, v) \in H_{q_0}^1(0, 1; \mathcal{S}) \times L^2(0, 1; V)$ tels que

$$\dot{q}(t) = \xi_{q(t)} X(t), \quad C_{q(t)} X(t) = 0, \quad p.p. t \in [0, 1].$$

Remark 6.2. Les estimées (6.2) sont des conditions suffisantes pour que J soit différentiable, de dérivée

$$dJ_{(q,X)}(\delta q, \delta X) = \int_0^1 \partial_q L(q(t), X(t)) \delta q(t) + \partial_X L(q(t), X(t)) \delta X(t) dt + dg(q(1)) \delta q(1).$$

Pour cela, on définit le hamiltonien augmenté $H : \mathcal{S} \times B^* \times V \times Y^* \rightarrow \mathbb{R}$ par

$$H(q, p, X, \lambda) = p \xi_q X - L(q, X) - \lambda C_q X.$$

On peut alors donner le principe du maximum de Pontryagin (PMP) suivant.

Théorème 6.2. *Supposons que pour tout $q \in \mathcal{S}$, C_q soit surjective. Soit $(q, X) \in H_{q_0}^1(0, 1; \mathcal{S}) \times L^2(0, 1; V)$ une solution du problème 3. Alors il existe $p \in H^1(0, 1; B^*)$ et $\lambda \in L^2(0, 1; Y^*)$ tels que $p(1) + dg(q(1)) = 0$ et*

$$\begin{cases} 0 = \partial_X H(q(t), p(t), X(t), \lambda(t)), \\ \dot{q}(t) = \partial_p H(q(t), p(t), X(t), \lambda(t)), \\ \dot{p}(t) = -\partial_q H(q(t), p(t), X(t), \lambda(t)), \end{cases} \quad (6.3)$$

pour presque tout $t \in [0, 1]$. De plus, λ est unique, appelé le **multiplieur de Lagrange** du problème.

Plusieurs variantes du PMP en dimension infinie sont disponibles dans la littérature, par exemple pour des problèmes de contrôle d'EDP, mais ces variantes sont généralement dépourvues de contraintes. Des versions du PMP avec contraintes sont nombreuses en dimension finie, mais peu d'entre elles concernent la dimension infinie, et même celles-là demandent des contraintes telles que $\ker C_q$ soit de codimension finie. La preuve fournie dans [ATTY13] contient certaines subtilités qui font du théorème 6.2 une extension non triviale du PMP habituel.

En effet, une des premières étapes de la preuve usuelle est de construire notre p dans $(L^2(0, 1; B))^*$. Lorsque B est réflexif, comme en dimension finie, $(L^2(0, 1; B))^* = L^2(0, 1; B^*)$ [Bou83], et on peut alors montrer que p est en fait de classe H^1 . Toutefois, lorsque B n'est pas réflexif, on n'a pas forcément $(L^2(0, 1; B))^* = L^2(0, 1; B)$. Il a donc d'abord fallut montrer que l'on pouvait identifier p à un élément de $L^2(0, 1; B^*)$.

6.4.2 Équation des géodésiques contraintes sur les espaces de formes

On revient aux notations habituelles pour la résolution du problème 1. Dans ce cas, $L(q, X) = \frac{1}{2} \langle X, X \rangle$. La condition

$$\partial_X H(q, p, X, \lambda) = 0$$

est donc vérifiée si et seulement si $X = K^V(\xi_q^* p - C_q^* \lambda)$. Ainsi, $C_q X = 0$ si et seulement si

$$C_q K^V C_q^* \lambda = C_q K^V \xi_q^* p.$$

Or, lorsque C_q est surjective pour tout q , cette équation a une unique solution $\lambda(q, p)$, de même régularité que ξ_q et C_q . Donc $X = X(q, p)$ est également déterminé par (q, p) . Cela nous permet de définir le hamiltonien normal $h : \mathcal{S} \times B^* \rightarrow \mathbb{R}$ par

$$h(q, p) = H(q, p, X(q, p), \lambda(q, p)),$$

ainsi que le résultat suivant.

Théorème 6.3. *Supposons que ξ , C et g sont de classe \mathcal{C}^2 , et que C_q est surjective pour tout q dans \mathcal{S} . Soit $(q, X) \in H_{q_0}^1(0, 1; \mathcal{S}) \times L^2(0, 1; V)$ une solution du problème 1. Alors il existe $p \in H^1(0, 1; B^*)$ tel que (q, p) satisfasse les équations géodésiques avec contraintes*

$$\begin{cases} \dot{q}(t) = \partial_p h(q(t), p(t)) &= \xi_{q(t)} X(t) = K_{q(t)} p(t) - \xi_{q(t)} K^V C_{q(t)}^* \lambda(t), \\ \dot{p}(t) = -\partial_q h(q(t), p(t)) &= -\partial_q (K_{q(t)} X(t))^* p(t) + \partial_q (C_{q(t)} X(t))^* \lambda(t), \end{cases} \quad (6.4)$$

sur $[0, 1]$, où le multiplicateur de Lagrange est donné par

$$C_{q(t)} K^V C_{q(t)}^* \lambda(t) = C_{q(t)} K^V \xi_{q(t)}^* p(t),$$

et

$$X(t) = K^V (\xi_{q(t)}^* p(t) - C_{q(t)}^* \lambda(t)).$$

De plus, la propriété " C_q surjectif pour tout q " est une condition suffisante pour que $(\ker C, \xi, \langle \cdot, \cdot \rangle)$ soit une structure sous-riemannienne lisse, où $\ker C$ est le fibré sur \mathcal{S} dont la fibre au-dessus de q est $\ker C_q$. On voit alors facilement que h est aussi le hamiltonien normal de cette nouvelle structure. On peut donc appliquer le théorème 2.1.

Proposition 6.4. *Si C et ξ sont de classe \mathcal{C}^2 , alors h aussi, et l'équation des géodésiques contraintes a une unique solution maximale $(q, p) : I \ni 0 \rightarrow \mathcal{S} \times B$ avec condition initiale $(q_0, p_0) \in \mathcal{S} \times B^*$ fixée. La courbe q est alors un point critique de l'action sur les courbes horizontales contraintes à extrémités fixées. De plus, si C et ξ sont \mathcal{C}^3 , alors q minimise l'action sur l'ensemble des courbes horizontales contraintes à extrémités fixées.*

Remark 6.3. Dans le cas de contraintes cinétiques $C_q \xi_q X = 0$, en notant $\tilde{p}(t) = p(t) - C_{q(t)}^* \lambda(t)$, les équations se simplifient en

$$\begin{cases} \dot{q}(t) = K_{q(t)} \tilde{p}(t), \\ \dot{\tilde{p}}(t) = -\frac{1}{2} \partial_q (K_{q(t)} \tilde{p}(t))^* \tilde{p}(t) + \partial_q (C_{q(t)} K_{q(t)} \tilde{p}(t))^* \lambda(t), \end{cases} \quad (6.5)$$

Dans le cas de contraintes d'états pures, c'est-à-dire le cas où l'on remplace les contraintes $C_{q(t)} X(t) = 0$ par $dC(q(t)) X(t) = 0$ presque partout, avec $C : \mathcal{S} \rightarrow Y$ de classe \mathcal{C}^2 , on peut également réécrire l'équation

$$\begin{cases} \dot{q}(t) = K_{q(t)} \tilde{p}(t), \\ \dot{\tilde{p}}(t) = -\partial_q (K_{q(t)} \tilde{p}(t))^* \tilde{p}(t) + dC(q(t))^* \tilde{\lambda}(t). \end{cases}$$

Dans ce cas, l'équation pour le multiplicateur de Lagrange λ devient

$$dC(q) K_q dC(q)^* \tilde{\lambda} = \frac{1}{2} \partial_q (K_q \tilde{p})^* \tilde{p} - d^2 C(q) (K_q \tilde{p}, K_q \tilde{p}) - dC(q) \partial_q (K_q \tilde{p}) K_q \tilde{p}.$$

Chapitre 7

Algorithmes de minimisation

On conserve les notations du chapitre précédent.

Dans ce chapitre, résumé des sections 4 et 5 de [ATTY13], on donne des algorithmes développés pendant cette thèse permettant de calculer approximativement des minimiseurs contraints de J_1 . On décrit les deux principales méthodes : la minimisation sur l'espace des contrôles, qui s'effectue dans $L^2(0, 1; V)$, et la minimisation par tir. Celle-ci s'appuie sur le principe du maximum (théorème 6.2), qui assure que les minimiseurs de J_1 sont des géodésiques normales contraintes. Il suffit donc de minimiser J_1 sur l'espace des géodésiques partant de q_0 , qui sont paramétrées par la condition initiale $p(0) \in T_{q_0}^* \mathcal{S}$. On réduit donc dans ce cas le problème à la minimisation d'une fonction sur $T_{q_0}^* \mathcal{S}$. Les différentes preuves, ainsi que quelques précisions, sont données en section 10.3.

On finit par la description d'une variante de l'algorithme de quasi-Newton qui permet d'effectuer plus rapidement certains calculs nécessaires à la méthode du tir. Il s'agit d'un résumé rapide de [Arg13], reproduit au chapitre 11.

7.1 Minimisation sur l'espace des contrôles par l'algorithme du gradient

On commence par calculer le gradient sans contraintes de la fonctionnelle J_1 . Ensuite, on rappelle la méthode du lagrangien augmenté pour minimiser une fonctionnelle sous contraintes, et on calcule le gradient correspondant pour cet algorithme.

7.1.1 Gradient sans contraintes

Il s'agit d'une méthode bien connue et déjà utilisée dans des exemples classiques d'espaces de formes [BMTY05, DGM98, GTY, GTY06, JM00, MTY02, MTY06, Tro95]. Nous la généralisons au cadre décrit en section 6.4 : on considère un espace de Banach B , \mathcal{S} un ouvert de B , et V un espace de Hilbert. On prend également une fonction $\xi : \mathcal{S} \rightarrow \mathcal{L}(V, B)$ telle que $(q, X) \mapsto \xi_q X$ soit de classe \mathcal{C}^2 , et un lagrangien $L : \mathcal{S} \times V \rightarrow \mathbb{R}$ de classe \mathcal{C}^2 qui satisfait les estimées 6.2. On définit enfin $g : \mathcal{S} \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 , et on fixe $q_0 \in \mathcal{S}$. On veut minimiser

$$J(q, X) = \int_0^1 L(q(t), X(t)) dt + g(q(1)),$$

sur l'ensemble des $(q, X) \in H_{q_0}^1(0, 1; \mathcal{S}) \times L^2(0, 1; V)$ tels que $\dot{q}(t) = \xi_{q(t)} X(t)$ presque partout.

On sait (voir par exemple [Tré08] ou l'appendice D de [Mon02]) que pour tout contrôle fixé X dans un ouvert \mathcal{U} de $L^2(0, 1; V)$, il existe une unique courbe $q(\cdot, X) = q^X$ horizontale partant de q_0 telle que $\dot{q}^X(t) = \xi_{q^X(t)} X(t)$. De plus, l'application $X \mapsto q^X$ est de même

régularité que ξ (voir les sections 1.4 et 2.4). On peut donc identifier J à la fonctionnelle $X \mapsto J(X, q^X)$ définie sur l'ouvert \mathcal{U} . Enfin, le hamiltonien H est défini sur $\mathcal{S} \times B^* \times V$ par

$$H(q, p, X) = p\xi_q X - L(q, X).$$

Proposition 7.1. *La différentielle en $X \in \mathcal{U}$ de J est donnée par*

$$dJ(X)\delta X = - \int_0^1 \partial_X H(q^X(t), p(t), X(t))\delta X(t)dt, \quad \delta v \in L^2(0, 1; V),$$

où $p \in H^1(0, 1; B^*)$ est la solution du problème de Cauchy adjoint $\dot{p}(t) = \partial_q H(q^X(t), p(t), X(t))$ avec condition initiale $p(1) + dg(q^X(1)) = 0$. On obtient ainsi

$$dJ(X)\delta X = \int_0^1 \langle \nabla J(X)(t), \delta X(t) \rangle dt,$$

où

$$\nabla J(X)(t) = -K^V \partial_X H(q^X(t), p(t), X(t)).$$

Dans le cas des espaces de formes, $L(q, X) = \frac{1}{2} \langle X, X \rangle$, et $\partial_X H(q, p, X) = \xi_q^* p - \langle X, \cdot \rangle$, d'où

$$\nabla J_1(X) = X - K^V \xi_q^* p, \quad X \in L^2(0, 1; V).$$

Remark 7.1. Une remarque particulièrement intéressante est que pour $u \in L^2(0, 1; B^*)$,

$$\nabla J_1(K^V \xi_q^* u) = K^V \xi_q^* (u - p).$$

On en déduit aisément que $\nabla J_2(u) = u - p$ est un gradient de

$$J_2(u) = J_2(q^u, u) = \frac{1}{2} \int_0^1 u(t) K_{q^u(t)} u(t) dt + g(q^u(1)),$$

où q^u satisfait $\dot{q}^u(t) = K_{q^u(t)} u(t)$ presque partout, par rapport à la forme bilinéaire semi-définie positive

$$(u_1, u_2) \mapsto \int_0^1 u_1(t) K_{q^u(t)} u_2(t) dt.$$

Cela permet de travailler directement dans $L^2(0, 1; B^*)$.

7.1.2 Méthode du lagrangien augmenté pour le cas contraint

On travaille dans le cadre d'un espace de formes de dimension finie ($B = \mathbb{R}^n$ et $Y = \mathbb{R}^k$), avec des contraintes cinétiques de la forme $C_q \dot{q} = 0$ presque partout. En particulier, on identifie les éléments des espaces \mathbb{R}^a , $a \in \mathbb{N}$, à des matrices colonnes de taille a , ceux de $(\mathbb{R}^a)^*$ à des matrices lignes, et on note $|w|$ la norme euclidienne d'un vecteur $w \in \mathbb{R}^a$.

Alors, par la proposition 6.1 le problème 1 est équivalent au problème 2 consistant à minimiser

$$J_2(u) = J_2(q^u, u) = \frac{1}{2} \int_0^1 u(t) K_{q^u(t)} u(t) dt + g(q^u(1))$$

sur l'ensemble des $u \in L^2(0, 1; (\mathbb{R}^n)^*)$ tels que $q^u(0) = q_0$, et, presque partout, on ait

$$\dot{q}^u(t) = K_{q^u(t)} u(t), \quad C_{q^u(t)} K_{q^u(t)} u(t) = 0.$$

Définissons alors le lagrangien augmenté $L_A : \mathcal{S} \times V \times (\mathbb{R}^k)^* \times \mathbb{R}_+ \rightarrow \mathbb{R}$ par

$$L_A(q, u, \lambda, \mu) = \frac{1}{2} u K_q u + \lambda C_q K_q u + \frac{1}{2\mu} |C_q K_q u|^2.$$

L'algorithme du lagrangien augmenté est défini par récurrence comme suit.

1. On prend un contrôle $u_0 \in L^2(0, 1; (\mathbb{R}^n)^*)$ (par exemple $u_0 = 0$). On fixe également λ_0 dans $L^2(0, 1; (\mathbb{R}^k)^*)$ (par exemple $\lambda_0 = 0$), et $\mu_0 > 0$. On pose $r = 0$ et on note $q^r = q^{u_0}$.
2. On minimise par rapport à $u \in L^2(0, 1; (\mathbb{R}^n)^*)$ la fonctionnelle sans contrainte

$$u \mapsto J_A(q^u, u, \lambda_r, \mu_r) = \int_0^1 L_A(q^u(t), u(t), \lambda_r(t), \mu_r) dt + g(q^u(1)),$$

par exemple en partant du point initial u_r , et en utilisant un algorithme du gradient. Le minimiseur obtenu est noté u_{r+1} .

3. On met à jour $q^{r+1} = q^{u_{r+1}}$, et

$$\lambda_{r+1} = \lambda_r - \frac{1}{\mu_r} C_{q^{r+1}} K_{q^{r+1}} u_{r+1}.$$

On fixe ensuite $\mu_{r+1} \leq \mu_r$. On remplace enfin r par $r + 1$ et on reprend l'étape 2.

Sous des hypothèses appropriées, si, à partir d'un certain r , μ_r est plus petit qu'une certaine constante, la suite de contrôles u_r converge vers un contrôle optimal \bar{u} . Plus précisément, un théorème du chapitre 3 de [IK08] permet d'assurer la proposition suivante.

Proposition 7.2. ([IK08, Chapitre 3]) *Supposons que toutes les applications apparaissant dans l'algorithme du lagrangien augmenté soient de classe \mathcal{C}^2 et que $C_q K_q$ soit surjective pour tout q dans \mathcal{S} . Soit \bar{u} un contrôle optimal contraint de trajectoire associée \bar{q} , et de multiplicateur de Lagrange associé $\bar{\lambda}$. On suppose qu'il existe c et μ des constantes strictement positives telles que*

$$\forall \delta u \in L^2(0, 1; (\mathbb{R}^n)^*), \quad \partial_u^2 J(\bar{q}, \bar{u}, \bar{\lambda}, \mu)(\delta u, \delta u) \geq c \|\delta u\|_{L^2(0, 1; (\mathbb{R}^n)^*)}^2.$$

Alors il existe un voisinage \mathcal{U} de \bar{u} dans $L^2(0, 1; (\mathbb{R}^n)^*)$ tel que, si le contrôle initial u_0 de l'algorithme du lagrangien augmenté appartient à \mathcal{U} , et si on prend $\mu_r \leq \mu$ pour tout r , alors la suite $(u_r)_{r \in \mathbb{N}}$ (resp. $(\lambda_r)_{r \in \mathbb{N}}$) obtenue converge fortement vers \bar{u} (resp. $\bar{\lambda}$).

Pour appliquer l'algorithme, il ne manque plus qu'à calculer le gradient en u de $J_A(q^u, u, \lambda, \mu)$ pour λ et μ fixés. La section 10.3.2 nous donne la formule suivante.

Proposition 7.3. *Pour $u \in L^2(0, 1; (\mathbb{R}^n)^*)$, et pour la forme bilinéaire semi-définie positive*

$$(u_1, u_2) \mapsto \int_0^1 u_1(t) K_{q^{u(t)}} u_2(t) dt, \quad u_1, u_2 \in L^2(0, 1; (\mathbb{R}^n)^*),$$

on a

$$\nabla_u J_A(q^u, u, \lambda, \mu) = u - p + \lambda C_{q^u} + \frac{1}{\mu} u K_{q^u} C_{q^u}^T C_{q^u},$$

avec

$$\dot{p}(t) = \left(\frac{u(t)}{2} - p(t) \right) \partial_q (K_{q^{u(t)}} u(t)) + \lambda(t) \partial_q (C_{q^{u(t)}} K_{q^{u(t)}} u(t)) + \frac{1}{2\mu} \partial_q |C_{q^{u(t)}} K_{q^{u(t)}} u(t)|^2.$$

presque partout et $p(1) + dg(q^u(1)) = 0$.

Remark 7.2. Le calcul de $\nabla_u J_A(q^u, u, \lambda, \mu)$ est relativement rapide. En effet, une fois une discrétisation en temps effectuée, K_q a généralement une forme suffisamment simple pour que l'expression

$$\left(\frac{u(t)}{2} - p(t) \right) \partial_q (K_{q^{u(t)}} u(t)) + \lambda(t) \partial_q (C_{q^{u(t)}} K_{q^{u(t)}} u(t)) - \frac{1}{2\mu} \partial_q |C_{q^{u(t)}} K_{q^{u(t)}} u(t)|^2$$

soit calculable avec un nombre de multiplications de l'ordre de n^2 (on rappelle que n est la dimension de \mathcal{S}). En revanche, un très grand nombre d'itérations est nécessaire pour que les contraintes soient approximativement satisfaites.

7.2 Minimisation sur le moment initial par la méthode du tir

On se place, ici encore, dans le cas d'un espace de formes de dimension finie, de telle sorte que $B = \mathbb{R}^n$. Alors le principe du maximum implique que minimiser J_1 sous contraintes revient à minimiser

$$\tilde{J}_1(p_0) = J(q, \xi_q^* p) = \frac{1}{2} p_0 K_{q_0} p_0 + g(q(1)),$$

où $(q, p) : [0, 1] \rightarrow \mathcal{S} \times (\mathbb{R}^n)^*$ satisfait l'équation des géodésiques

$$\dot{q} = \partial_p h(q, p), \quad \dot{p} = -\partial_q h(q, p)$$

sur $[0, 1]$, avec h le hamiltonien normal défini en section 6.4.2. Autrement dit, plutôt que de minimiser J_1 sur l'espace des contrôles, on se restreint à l'espace des géodésiques contraintes partant de q_0 et identifié à $(\mathbb{R}^n)^* = T_{q_0}^* \mathcal{S}$. Il faut alors calculer un gradient de \tilde{J}_1 . On utilise la proposition bien connue suivante, prouvée en section 10.3.1.

Proposition 7.4. *Soit $X : \mathbb{R}^d \rightarrow \mathbb{R}^d$ un champ de vecteurs de classe \mathcal{C}^1 , et $x_0 \in \mathbb{R}^d$ tel que le flot φ^X de X soit défini sur $[0, 1] \times U$, avec U un voisinage de x_0 . On définit $G : U \rightarrow \mathbb{R}$ par $G(x) = g(\varphi^X(1, x))$, où $g : \mathbb{R}^d \rightarrow \mathbb{R}$ est de classe \mathcal{C}^1 . Alors G est aussi de classe \mathcal{C}^1 et $dG(x_0) = Z(1)$, où $Z(0) = dg(\varphi(1, x_0))$, et*

$$\forall t \in [0, 1], \quad \dot{Z}(t) = Z(t) dX(\varphi(1 - t, x_0)).$$

Au niveau algorithmique, cette méthode a un défaut : pour calculer ZdX il faut calculer toutes les dérivées partielles de X . Le calcul de ZdX^T aura donc une complexité d fois supérieure à celui de X . Une exception à cette règle survient lorsque X est un champ de gradient, auquel cas dX est symétrique, et $ZdX = (dX Z^T)^T$. Le calcul de ZdX est alors du même ordre de complexité que celui de X , en utilisant par exemple une différence finie

$$ZdX(x) = (dX(x) Z^T)^T \simeq \frac{1}{h} (X(x + hZ^T) - X(x)).$$

Dans notre cas, on a $E = \mathbb{R}^n \times (\mathbb{R}^n)^*$ et $X(q, p) = \nabla^\omega h(q, p)$. Calculer $X(q, p)$ requiert un nombre de multiplications de l'ordre de $n^2 + k^{2,8}$, où $k = \dim(Y)$ est la dimension de l'espace des contraintes. Le terme en $k^{2,8}$ provient de la résolution du système linéaire

$$C_q K^V C_q^* \lambda = C_q K^V \xi_q^* p,$$

nécessaire pour le calcul de h . Donc, à chaque étape de temps, le calcul de \dot{Z} nécessite un nombre de multiplications de l'ordre de $n^3 + nk^{2,8}$, ce qui est beaucoup trop important si k est du même ordre de grandeur que n .

Toutefois, comme on est dans le cas d'un champ de gradient symplectique, une simplification similaire à celle du cas où dX est symétrique est possible, et on obtient

$$(Zd\nabla^\omega h)^T = \begin{pmatrix} \partial_p(\partial_q h) z^T - \partial_q(\partial_q h) \alpha^T \\ \partial_p(\partial_p h) z^T - \partial_q(\partial_p h) \alpha^T \end{pmatrix},$$

dont le calcul est d'une complexité du même ordre que celui de $\nabla^\omega h$ (voir section 10.3.2 pour plus de détails). Ces formules permettent de calculer le gradient de \tilde{J}_1 avec une complexité de l'ordre de $n^2 + nk^{2,8}$, et donc de minimiser J_1 par la méthode du tir.

Remark 7.3. Lorsque le nombre de contraintes k est du même ordre de grandeur que la dimension n de l'espace de formes, le calcul du gradient de \tilde{J}_1 est sensiblement plus long que celui du gradient de la fonctionnelle J_A pour la méthode du lagrangien augmenté. En revanche,

le nombre d'étapes nécessaires est bien moindre, car les contraintes sont par construction toujours satisfaites avec une grande précision : on ne travaille qu'avec des courbes contraintes dans la méthode du tir.

Lorsque le nombre de contraintes est petit, la méthode du tir est supérieure, alors que les deux méthodes ont tendance à se valoir dans le cas général si on ne demande pas aux contraintes d'être satisfaites avec une trop grande précision. En revanche, la matrice $C_q K_q C_q^*$ peut être très mal conditionnée dans certains cas pratiques. Il est alors préférable d'utiliser la méthode du lagrangien augmenté.

7.3 Simulations numériques

7.3.1 Courbes englobant une aire constante

Commençons par considérer un problème avec des contraintes très simples. On prend

$$\mathcal{S} = \text{Emb}^1(S^1, \mathbb{R}^2),$$

l'espace des courbes simples fermées de classe \mathcal{C}^1 dans le plan. On considère l'espace à noyau reproduisant V associé au noyau gaussien

$$K(x, y)p = \exp\left(-\frac{|x - y|^2}{2}\right)p^T.$$

Une courbe $q \in \mathcal{S}$ est le bord d'un domaine compact $U(q)$ dont l'aire est donnée par

$$A(U(q)) = \int_{S^1} x(s)\dot{y}(s)ds.$$

Soit $q_0 \in \mathcal{S}$ un point initial fixé, q_1 une cible telle que $A(U(q_1)) = A(U(q_0))$, et $g : \mathcal{S} \rightarrow \mathbb{R}$ avec

$$g(q) = d_{\mathcal{S}}(q, q_1)^2,$$

où $d_{\mathcal{S}}$ est une distance sur l'espace des courbes avec de bonnes propriétés (voir [VG05, GTY, GTY06] pour des exemples de telles distances).

On fait une implémentation numérique de la minimisation de J_1 sous les contraintes

$$A(U(q(t))) - A(U(q_0)) = 0.$$

Pour cela, on discrétise \mathcal{S} en l'espace des landmarks $Lmk_n(\mathbb{R}^2)$, dont les éléments

$$q = ((x_1, y_1), \dots, (x_n, y_n))$$

représentent une discrétisation d'une courbe simple fermée. L'aire est alors celle du polygone $U(q)$ de sommets $(x_1, y_1), \dots, (x_n, y_n)$, calculée à l'aide du polynôme

$$A(U(q)) = \frac{1}{2}(x_1 y_2 - x_2 y_1 + \dots + x_n y_1 - x_1 y_n).$$

On a pris q_0 un cercle de centre $(0, 0)$ et de rayon 1 et q_1 un cercle de centre $(3, 0)$ et de rayon 1. La figure 7.3.1 compare la trajectoire optimale sans contraintes et la trajectoire optimale avec contraintes de volume constant. Enfin, la figure 7.2 montre le difféomorphisme final sous contraintes.

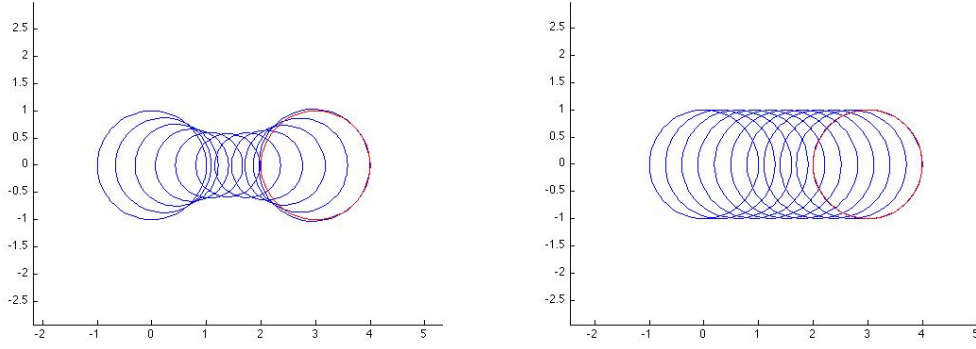


FIGURE 7.1 – Trajectoire optimale sans contraintes (à gauche) et avec volume constant (sur la droite).

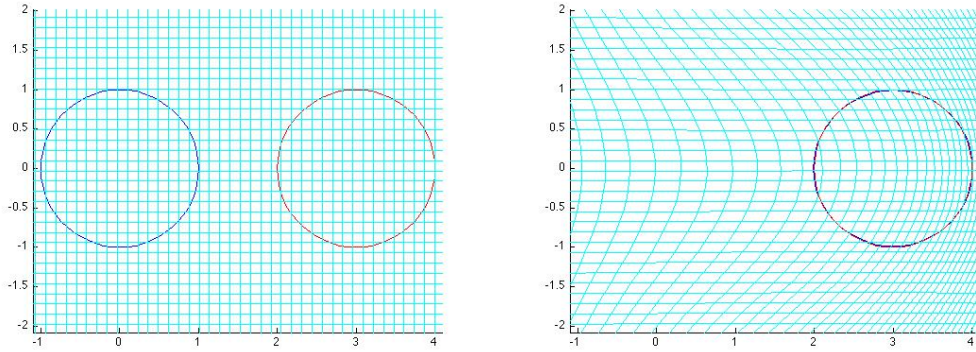


FIGURE 7.2 – Diffeomorphisme correspondant à la trajectoire optimale. Figure gauche : condition initiale (en bleu) et cible (en rouge). Figure droite : diffeomorphisme résultant.

7.3.2 Cas des multiformes

On considère maintenant un problème de multiformes. On prend $\mathcal{S}_1 = \mathcal{S}_2 = \text{Emb}^1(S^1, \mathbb{R}^2)$ et $\mathcal{S}_3 = \mathcal{S}_1 \times \mathcal{S}_2$. On considère également trois noyaux K_1, K_2 et K_3 correspondant à des espaces de Hilbert de champs de vecteurs sur \mathbb{R}^2 . On définit $g : \mathcal{S}_1 \times \mathcal{S}_2 \times \mathcal{S}_3 \rightarrow \mathbb{R}$ par

$$g(q_1, q_2, q_3^1, q_3^2) = \sum_{i=1}^2 d(q_i, q_{(i)})^2 + d(q_3^i, q_{(i)}),$$

où les $q_{(i)} \in \mathcal{S}_i$ sont des courbes cibles fixées.

Les courbes q_1 et q_2 représentent le bord de domaines respectifs $U(q_1)$ et $U(q_2)$ dans \mathbb{R}^2 , tandis que $q_3 = (q_3^1, q_3^2)$ représente celui du milieu continu $(U(q_1) \cup U(q_2))^c$ dans lequel on les suppose plongés. On veut conserver cette configuration le long d'une déformation $t \mapsto (q_1(t), q_2(t), q_3(t))$ induite par des flots de champs de vecteurs $X_i \in L^2(0, 1; V_i)$, $i = 1, 2, 3$.

On regarde deux types de contraintes : les contraintes collées, qui s'écrivent $q_i(t) = q_3^i(t)$, $i = 1, 2$, $t \in [0, 1]$, et les contraintes glissées, qui signifient que $q_3^i(t)$ se déduit de $q_i(t)$ par une reparamétrisation $f_i(t) \in \text{Diff}(S^1)$ dépendant du temps : $q_3^i(t) = q_i(t) \circ f_i(t)$, $i = 1, 2$, $t \in [0, 1]$. Ces contraintes sont équivalentes au fait que

$$\forall x \in \text{im}(q_i(t)), \quad t \in [0, 1], \quad (X_i(t, x) - X_3(t, x))^T \nu_i(t, x) = 0, \quad i = 1, 2,$$

où $\nu_i(t, x)$ est la normale à $\text{Im}(q_i(t))$ en x .

Pour l'implémentation numérique, on a discrétisé q_i en $(x_{i,1}, \dots, x_{i,n}) \in (\mathbb{R}^2)^n$ pour $i = 1, 2$, et q_3 en $(x_3^{1,1}, \dots, x_3^{1,n}; x_3^{2,1}, \dots, x_3^{2,n}) \in (\mathbb{R}^2)^{2n}$. On peut alors calculer la normale $\nu_{i,k}$ approchée en $x_{i,k}$ à la courbe discrétisée q_i en prenant un vecteur unitaire orthogonal à $x_{i,k+1} - x_{i,k-1}$.

La figure 7.3 montre la configuration initiale en bleu et la cible en rouge. La figure 7.4 montre le résultat final, après avoir appliqué un algorithme de minimisation par lagrangien augmenté. Les trois difféomorphismes correspondant sont également représentés. Cette méthode se généralise facilement au cas de surfaces triangulées dans \mathbb{R}^3 .

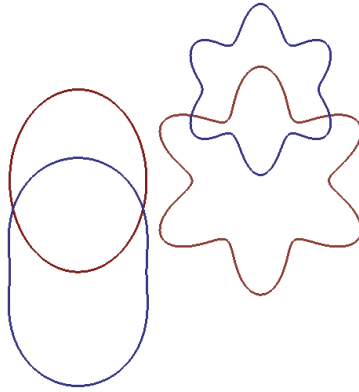


FIGURE 7.3 – Multiformes : états initiaux (en bleu) et cible (en rouge).



FIGURE 7.4 – Multiformes : contraintes collées (à gauche) et glissées (à droite).

7.4 Un algorithme de quasi-Newton

Les résultats de cette section sont prouvés dans [Arg13], reproduit dans le chapitre [].

Le problème principal au niveau de la complexité de calcul de la méthode du tir est la résolution du système linéaire permettant d'obtenir le multiplicateur de Lagrange associé aux contraintes, qui est de la forme

$$A\lambda = c, \quad \lambda, c \in \mathbb{R}^k,$$

avec A une matrice symétrique définie positive. Le but de cette section est de donner une méthode plus rapide permettant de résoudre ce système de manière approximative à chaque étape de gradient.

La méthode est inspirée des algorithmes dits de quasi-Newton. Une façon efficace de minimiser une fonction $f : \mathbb{R}^d \rightarrow \mathbb{R}$ de classe \mathcal{C}^2 est de construire par la méthode de Newton une suite $(x_k)_{k \in \mathbb{N}}$ par la relation de récurrence

$$x_{k+1} = x_k - h_k H(f)^{-1}(x_k) \nabla f(x_k),$$

où $H(f)$ est la hessienne de f , ∇f son gradient, et $h_k > 0$ est le pas. Toutefois, des cas où le calcul de $H(f)$ est trop compliqué ont conduit à introduire des algorithmes dits de *quasi-Newton* : il s'agit de construire x_{k+1} par la relation de récurrence

$$x_{k+1} B_k^{-1} \nabla f(x_k),$$

où (B_k) est une suite de matrices symétriques définies positives de taille $d \times d$ telles que

$$B_{k+1}(x_{k+1} - x_k) = \nabla f(x_{k+1}) - \nabla f(x_k),$$

de telle sorte que

$$B_{k+1}(x_{k+1} - x_k) \simeq H(f)(x_k)(x_{k+1} - x_k).$$

Il existe de nombreuses façons de construire de manière simple les matrices B_k [DM77, DS96, GMW81]. Toutefois, il a été prouvé dans [CGT91, Sch74] que certaines de ces méthodes permettent de construire une suite de matrices (B_k) qui approchent la suite de matrices $H(f)(x_k)$ dans toutes les directions plutôt que dans une seule (la direction $(x_{k+1} - x_k)$), si on suppose en plus que la suite des

$$s_k = x_{k+1} - x_k$$

est *uniformément linéairement indépendante* (voir définition 7.1). Par exemple, en posant

$$y_k = A_k s_k, \quad r_k = y_k - B_k s_k, \quad B_{k+1} = B_k + \frac{r_k r_k^T}{r_k^T s_k}, \quad (7.1)$$

avec

$$A_k = \int_0^1 H(f)(x_k + t s_k) dt,$$

on obtient, sous la condition que la suite (x_k) converge vers un certain \bar{x} , que (B_k) a la même limite que (A_k) , c'est-à-dire $H(f)(\bar{x})$ [CGT91].

Toutefois, la suite (B_k) peut être définie pour toutes suites (A_k) de matrices symétriques et (s_k) de vecteurs dans \mathbb{R}^d . Notre résultat montre alors que, sous certaines hypothèses, la suite (B_k) obtenue converge vers la même limite que la suite (A_k) .

7.4.1 Cadre, définitions et énoncé du résultat

Commençons par introduire le cadre de notre problème et quelques notations. On définit d'abord une suite $(A_k)_{k \in \mathbb{N}}$ de matrices symétriques et inversibles de taille $d \times d$. On suppose que cette suite converge vers une limite A_* , également symétrique et inversible. On note alors pour $k \in \mathbb{N}$

$$\eta_{k,*} = \sup_{i \geq k} \|A_i - A_k\|,$$

où $\|\cdot\|$ est la norme d'opérateur.

Soit également $(s_k)_{k \in \mathbb{N}}$ une suite de vecteurs de \mathbb{R}^d . On fixe B_0 une matrice carrée de taille $d \times d$ symétrique définie positive, et pour $k \in \mathbb{N}$, on définit

$$y_k = A_k s_k, \quad r_k = (A_k - B_k) s_k = y_k - B_k s_k,$$

où (B_k) est calculée par la méthode dite de *mise à jour de rang 1*, qui correspond à la récurrence

$$B_{k+1} = B_k + \frac{r_k r_k^T}{r_k^T s_k},$$

avec $r_k^T s_k \neq 0$ pour tout k .

Définition 7.1. *La suite $(s_k)_{k \in \mathbb{N}}$ est dite (m, β) -uniformément linéairement indépendante, pour $m \geq d \in \mathbb{N}^*$ et $\beta > 0$, si*

$$\forall k \in \mathbb{N}, \exists k+1 \leq k_1 < \dots < k_d \leq k+m, \quad \sigma \left(\frac{s_{k_1}}{|s_{k_1}|}, \dots, \frac{s_{k_d}}{|s_{k_d}|} \right) \geq \beta,$$

où $\sigma(M)$ est la plus petite valeur singulière d'une matrice M .

On rappelle que les valeurs singulières d'une matrice M sont les racines des valeurs propres de $M^T M$.

Remark 7.4. Cela signifie que pour tout k , on peut trouver $k+1 \leq k_1 < \dots < k_d \leq k+m$ tels que le volume du parallélépipède formé par les $\frac{s_{k_i}}{|s_{k_i}|}$ soit minoré par une constante positive indépendante de k .

Notre résultat est le suivant.

Théorème 7.1. *Supposons qu'il existe $c > 0$ tel que*

$$\forall k \in \mathbb{N}, \quad |r_k^T s_k| \geq c |r_k| |s_k|.$$

Alors, pour tout $(m, \beta) \in \mathbb{N} \times \mathbb{R}_+^$ tel que (s_k) soit (m, β) -uniformément linéairement indépendante, on a*

$$\forall k \in \mathbb{N}, \quad \|B_{k+m} - A_*\| \leq \left(1 + \left(\frac{2+c}{c} \right)^{m+1} \right) \frac{\sqrt{d}}{\beta} \eta_{k,*}.$$

En particulier, B_k tend vers A_ quand k tend vers l'infini.*

7.4.2 Application au calcul d'une suite d'inverses

On peut utiliser cet algorithme pour approcher une suite d'inverses de matrices (A_k^{-1}) . Il suffit pour cela de poser

$$s_k = A_k e_{k \bmod d}, \quad y_k = A_k^{-1} s_k = e_{k \bmod d},$$

où les (e_0, \dots, e_{d-1}) est la base canonique de \mathbb{R}^d , et $k \bmod d$ est le reste de la division euclidienne de k par d .

On peut alors utiliser la méthode de mise à jour de rang 1 pour ces deux suites pour obtenir une suite (B_k) de matrices symétriques qui converge vers $(A_*)^{-1}$. En effet, la suite (s_k) sera bien (d, β) -uniformément linéairement indépendante à partir d'un certain rang, puisque la famille $(s_{k+1}, \dots, s_{k+d})$ convergera vers $(A_*e_0, \dots, A_*e_{d-1})$, qui est une base de \mathbb{R}^d . On peut donc appliquer le théorème 7.1 et montrer que B_k converge bien vers A_*^{-1} quand k tend vers l'infini.

Dans [Arg13], reproduit dans le chapitre 11, de nombreuses simulations numériques ont été effectuées pour tester l'efficacité de cet algorithme. Il en ressort qu'il fonctionne extrêmement bien dans le cas où la matrice limite A_* est bien conditionnée, mais peut nécessiter de nombreuses itérations dans le cas contraire.

On peut utiliser cette variante pour calculer plus rapidement des valeurs approchées du multiplicateur de Lagrange. Les détails sont donnés en section 11.7.

Chapitre 8

Conclusion et perspectives

Dans cette thèse, on a d'abord défini la géométrie sous-riemannienne en dimension infinie dans un cadre très général (distribution horizontale non fermée, et équivalent d'une distribution à rang variable), on a obtenu plusieurs résultats sur l'équation géodésique. On a également donné une idée de la difficulté du sujet par rapport au cas sous-riemannien de dimension finie et au cas riemannien de dimension infinie. On a également procédé à une étude complète des structures sous-riemanniennes fortes sur le groupe des difféomorphismes $\mathcal{D}^s(M)$ d'une variété (M, g) à géométrie bornée, pour lesquelles l'équation géodésique a une forme particulièrement simple si on utilise le noyau reproduisant du sous-espace de champs de vecteurs qui engendre la structure.

On a ensuite donné une définition des espaces de formes qui englobe les exemples connus dans la littérature, unifiant et généralisant les méthodes usuelles. On a de plus construit des structures sous-riemanniennes particulières sur ces espaces de formes, les structures relevées, qui promettent de nouvelles applications. Enfin, on a ajouté des contraintes dépendant de l'état sur l'ensemble des transformations possibles, prouvé un nouveau principe du maximum de Pontryagin adapté à ce problème, et développé des méthodes LDDMM à ce nouveau cadre, elles-mêmes appuyées par des simulations numériques sur des exemples ayant de nombreuses applications potentielles.

Ce travail de thèse débouche sur les perspectives suivantes.

Étude des structures sous-riemanniennes fortes

Soit (\mathcal{H}, ξ, g) une structure sous-riemannienne forte sur une variété de Hilbert M . Plusieurs questions fondamentales peuvent se poser.

Comportement de la distance au voisinage d'un point. Il s'agit de décrire l'allure de l'orbite \mathcal{O}_q au voisinage de $q \in M$. Pour mieux situer la question, rappelons qu'en dimension finie, le théorème de la boule emboîtée nous donne une bonne description de la distance sous-riemannienne d à un point : autour de chaque point q_0 , il existe $C, C' > 0$ et des coordonnées

$$q = \left(\underbrace{x_1^1, \dots, x_1^{k_1}}_{x^1}, \dots, \underbrace{x_r^1, \dots, x_r^{k_r}}_{x^r} \right)$$

centrées en $q_0 = 0$ telles que

$$C \sum_{i=1}^r |x_i|^{2/i} \leq d(q_0, q)^2 \leq C' \sum_{i=1}^r |x_i|^{2/i}.$$

La première partie de l'inégalité se prouve heuristiquement de la façon suivante : on prend X, Y des champs de vecteurs horizontaux et normés, de flots φ^X, φ^Y , et on note $Z = [X, Y]$.

Alors pour $t \geq 0$,

$$\varphi^{-Y}(\sqrt{t}) \circ \varphi^{-X}(\sqrt{t}) \circ \varphi^Y(\sqrt{t}) \circ \varphi^X(\sqrt{t})(x) \underset{t \rightarrow 0}{=} tZ(q) + o(t).$$

Intuitivement, en coordonnées, on peut donc joindre q et $q + tZ(q)$ par une courbe horizontale de longueur $4\sqrt{t}$. En itérant le procédé, on voit que si $Z_i = [X_i, Y_i]$, $i = 1, \dots, l$, avec les X_i et Y_i horizontaux et normés, on peut joindre q et $q + t_1Z_1(q) + \dots + t_lZ_l(q)$ par une courbe de longueur $4l \sum_{i=1}^l \sqrt{t_i}$, et en déduire le théorème de la boule emboîtée si $r = 2$. On peut généraliser très facilement ce raisonnement pour r plus grand.

On a vu que dans le cas d'un produit infini de groupes de Carnot, ce résultat reste vrai au voisinage de l'origine. Toutefois, dans un tel espace, la distribution horizontale Δ est engendrée par des champs de vecteurs X tels que $\dim([X, \Delta]/\Delta) \leq \ell$, où $\ell \in \mathbb{N}$ fixé. On voit facilement que la méthode décrite devrait encore marcher, avec quelques subtilités. Mais qu'en est-il sur des variétés sous-riemanniennes plus générales ?

Dans ce cas, on doit donc a priori effectuer un raisonnement similaire avec une infinité de Z_i . En conséquence, en suivant la même méthode, pour pouvoir atteindre

$$q + \sum_{i=0}^{\infty} t_i Z_i(q)$$

avec une courbe de longueur finie, il faut que

$$\sum_{i=1}^{\infty} \sqrt{t_i} < +\infty \implies (t_i)_{i \in \mathbb{N}} \in \ell^{1/2}(\mathbb{N}, \mathbb{R}).$$

La méthode ne donne donc pas un résultat optimal (appliquée telle quelle au produit infini de groupes de Heisenberg, on obtient seulement $\ell^2(\mathbb{N}, \mathbb{R}^2) \times \ell^{1/2}(\mathbb{N}, \mathbb{R}) \subset \mathcal{O}_0$). Il faut donc développer de nouveaux outils pour répondre à ce problème.

Recherche du "bon" espace cotangent pour la formulation d'un principe du maximum de Pontryagin. Rappelons que l'apparition de géodésiques élusives (au sens du lemme 2.1) est due à un défaut de multiplicateurs de Lagrange : comme on l'a vu en remarque 2.6, dans le cas sous-riemannien fort, l'espace cotangent n'est pas trop gros (contrairement au cas faible, qui nécessite de trouver un sous-fibré adapté), mais trop petit : en restreignant la structure à une sous-variété dense et stable de l'espace ambiant, on obtient de nouveaux covecteurs initiaux qui "normalisent" certaines des géodésiques élusives. Cela amène naturellement à se demander s'il existe un espace cotangent suffisamment gros pour obtenir tous les multiplicateurs de Lagrange possibles, c'est-à-dire pour lequel il n'existe pas de géodésique élusive.

Dans un produit infini de groupes de Heisenberg, la distance sous-riemannienne sur l'orbite de 0 est équivalente à la distance du sous-espace dense $\ell^2(\mathbb{N}, \mathbb{R}^2) \times \ell^1(\mathbb{N}, \mathbb{R}) \subset \ell^2(\mathbb{N}, \mathbb{R}^3)$. On voit alors aisément qu'une fois que l'on se restreint à ce sous-espace, dont le cotangent en 0 est $\ell^2(\mathbb{N}, \mathbb{R}^2) \times \ell^\infty(\mathbb{N}, \mathbb{R})$, toutes les géodésiques sont normales : $\ell^2(\mathbb{N}, \mathbb{R}^2) \times \ell^\infty(\mathbb{N}, \mathbb{R})$ est le "bon" espace cotangent pour cette structure sous-riemannienne.

En revanche, dans le cas d'un produit infini de groupe de Engel, la distance (et l'orbite) est équivalente à celle d'un espace vectoriel topologique qui n'est pas localement convexe, et la même méthode ne fonctionne pas, car le dual topologique d'espaces non localement convexes est en général "trop petit" par rapport à ce qu'on voudrait (Il en existe d'ailleurs dont le dual topologique est nul [Day40]).

Groupe nilpotent libre d'ordre 2 sur un espace de Hilbert Pour commencer à répondre partiellement à ces deux questions, on pourrait pour commencer étudier des exemples simples mais moins triviaux. Par exemple, $M = H \times AS(H)$, $(H, \langle \cdot, \cdot \rangle)$ un espace de Hilbert séparable et $AS(H)$ les opérateurs anti-symétriques de Hilbert-Schmidt sur H , muni de la structure $(M \times H, \xi, \langle \cdot, \cdot \rangle)$ avec

$$\xi_{(q,A)}u = (u, \frac{1}{2}q \wedge u).$$

Il s'agit d'une structure invariante à gauche pour la loi de groupe de Lie

$$(q, A) \cdot (q', A') = (q + q', A + A' + \frac{1}{2}q \wedge q').$$

C'est une extension à la dimension infinie des groupes nilpotents libres d'ordre 2 [Mon02, Chapitre 6], étudiés en détails en dimension finie dans [MPAM06].

Applications à l'analyse mathématique des formes

Réduction des équations géodésiques sur un espace de formes. Un autre sujet que nous n'avons pas abordé, et qui est très naturel à étudier, est celui de la réduction des équations géodésiques dans le cas où il existe des symétries dans la structure. Un exemple simple est celui d'un espace de courbes dans \mathbb{R}^d : on prend S^1 le cercle unité, et $\mathcal{S} = \text{Emb}^1(S^1, \mathbb{R}^d)$ l'ensemble des arcs paramétrés \mathcal{C}^1 , réguliers, fermés, et simples dans \mathbb{R}^d . On considère $(V, \langle \cdot, \cdot \rangle)$ un sous-espace de champs de vecteurs avec inclusion continue dans $H^{s+2}(\mathbb{R}^d, \mathbb{R}^d)$ avec $s > d/2 + 1$, et on munit \mathcal{S} de la structure sous-riemannienne habituelle $(\mathcal{S} \times V, \xi, \langle \cdot, \cdot \rangle)$, où pour $q \in \mathcal{S}$ et $X \in V$, $\xi_q X = X \circ q \in \mathcal{C}^1(S^1, \mathbb{R}^d)$.

Alors, pour tout difféomorphisme ψ de classe \mathcal{C}^1 de S^1 , la transformation $q \mapsto q \circ \psi$ est une isométrie pour la structure sous-riemannienne. Cela signifie, essentiellement, que cette transformation envoie toute courbe horizontale sur une courbe horizontale de même longueur et de même action. Elle conserve donc la distance sous-riemannienne. De plus, on voit aisément que le groupe $\text{Diff}^1(M)$ des difféomorphismes \mathcal{C}^1 de S^1 est une variété de Banach et un groupe topologique dont l'espace tangent en l'identité est $\Gamma^1(TS^1)$.

Pour $q \in \mathcal{S}$, $p \in \mathcal{M}(S^1) \otimes (\mathbb{R}^d)^* \subset T_q^* \mathcal{S}$, définissons $\eta(q, p) \in \Gamma^1(TS^1)^*$ par

$$\eta(q, p)Y = \int_{S^1} \partial_s q(s) Y(s) dp(s).$$

Alors on peut voir ([MM07, TY11]) que si $t \mapsto (q(t), p(t))$ satisfait l'équation hamiltonienne des géodésiques avec $p(0) \in \mathcal{M}(S^1) \otimes (\mathbb{R}^d)^*$, alors, pour tout t , $p(t) \in \mathcal{M}(S^1) \otimes (\mathbb{R}^d)^*$ et

$$\eta(q(t), p(t)) = \eta(q(0), p(0)).$$

Notons qu'il s'agit essentiellement du théorème de Noether sur les systèmes mécaniques à symétries. On obtient ainsi des lois de conservation le long des géodésiques, ce qui permet de mieux comprendre leur structure, et d'étudier leur caractère intégrable au sens de Liouville.

En particulier, pour minimiser

$$\frac{1}{2} \int_0^1 \langle X(t), X(t) \rangle dt + g(\varphi^X \circ q_0),$$

lorsque $g(q \circ \psi) = g(q)$ pour tous $q \in \mathcal{S}$, $\psi \in \text{Diff}^1(M)$, on peut voir que le covecteur initial $p(0)$ de la géodésique est tel que $\eta(q_0, p(0)) = 0$. Cela permet de restreindre l'espace dans lequel il faut choisir le moment initial dans la méthode du tir.

Formes contraintes et espaces de multiformes. La puissance des méthodes LDDMM est de produire de façon générique des structures (sous)-riemanniennes induites sur des espaces de formes très variés avec un minimum d'a priori. La construction de cartes exponentielles autour d'un état initial fixé permet ensuite le déploiement de systèmes de coordonnées pour la modélisation de la variabilité sur lequel une approche statistique va pouvoir par exemple estimer une densité de probabilités ou des approches de type machine learning pour construire des fonctions de discriminations.

Cependant cette approche ne considère pas la nature physique des formes sous-jacentes mais plutôt de l'espace qui les contient, et n'est pas toujours un point de départ suffisant pour construire un système de représentation qui soit assez proche des données. L'introduction de contraintes comme au chapitre 6 permet de *conditionner* un modèle dans une situation particulière par des connaissances a priori. Nous en avons donné un exemple très simple par la prise en compte d'une contrainte globale de volume constant sur les courbes horizontales, qui génère des trajectoires géodésiques très différentes du cas sans contraintes. Le conditionnement sur une courbe d'englober un volume prescrit tout au long de sa déformation peut d'ailleurs être très utile dans l'étude de phénomènes d'évolution dans les situations où un même patient est observé sur plusieurs échelles de temps dans des modélisations de phénomènes de croissance. D'autres contraintes locales décrivant des contraintes sur l'évolution d'un objet au cours du temps peuvent être prises en comptes, comme par exemple des contraintes isométriques entre certains points de la forme. De tels constructions ne peuvent être obtenues par les modèles génériques invariants à droite. Enfin la situation des multi-objets que nous avons décrite dans les chapitres 6 et 7 montre la richesse potentielle de telles approches en terme de modélisation pour l'étude de complexes de formes, qui sont pourtant des situations courantes en anatomie computationnelle pour étudier simultanément plusieurs parties du cerveau lors de processus neurodégénératifs. L'optimisation du choix des contraintes dans une situation particulière à partir d'une base de données reste un sujet encore très largement ouvert qui pourrait passer par la compréhension de la géométrie des modèles induits.

De plus, on n'a pas étudié le cas de contraintes non-linéaires, ni de contraintes d'inégalités. Ces dernières tout particulièrement mèneraient à encore plus d'applications intéressantes, comme imposer au volume de courbes de diminuer au cours de la transformation plutôt que de rester constant, afin de mieux modéliser des phénomènes de dégénération du cerveau d'un patient atteint de certaines maladies neurologiques. Il s'agit toutefois (pour le moins en dimension infinie) d'un cas très difficile à étudier, bien plus que les contraintes d'égalités.

Un exemple d'application des structures relevées : modélisation d'un muscle.

Les formes relevées sont très utiles pour la modélisation de mouvement de muscles. Plus précisément, on peut actuellement étudier un tel mouvement de deux façons. La première utilise un point de vue dynamique, et consiste à modéliser avec précision la façon dont le muscle bouge par un système d'équations aux dérivées partielles. Il s'agit d'un processus extrêmement compliqué, développé dans le cas du cœur dans [BCF⁺13] par exemple. Les algorithmes nécessaires sont très coûteux, et il est difficile d'utiliser de tels modèles pour une analyse statistique comparant le mouvement du muscle de différents patients. La deuxième manière est de procéder par une méthode classique d'analyse de formes (comme la méthode LDDMM). Toutefois, les mouvements obtenus ne cadrent alors pas du tout à ceux observés en pratiques, car aucun des aspects dynamiques du mouvement à modéliser n'est pris en compte. Pour remédier à ce problème, on peut utiliser des structures relevées.

En effet, pour discrétiser une forme (notre muscle) dans \mathbb{R}^d , on prend un ensemble fini de points $q = (x_1, \dots, x_n) \in (\mathbb{R}^d)^n = \mathcal{S}$. Le groupe $\mathcal{D}^s(\mathbb{R}^d)$, $s > d/2 + 1$ agit diagonalement sur \mathcal{S} par

$$\varphi \cdot \hat{q} = (\varphi(x_1), \dots, \varphi(x_n)).$$

Soit $(V, \langle \cdot, \cdot \rangle)$ un espace de Hilbert de champs de vecteurs de classe H^{s+2} et K son noyau reproduisant. Les équations hamiltoniennes pour la structure sous-riemannienne associée sur \mathcal{S} sont données sur une courbe

$$t \mapsto (q(t), p(t)) = (x_1(t), \dots, x_n(t), p_1(t), \dots, p_n(t)) \in T^*\mathcal{S} = (\mathbb{R}^d)^n \times ((\mathbb{R}^d)^*)^n$$

par

$$\begin{cases} \dot{x}_i(t) = \sum_{j=1}^n K(x_i(t), x_j(t)) p_j(t), & i = 1, \dots, n, \\ \dot{p}_i(t) = \sum_{j=1}^n p_j(t) \partial_1 K(x_i(t), x_j(t)) p_j(t), & i = 1, \dots, n. \end{cases}$$

Pour modéliser un muscle, on va alors ajouter en chaque x_i un vecteur w_i représentant la direction de la fibre musculaire :

$$\hat{q} = (q, w) = (x_1, \dots, x_n, w_1, \dots, w_n) \in \hat{\mathcal{S}} = T(\mathbb{R}^d)^n \simeq (\mathbb{R}^d)^n \times (\mathbb{R}^d)^n.$$

Le groupe $\mathcal{D}^{s+1}(\mathbb{R}^d)$ agit sur $\hat{\mathcal{S}}$ par

$$\varphi \cdot \hat{q} = (\varphi(x_1), \dots, \varphi(x_n), d\varphi(x_1)w_1, \dots, d\varphi(x_n)w_n).$$

L'action infinitésimale d'un champ de vecteur $X \in H^s(\mathbb{R}^d, \mathbb{R}^d)$ est donc

$$\xi_{\hat{q}} X = (X(x_1), \dots, X(x_n), dX(x_1)w_1, \dots, dX(x_n)w_n).$$

Si on relève la géodésique $t \mapsto (q(t), p(t))$ ci-dessus à $\hat{\mathcal{S}}$, on obtient

$$\dot{w}_i(t) = \sum_{j=1}^n \partial_1 (K(x_i(t), x_j(t)) p_j(t)) w_i, \quad i = 1, \dots, n.$$

Or, un muscle agit sur lui-même en contractant ses fibres, c'est-à-dire en exerçant une *force* en x_i dans la direction w_i . Cela se traduit par l'ajout d'un terme parallèle à w_i^T dans la dérivée du covecteur p_i :

$$\dot{p}_i(t) = \sum_{j=1}^n p_j(t) \partial_1 K(x_i(t), x_j(t)) p_j(t) + u_i(t) w_i(t)^T, \quad u_i(t) \in \mathbb{R}, \quad i = 1, \dots, n.$$

On modélise finalement le mouvement d'un muscle par le système de contrôle

$$\begin{cases} \dot{x}_i(t) = \sum_{j=1}^n K(x_i(t), x_j(t)) p_j(t), & i = 1, \dots, n, \\ \dot{p}_i(t) = \sum_{j=1}^n p_j(t) \partial_1 K(x_i(t), x_j(t)) p_j(t) + u_i(t) w_i(t)^T, & i = 1, \dots, n, \\ \dot{w}_i(t) = \sum_{j=1}^n \partial_1 (K(x_i(t), x_j(t)) p_j(t)) w_i, & i = 1, \dots, n. \end{cases}$$

Ce modèle, inspiré d'une remarque de [TV12], pourrait permettre d'étudier, par exemple, le battement d'un cœur. Il permet en effet une meilleure modélisation du mouvement au cours du temps, tout en restant simple et peu coûteux du point de vue algorithmique.

Géométries sous-riemanniennes lisses sur un groupes de difféomorphismes et application à la recherche de solutions d'équations conservatives

Structures invariantes à droites lisses et faibles. Il reste plusieurs problèmes que nous n'avons pas abordés concernant les structures sous-riemanniennes invariantes à droite sur le groupe des difféomorphismes d'une variété à géométrie bornée (M, g) . Par exemple, si le sous-espace de champs de vecteurs V a une inclusion continue dans $\Gamma^{s+k}(M)$, $k \in \mathbb{N}^*$, il semble plus logique d'étudier la structure sous-riemannienne forte et invariante à droite induite sur $D^{s+k}(M)$ plutôt que celle sur $\mathcal{D}^s(M)$. Le problème est que cette structure est seulement continue, et l'équation hamiltonienne des géodésiques n'a donc pas forcément de solution. Nous avons également négligé le cas des structures invariantes à droite *faibles*, qui est plus compliqué, car il peut être difficile de trouver un espace cotangent relatif adapté qui soit lisse.

Toutefois, ces deux problèmes peuvent être abordés dans le cas où :

1. V est engendré par des champs de vecteurs $X_1, \dots, X_k \in \Gamma^s(TM)$ de classe \mathcal{C}^∞ et bornés à tout ordre, comme dans la section 2.9. Autrement dit,

$$V = \left\{ \sum_{i=1}^k u_i X_i, u_i \in H^s(M, \mathbb{R}) \right\}.$$

Dans ce cas, V est un sous-espace fermé de $\Gamma^s(TM)$. On définit alors le sous-fibré $\mathcal{H}' \subset T\mathcal{D}^s(M)$ par

$$\mathcal{H}'_\varphi = \left\{ \sum_{i=1}^k u_i X_i \circ \varphi, u_i \in H^s(M) \right\}.$$

En prenant l'inclusion pour $\xi' : \mathcal{H}' \rightarrow T\mathcal{D}^s(M)$, on obtient alors un espace tangent relatif de classe \mathcal{C}^∞ . De plus, l'application $\mathcal{H} = \mathcal{D}^s(M) \times V \rightarrow \mathcal{H}'$ définie par $X \mapsto X \circ \varphi$ est un isomorphisme continu d'espaces tangents relatifs.

2. On remplace la norme de Hilbert sur V par la norme pré-hilbertienne

$$\langle X, X \rangle_{L_r} = \int_M g(L_r(X)(x), X(x)) dx,$$

où $L_r : \Gamma^s(TM) \rightarrow \Gamma^{s-2r}(TM)$ est un opérateur différentiel linéaire, symétrique, elliptique à coefficients lisses et uniformément bornés à tout ordre. On a une structure forte lorsque $r = s$. Sinon la structure sera faible. On définit alors, pour $(\varphi, X) \in T\mathcal{D}^s(M)$,

$$g_{\varphi, L_r}(X, X) = \int_M g_x(L_r(X \circ \varphi^{-1})(x), X \circ \varphi^{-1}(x)) dx.$$

Le changement de variable $x = \varphi(y)$ montre que g est en fait de classe \mathcal{C}^∞ .

Dans ce cas, $(\mathcal{H}', \xi', g_{L_r})$ est une structure sous-riemannienne sur $\mathcal{D}^s(M)$ équivalente à la structure sous-riemannienne invariante à droite engendrée par $(V, \langle \cdot, \cdot \rangle_{L_r})$.

On peut alors travailler avec cette nouvelle structure pour appliquer les théorèmes de la section 2. Un candidat d'espace cotangent en e adapté serait $\Gamma^{s-2r}(T^*M)$, comme dans la section 9.1.5. En plus de s'inscrire naturellement dans la continuation de nos travaux, les structures faibles invariantes à droite sur des groupes de difféomorphismes ont des applications naturelles dans la recherche de solution d'équations conservatives (équation d'Euler des fluides parfaits, équations de Burgers, KdV...), voir [EM70, BV13, MP10] et leurs références par exemple. Nous donnons un exemple d'application possible au paragraphe suivant.

Fluides incompressibles dans des variétés sous-riemanniennes. L'équation des fluides incompressibles d'Euler sur une variété compacte orientable (M, g) est donnée par

$$\begin{cases} \dot{X}(t) + \nabla_{X(t)}^g X(t) = \nabla^g p(t), \\ \operatorname{div} X(t) = 0, \end{cases} \quad (8.1)$$

où $X(t) \in \Gamma^s(TM)$, $p : M \rightarrow \mathbb{R}$ et ∇^g est la connection de Levi-Civita. On sait depuis les travaux d'Arnold [Arn66], puis d'Ebin et Marsden [EM70], que c'est la forme réduite de l'équation des géodésiques sur $\mathcal{D}^s(M)$ (i.e. l'équation des moments, voir section 3.5.4) pour la métrique riemannienne faible

$$g_{\varphi,1}(X, X) = \int_M g_{\varphi(x)}(X(x), X(x)) dx, \quad X \in T_\varphi \mathcal{D}^s(M),$$

sous la contrainte d'état pure

$$\operatorname{div}(X \circ \varphi^{-1}) = 0.$$

Cette condition, qui est continue seulement en φ , devient lisse lorsqu'on la remplace par la condition équivalente

$$\operatorname{div}(X \circ \varphi^{-1}) \circ \varphi = 0.$$

La structure

$$(\mathcal{H}', \xi', g_{\varphi,1}), \quad \mathcal{H}'_\varphi = \ker C_\varphi, \quad \xi'_\varphi(X) = X, \quad \varphi \in \mathcal{D}^s(M),$$

se restreint à $\mathcal{D}_{\text{vol}}^s(M)$, le groupe des difféomorphismes qui préservent la forme volume. Cette restriction est équivalente à la structure invariante à droite sur $\mathcal{D}_{\text{vol}}^s(M)$ engendrée par l'espace pré-hilbertien $(V, g_{\text{Id}_M,1})$, où

$$V = \ker \operatorname{div} \cap \Gamma^s(TM),$$

de la même manière qu'on a vu que g_{L_r} était lisse dans la discussion précédente, et ce car φ préserve la forme volume. Ebin et Marsden ont ensuite prouvé que cette métrique admettait une connection de Levi-Civita et un flot géodésique, ce qui a montré l'existence locale et l'unicité des solutions à l'équation d'Euler pour toute condition initiale $X(0) \in \Gamma^s(TM)$, $s > \dim(M)/2 + 1$.

Il est alors naturel de se demander comment se comporte un fluide incompressible sur une variété sous-riemannienne. Un des développements possibles de cette thèse est la démonstration d'un résultat similaire d'existence et d'unicité des solutions à une telle équation. Considérons $(\mathcal{H}^0, \xi^0, g^0)$, une structure sous-riemannienne sur une variété M compacte de dimension finie. Supposons, pour simplifier, que \mathcal{H}^0 soit trivialisable, et soit alors (e_1, \dots, e_k) un repère orthonormé global de \mathcal{H}^0 . Soit également $X_i : x \mapsto \xi_x^0 e_i(x)$, champ de vecteurs horizontal de classe \mathcal{C}^∞ sur M .

On considère, comme dans la section 2.9, la structure sous-riemannienne (\mathcal{H}, ξ, g) sur $\mathcal{D}^s(M)$ définie par

$$\mathcal{H} = \mathcal{D}^s(M) \times H^s(M, \mathbb{R}^k), \quad \xi_\varphi(u) = \sum_{i=1}^k u_i X_i \circ \varphi, \quad g(u, u) = \sum_{i=1}^k \int_M u_i^2(x) dx, \quad (\varphi, u) \in \mathcal{H}.$$

On rajoute les contraintes $C_\varphi : H^s(M, \mathbb{R}^k) \rightarrow H^{s-1}(M, \mathbb{R})$ définies par

$$C_\varphi u = \operatorname{div}((\xi_\varphi u) \circ \varphi^{-1}) \circ \varphi.$$

Cette structure, de classe \mathcal{C}^∞ , se restreint à $\mathcal{D}_{\text{vol}}^s(M)$ en une structure équivalente à une structure sous-riemannienne faible et invariante à droite.

L'objectif est alors de trouver un espace cotangent relatif adapté sur lequel le hamiltonien possède un gradient symplectique suffisamment régulier, et expliciter ce gradient. Cela peut

se faire en étendant les résultats de la section 6 au cas des métriques faibles. Le but serait alors d'écrire la forme réduite de l'équation associée pour obtenir l'équation d'Euler des fluides incompressibles sur une variété sous-riemannienne, ainsi qu'un espace de conditions initiales pour lesquelles elle admet une unique solution locale. En plus d'être un problème très naturel de géométrie sous-riemannienne, ces considérations ouvrent naturellement la voie à des considérations de fluides incompressibles à mouvements contraints.

Équation d'Euler avec contraintes. Il s'agit d'ajouter au fluide des contraintes de la forme

$$C_{\varphi^X(t)} \dot{\varphi}^X(t) = 0, \quad p.p. t,$$

où φ^X représente le mouvement du fluide, c'est-à-dire le flot du champ de vitesse de déplacement X du fluide sur la variété M . On obtient alors le mouvement d'un fluide contraint en cherchant l'équation des géodésiques *contraintes* sur $\mathcal{D}^s(M)$ pour la norme

$$\int \left(\int_M g_x(X(x), X(x)) dx \right) dt$$

pour $X(t) \in V = \ker \operatorname{div} \cap \Gamma^s(TM)$ en combinant les résultats des chapitres 3 et 6, à condition bien sûr de les généraliser au cas où $\langle \cdot, \cdot \rangle$ est une métrique pré-Hilbertienne sur V , plutôt qu'une métrique forte. Un tel point de vue ouvre de nombreuses applications potentielles.

Le cadre des multiformes, par exemple, semble particulièrement adapté à l'étude d'une interaction solide-fluide (ou même une interaction entre deux fluides inhomogènes). En effet, on peut repérer un solide compact par un domaine compact $U_0 \subset M$, plongé dans un fluide parfait repéré par le complémentaire U_0^c . On peut alors étudier le mouvement qu'une déformation du solide, modélisée par la restriction à U_0 d'une courbe $t \mapsto \varphi_1(t) \in \mathcal{D}^s(M)$, induit sur le fluide. Ce mouvement est repéré par la restriction à U_0^c d'une seconde courbe $t \mapsto \varphi_2(t) \in \mathcal{D}^s(M)$.

Par exemple, on peut imposer des contraintes collées

$$\varphi_1(t, x) = \varphi_2(t, x), \quad x \in \partial U_0$$

pour tout t , ou, plus logiquement, des contraintes glissés,

$$\varphi_1(\partial U_0) = \varphi_2(\partial U_0),$$

ce qui laisse la possibilité au fluide de glisser le long du solide.

Deuxième partie

Compilation des articles

Chapter 9

Sub-Riemannian geodesic equations on groups of diffeomorphisms and shape spaces

The contents of this chapter were presented in Chapters 3, 4 and 5.

Abstract

In this paper, we define and study strong right-invariant sub-Riemannian structures on groups $\mathcal{D}^s(M)$ of Sobolev diffeomorphisms of class H^s , of a manifold M of bounded geometry, $s > \dim(M)/2 + 1$, and their applications to mathematical shape analysis. We derive the Hamiltonian geodesic equations for normal geodesics and singular curves for such structures. We then define a shape space in M as an infinite-dimensional manifold \mathcal{S} on which the group of diffeomorphisms of M has an action that is compatible with the particular group structure of $\mathcal{D}^s(M)$, and investigate the invariant sub-Riemannian structures induced on \mathcal{S} by this action and a right-invariant sub-Riemannian structure on $\mathcal{D}^s(M)$. Then, we introduce a new kind of structure on shape spaces: sub-Riemannian that are lifted through equivariant submersions between shape spaces. We give the geodesic equations in each case, along with several examples.

Introduction

The general purpose of shape analysis is to compare several shapes while keeping track of their geometric properties. This is done by finding a deformation, mapping one shape onto the others, which minimizes a certain action that depends on the properties of the shape. This implies that an action has been assigned to every possible deformation of a shape, the design of this cost function being a crucial step in the method. This technique has been used in the analysis of anatomical organs from medical images (see [GM98]).

A recent approach represents deformations as flows of diffeomorphisms generated by time-dependent vector fields [DGM98, Tro95, Tro05]. Indeed, when considering the studied shapes as embedded in an ambient manifold M of dimension d , diffeomorphisms induce deformations of the shape itself. Such deformations preserve local (such as the smoothness) and global (such as the number of self-intersections) geometric properties of the shape. The set of all possible deformations is defined as the set of flows of time-dependent vector fields on the space of

”infinitesimal transformations”, which is a subset of the space of all vector fields on M of Sobolev class at least H^{s_0+1} , where s_0 is the smallest integer such that $s_0 > d/2$. Then, defining a pre-Hilbert norm on those vector fields that are infinitesimal transformations, the total action of the final deformation is given by the time-integral of the squared norm of the time-dependent vector field by which it is generated.

In order to define the space of infinitesimal transformations and the corresponding norm, two ways are classically considered in the existing literature.

The first, used in [BHM11, BBHM13, MM07], consists of considering C^∞ vector fields with compact support, and weak H^k norms with $k \in \mathbb{N}$. The resulting structure is then a weak right-invariant Riemannian structure on the Fréchet Lie group of smooth diffeomorphisms with compact support on M .

The second, that we will use in this paper, consists of considering a Hilbert space $(V, \langle \cdot, \cdot \rangle)$ of vector fields on M with continuous injection in the space of H^{s_0+1} -vector fields. This framework has led to the development of the so-called *Large Deformation Diffeomorphic Metric Mapping* (LDDMM) method (see [AG04, BMTY05, JM00, MTY02, MTY06]). Actually, in this framework, we will show that the group of diffeomorphisms of class H^{s_0+1} becomes a *sub-Riemannian manifold*, equipped with a right-invariant sub-Riemannian structure (provided this group can be equipped with a manifold structure compatible with its group law), under the additional assumption that M has bounded geometry (see [Sch04]). In this case, all groups of diffeomorphisms of class H^s , $s > s_0 + 1$, are both Hilbert manifolds and topological groups such that composition on the right is smooth, while composition on the left is only continuous. By projection, the shape space inherits of a sub-Riemannian structure in the sense of [ABC⁺10] thanks to the restriction of the infinitesimal action to the space of infinitesimal deformations. In the existing literature, most authors were not familiar with sub-Riemannian geometry, and the subject was actually treated from the point of view of Riemannian geometry. This was possible because of two facts: firstly, they mostly focused on finite-dimensional shape spaces (spaces of landmarks) on which the structure induced by the infinitesimal action is indeed Riemannian. Secondly, even when that was not the case, they used the Hamiltonian characterization of geodesics, which does not show that a structure might not be Riemannian, especially when the horizontal distribution is dense.

The first purpose of this paper is therefore to establish in a complete way the point of view of sub-Riemannian geometry applied to the study of shapes, and then to determine the Hamiltonian geodesic equations on the group of diffeomorphisms and on shape spaces. This approach provides an appropriate geometric point of view for shape deformation analysis and yields new interesting examples. In the last section of this paper, we will also introduce a new category of shape spaces: the lifted shape spaces, obtained by lifting the deformation on a small shape to a bigger shape space. This new approach to shape analysis explicitly requires the use of sub-Riemannian geometry.

Recall that a sub-Riemannian manifold is a triple (M, \mathcal{H}, h) , where M is a finite-dimensional manifold and (\mathcal{H}, h) is a smooth Riemannian subbundle of the tangent space TM of M , called *horizontal distribution*, equipped with a Riemannian metric h . Horizontal curves are defined as curves on M with velocity in \mathcal{H} . Their length can be defined according to the metric h , which leads to define the corresponding sub-Riemannian distance between two points. We refer the reader to [BR96, Mon02] for a survey on sub-Riemannian geometry in finite dimension. Note that, in [ABC⁺10], rank-varying horizontal distributions are considered by including horizontal distributions of the form $\xi(E)$, where E is a vector bundle on M and $\xi : E \rightarrow TM$ is a vector bundle morphism, which is rather the point of view that we need here.

Since the shape spaces we are dealing with are infinite dimensional, we need here to consider infinite-dimensional sub-Riemannian manifolds. To the best of our knowledge, sub-

Riemannian geometry has been very much unexplored in the infinite-dimensional context. Probably, the first result established in an approaching context is the controllability result on diffeomorphisms obtained in [AC09]. We also mention the recent preprint [GMV12] in which the authors establish the geodesic equations for infinite-dimensional weak sub-Riemannian geometries. In this reference, however, while the manifolds considered are modeled on more general *convenient spaces* (see [KM97]), only closed horizontal distributions with closed complement are considered, which does not include the dense distributions that usually emerge in shape deformation analysis. Whereas there is no canonical way to define the concept of sub-Riemannian geometry in infinite dimension, and whereas a systematic study of infinite-dimensional sub-Riemannian geometries is still to be done, in this paper we will define and develop all corresponding geometric notions that we think appropriate in order to model and investigate shape spaces in this context.

We will characterize the geodesics, which are critical points of the action A between two end-points. Contrarily to the finite-dimensional setting where any minimizer of the action is a critical point, in infinite dimension a serious difficulty emerges, due to the fact that the range of the differential of the end-point mapping need not be closed. This problem is particularly accurate in our shape analysis setting, because the horizontal distribution itself is never closed in general. In this paper, instead of deriving necessary conditions for minimizing the action, we will provide sufficient conditions for being a critical point of the action, showing that any solution of certain Hamiltonian equations is a critical point of the action.

The paper is organized as follows.

In Section 9.1, we recall some results on manifolds of bounded geometry and on their groups of diffeomorphisms, and we define strong right-invariant sub-Riemannian structures on these groups. We establish the metric and geodesic completeness for the corresponding sub-Riemannian distance. We compute the derivative and the adjoint of the end-point mapping, and we infer the Hamiltonian geodesic equations. We provide several examples and comment on the difficulty of finding reasonable necessary conditions for optimality.

In Section 9.2, we consider general shape spaces and investigate the sub-Riemannian structure induced by the action of the group of diffeomorphisms. We establish the Hamiltonian geodesic equations and give some examples.

Finally, in Section 9.3, we define lifted shape spaces, which are shape spaces on which the sub-Riemannian structure is induced by lifting that of another shape space through an equivariant mapping. The corresponding Hamiltonian geodesic equations are harder to establish in general, but we will provide a simpler version by using a slightly different point of view. We establish the Hamiltonian geodesic equations on some finite-dimensional examples.

9.1 Sub-Riemannian structure on groups of diffeomorphisms

Let $d \in \mathbb{N}^*$, and let (M, g) be a smooth oriented Riemannian manifold, where M is a smooth manifold of dimension d equipped with a Riemannian metric g . We assume that M has bounded geometry, that is, we assume that its global injectivity radius $\text{inj}(M)$ is positive and that one of the following equivalent conditions is satisfied:

1. For every $i \in \mathbb{N}$, the Riemannian norm of the i -th covariant derivative of the curvature tensor of M is bounded.
2. For every $i \in \mathbb{N}$, there exists $C_i > 0$ such that $|d^i g| \leq C_i$ in every normal coordinate chart (defined by the Riemannian exponential map of M) of radius $\text{inj}(M)/2$ on M .

In this section, we are going to define a sub-Riemannian structure on the manifold $\mathcal{D}^s(M)$, which is the connected component of the neutral element of the group of diffeomorphisms of

M of Sobolev class H^s . We will then derive Hamiltonian equations for normal geodesics of this sub-Riemannian manifold, and provide several examples with $M = \mathbb{R}^d$.

9.1.1 Definition of the manifold $\mathcal{D}^s(M)$

Let us first settle some notations and recall some results about the manifold structure of the space $H^s(M, N)$ of mappings from M to N of Sobolev class H^s , for s large enough. Here, (N, h) is another smooth Riemannian manifold with bounded geometry.

Let $f : M \rightarrow N$ be a smooth mapping. Its differential is a section of the vector bundle $T^*M \otimes_M f^*TN$, where $f^*TN = \{(x, v) \in M \times TN \mid v \in T_{f(x)}N\}$. The Levi-Civita connections on M and N induce a Riemannian connection on $T^*M \otimes f^*TN$, which allows us to set $|df|_s^2 = \sum_{i=0}^s \int_M |\nabla^i f(x)|^2 dx$, for every integer $s \geq d/2 + 1$. Then, we define the set $\mathcal{C}^{\infty, s}(M, N)$ of smooth functions from M to N of Sobolev class H^s as that of smooth functions from M to N such that $|df|_s^2 < +\infty$. The tangent space $T_f \mathcal{C}^{\infty, s}(M, N)$ at f is defined as the set of smooth sections X of f^*TN such that

$$|X|_s^2 = \sum_{i=0}^s \int_M |\nabla^i X(x)|^2 dx < +\infty. \quad (9.1)$$

The norm $|\cdot|_s$ is a pre-Hilbert one. Note that any $X \in T_f \mathcal{C}^{\infty, s}(M, N)$ is bounded and there exists $C > 0$ such that $h_x(X(x), X(x)) \leq C|X|_s$ for every $x \in N$.

For $\delta < \text{inj}(N)/2$, we say that two elements f_1 and f_2 of $\mathcal{C}^{\infty, s}(M, N)$ are δ -close if there exists $X \in T_{f_1} \mathcal{C}^{\infty, s}(M, N)$ such that $|X|_s \leq \delta$ and $f_2 = \exp_N \circ X$, where \exp_N is the Riemannian exponential map on N . We denote \mathcal{V}_δ the set of all such couples.

It is known (see [ES96, Sch04]) that the set $\{\mathcal{V}_\delta \mid 0 < \delta < \text{inj}(N)/2\}$ is a basis for a metrizable uniform structure on $\mathcal{C}^{\infty, s}(M, N)$, with distance $d_{H^s(M, N)}$. Its completion $H^s(M, N)$ is a smooth Hilbert manifold, with $T_f H^s(M, N) = \{X \in H^s(M, TN) \mid \forall x \in M \quad g(x) \in T_{f(x)}N\}$, and on which the norm $|\cdot|_s$ defined by (9.1) is a Hilbert norm.

If $s > d/2 + \ell$, then we have a continuous inclusion $H^s(M, N) \hookrightarrow \mathcal{C}^\ell(M, N)$. Moreover, if M is compact, then the inclusion $H^{s+1}(M, N) \hookrightarrow H^{s'+1}(M, N)$, $s' < s$, is compact.

Taking $s > d/2 + 1$ and $M = N$, we are now in a position to define $\mathcal{D}^s(M)$. We denote by $H_0^s(M, M)$ the connected component of $e = \text{id}_M$ in $H^s(M, M)$, and by $\text{Diff}(M)$ the set of C^1 diffeomorphisms on M .

Definition 9.1. For every integer $s > d/2 + 1$, we define $\mathcal{D}^s(M) = H_0^s(M, M) \cap \text{Diff}(M)$ as the connected component of the space of diffeomorphisms of class H^s on M .

The Hilbert manifold $\mathcal{D}^s(M)$ is an open subset of $H^s(M, M)$. In what follows, we denote by $\Gamma^s(TM) = T_e \mathcal{D}^s(M)$ the tangent space of $\mathcal{D}^s(M)$ at $e = \text{id}_M$, which coincides with the space of vector fields of class H^s on M .

The set $\mathcal{D}^s(M)$ is a topological group for the composition $(\varphi, \psi) \mapsto \varphi \circ \psi$, and for every $\psi \in \mathcal{D}^s(M)$, the right multiplication $R_\psi : \varphi \mapsto \varphi \circ \psi$ is smooth. Moreover, for every $k \in \mathbb{N}$, the mapping $(\varphi, \psi) \in \mathcal{D}^{s+k}(M) \times \mathcal{D}^s(M) \mapsto \varphi \circ \psi \in \mathcal{D}^s(M)$ is of class \mathcal{C}^k . The derivative of R_ψ is given by $dR_\psi \cdot X = X \circ \psi$ (and this, at any point φ). Similarly, the derivative of the left multiplication is given by $dL_\varphi \cdot X = d\varphi(\psi) \cdot X$ whenever φ is of class H^{s+1} (see [ES96, Sch04]).

Using the smoothness of the right multiplication, we define the set of right-invariant vector fields on $\mathcal{D}^s(M)$ as the set of vector fields $X : \mathcal{D}^s(M) \rightarrow T\mathcal{D}^s(M)$ that satisfy $X(\varphi) = X(e) \circ \varphi$. The set of right-invariant vector fields is therefore identified to $T_e \mathcal{D}^s(M) = \Gamma^s(TM)$.

Note that $\mathcal{D}^s(M)$ is not a Lie group because, although the right multiplication is smooth, the left composition in $\mathcal{D}^s(M)$ is only continuous. It can however be noticed that the set $\mathcal{D}^\infty(M) = \bigcap_{s > d/2+1} \mathcal{D}^s(M)$ is a Lie group; more precisely, endowed with the inverse limit topology, it is an *inverse limit Hilbert Lie group*, which is a particular type of a Fréchet Lie group (see [Omo74]).

Curves on $\mathcal{D}^s(M)$. For every $\varphi(\cdot) \in H^1(0, 1; \mathcal{D}^s(M))$, the time-dependent vector field $X(\cdot) = \dot{\varphi}(\cdot) \circ \varphi(\cdot)^{-1} \in L^2(0, 1; \Gamma^s(TM))$ is called the *logarithmic velocity* of $\varphi(\cdot)$. Note that, by definition, we have $\dot{\varphi}(t) = X(t) \circ \varphi(t)$ for almost every $t \in [0, 1]$.

Such curves of diffeomorphisms are flows of time-dependent right-invariant vector fields on $\mathcal{D}^s(M)$ whose norm is square-integrable in time. Conversely, thanks to a generalized version of the Cauchy-Lipschitz theorem (see [Tré08]), any time-dependent vector field $X(\cdot) \in L^2(0, 1; \Gamma^s(TM))$ generates a unique flow $\varphi_e^X(\cdot) \in H^1(0, 1; \mathcal{D}^s(M))$ such that $\varphi_e^X(0) = e$, and therefore defines a unique curve $\varphi^X(\cdot) = \varphi_e^X(\cdot) \circ \varphi_0$ for any fixed $\varphi_0 \in \mathcal{D}^s(M)$.

In other words, for any fixed $\varphi_0 \in \mathcal{D}^s(M)$, there is a one-to-one correspondence $X(\cdot) \leftrightarrow \varphi^X(\cdot)$ between time-dependent vector fields $X(\cdot) \in L^2(0, 1; \Gamma^s(TM))$ and curves $\varphi(\cdot) \in H^1(0, 1; \mathcal{D}^s(M))$ such that $\varphi(0) = \varphi_0$.

9.1.2 Sub-Riemannian structure on $\mathcal{D}^s(M)$

A sub-Riemannian manifold is usually defined as a triple (\mathcal{M}, Δ, h) , where \mathcal{M} is a manifold and (Δ, h) is a smooth Riemannian subbundle of the tangent space $T\mathcal{M}$ of \mathcal{M} , called *horizontal distribution*, equipped with a Riemannian metric h .

Here, keeping the framework and notations of the previous section, we take $\mathcal{M} = \mathcal{D}^s(M)$, and we are going to define a (right-invariant) horizontal distribution on the manifold $\mathcal{D}^s(M)$, endowed with a (right-invariant) Riemannian metric, that is, to define a (right-invariant) sub-Riemannian structure on the infinite-dimensional manifold $\mathcal{D}^s(M)$.

Let s and \tilde{s} be integers such that $\tilde{s} \geq s > d/2 + 1$. In what follows, we consider a Hilbert space $(V, \langle \cdot, \cdot \rangle)$ of vector fields of class $H^{\tilde{s}}$ on M such that the inclusion $V \hookrightarrow \Gamma^{\tilde{s}}(TM)$ is continuous.

Definition 9.2. We consider the subbundle \mathcal{H}^s of $T\mathcal{D}^s(M)$ defined by $\mathcal{H}_\varphi^s = R_\varphi V = V \circ \varphi$ for every $\varphi \in \mathcal{D}^s(M)$, endowed with the (fibered) metric $\langle X, Y \rangle_\varphi = \langle X \circ \varphi^{-1}, Y \circ \varphi^{-1} \rangle$. This subbundle induces a sub-Riemannian structure on $\mathcal{D}^s(M)$, that we refer to as the *strong right-invariant sub-Riemannian structure defined by V on $\mathcal{D}^s(M)$* .

Note that the mapping $(\varphi, X) \mapsto X \circ \varphi$ gives a parametrization of \mathcal{H}^s by $\mathcal{D}^s(M) \times V$. This parametrization is of class \mathcal{C}^k with $k = \tilde{s} - s$.

Note that, if $s < s' < \tilde{s}$, then the restriction of \mathcal{H}^s to $\mathcal{D}^{s'}(M)$ coincides with $\mathcal{H}^{s'}$. In particular, for a fixed diffeomorphism $\varphi \in \mathcal{D}^{s'}(M)$, we have $\mathcal{H}_\varphi^s = \mathcal{H}_\varphi^{s'} = V \circ \varphi$. Hence \mathcal{H}_φ^s does not depend on s and we will simply write it as \mathcal{H}_φ .

Example 9.1. The simplest example of such a structure on the group of diffeomorphisms is obtained by taking $V = \Gamma^{\tilde{s}}(TM)$ and

$$\langle X, Y \rangle = \sum_{i=1}^{\tilde{s}} \int_M g_x(\nabla^i X(x), \nabla^i Y(x)) \, dx.$$

Then \mathcal{H}_φ^s is the space of $X \in H^s(M, TM)$ such that $X(x) \in T_{\varphi(x)}M$ and $X \circ \varphi^{-1} \in \Gamma^{\tilde{s}}(TM)$.

For $s = \tilde{s}$, we have $\mathcal{H}_\varphi^s = T\mathcal{D}^s(M)$ and we obtain a Riemannian structure on $\mathcal{D}^s(M)$. Moreover, using a careful computation and the change of variable $x = \varphi(y)$ in the integral, we can see that the Riemannian structure is smooth.

Such metrics have been studied in [BHM11, BBHM13, MM07], seen as "weak" Riemannian metrics on the group of smooth diffeomorphisms (in contrast with "strong" metrics as in [BV13]).

Example 9.2. Let Δ be a smooth subbundle of TM , endowed with the restriction of the metric g to Δ . We define the space $V = \{X \in \Gamma^s(TM) \mid \forall x \in M \quad X(x) \in \Delta_x\}$ of all horizontal vector fields of class H^s . Then \mathcal{H}_φ is the space of H^s mappings $X : M \rightarrow TM$ such that

$X(x) \in \Delta_{\varphi(x)}$. Note that, although the parametrization $(X, \varphi) \mapsto X \circ \varphi$ is only continuous, \mathcal{H} is a smooth subbundle of $T\mathcal{D}^s(M)$ (indeed the subbundle Δ itself is smooth).

These examples show that the parametrization $(X, \varphi) \mapsto X \circ \varphi$ may not be the most appropriate one in some particular situations. We will however keep this general point of view throughout the paper.

Remark 9.1. In Definition 9.2, we consider "strong" sub-Riemannian structures. We can define "weak" sub-Riemannian structures on $\mathcal{D}^s(M)$, for which the norm induced by $\langle \cdot, \cdot \rangle$ on V is not complete. This allows to deal with more general metrics such as the L^2 norm on vector fields, $\langle X, Y \rangle = \int_M g_x(X(x), Y(x)) \, d\text{Vol}_g(x)$, where Vol_g is the canonical volume form on (M, g) . When V is the set of volume-preserving vector fields, this metric induces a smooth Riemannian metric on the group of volume-preserving diffeomorphisms of M , whose geodesics are the solutions of the Euler equations on M (see [EM70]).

Remark 9.2. Any Hilbert space V of vector fields of class at least H^s , with $s > d/2$, has a reproducing kernel. This means that the operator $V^* \rightarrow V$ given by the inverse of the isometry $X \mapsto \langle X, \cdot \rangle$, is the convolution with a section K of the bundle $L(T^*M, TM) = TM \otimes TM \rightarrow M \times M$, called the reproducing kernel of V .

Any element P of the dual $\Gamma^{-s}(T^*M)$ of $\Gamma^s(TM)$ can be represented by a one-form with (distributional) coefficients of class H^{-s} , so that $P(X) = \int_M P(x)(X(x)) \, dx$, for every $X \in \Gamma^s(TM)$ (the integral is computed in coordinates by means of a partition of unity). Such a P is called a *co-current*. By restriction, any co-current $P \in \Gamma^{-s}(T^*M)$ belongs as well to the dual V^* of V . In particular, for any $(x, p) \in T^*M$, the linear form $p \otimes \delta_x : X \mapsto p(X(x))$ belongs to V^* , and the unique element Y such that $\langle Y, \cdot \rangle = p \otimes \delta_x$ on V is denoted by $K(\cdot, x)p$. Then $K(x, y)$ is a linear mapping from T_y^*M to T_xM , that is an element of $T_yM \otimes T_xM$. Moreover, for every co-current $P \in \Gamma^{-s}(T^*M)$, the unique element Y such that $\langle Y, \cdot \rangle = P$ on V is given by $Y(x) = \int_M K(x, y)P(y) \, d\text{Vol}_g(y)$. In particular, K is of class at least H^s on $M \times M$.

We are now in a position to consider horizontal curves on $\mathcal{D}^s(M)$ for the strong right-invariant sub-Riemannian structure induced by V , and to define the associated concept of sub-Riemannian distance.

Horizontal curves.

Definition 9.3. An horizontal curve for the strong right-invariant sub-Riemannian structure induced by $(V, \langle \cdot, \cdot \rangle)$ on $\mathcal{D}^s(M)$ is a curve $\varphi(\cdot) \in H^1(0, 1; \mathcal{D}^s(M))$ such that $\dot{\varphi}(t) \in \mathcal{H}^s$ for almost every $t \in [0, 1]$. Equivalently, it is a curve with logarithmic velocity $X(\cdot) \in L^2(0, 1; V)$.

For every $\varphi_0 \in \mathcal{D}^s(M)$, we define Ω_{φ_0} as the set of all horizontal curves $\varphi(\cdot) \in H^1(0, 1; \mathcal{D}^s(M))$ such that $\varphi(0) = \varphi_0$. We define the mapping $\Phi_{\varphi_0} : L^2(0, 1; V) \rightarrow \Omega_{\varphi_0}$ by $\Phi_{\varphi_0}(X(\cdot)) = \varphi^X(\cdot)$, where $\varphi^X(\cdot)$ is the unique solution of the Cauchy problem $\dot{\varphi}^X(\cdot) = X(\cdot) \circ \varphi^X(\cdot)$, $\varphi^X(0) = \varphi_0$. We have $\Omega_{\varphi_0} = \Phi_{\varphi_0}(L^2(0, 1; V))$.

Lemma 9.1. We assume that $k = \tilde{s} - s \geq 1$. For every $\varphi_0 \in \mathcal{D}^s(M)$, the mapping $\Phi_{\varphi_0} : L^2(0, 1; V) \rightarrow \Omega_{\varphi_0}$ is a C^k diffeomorphism, and the set Ω_{φ_0} is a C^k submanifold of $H^1(0, 1; \mathcal{D}^s(M))$.

Proof. Using the correspondence $\varphi(\cdot) \leftrightarrow X(\cdot)$ explained in Section 9.1.1 (for φ_0 fixed), it suffices to prove that the graph in $H^1(0, 1; \mathcal{D}^s(M)) \times L^2(0, 1; V)$ of the C^k mapping $X(\cdot) \mapsto \varphi^X(\cdot)$ is a C^k manifold, globally parametrized by $X(\cdot) \mapsto (\varphi^X(\cdot), X(\cdot) \circ \varphi^X(\cdot))$.

We denote by $H_{\varphi_0}^1(0, 1; \mathcal{D}^s(M))$ the set of $\varphi(\cdot) \in H^1(0, 1; \mathcal{D}^s(M))$ such that $\varphi(0) = \varphi_0$. We define $H_{\varphi_0}^1 \times L^2(0, 1; T\mathcal{D}^s(M))$ as the fiber bundle over $H_{\varphi_0}^1(0, 1; T\mathcal{D}^s(M))$ defined by

$\{(\varphi(\cdot), \delta\varphi(\cdot)) \in H^1_{\varphi_0}(0, 1; \mathcal{D}^s(M)) \times L^2(0, 1; T\mathcal{D}^s(M)) \mid \delta\varphi(t) \in T_{\varphi(t)}\mathcal{D}^s(M) \text{ for a.e. } t \in [0, 1]\}$. We consider the affine vector bundle morphism $C : H^1(0, 1; \mathcal{D}^s(M)) \times L^2(0, 1; V) \rightarrow H^1_{\varphi_0} \times L^2(0, 1; T\mathcal{D}^s(M))$ defined by $C(\varphi(\cdot), X(\cdot))(t) = \dot{\varphi}(t) - X(t) \circ \varphi(t)$. Then C is of class \mathcal{C}^k and $\Omega_{\varphi_0} = C^{-1}(\{0\})$. In coordinates on $\mathcal{D}^s(M)$, we have $\partial_{\varphi}C(\varphi(\cdot), X(\cdot)).\delta\varphi(\cdot) = \delta\dot{\varphi}(\cdot) - d(X(\cdot) \circ \varphi(\cdot)).\delta\varphi(\cdot)$, which is a continuous linear differential operator of order one, and the linear Cauchy-Lipschitz theorem in Banach spaces implies that it is an isomorphism. The lemma follows from the implicit function theorem. \square

Definition 9.4. For every $\varphi_0 \in \mathcal{D}^s(M)$, the end-point mapping $\text{end}_{\varphi_0} : \Omega_{\varphi_0} \rightarrow \mathcal{D}^s(M)$ is defined by $\text{end}_{\varphi_0}(\varphi(\cdot)) = \varphi(1)$, for every $\varphi(\cdot) \in \Omega_{\varphi_0}$.

The end-point mapping is of class \mathcal{C}^k , with $k = \tilde{s} - s \geq 1$.

Definition 9.5. An horizontal curve $\varphi(\cdot) \in H^1(0, 1; \mathcal{D}^s(M))$ is said to be singular whenever the codimension of $\text{Range}(d\text{end}_{\varphi_0}(\varphi(\cdot)))$ in $T_{\varphi_1}\mathcal{D}^s(M)$ is positive.

Remark 9.3. An horizontal curve $\varphi(\cdot) \in H^1(0, 1; \mathcal{D}^s(M))$ is singular if and only if there exists $P_{\varphi_1} \in T_{\varphi_1}^*\mathcal{D}^s(M) \setminus \{0\}$ such that $(d\text{end}_{\varphi_0}(\varphi(\cdot)))^*.P_{\varphi_1} = 0$. Examples of singular curves can be built by considering a sub-Riemannian manifold M on which there exists a nontrivial singular curve $\gamma(\cdot)$ (such manifolds do exist, see, e.g., [Mon02]), and by taking V as the set of horizontal vector fields of class $H^{\tilde{s}}$ on M , as explained in Example 9.2. Then any autonomous vector field X such that $X \circ \gamma(\cdot) = \dot{\gamma}(\cdot)$ generates a singular curve.

Note that, in Theorem 9.4 further, we provide an Hamiltonian characterization of singular curves.

Given φ_0 and φ_1 in $\mathcal{D}^s(M)$, we consider the set $\Omega_{\varphi_0, \varphi_1} = \text{end}_{\varphi_0}^{-1}(\{\varphi_1\})$ of horizontal curves steering φ_0 to φ_1 .

Remark 9.4. The set $\Omega_{\varphi_0, \varphi_1}$ need not be a submanifold of Ω_{φ_0} , due to the fact that end_{φ_0} need not be a submersion. In the finite-dimensional context a singularity of this set is exactly a singular curve, that is a critical point of the end-point mapping, or equivalently, a projection of an abnormal extremal (see [CJT06, Mon02]). In infinite dimension the situation is more complicated because of the possible existence of proper subsets that are dense. More precisely, since we are in infinite dimension, we can have either of the three following possibilities:

1. $\text{Range}(d\text{end}_{\varphi_0}(\varphi(\cdot))) = T_{\varphi_1}\mathcal{D}^s(M)$;
2. the codimension of $\text{Range}(d\text{end}_{\varphi_0}(\varphi(\cdot)))$ in $T_{\varphi_1}\mathcal{D}^s(M)$ is positive;
3. $\text{Range}(d\text{end}_{\varphi_0}(\varphi(\cdot)))$ is a proper dense subset of $T_{\varphi_1}\mathcal{D}^s(M)$.

In finite dimension, only the two first possibilities do occur. In the first case it is usually said that $\varphi(\cdot)$ is *regular*, and in that case $\Omega_{\varphi_0, \varphi_1}$ is locally at $\varphi(\cdot)$ a submanifold of Ω_{φ_0} .

However, in the present infinite-dimensional framework (adapted to shape analysis), since the horizontal distribution \mathcal{H} itself is not closed in $T\mathcal{D}^s(M)$, actually the first possibility never occurs, and in particular we have to deal with the third possibility. In the context of controlled partial differential equations, this third possibility corresponds to a situation where the control system is approximately controllable but not exactly controllable.

Sub-Riemannian distance and action.

Definition 9.6. The sub-Riemannian length $L(\varphi(\cdot))$ and the sub-Riemannian action $A(\varphi(\cdot))$ of an horizontal curve $\varphi(\cdot) \in H^1(0, 1; \mathcal{D}^s(M))$ with logarithmic velocity $X(\cdot) = \dot{\varphi}(\cdot) \circ \varphi(\cdot)^{-1} \in L^2(0, 1; V)$ are respectively defined by

$$L(\varphi(\cdot)) = \int_0^1 \sqrt{\langle X(t), X(t) \rangle} dt \quad \text{and} \quad A(\varphi(\cdot)) = \frac{1}{2} \int_0^1 \langle X(t), X(t) \rangle dt.$$

The sub-Riemannian distance d_{SR} between two elements φ_0 and φ_1 of $\mathcal{D}^s(M)$ is defined as the infimum of the length of horizontal curves steering φ_0 to φ_1 , with the agreement that $d_{SR}(\varphi_0, \varphi_1) = +\infty$ whenever there is no horizontal curve steering φ_0 to φ_1 .

An horizontal curve $\varphi(\cdot) : [0, 1] \rightarrow \mathcal{D}^s(M)$ is said to be minimizing if $d_{SR}(\varphi(0), \varphi(1)) = L(\varphi(\cdot))$.

We have $d_{SR}(\varphi_0, \varphi_1) = d_{SR}(e, \varphi_1 \circ \varphi_0^{-1})$, that is, d_{SR} is right-invariant. Concatenation and time-reversals of horizontal curves are horizontal as well, hence d_{SR} is at least a semi-distance, and the subset $\{\varphi_1 \in \mathcal{D}^s(M) \mid d_{SR}(e, \varphi_1) < \infty\}$ is a subgroup of $\mathcal{D}^s(M)$ that does not depend on s (it is proved in Theorem 9.1 further that d_{SR} is a distance). Moreover, it follows from the Cauchy-Schwarz inequality that $L^2 \leq 2A$, and therefore $d_{SR}(\varphi_0, \varphi_1)$ is equal to $\sqrt{2}$ times the infimum of the action over all horizontal curves steering φ_0 to φ_1 . Therefore, minimizing the length over horizontal curves between two end-points is equivalent to minimizing the action over this set.

Theorem 9.1. *The sub-Riemannian distance d_{SR} is indeed a distance (taking its values in $[0, +\infty]$), that is, $d_{SR}(\varphi_0, \varphi_1) = 0$ implies $\varphi_0 = \varphi_1$. Moreover, any two elements φ_0 and φ_1 of $\mathcal{D}^s(M)$ with $d_{SR}(\varphi_0, \varphi_1) < +\infty$ can be connected by a minimizing horizontal curve, and $(\mathcal{D}^s(M), d_{SR})$ is a complete metric space.*

Proof. The proof follows, in the context of sub-Riemannian geometry, some arguments of [Tro95].

Firstly, in order to prove that $d_{SR}(\varphi_0, \varphi_1) = 0$ implies $\varphi_0 = \varphi_1$, since d_{SR} is right-invariant, it suffices to prove that $\varphi_1 \neq e$ implies $d_{SR}(e, \varphi_1) > 0$. Let $\varphi(\cdot) \in \Omega_e$ be an horizontal curve such that $\varphi(0) = e$ and $\varphi(1) = \varphi_1 \neq e$, and let $X(\cdot)$ be its logarithmic velocity. Let $x \in M$ such that $\varphi_1(x) \neq x$. Setting $x(t) = \varphi(t, x)$, we have $\dot{x}(t) = X(t, x(t))$, and then, $0 < d_M(x, \varphi_1(x))^2 \leq \int_0^1 g_{x(t)}(X(t, x(t)), X(t, x(t))) dt$, where d_M is the Riemannian distance on M . Since there exist positive constants C_1 and C_2 such that $g_y(Y(y), Y(y)) \leq C_1 \|Y\|_{\Gamma^s(TM)}^2 \leq C_2 \langle Y, Y \rangle$, for every $y \in M$ and for every $Y \in V$, it follows that $0 < d_M(x, \varphi_1(x))^2 \leq C_2 \int_0^1 \langle X(t), X(t) \rangle dt = 2C_2 A(\varphi(\cdot))$. Since this inequality is true for every horizontal curve $\varphi(\cdot)$ steering e to φ_1 , we get that $0 < d_M(x, \varphi_1(x)) \leq d_{SR}(e, \varphi_1)$ (note by the way that this inequality holds true for every $x \in M$).

Secondly, since the structure is right-invariant, it suffices to prove that any $\varphi_1 \in \mathcal{D}^s(M)$ such that $d_{SR}(e, \varphi_1) < \infty$ can be reached from e by a minimizing horizontal curve. In order to prove this fact, we will use the following lemma, which itself mainly follows from the Sobolev embedding theorem (see [ATTY13] for the proof).

Lemma 9.2. *Let $(X^n(\cdot))_{n \in \mathbb{N}}$ be a bounded sequence of $L^2(0, 1; V)$, consisting of logarithmic velocities of horizontal curves $(\varphi^n(\cdot))_{n \in \mathbb{N}}$ such that $\varphi^n(0) = e = \text{id}_M$. Then there exist $\bar{X}(\cdot) \in L^2(0, 1; V)$, corresponding to the horizontal curve $\bar{\varphi}(\cdot)$, and an increasing sequence $(n_j)_{j \in \mathbb{N}}$ of integers such that $(X^{n_j}(\cdot))_{j \in \mathbb{N}}$ converges weakly to $\bar{X}(\cdot)$, and such that $\sup_{t \in [0, 1]} d_{H^{\bar{s}-1}(U, M)}(\varphi^{n_j}(t), \bar{\varphi}(t))$ converges to 0 as j tends to $+\infty$, for every compact subset U of M .*

Let $(X^n(\cdot))_{n \in \mathbb{N}}$ be a minimizing sequence of $L^2(0, 1; V)$, associated with horizontal curves $(\varphi^n(\cdot))_{n \in \mathbb{N}}$, for the problem of minimizing the action over all horizontal curves steering e to φ_1 . By Lemma 9.2, up to some subsequence, the sequence $(X^n(\cdot))_{n \in \mathbb{N}}$ converges weakly to $\bar{X}(\cdot)$, associated with an horizontal curve $\bar{\varphi}(\cdot)$ such that $\bar{\varphi}(1) = \varphi_1$ (note that $\bar{s} - 1 > d/2$). Hence $\bar{\varphi}(\cdot)$ is an horizontal curve steering e to φ_1 , and by lower semi-continuity of the action, we have $A(\bar{\varphi}(\cdot)) = \frac{1}{2} \int_0^1 \langle \bar{X}(t), \bar{X}(t) \rangle dt \leq \liminf_n \frac{1}{2} \int_0^1 \langle X^n(t), X^n(t) \rangle dt$, and hence $\bar{\varphi}(\cdot)$ is a minimizing horizontal curve steering e to φ_1 .

Let us finally prove that $(\mathcal{D}^s(M), d_{SR})$ is complete. Let $(\varphi_1^n)_{n \in \mathbb{N}}$ be a Cauchy sequence of $\mathcal{D}^s(M)$. Then $(\varphi_1^n)_{n \in \mathbb{N}}$ is a Cauchy sequence as well in $\mathcal{C}^0(M, M)$ for the metric topology of

uniform convergence on compact subsets, which is complete, and therefore converges to some $\bar{\varphi}_1 \in \mathcal{C}^0(M, M)$. To conclude, it suffices to prove that $\bar{\varphi}_1 \in \mathcal{D}^s(M)$ and that $d_{SR}(\varphi_1^n, \bar{\varphi}_1) \rightarrow 0$ as $n \rightarrow +\infty$.

For every integer $m \geq n$, let $X^{n,m}(\cdot)$ be the logarithmic velocity of a minimizing horizontal curve $\varphi^{n,m}(\cdot)$ steering φ_1^n to φ_1^m . For every n , the sequence $(X^{n,m}(\cdot))_{m \geq n}$ (indexed by m) is bounded in $L^2(0, 1; V)$, and hence from Lemma 9.2, up to some subsequence it converges weakly to some $\bar{X}^n(\cdot) \in L^2(0, 1; V)$, which is the logarithmic velocity of an horizontal curve $\bar{\varphi}^n(\cdot)$. Moreover $\varphi_1^m = \varphi^{n,m}(1) \rightarrow \bar{\varphi}^n(1)$ in $\mathcal{C}^0(M, M)$ as $m \rightarrow +\infty$. But since $\varphi_1^m \rightarrow \bar{\varphi}_1$, it follows that $\bar{\varphi}^n(1) = \bar{\varphi}_1$. In particular, $\bar{\varphi}_1 \in \mathcal{D}^s(M)$ and $\bar{\varphi}_1^n(\cdot)$ is an horizontal curve steering φ_1^n to $\bar{\varphi}_1$. By weak convergence of $X^{n,m}(\cdot)$ to $\bar{X}^n(\cdot)$ as $m \rightarrow +\infty$, and by lower semi-continuity, we infer that

$$\begin{aligned} d_{SR}(\bar{\varphi}_1, \varphi_1^n)^2 &\leq 2A(\bar{\varphi}^n(\cdot)) = \int_0^1 \langle \bar{X}^n(t), \bar{X}^n(t) \rangle dt \\ &\leq \liminf_{m \rightarrow \infty} \int_0^1 \langle X^{n,m}(t), X^{n,m}(t) \rangle dt = \liminf_{m \rightarrow \infty} 2A(\varphi^{n,m}(\cdot)) = \liminf_{k \rightarrow \infty} d_{SR}(\varphi_1^n, \varphi_1^k)^2. \end{aligned}$$

The equality in the last part is due to the fact that $\varphi^{n,m}(\cdot)$ is a minimizing horizontal curve steering φ_1^n to φ_1^m . Since $(\varphi_1^n)_{n \in \mathbb{N}}$ is a Cauchy sequence, the right-hand side of the above inequality tends to 0 as $n \rightarrow +\infty$, and hence $d_{SR}(\bar{\varphi}_1, \varphi_1^n) \rightarrow 0$ as $n \rightarrow +\infty$. \square

9.1.3 The accessibility property on $\mathcal{D}^s(M)$

In the rest of this section, $(V, \langle \cdot, \cdot \rangle)$ is a Hilbert space of vector fields of class H^s on a manifold M of bounded geometry and dimension d . We consider the corresponding right-invariant sub-Riemannian structure $(\mathcal{H}^s, \langle \cdot, \cdot \rangle)$ on $\mathcal{D}^s(M)$ as in Definition 9.2.

The purpose of this section is to give conditions on V ensuring the approximate accessibility or (resp. exact accessibility) property from e . This means, giving conditions ensuring that the set $\mathcal{D}^{\mathcal{H}}$, defined in Theorem 9.1, is dense in (resp. equal to) the connected component $\mathcal{D}_0^s(M)$ of e in $\mathcal{D}^s(M)$.

Approximate accessibility There are some general results of approximate accessibility in infinite dimensional sub-Riemannian geometry, though they cannot be applied directly here because our structures are not smooth. The property of exact accessibility from a given point in infinite dimensional sub-Riemannian manifolds, however, is almost never satisfied. Therefore, in most cases, we cannot expect to have better than approximate accessibility. For example, we have the following result.

Proposition 1. *If V is dense in $\Gamma^s(TM)$ then $\mathcal{D}^{\mathcal{H}}$ is dense in $\mathcal{D}_0^s(M)$.*

Proof. The assumption implies that \mathcal{H}^s is dense in $T\mathcal{D}^s(M)$. The proof of [DS80] then implies the result. \square

A more general result can be deduced from a result by Agrachev and Caponigro [AC09] in the case where M is compact.

Proposition 2. (Agrachev-Caponigro) *Assume that M is compact and let X_1, \dots, X_r be smooth vector fields on M .*

Assume that any two points $x, y \in M$ can be connected by a smooth curve $t \mapsto x(t) \in M$ whose velocity belongs to the span of $(X_i(x(t)))_{i=1, \dots, r}$ for every t .

Finally, assume that

$$\left\{ \sum_{i=1}^r u_i X_i \mid u_1, \dots, u_r \in \mathcal{C}^\infty(M) \right\} \subset V.$$

Then $\mathcal{D}^{\mathcal{H}}$ contains the set $\mathcal{D}_0^\infty(M)$ of all smooth diffeomorphisms of $\mathcal{D}_0^s(M)$. In particular, it is dense in $\mathcal{D}_0^s(M)$.

Proof. The main result of [AC09] is actually stronger. It says that, if M is compact and if any two points $x, y \in M$ can be connected by a smooth curve $t \mapsto x(t) \in M$ whose velocity belongs to the span of $(X_i(x(t)))_{i=1, \dots, r}$ for every t , then

$$\begin{aligned} \exists m \in \mathbb{N}, \forall \varphi \in \mathcal{D}_0^\infty(M), \exists u_1, \dots, u_m \in \mathcal{C}^\infty(M), i_1, \dots, i_m \in \{1, \dots, r\} \\ \varphi = \varphi^{u_1 X_{i_1}}(1) \circ \dots \circ \varphi^{u_m X_{i_m}}(1). \end{aligned}$$

This directly implies the result. \square

Remark 9.5. It is straight-forward to generalize the results of [AC09] to the case where

$$\left\{ \sum_{i=1}^r u_i X_i \mid u_1, \dots, u_r \in H^s(M) \right\} \subset V,$$

to obtain exact controllability. However, we found a shorter and simpler proof for the slightly less general case where the family (X_1, \dots, X_r) is bracket-generating, which we give in the appendix. The main interest of this new proof is that it also lets us give estimates similar to that of the ball-box theorem, see [Mon02] for the ball-box theorem, and the next section for a complete statement and implications of this new result.

Exact controllability As mentioned, exact accessibility is very rare in infinite dimensional sub-Riemannian manifolds. However, there is a special case for which we do have $\mathcal{D}^{\mathcal{H}} = \mathcal{D}_0^s(M)$. For this, we will first need to define the notion of a bracket-generating family of smooth vector fields.

Let (M, g) be a Riemannian manifold with bounded geometry and dimension d . Let X_1, \dots, X_r be smooth vector fields on M , bounded at every order.

Remark 9.6. The family (X_1, \dots, X_r) defines a rank-varying sub-Riemannian structure Δ on M (see [ABC⁺10]).

Define the following space of vector field

$$V = \{X \in \Gamma^s(TM) \mid \forall x \in M, X(x) \in \text{span}(X_1(x), \dots, X_r(x))\},$$

on which we define a Hilbert product $\langle \cdot, \cdot \rangle$ whose norm is equivalent to the H^s -norm defined in Section 9.1.1. Note that V is a $H^s(M)$ -module generated by X_1, \dots, X_r , so that any $X \in V$ can be written

$$X = \sum_{i=1}^r u^i X_i, \quad u = (u^1, \dots, u^r) \in H^s(M, \mathbb{R}^r),$$

and

$$\langle X, X \rangle \leq \sum_{i=1}^r \|u^i\|_{H^s(M)}^2.$$

In this section, we consider the right-invariant sub-Riemannian structure $(\mathcal{H}^s, \langle \cdot, \cdot \rangle)$ induced by the Hilbert space $(V, \langle \cdot, \cdot \rangle)$ on $\mathcal{D}^s(M)$, and denote by d the corresponding sub-Riemannian distance. Note that we have

$$\mathcal{H}_\varphi = \left\{ \sum_{i=1}^r u^i \circ \varphi X_i \circ \varphi \mid u^i \in H^s(M) \right\}.$$

Remark 9.7. It is easy to see that a curve $t \mapsto \varphi(t) \in H^1(0, 1; \mathcal{D}^s(M))$ is horizontal if and only if each curve $t \mapsto \varphi(t, x) \in M$, with $x \in M$, is horizontal for the sub-Riemannian structure Δ on M generated by (X_1, \dots, X_r) .

Recall that we defined $\mathcal{D}^{\mathcal{H}}$ as the set of diffeomorphisms $\varphi \in \mathcal{D}^s(M)$ with finite sub-Riemannian distance from $e = \text{Id}_M$, that is, those that can be connected to e through a horizontal curve $t \mapsto \varphi(t)$.

Definition 9.7. The family (X_1, \dots, X_r) is said to be **bracket-generating** if, for every $x \in M$,

$$T_x M = \text{span} \left\{ [X_{i_1} [\dots, [X_{i_{j-1}}, X_{i_j}] \dots]](x) \mid j \in \mathbb{N}^*, 1 \leq i_1, \dots, i_j \leq r \right\}.$$

In other words, if any tangent vector at $x \in M$ is a linear combination of Lie brackets of X_1, \dots, X_k [Mon02].

In this case, any two points of M can be connected by a curve whose velocity is a linear combination of the vector fields (X_1, \dots, X_r) at each time. This is the so-called Chow-Rashevski theorem [Mon02, BR96].

Recall that we took the X_i to be smooth, so that their iterated Lie brackets are well-defined. For more simplicity, if $I = (i_1, \dots, i_j)$ with $1 \leq i_1, \dots, i_j \leq r$, we shall denote

$$X_I = [X_{i_1} [\dots, [X_{i_{j-1}}, X_{i_j}] \dots]]. \tag{9.2}$$

When M is compact, it is well-known [Mon02, BR96] that (X_1, \dots, X_r) is bracket-generating if and only if one can find an integer $m \geq d$ and I_1, \dots, I_m , with $I_l = (i_1^l, \dots, i_{j_l}^l) \subset \{1, \dots, r\}^{j_l}$ such that for every $x \in M$,

$$T_x M = \text{span}(X_{I_1}(x), \dots, X_{I_m}(x)).$$

In particular, in this case, for every $X \in \Gamma^s(TM)$, we can write

$$X = \sum_{i=1}^m u^i X_{I_i}, \quad (u^1, \dots, u^m) \in H^s(M, \mathbb{R}^m)$$

with

$$\|X\|_s^2 \leq \sum_{i=1}^m \|u^i\|_s^2.$$

Theorem 9.2. Suppose that M is compact and that the family (X_1, \dots, X_k) is **bracket-generating**. Then there exists a \mathcal{C}^1 -mapping $\phi : H^s(M, \mathbb{R}^m) \rightarrow \mathcal{D}^s(M)$ with $\phi(0) = e$ and a constant $C > 0$ such that

$$d_{SR}(e, \phi(u_1, \dots, u_m)) \leq C \sum_{i=1}^m \|u_i\|_s^{1/j^i}.$$

In particular, $\mathcal{D}^{\mathcal{H}} = \mathcal{D}_0^s(M)$ and the topology induced by the sub-riemannian distance d_{SR} coincides with the intrinsic manifold topology on $\mathcal{D}^s(M)$.

Remark 9.8. This theorem is an H^s version of [AC09]. The part about the induced topology, however, is completely new and could not be easily deduced from Agrachev and Caponigro’s proof. This theorem is of particular interest in the context of sub-Riemannian geometry in infinite dimensions. Indeed, it is known that, in almost every cases, only a dense subset can be reached from a given point through horizontal curve, and the topology does not coincide. Indeed, the bracket-generating condition is an algebraic condition, and therefore, in infinite

dimensions, the linear combinations of all brackets of horizontal vector fields is only dense in the tangent space to the manifold. However, in our particular case, we will see that we can adapt the finite dimensional proof, although it will require a bit of care.

Also of note, this theorem and more was proved in [BV13] for the right-invariant Riemannian case ($V = \Gamma^s(TM)$), even in noncompact manifolds.

Proof. The proof follows a method similar to that of Chow-Rashevski's theorem from [Mon02], with a few extra tricks.

– *Step 1: Reduction to a neighbourhood of $e = \text{Id}_m$.*

The first step is a consequence of the following result, which is an immediate consequence of the right-invariance of the sub-Riemannian distance.

Lemma 9.3. $\mathcal{D}^{\mathcal{H}}$ is a subgroup of the connected component $\mathcal{D}_0^s(M)$ of e in $\mathcal{D}^s(M)$.

Proof. Let $\varphi, \psi \in \mathcal{D}^{\mathcal{H}}$. Then $\varphi \circ \psi^{-1} \in \mathcal{D}^{\mathcal{H}}$ since

$$d_{SR}(\varphi \circ \psi^{-1}, e) = d_{SR}(\varphi, \psi) \leq d_{SR}(\varphi, e) + d(\psi, e) < +\infty.$$

□

Then, to prove that $\mathcal{D}^{\mathcal{H}} = \mathcal{D}_0^s(M)$, we just need to show that $\mathcal{D}^{\mathcal{H}}$ contains a neighbourhood of e .

The distance being right-invariant, to prove that the topologies coincide, it is sufficient to show that, inside any sub-Riemannian ball centered at e is included a neighbourhood of e for the intrinsic topology. This would also prove that $\mathcal{D}^{\mathcal{H}}$ contains a neighbourhood of e , and therefore finish the proof.

– *Step 2: Smooth parametrization of the horizontal distribution.* The next problem we need to settle in order to apply the method of the proof for the finite dimensional case from [Mon02], compared to the finite dimensional case, is that our parametrization of \mathcal{H} , which we recall is given by

$$(\varphi, X) \in \mathcal{D}^s(M) \times V \mapsto X \circ \varphi \in T_{\varphi}\mathcal{D}^s(M)$$

is only continuous. This makes it impossible to blindly compute Lie brackets of horizontal vector fields on $\mathcal{D}^s(M)$ (recall that a vector field $\mathbf{X} : \mathcal{D}^s(M) \rightarrow T\mathcal{D}^s(M)$ is *horizontal* if $\mathbf{X}(\varphi) \in \mathcal{H}_{\varphi}$ for every φ).

Remark 9.9. One should be careful not to confuse vector fields on M and vector fields on $\mathcal{D}^s(M)$. To avoid any confusion, we will write vector fields on the infinite dimensional manifold $\mathcal{D}^s(M)$ using emboldened letters, as above.

Lemma 9.4. The mapping $\mathcal{D}^s(M) \times H^s(M, \mathbb{R}^r) \rightarrow T\mathcal{D}^s(M)$ defined by

$$(\varphi, u^1, \dots, u^r) \mapsto \mathbf{X}^u(\varphi) = \sum_{i=1}^r u^i X_i \circ \varphi$$

is smooth, and its range is equal to \mathcal{H} . In particular, any $\varphi \mapsto \mathbf{X}^u(\varphi)$ is a smooth horizontal vector field on $\mathcal{D}^s(M)$.

Proof. The smoothness is obvious: in the expression $\sum_{i=1}^r u^i X_i \circ \varphi$, only the terms $X_i \circ \varphi$ depend on φ , and those are smooth because we assumed the X_i s were smooth. Moreover,

$$\sum_{i=1}^r u^i X_i \circ \varphi = \left(\sum_{i=1}^r \underbrace{(u^i \circ \varphi^{-1})}_{\in H^s(M)} X_i \right) \circ \varphi \in V \circ \varphi = \mathcal{H}_{\varphi},$$

so that any \mathbf{X}^u is indeed horizontal. Conversely, if $Y_\varphi \in \mathcal{H}_\varphi$, we can write

$$Y_\varphi = \sum_{i=1}^r u^i \circ \varphi X_i \circ \varphi = \mathbf{X}^{u \circ \varphi}(\varphi).$$

The range of the map is indeed \mathcal{H} . □

– *Step 3: Length of integral curves of smooth horizontal vector fields.*

A consequence of Lemma 9.4 is that each \mathbf{X}^u has a unique local flow $(t, \varphi, u) \mapsto \Phi(t, u\varphi)$ on $\mathcal{D}^s(M)$, which is smooth. Moreover, if $t \mapsto \varphi(t) = \Phi(t, u\varphi(0))$ is an integral curve of this flow, it is a smooth horizontal curve for the right-invariant sub-Riemannian structure induced by V . Moreover, we can estimate the length of this curve as follows.

Lemma 9.5. *On any open subset \mathcal{U} of $\mathcal{D}^s(M)$ small enough, there is a constant $C > 0$ such that, for every $u \in H^s(M, \mathbb{R}^r)$ and $\varphi \in \mathcal{U}$*

$$\left\langle \sum_{i=1}^r u_i \circ \varphi^{-1} X_i, \sum_{i=1}^r u_i \circ \varphi^{-1} X_i \right\rangle \leq C \sum_{i=1}^r \|u_i\|_{H^s}^2.$$

Consequently, the length of the curve $t \in [0, 1] \mapsto \Phi(t, u, \varphi)$ is bounded from above by

$$C \sqrt{\sum_{i=1}^r \|u_i\|_{H^s}^2}.$$

Proof. Indeed, the mapping $\mathcal{D}^s(M) \times H^s(M, \mathbb{R}^r) \rightarrow H^s(M, \mathbb{R}^r)$ defined by $(\varphi, u) \mapsto u \circ \varphi^{-1} = R_{\varphi^{-1}} u$ is continuous. This implies [Omo74] that the map $\varphi \mapsto R_{\varphi^{-1}}$, defined on $\mathcal{D}^s(M)$ with values in the space of continuous linear operators on $H^s(M, \mathbb{R}^r)$ is locally bounded, though it may not be continuous. Since the Hilbert norm on V is equivalent to the H^s norm, this immediately yields the result. □

– *Step 4: Lie brackets of horizontal vector fields.*

For $i \in \{1, \dots, r\}$, and $u \in H^s(M)$, let \mathbf{X}_i^u be defined by $\mathbf{X}_i^u(\varphi) = u X_i(\varphi)$, so that

$$\mathbf{X}_i^u = \mathbf{X}^{(0, \dots, 0, u, 0, \dots, 0)}.$$

Each \mathbf{X}_i^u is a smooth vector field for every $u \in H^s(M)$. We can then compute the Lie bracket of \mathbf{X}_i^u with \mathbf{X}_j^v . First of all, for every $(\varphi, \delta\varphi) \in T\mathcal{D}^s(M)$, and $i \in \{1, \dots, r\}$, we have

$$d\mathbf{X}_i^u(\varphi)\delta\varphi = \partial_\varphi(u X_i \circ \varphi)(\varphi)\delta\varphi = u^i(dX_i \circ \varphi)\delta\varphi.$$

Hence, for every $\varphi \in \mathcal{D}^s(M)$, $u, v \in H^s(M)$, and $i, j \in \{1, \dots, r\}$,

$$\begin{aligned} [\mathbf{X}_i^u, \mathbf{X}_j^v](\varphi) &= d\mathbf{X}_j^v(\varphi)\mathbf{X}_i^u(\varphi) - d\mathbf{X}_i^u(\varphi)\mathbf{X}_j^v(\varphi) \\ &= v(dX_j \circ \varphi)(u X_i \circ \varphi) - u(dX_i \circ \varphi)(v X_j \circ \varphi) \end{aligned}$$

Since for every $x \in M$,

$$dX_i(\varphi(x))(v(x)X_j(\varphi(x))) = v(x)dX_i(\varphi(x))X_j(\varphi(x)),$$

we obtain

$$\begin{aligned} [\mathbf{X}_i^u, \mathbf{X}_j^v](\varphi) &= uv(dX_j(\varphi(x))X_i(\varphi(x)) - dX_i(\varphi(x))X_j(\varphi(x))) \\ &= uv[X_i, X_j] \circ \varphi \\ &= uvX_{i,j} \circ \varphi. \end{aligned}$$

A trivial induction reasoning yields the next lemma.

Lemma 9.6. *Let $j \in \mathbb{N}^*$, $I = (i_1, \dots, i_j) \in \{1, \dots, r\}^j$, $\varphi \in \mathcal{D}^s(M)$, and $u^1, \dots, u^j \in H^s(M)$. Then*

$$\left[\mathbf{X}_{i_1}^{u_1}, [\dots, [\mathbf{X}_{i_{j-1}}^{u_{j-1}}, \mathbf{X}_{i_j}^{u_j}] \dots] \right] (\varphi) = u_{i_1} \dots u_{i_j} X_I \circ \varphi,$$

where the smooth vector field X_I on M was defined in (9.2) as

$$X_I = \left[X_{i_1}, [\dots, [X_{i_{j-1}}, X_{i_j}] \dots] \right].$$

In particular, the Lie bracket is a continuous, multilinear operator with j entries with respect to (u_1, \dots, u_j) .

- *Step 5: Taylor expansions of commutators of horizontal flows and corresponding length.* For $i \in \{1, \dots, r\}$, and $u_i \in H^s(M)$, let $(t, \varphi \mapsto \Phi_i^{u_i}(t, \varphi)$ denote the flow of $\mathbf{X}_i^{u_i}$ on $\mathcal{D}^s(M)$. Moreover, for $j \in \mathbb{N}^*$ and $I = (i_1, \dots, i_j) \in \{1, \dots, r\}^j$, we shall denote

$$\Phi_{i_j}^{-u_j}(t) \circ \dots \circ \Phi_{i_1}^{-u_1}(t) \circ \Phi_{i_j}^{u_j}(t) \circ \dots \circ \Phi_{i_1}^{u_1}(t) = \Phi_I(t, u_1, \dots, u_j).$$

Remark 9.10. Lemma 9.5 implies that, for every φ , there exists $C > 0$ such that, for small enough u_1, \dots, u_j, φ and $\Phi_I(1, u_1, \dots, u_j)(\varphi)$ can be connected by a curve of length at most $2C(\|u_1\|_{H^s} + \dots + \|u_j\|_{H^s})$. In particular,

$$d_{SR}(\varphi, \Phi_I(1, u_1, \dots, u_j)(\varphi)) \leq 2C(\|u_1\|_{H^s} + \dots + \|u_j\|_{H^s}).$$

Our key lemma is the following.

Lemma 9.7. *Let $\varphi \in \mathcal{D}^s(M)$, $j \geq 2$, and $I = (i_1, \dots, i_j) \in \{1, \dots, r\}^j$. Then*

$$\Phi_I(1, u_1, \dots, u_j)(\varphi) = \varphi + u_1 \dots u_j X_I \circ \varphi + o(\|(u_1, \dots, u_j)\|_{H^s}^j),$$

as (u_1, \dots, u_j) go to 0.

Proof. It is well-known that, for fixed u_1, \dots, u_j and for small $t \in \mathbb{R}$,

$$\Phi_I(t, u_1, \dots, u_j)(\varphi) = \varphi + t^j \left[\mathbf{X}_{i_1}^{u_1}, [\dots, [\mathbf{X}_{i_{j-1}}^{u_{j-1}}, \mathbf{X}_{i_j}^{u_j}] \dots] \right] (\varphi) + o(t^j) = \varphi + t^j u_1 \dots u_j X_I \circ \varphi + o(t^{j+1}).$$

But, since we clearly have $t\mathbf{X}_{i_k}^{u_k} = \mathbf{X}_{i_k}^{tu_k}$ for each k ,

$$\Phi_I(t, u_1, \dots, u_j)(\varphi) = \Phi_I(1, tu_1, \dots, tu_j)(\varphi).$$

As a consequence, if $\Phi_I^{u_1, \dots, u_j}(1, \varphi)$ admits a Taylor expansion in (u_1, \dots, u_j) , this expansion is given by

$$\Phi_I(1, u_1, \dots, u_j)(\varphi) = \varphi + u_1 \dots u_j X_I \circ \varphi + o(\|(u_1, \dots, u_j)\|_{H^s}^j).$$

We already know that the term on the left is smooth in (φ, u) , so it does admit a Taylor expansion in u of order j . Hence we obtain our result. \square

- *Step 6: Proof of the theorem.*

First of all, recall that, since the family (X_1, \dots, X_r) is bracket-generating, there exists I_1, \dots, I_m , families of elements of $\{1, \dots, r\}$ of increasing cardinal, such that any $X \in \Gamma^s(TM)$ is an H^s -combination of X_{I_1}, \dots, X_{I_m} . Then, for $k \in \{1, \dots, m\}$ and j_k the size of I_k consider the map ϕ_k defined from a neighbourhood of $(0, e)$ in $H^s(M) \times \mathcal{D}^s(M)$ onto a neighbourhood of e in $\mathcal{D}^s(M)$ given by

$$\phi_k(u, \varphi) = \Phi_{I_k} \left(1, \|u\|_{H^s}^{\frac{1-j_k}{j_k}} u, \|u\|_{H^s}^{1/j_k}, \dots, \|u\|_{H^s}^{1/j_k} \right) (\varphi).$$

This map is smooth in u outside of 0, and it is of class \mathcal{C}^1 around 0 since Lemma 9.7 shows that for every $\varphi \in \mathcal{D}^s(M)$,

$$\phi_k(u, \varphi) = \varphi + uX_{I_k} \circ \varphi + o(u)$$

as $\|u\|_{H^s} \rightarrow 0$. Moreover, Remark 9.10 shows that, for every φ , there exists a constant $C > 0$ such that for small u ,

$$d_{SR}(\varphi, \phi_k(u, \varphi)) \leq C\|u\|_{H^s}^{1/j_k}. \quad (9.3)$$

Finally, consider the map ϕ from a neighbourhood of 0 in $H^s(M, \mathbb{R}^m)$ into $\mathcal{D}^s(M)$, defined by

$$\phi(u_1, \dots, u_m) = \phi(u_m) \circ \dots \circ \phi(u_1)(e).$$

Then ϕ is of class \mathcal{C}^1 around 0, and its derivative is given by

$$d\phi(0)(\delta u_1, \dots, \delta u_m) = \sum_{k=1}^m u_k X_{I_k} \in \Gamma^s(TM), \quad (\delta u_1, \dots, \delta u_m) \in H^s(M, \mathbb{R}^m),$$

and is therefore surjective, as we assumed that $(X_{I_1}, \dots, X_{I_n})$ generated all of TM . The submersion theorem proves that the image through ϕ of any neighbourhood of 0 in $H^s(M, \mathbb{R}^m)$ is a neighbourhood of e in $\mathcal{D}^s(M)$. Let $\varepsilon > 0$ be small enough that ϕ is defined and smooth on the ball B of center 0 and radius ε^{j_m} in $H^s(M, \mathbb{R}^m)$. Then, there exists a neighbourhood \mathcal{U}_ε of e in $\mathcal{D}^s(M)$ such that

$$\mathcal{U} \subset \phi(B(0, \varepsilon^{j_m})).$$

But, for ε small enough, (9.3) also implies that there is a constant $C > 0$ such that for every $(u_1, \dots, u_m) \in B(0, \varepsilon^{j_m})$,

$$d_{SR}(e, \phi(u_1, \dots, u_m)) \leq 2C \sum_{k=1}^m \|u_k\|_{H^s}^{1/j_k} \leq 2mC\varepsilon.$$

Hence any sub-Riemannian ball of center e contains a neighbourhood \mathcal{U} of e for the intrinsic topology. Thanks to our reasoning in Step 1, this concludes the proof. \square

9.1.4 Geodesics on $\mathcal{D}^s(M)$

We keep the framework and notations used in the previous sections.

Definition 9.8. *A geodesic $\varphi(\cdot)$ is an horizontal curve which is a critical point of the action mapping A restricted to $\Omega_{\varphi(0), \varphi(1)}$. In other words, for any \mathcal{C}^1 family of horizontal curves $s \in (-\varepsilon, \varepsilon) \mapsto \varphi^s(\cdot) \in \Omega_{\varphi(0), \varphi(1)}$, with $\varepsilon > 0$ and $\varphi^0(\cdot) = \varphi(\cdot)$, we have $dA(\varphi(\cdot)) \cdot \partial_s \varphi^s(\cdot)|_{s=0} = 0$. With a slight abuse of notation, we will denote by $T_{\varphi(\cdot)}\Omega_{\varphi(0), \varphi(1)}$ the set of all such infinitesimal variations $\partial_s \varphi^s(\cdot)|_{s=0}$.*

A geodesic $\varphi(\cdot)$ is said to be minimizing if $L(\varphi(\cdot)) = d_{SR}(\varphi(0), \varphi(1))$.

Note that, obviously, any minimizing horizontal curve is a geodesic.

Preliminary discussion: Lagrange multipliers. In finite dimension, the critical point property usually leads to a Lagrange multipliers relation, which provides a first-order necessary condition for optimality, itself allowing us to derive Hamiltonian geodesic equations. Here, since we are in infinite dimension, the situation is more complex and we do not have necessarily a nontrivial Lagrange multiplier. Let us be more precise with this important difficulty, because it justifies the point of view that we are going to adopt in the sequel. The discussion goes as follows.

Let $\varphi(\cdot) \in H^1(0, 1; \mathcal{D}^s(M))$ be a minimizing horizontal curve steering $\varphi_0 \in \mathcal{D}^s(M)$ to $\varphi_1 \in \mathcal{D}^s(M)$. Then $\varphi(\cdot)$ is a geodesic, that is, a critical point of the action A restricted to $\Omega_{\varphi_0, \varphi_1} = \text{end}_{\varphi_0}^{-1}(\{\varphi_1\})$. Defining the mapping $E_{\varphi_0} : L^2(0, 1; V) \rightarrow \mathcal{D}^s(M)$ by the composition $E_{\varphi_0} = \text{end}_{\varphi_0} \circ \Phi_{\varphi_0}$, and defining the mapping $F_{\varphi_0} : L^2(0, 1; V) \rightarrow \mathcal{D}^s(M) \times \mathbb{R}$ by

$$F_{\varphi_0}(X(\cdot)) = (E_{\varphi_0}(X(\cdot)), A(\Phi_{\varphi_0}(X(\cdot)))) = (\text{end}_{\varphi_0}(\varphi^X(\cdot)), A(\varphi^X(\cdot))),$$

it follows that the logarithmic velocity $X(\cdot)$ of $\varphi(\cdot)$ is a critical point of F_{φ_0} ; or, in other words, the differential $dF_{\varphi_0}(X(\cdot))$ is not surjective, that is, $\text{Range}(dF_{\varphi_0}(X(\cdot)))$ is a proper subset of $T_{\varphi_1} \mathcal{D}^s(M) \times \mathbb{R}$. Then, there are two possible cases:

1. either the codimension of $\text{Range}(dF_{\varphi_0}(X(\cdot)))$ in $T_{\varphi_1} \mathcal{D}^s(M) \times \mathbb{R}$ is positive, which is equivalent to the fact that $\ker((dF_{\varphi_0}(X(\cdot)))^*) \neq \{0\}$,
2. or the space $\text{Range}(dF_{\varphi_0}(X(\cdot)))$ is dense in $T_{\varphi_1} \mathcal{D}^s(M) \times \mathbb{R}$, which is equivalent to the fact that $\ker((dF_{\varphi_0}(X(\cdot)))^*) = \{0\}$.

The first case means that we have a nontrivial Lagrange multiplier, and the second case means that there does not exist any nontrivial Lagrange multiplier. Note that the second case can never occur in finite dimension. Here, since we are in infinite dimension, we have to face with this additional difficulty.

Since $d\Phi_{\varphi_0} : L^2(0, 1; V) \rightarrow T_{\varphi(\cdot)} \Omega_{\varphi_0}$ is an isomorphism, it follows that, for the geodesic $\varphi(\cdot) = \varphi^X(\cdot)$, there are two possible issues:

1. There exists $(P_{\varphi_1}, p^0) \in T_{\varphi_1}^* \mathcal{D}^s(M) \times \mathbb{R} \setminus \{(0, 0)\}$ such that

$$(d \text{end}_{\varphi_0}(\varphi(\cdot)))^* . P_{\varphi_1} + p^0 dA(\varphi(\cdot)) = 0. \tag{9.4}$$

This is a Lagrange multipliers relation. In finite dimension, only this first possibility does occur, and leads to the Pontryagin maximum principle (see, e.g., [Tré08] for this point of view). Note that the Lagrange multiplier (P_{φ_1}, p^0) is defined up to some multiplying scalar, and usually it is normalized by distinguishing between two subcases:

- (a) *Normal case:* $p^0 \neq 0$. In that case, we normalize the Lagrange multiplier so that $p^0 = -1$. Then (9.4) implies that $dA(\varphi(\cdot)) = (d \text{end}_{\varphi_0}(\varphi(\cdot)))^* . P_{\varphi_1}$, and in that case we will then derive the so-called *normal geodesic equations*.
- (b) *Abnormal case:* $p^0 = 0$. In that case, (9.4) implies that $(d \text{end}_{\varphi_0}(\varphi(\cdot)))^* . P_{\varphi_1} = 0$ (and for instance we can normalize the Lagrange multiplier by normalizing P_{φ_1}). This is equivalent to saying that the corank of $d \text{end}_{\varphi_0}(\varphi(\cdot))$ is positive: in other words, according to Definition 9.5, $\varphi(\cdot)$ is a singular curve. In that case, we will then derive the co-called *abnormal geodesic equations*, which are the Hamiltonian characterization of singular curves.

Remark 9.11. If $\varphi(\cdot)$ is a singular curve, then there exists $P_{\varphi_1} \in T_{\varphi_1}^* \mathcal{D}^s(M) \setminus \{0\}$ such that $(d \text{end}_{\varphi_0}(\varphi(\cdot)))^* . P_{\varphi_1} = 0$. In other words, there exists an abnormal Lagrange multiplier. This Lagrange multiplier is not necessarily unique (up to some multiplying scalar), and the dimension of the space of such Lagrange multipliers is usually called the corank of the singular curve (see [CJT06, CJT08] where generic properties are established for singular curves in finite dimension).

2. The mapping $(dF_{\varphi_0}(X(\cdot)))^*$ is injective, that is, if we have (9.4) for some $(P_{\varphi_1}, p^0) \in T_{\varphi_1}^* \mathcal{D}^s(M)$, then $(P_{\varphi_1}, p^0) = (0, 0)$.

This case is peculiar to the infinite-dimensional setting, and can never occur in finite dimension. In that case a necessary condition for optimality in the form of a Pontryagin maximum principle cannot be derived (see [LY95, Chapter 4]).

Definition 9.9. A geodesic $\varphi(\cdot) \in H^1(0, 1; \mathcal{D}^s(M))$ is said to be elusive whenever the mapping

$$(dF_{\varphi_0}(X(\cdot)))^*$$

is injective.

In what follows, we are going to state a Pontryagin maximum principle for non-elusive geodesics (that is, for the first case of the above discussion), and derive the normal and abnormal geodesic equations.

Remark 9.12. The concept of elusive geodesic is new, and is specific to the infinite dimension. What is important to understand is that elusive geodesics escape to the dual analysis in terms of Lagrange multipliers, due to the topology of the ambient space.

As it follows from the definition of an elusive curve, a Lagrange multiplier cannot exist because, although the mapping $dF_{\varphi_0}(X(\cdot))$ is not surjective, its range is however dense in the target space $T_{\varphi_1} \mathcal{D}^s(M) \times \mathbb{R}$. This difficulty, which is specific to the infinite-dimensional setting, is actually well known in constrained optimization. In [Kur76] the author provided some weak regularity conditions under which the existence of Lagrange multipliers can be established for a general nonlinear programming problem. He showed that the topology of the target space plays an important role, and he established a connection between the choice of suitable function spaces and the existence of Lagrange multipliers.

Before commenting on this choice, let us first provide an easy way to exhibit elusive geodesics. The idea relies on the fact that V is not closed in $\Gamma^s(TM) = T_e \mathcal{D}^s(M)$, which results in “missing” some initial momenta. For example, if $\varphi_0 = e$ and if $\tilde{s} \geq s + 3$, then restricting \mathcal{H}^s to $\mathcal{D}^{s+1}(M)$ gives new initial momenta P_0 , namely, those belonging to $\Gamma^{-s-1}(T^*M) \setminus \Gamma^{-s}(T^*M)$ (where $\Gamma^{-s}(T^*M)$ is the dual of $\Gamma^s(TM) = T_e \mathcal{D}^s(M)$).

Based on this idea, the method to exhibit elusive geodesics consists of “decreasing the order” of the cotangent space, in the following sense. The Hilbert space $(V, \langle \cdot, \cdot \rangle)$, with $V = \Gamma^{s+3}(TM)$, induces a right-invariant sub-Riemannian structure on $\mathcal{D}^{s+1}(M)$. Anticipating a bit, let $t \mapsto (\varphi(t), P(t))$ be a normal geodesic on $T^* \mathcal{D}^{s+1}(M)$ with $\varphi(0) = e$ and $P(0) \in \Gamma^{-s-1}(T^*M) \setminus \Gamma^{-s}(T^*M)$. Then, we claim that $\varphi(\cdot)$ is an elusive horizontal curve for the sub-Riemannian structure induced by $(V, \langle \cdot, \cdot \rangle)$ on $\mathcal{D}^s(M)$.

Indeed, it is clear horizontal curves starting at e coincide for both structures. Let us prove that $\varphi(\cdot)$ can have neither a normal nor an abnormal Lagrange multiplier, in the sub-Riemannian structure on $\mathcal{D}^s(M)$. If there would exist a normal Lagrange multiplier, then the corresponding initial momentum $P'(0) \in \Gamma^{-s}(T^*M)$ would satisfy $P(0) = P'(0)$ on $V = \Gamma^{s+3}(TM)$ which is dense in $\Gamma^s(TM)$, and we would have $P'(0) = P(0)$, which is impossible. There cannot exist an abnormal Lagrange multiplier, because the range of $\text{dend}_e(\varphi(\cdot))$ contains $\mathcal{H}_{\varphi(1)} = T_{\varphi_1} \mathcal{D}^{s+3}(M)$, which is dense in $T_{\varphi_1} \mathcal{D}^s(M)$. This proves the claim.

Conversely, we can get new normal geodesics for a right-invariant sub-Riemannian structure on $\mathcal{D}^s(M)$ induced by a Hilbert subspace of vector fields with continuous inclusion in $\Gamma^{s+3}(TM)$, by restricting it to $\mathcal{D}^{s+1}(M)$, adding extra initial momenta by increasing the order of the cotangent bundle. In such a way, some elusive geodesics become normal geodesics. Note that it may happen that some elusive geodesics become abnormal curves because of the

increased range in the choice of momenta. However, note also that this simple process does not turn every elusive geodesic into either a normal or an abnormal geodesic.

We conclude that, in accordance with [Kur76], the choice of the cotangent space (and thus, the choice of the topology of the target space) is important. The stronger is the topology in the target space, and the larger is the dual, but then Lagrange multipliers become more and more irregular. It is therefore reasonable to avoid choosing a too strong topology in the target space.

In our setting there does not seem to exist a best possible choice for the cotangent bundle (better in the sense that, by adding new possibilities for the initial momenta, we would turn every possible elusive geodesic into either a normal or an abnormal geodesic). The question of finding a “good” space of initial or final momenta (implying the absence of elusive geodesics) is open and seems quite difficult.¹ For instance, the authors found examples of Hilbert Lie groups (more precisely, $\ell^2(\mathbb{N}, \mathbb{R}^4)$, with \mathbb{R}^4 equipped with the Engel group structure) for which the natural space of momenta is not even locally convex.

Following this preliminary discussion, we are now going to derive the normal and abnormal geodesic equations, which are Hamiltonian characterizations of non-elusive geodesics. We stress again that, in general, no such Hamiltonian characterization can be derived for an elusive geodesic.

We will prove that any solution of the normal geodesic Hamiltonian equations, if it is well defined on $[0, 1]$, projects onto a geodesic (critical point of the action). Therefore, instead of giving necessary conditions to minimize the action, we are rather going to provide sufficient conditions under which we have a non-elusive geodesic.

Hereafter, we first establish the normal geodesic equations, then the abnormal geodesic equations, and we finally provide necessary conditions for optimality. The three theorems are then proved together.

Normal geodesic equations. Let $K^V : V^* \rightarrow V$ be the inverse of $X \mapsto \langle X, \cdot \rangle$. We define the \mathcal{C}^k vector bundle morphism $K_{\mathcal{H}_\varphi} : T_\varphi^* \mathcal{D}^s(M) \rightarrow T_\varphi \mathcal{D}^s(M)$ by $K_{\mathcal{H}_\varphi} = dR_\varphi K^V dR_\varphi^*$ for every $\varphi \in \mathcal{D}^s(M)$.

Definition 9.10. We define the normal Hamiltonian $h : T^* \mathcal{D}^s(M) \rightarrow \mathbb{R}$ by

$$h(\varphi, P) = \frac{1}{2} P(K_{\mathcal{H}_\varphi} P), \quad (9.5)$$

for every $(\varphi, P) \in T^* \mathcal{D}^s(M)$.

The expression (9.5) means that, as in classical sub-Riemannian geometry (see [Mon02]), $h(\varphi, P)$ is the squared norm of P for the cometric induced on $T^* \mathcal{D}^s(M)$ by the sub-Riemannian structure. This is the usual way to define the normal Hamiltonian. Any element $P \in T_\varphi^* \mathcal{D}^s(M)$ of the cotangent bundle is called a *momentum*.

1. Note that these difficulties are also due to the fact that the sub-Riemannian problem consists of minimizing the action $A(\varphi(\cdot))$ over all horizontal curves $\varphi(\cdot)$ such that $\text{end}_{\varphi_0}(\varphi(\cdot)) = \varphi_1$, that is, of minimizing a functional under an infinite number of constraints. Assume that, instead, we consider the problem of minimizing the penalized functional $J(X(\cdot)) = \int_0^1 \langle X(t), X(t) \rangle dt + G(\text{end}_{\varphi_0}(\varphi^X(\cdot)))$ over the set $L^2(0, 1; V)$, for some \mathcal{C}^1 function G . Then, we claim that, if G is bounded below then this (unconstrained) penalized problem has at least one solution $X(\cdot) \in L^2(0, 1; V)$, and there exists a momentum mapping $t \mapsto P(t)$ such that $P(1) + dG_{\varphi_1} = 0$ and such that $(\varphi(\cdot), P(\cdot))$ is solution of the normal geodesic Hamiltonian geodesic equations (9.6). This claim follows from an easy adaptation of the proofs of the results in the present paper (see also [ATTY13]). This framework can be used in order to “approach” a target diffeomorphism φ_1 with an horizontal curve, by choosing a smooth penalization function G .

In canonical coordinates, we have $K_{\mathcal{H}_\varphi}P = X(\varphi, P) \circ \varphi$, with $X(\varphi, P) = K^V(dR_\varphi)^*.P$, and hence

$$h(\varphi, P) = \frac{1}{2}P(X(\varphi, P) \circ \varphi) = \frac{1}{2}\langle X(\varphi, P), X(\varphi, P) \rangle.$$

Note that $(\varphi, P) \mapsto X(\varphi, P)$ is of class \mathcal{C}^k .

Denoting by ω the canonical strong symplectic form on $T^*\mathcal{D}^s(M)$, the symplectic gradient $\nabla^\omega h : T^*\mathcal{D}^s(M) \rightarrow TT^*\mathcal{D}^s(M)$ of h (which is of class \mathcal{C}^{k-1}) is defined by the relation $dh = \omega(\nabla^\omega h, \cdot)$. In canonical coordinates, we have $\nabla^\omega h = (\partial_P h, -\partial_\varphi h)$, where $\partial_P h \in T_{\varphi}^{**}\mathcal{D}^s(M) = T_\varphi\mathcal{D}^s(M)$ thanks to the natural isomorphism between a Hilbert space and its bidual space, and we have

$$\nabla^\omega h(\varphi, P) = (K_{\mathcal{H}_\varphi}P, -(\partial_\varphi K_{\mathcal{H}_\varphi}P)^*.P) = (X(\varphi, P) \circ \varphi, -(\partial_\varphi(X(\varphi, P) \circ \varphi))^*.P).$$

Theorem 9.3. *We assume that $s > d/2 + 1$ and $k = \tilde{s} - s \geq 2$. For every $\varphi_0 \in \mathcal{D}^s(M)$ and every $P_0 \in T_{\varphi_0}^*\mathcal{D}^s(M)$, there is a unique absolutely continuous fibered² mapping $(\varphi(\cdot), P(\cdot)) : I \rightarrow T^*\mathcal{D}^s(M)$, defined on a maximal interval $I \subset \mathbb{R}$ with $0 \in I$, solution of the Cauchy problem*

$$(\dot{\varphi}(t), \dot{P}(t)) = \nabla^\omega h(\varphi(t), P(t)), \quad (9.6)$$

for almost every $t \in I$, and $(\varphi(0), p(0)) = (\varphi_0, P_0)$.

Moreover, if $I \subset [0, 1]$ then $\varphi(\cdot)$ is a geodesic.

Definition 9.11. *In the conditions of Theorem 9.3, if $I \subset [0, 1]$ then $\varphi(\cdot)$ is said to be a normal geodesic.*

Remark 9.13. In canonical coordinates, the normal geodesic equations (9.6) are written as

$$\dot{\varphi}(t) = K_{\mathcal{H}_{\varphi(t)}}P(t), \quad \dot{P}(t) = -(\partial_\varphi K_{\mathcal{H}_{\varphi(t)}}P(t))^*.P(t).$$

Note that $t \mapsto h(\varphi(t), p(t))$ is constant and that, if $I \subset [0, 1]$ then $A(\varphi) = \frac{1}{2} \int_0^1 h(\varphi(t), P(t)) dt = \frac{1}{2}h(\varphi(0), P(0))$.

Remark 9.14. According to the notations above, the logarithmic velocity of any normal geodesic $\varphi(\cdot)$ is given by $X(\cdot) = X(\varphi(\cdot), P(\cdot)) = K^V(dR_{\varphi(\cdot)})^*.P(\cdot)$.

Remark 9.15. If V has a continuous inclusion in $\Gamma^s(TM)$ for every $s \in \mathbb{N}$ (which implies that V has a continuous injection in the Fréchet space of smooth vector fields), then, since any compactly supported co-current P (that is, any one-form with distributional coefficients) belongs to $\Gamma^{-s}(T^*M)$ for some $s \in \mathbb{N}$, it follows that any such P generates a locally minimizing normal geodesic starting at e . Therefore, the Fréchet Lie group $\mathcal{D}^\infty(M) = \cap_{s > d/2+1} \mathcal{D}^s(M)$ inherits of a strong right-invariant sub-Riemannian structure.

Remark 9.16 (Momentum formulation). We define the *momentum map* $\mu : T^*\mathcal{D}^s(M) \rightarrow \Gamma^{-s}(T^*M)$ by $\mu(\varphi, P) = (dR_\varphi)^*.P$. Let us assume that $k = \tilde{s} - s \geq 2$. Let $(\varphi(\cdot), P(\cdot))$ be a solution of the normal geodesic equations (9.6). We set $\mu(t) = \mu(\varphi(t), P(t))$ and $X(t) = K^V(dR_{\varphi(t)})^*.P(t) = K^V\mu(t)$.

Let $Y \in \Gamma^{s+1}(TM) \subset \Gamma^s(TM)$. Then $\mu(t)(Y) = P(t)(Y \circ \varphi(t))$. Omitting the dependence in t for the readability, we have

$$\dot{\mu}(Y) = P(\partial_\varphi(Y \circ \varphi)(X \circ \varphi) - \partial_\varphi(X \circ \varphi)(Y \circ \varphi)) = P([X, Y] \circ \varphi) = \mu([X, Y]) = \text{ad}_X^*\mu(Y).$$

In other words, $t \mapsto \mu(t)$ has a derivative $\dot{\mu}(t) = \text{ad}_{K^V\mu(t)}^*\mu(t)$ in the bigger space $\Gamma^{-s-1}(T^*M)$. Conversely, if a map $t \mapsto \mu(t) \in \Gamma^{-s}(T^*M)$ satisfies this equation, then the horizontal curve

2. This means, here, that $P(t) \in T_{\varphi(t)}^*\mathcal{D}^s(M)$ for almost every t .

$\varphi(\cdot)$ with logarithmic velocity $X(\cdot) = K^V \mu(\cdot)$ such that $\varphi(0) = e$ is a normal geodesic with initial momentum $\mu(0)$.

We recover in such a way the classical formula for critical points of the action on Lie groups (see [MR99]). The difficulty in this equation is due to the loss of a derivative. Indeed, since $\text{ad}_{K^V \mu}^*$ is an unbounded operator on $\Gamma^{-s}(T^*M)$, it may be hard to prove the existence of solutions of this differential equation without using a priori the result of Theorem 9.3.

Abnormal geodesic equations. The abnormal geodesic equations actually provide a Hamiltonian characterization of singular curves.

Definition 9.12. We define the abnormal Hamiltonian $H^0 : T^*\mathcal{D}^s(M) \times V \rightarrow \mathbb{R}$ by

$$H^0(\varphi, P, X) = P(X \circ \varphi) = P(\text{d}R_\varphi.X). \tag{9.7}$$

Since X is of class H^{s+k} , it follows that H^0 is of class \mathcal{C}^k . We have $\partial_X H^0(\varphi, P, X) = (\text{d}R_\varphi)^*.P$, where the partial derivative $\partial_X H^0 : T^*\mathcal{D}^s(M) \times V \rightarrow V^*$ is understood as a partial derivative along the fibers of a vector bundle. The symplectic gradient $\nabla^\omega H^0 : T^*\mathcal{D}^s(M) \times V \rightarrow TT^*\mathcal{D}^s(M)$ of H^0 , defined by the relation $\text{d}H^0 = \omega(\nabla^\omega H^0, \cdot)$, is given in canonical coordinates (φ, P) on $T^*\mathcal{D}^s(M)$ by $\nabla^\omega H^0(\varphi, P, X) = (X \circ \varphi, -(\partial_\varphi(X \circ \varphi))^*.P)$.

Theorem 9.4. We assume that $s > d/2 + 1$ and $k = \tilde{s} - s \geq 2$. Let $\varphi(\cdot) \in H^1(0, 1; \mathcal{D}^s(M))$ be an horizontal curve with logarithmic velocity $X(\cdot) = \dot{\varphi}(\cdot) \circ \varphi(\cdot)^{-1}$. Then $\varphi(\cdot)$ is a singular curve if and only if there exists an absolutely continuous mapping $P(\cdot)$ on $[0, 1]$ such that $P(t) \in T_{\varphi(t)}^*\mathcal{D}^s(M) \setminus \{0\}$ and

$$(\dot{\varphi}(t), \dot{P}(t)) = \nabla^\omega H^0(\varphi(t), P(t), X(t)), \tag{9.8}$$

$$\partial_X H^0(\varphi(t), P(t), X(t)) = (\text{d}R_{\varphi(t)})^*.P(t) = 0, \tag{9.9}$$

for almost every $t \in [0, 1]$.

Necessary conditions for optimality. The following result is a version of the Pontryagin maximum principle, which is adapted to our infinite-dimensional setting.

Theorem 9.5. We assume that $s > d/2 + 1$ and $k = \tilde{s} - s \geq 1$. Let $\varphi(\cdot) \in H^1(0, 1; \mathcal{D}^s(M))$ be a minimizing horizontal curve with logarithmic velocity $X(\cdot) = \dot{\varphi}(\cdot) \circ \varphi(\cdot)^{-1}$. Then $\varphi(\cdot)$ is a geodesic, and:

- either $\varphi(\cdot)$ is a normal geodesic, and in that case, there exists an absolutely continuous mapping $P(\cdot)$ on $[0, 1]$ such that $P(t) \in T_{\varphi(t)}^*\mathcal{D}^s(M)$ and (9.6) is satisfied almost everywhere on $[0, 1]$;
- or $\varphi(\cdot)$ is a singular curve, and in that case, there exists an absolutely continuous mapping $P(\cdot)$ on $[0, 1]$ such that $P(t) \in T_{\varphi(t)}^*\mathcal{D}^s(M) \setminus \{0\}$ and (9.8)-(9.9) are satisfied almost everywhere on $[0, 1]$;
- or $\varphi(\cdot)$ is elusive.

Remark 9.17. In finite dimension, it has been established in [RT05] and in [Agr09] that the set of end-points of normal geodesics is an open dense subset of the ambient manifold. Although such a result is not established in our infinite-dimensional context (it is all the more difficult than one has also to deal with elusive curves), it is however expected that, in some appropriate sense, the "generic" case of the above theorem is the first one (normal geodesics).

Besides, it has been established in [CJT06, CJT08] that, in finite-dimensional sub-Riemannian geometry, for generic (in a strong Whitney sense) horizontal distributions of rank greater than or equal to three, any singular curve cannot be minimizing. Although such a result seems currently out of reach in our infinite-dimensional setting, we expect however that, since our distribution is infinite-dimensional, there is no minimizing singular curve.

Proof of Theorems 9.3, 9.4 and 9.5. Let us first compute the adjoint of the derivative of the end-point mapping.

Lemma 9.8. *Let $\varphi(\cdot) \in \Omega_{\varphi_0}$ be an horizontal curve with logarithmic velocity $X(\cdot) = \dot{\varphi}(\cdot) \circ \varphi(\cdot)^{-1}$. We set $\varphi_1 = \varphi(1) = \text{end}_{\varphi_0}(\varphi(\cdot))$. For every $P_{\varphi_1} \in T_{\varphi_1}^* \mathcal{D}^s(M)$, the pull-back $(\text{d end}_{\varphi_0}(\varphi(\cdot)))^* . P_{\varphi_1}$ of P_{φ_1} by $\text{d end}_{\varphi_0}(\varphi(\cdot))$ is a continuous linear form on $L^2(0, 1; V)$, and can therefore be identified to an element of $L^2(0, 1; *)^* = L^2(0, 1; *)$, given by*

$$((\text{d end}_{\varphi_0}(\varphi(\cdot)))^* . P_{\varphi_1})(t) = \partial_X H^0(\varphi(t), P(t), X(t)) = (\text{d}R_{\varphi(t)})^* . P(t), \quad (9.10)$$

for almost every $t \in [0, 1]$, where $(\varphi(\cdot), P(\cdot)) : [0, 1] \rightarrow T^* \mathcal{D}^s(M)$ is the unique absolutely continuous mapping solution of $(\dot{\varphi}(t), \dot{P}(t)) = \nabla^\omega H^0(\varphi(t), P(t), X(t))$ on $[0, 1]$ and $P(1) = P_{\varphi_1}$.

Proof of Lemma 9.8. In local coordinates, the fibered part of the differential equation of the lemma is $\dot{P}(t) = -(\partial_\varphi(X(t) \circ \varphi(t)))^* . P(t)$, which is a linear differential equation. The Cauchy-Lipschitz theorem for linear differential equations therefore ensures global existence and uniqueness of a solution $P(\cdot)$ of class H^1 (and thus, absolutely continuous) such that $P(1) = P_{\varphi_1}$. Let us now prove the formula (9.10). The mapping $X(\cdot) \mapsto \varphi^X(\cdot)$ is of class \mathcal{C}^k , with $k \geq 1$, and is defined implicitly by the differential equation $\dot{\varphi}^X(t) - X(t) \circ \varphi^X(t) = 0$ for almost every $t \in [0, 1]$, with $\varphi^X(0) = \varphi_0$. To compute its derivative $\delta\varphi(\cdot) = \text{d}\varphi^X(X(\cdot)) . \delta X(\cdot)$ in the direction $\delta X(\cdot) \in L^2(0, 1; V)$, we differentiate this differential equation in local coordinates, and obtain that $\delta\dot{\varphi}(t) - \delta X(t) \circ \varphi^X(t) - \partial_\varphi(X(t) \circ \varphi^X(t)) . \delta\varphi(t) = 0$ for almost every $t \in [0, 1]$, with $\delta\varphi(0) = 0$. For every $\delta X(\cdot) \in L^2(0, 1; V)$, this Cauchy problem has a unique solution $\delta\varphi(\cdot)$, and we have $\text{d end}_{\varphi_0}(\varphi(\cdot)) . \delta X(\cdot) = \delta\varphi(1)$. Moreover, we have

$$\begin{aligned} \int_0^1 \partial_X H^0(\varphi(t), P(t), X(t)) . \delta X(t) dt &= \int_0^1 P(t) (\delta X(t) \circ \varphi(t)) dt \\ &= \int_0^1 P(t) (\delta\dot{\varphi}(t)) dt - \int_0^1 P(t) (\partial_\varphi(X(t) \circ \varphi(t)) . \delta\varphi(t)) dt \\ &= \int_0^1 (P(t) (\delta\dot{\varphi}(t)) + \dot{P}(t) (\delta\varphi(t))) dt \\ &= P_{\varphi_1} (\delta\varphi(1)) = (\text{d end}_{\varphi_0}(\varphi(\cdot)))^* . P_{\varphi_1}, \end{aligned}$$

which yields (9.10). □

Theorem 9.4 follows from Lemma 9.8 because $\varphi(\cdot)$ is a singular curve if and only if then there exists $P_{\varphi_1} \in T_{\varphi_1}^* \mathcal{D}^s(M) \setminus \{0\}$ such that $(\text{d end}_{\varphi_0}(\varphi(\cdot)))^* . P_{\varphi_1} = 0$.

Let φ_0 and φ_1 be two elements of $\mathcal{D}^s(M)$, and let $\varphi(\cdot) \in \Omega_{\varphi_0, \varphi_1}$. Since $\Omega_{\varphi_0, \varphi_1} = \text{end}_{\varphi_0}^{-1}(\{\varphi_1\})$, we have $T_{\varphi(\cdot)} \Omega_{\varphi_0, \varphi_1} \subset \ker(\text{d end}_{\varphi_0}(\varphi(\cdot)))$ (see Definition 9.8 for the definition of the set of all infinitesimal variations). Note that, if $\text{d end}_{\varphi_0}(\varphi(\cdot))$ were surjective, then $\Omega_{\varphi_0, \varphi_1}$ would be, locally at $\varphi(\cdot)$, a \mathcal{C}^k submanifold of Ω_{φ_0} , and then $T_{\varphi(\cdot)} \Omega_{\varphi_0, \varphi_1} = \ker(\text{d end}_{\varphi_0}(\varphi(\cdot)))$. But, as already said, in our context only the inclusion is true.

Since $T_{\varphi(\cdot)} \Omega_{\varphi_0, \varphi_1} \subset \ker(\text{d end}_{\varphi_0}(\varphi(\cdot)))$, if there is some $P_{\varphi_1} \in T_{\varphi_1}^* \mathcal{D}^s(M)$ such that $\text{d}A(\varphi(\cdot)) = (\text{d end}_{\varphi_0}(\varphi(\cdot)))^* . P_{\varphi_1}$, then $\varphi(\cdot)$ is a critical point of A restricted to $\Omega_{\varphi_0, \varphi_1}$ (and hence $\varphi(\cdot)$ is a geodesic steering φ_0 to φ_1). Conversely, according to the discussion done at the beginning of Section 9.1.4, this Lagrange multiplier relation is satisfied whenever $\varphi(\cdot)$ is a geodesic steering φ_0 to φ_1 which is neither singular nor elusive.

Besides, the differential $\text{d}A(\varphi(\cdot)) \in L^2(0, 1; V^*)$ is given by $\text{d}A(\varphi)(t) = \langle X(t), \cdot \rangle$. It follows from Lemma 9.8 that

$$(\text{d}A(\varphi(\cdot)) - \text{d end}_{\varphi_0}(\varphi(\cdot))^* . P_{\varphi_1})(t) = \langle X(t), \cdot \rangle - \partial_X H^0(\varphi(t), P(t), X(t)),$$

for almost every $t \in [0, 1]$, where $P(\cdot) : [0, 1] \rightarrow T^*\mathcal{D}^s(M)$ is the unique solution of $(\dot{\varphi}(t), \dot{P}(t)) = \nabla^\omega H^0(\varphi(t), P(t), X(t))$ on $[0, 1]$ and $P(1) = P_{\varphi_1}$. Defining the total Hamiltonian by

$$\begin{aligned} H : T^*\mathcal{D}^s(M) \times V &\rightarrow \mathbb{R} \\ (\varphi, P, X) &\mapsto P(X \circ \varphi) - \frac{1}{2}\langle X, X \rangle, \end{aligned}$$

we have $\nabla^\omega H = \nabla^\omega H^0$, and $(dA(\varphi(\cdot)) - (d\text{end}_{\varphi_0}(\varphi(\cdot)))^*.P_{\varphi_1})(t) = -\partial_X H(\varphi(t), P(t), X(t))$ for almost every $t \in [0, 1]$. We have obtained the following lemma.

Lemma 9.9. *Let $\varphi_0 \in \mathcal{D}^s(M)$. Let $X \in L^2(0, 1; V)$ be the logarithmic velocity of an horizontal curve $\varphi(\cdot) \in H^1(0, 1; \mathcal{D}^s(M))$ starting at φ_0 . The two following statements are equivalent:*

- *There exists an absolutely continuous fibered mapping $(\varphi(\cdot), P(\cdot)) : [0, 1] \rightarrow T^*\mathcal{D}^s(M)$ such that $(\dot{\varphi}(t), \dot{P}(t)) = \nabla^\omega H(\varphi(t), P(t), X(t))$ for almost every $t \in [0, 1]$ and $P(1) = P_{\varphi_1}$, and such that*

$$0 = \partial_X H(\varphi(t), P(t), X(t)) = (dR_{\varphi(t)})^*.P(t) - \langle X(t), \cdot \rangle,$$

for almost every $t \in [0, 1]$.

- *There exists $P_{\varphi_1} \in T_{\varphi_1}^*\mathcal{D}^s(M) \setminus \{0\}$ such that $dA(\varphi(\cdot)) = (d\text{end}_{\varphi_0}(\varphi(\cdot)))^*.P_{\varphi_1}$.*

Under any of those statements, $\varphi(\cdot)$ is a (normal) geodesic.

For every fixed $(\varphi, P) \in T^*\mathcal{D}^s(M)$, the equation $\partial_X H(\varphi, P, X) = 0$ yields $\langle X, \cdot \rangle = (dR_\varphi)^*.P$, whose unique solution is given by $X(\varphi, P) = K^V(dR_\varphi)^*.P$. Then we obtain the normal Hamiltonian $h : T^*\mathcal{D}^s(M) \rightarrow \mathbb{R}$ by setting

$$h(\varphi, P) = H(\varphi, P, X(\varphi, P)) = \frac{1}{2}P(X(\varphi, P) \circ \varphi) = \frac{1}{2}\langle X(\varphi, P), X(\varphi, P) \rangle.$$

Theorems 9.3 and 9.5 follow.

9.1.5 Examples of geodesic equations in $\mathcal{D}(\mathbb{R}^d)$

In this section, we assume that $M = \mathbb{R}^d$. Let s_0 be the smallest integer such that $s_0 > d/2$. It is easy to prove that for every integer $s \geq s_0 + 1$, the group $\mathcal{D}^s(\mathbb{R}^d)$ coincides with the set of diffeomorphisms φ of \mathbb{R}^d such that $\varphi - \text{Id}_{\mathbb{R}^d} \in H^s(\mathbb{R}^d, \mathbb{R}^d)$, and is an open subset of the affine Hilbert space $\text{Id}_{\mathbb{R}^d} + H^s(\mathbb{R}^d, \mathbb{R}^d)$ (endowed with the induced topology, see Section 9.1.1).

Since we are in \mathbb{R}^d , we have $T\mathcal{D}^s(\mathbb{R}^d) = \mathcal{D}^s(\mathbb{R}^d) \times H^s(\mathbb{R}^d, \mathbb{R}^d)$ and $T^*\mathcal{D}^s(\mathbb{R}^d) = \mathcal{D}^s(\mathbb{R}^d) \times H^{-s}(\mathbb{R}^d, (\mathbb{R}^d)^*)$. Therefore, a momentum $P \in T_{\varphi}^*\mathcal{D}^s(\mathbb{R}^d)$ is a 1-form on \mathbb{R}^d with distributional coefficients in $H^{-s}(\mathbb{R}^d)$, denoted by $P = P_1 dx^1 + \dots + P_d dx^d = (P_1, \dots, P_d)$. For a vector field X , we can also write $X = X^1 e_1 + \dots + X^d e_d = (X^1, \dots, X^d)$, with (e_i) the canonical frame of \mathbb{R}^d .

The Euclidean inner product of two vectors v and w of \mathbb{R}^d is denoted by $v \cdot w$. The notation v^T stands for the linear form $w \mapsto v \cdot w$. Conversely, for a linear form $p \in (\mathbb{R}^d)^*$, we denote by p^T the unique vector v in \mathbb{R}^d such that $p = v^T$.

These notations are extended to vector fields and to 1-forms with distributional coefficients, by setting $(X^1 e_1 + \dots + X^d e_d)^T = X^1 dx^1 + \dots + X^d dx^d$ and $(P_1 dx^1 + \dots + P_d dx^d)^T = P_1 e_1 + \dots + P_d e_d$.

Spaces of vector fields with Gaussian kernels. Let V be the Hilbert space of vector fields on \mathbb{R}^d associated with the reproducing kernel $K : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \text{End}((\mathbb{R}^d)^*, \mathbb{R}^d)$ defined by $K(x, y)p = e(x, y)p^T$, for every $p \in (\mathbb{R}^d)^*$, with $e(x, y) = e^{-\frac{|x-y|^2}{2\sigma}}$, for some $\sigma > 0$. This

space is widely used in shape deformation analysis (see [TY05, TY11, You10]). Note that the mapping $(x, y) \mapsto e(x - y)$ is equal (up to a multiplying scalar) to the heat kernel at time σ .

The elements of V are analytic, and all their derivatives decrease exponentially at infinity, hence $V \subset H^s(\mathbb{R}^d, \mathbb{R}^d)$ for every $s \in \mathbb{N}$. Moreover, for every 1-form with tempered distributional coefficients P , the vector field X such that $\langle X, \cdot \rangle = P$ on V is given by $X(x) = \int_{\mathbb{R}^d} e(x-y)P(y)^T dy$, for every $x \in \mathbb{R}^d$. For every $(\varphi, P) \in T^*\mathcal{D}^s(\mathbb{R}^d) = \mathcal{D}^s(\mathbb{R}^d) \times H^{-s}(\mathbb{R}^d, \mathbb{R}^d)$, the solution of $\partial_X H(\varphi, P, X) = 0$ is $X(\varphi, P)(x) = \int_{\mathbb{R}^d} e(x - \varphi(y))P(y)^T dy$. Therefore, the normal Hamiltonian is given by

$$h(\varphi, P) = \frac{1}{2}P(X(\varphi, P) \circ \varphi) = \frac{1}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} e(\varphi(x) - \varphi(y))P(x) \cdot P(y) dy dx.$$

Since $e(x - y) = e(y - x)$ and $de(x) \cdot v = -\frac{1}{\sigma}e(x)(x \cdot v)$, we get that the normal geodesic equations are written in the distributional sense as

$$\begin{aligned} \partial_t \varphi(t, x) &= \int_{\mathbb{R}^d} e(\varphi(t, x) - \varphi(t, y))P(t, y)^T dy, \\ \partial_t P(t, x) &= \frac{1}{\sigma} \int_{\mathbb{R}^d \times \mathbb{R}^d} e(\varphi(t, x) - \varphi(t, y))(\varphi(t, x) - \varphi(t, y))^T P(t, x) \cdot P(t, y) dy. \end{aligned} \tag{9.11}$$

By inspecting the above equations, it is not difficult to prove that the support of $\partial_t P$ is contained in that of P , and that, if $P \in H^{-s'}(\mathbb{R}^d, \mathbb{R}^d)$ for some $s' \in \mathbb{N}$, then so is $\partial_t P$. In other words, the Hamiltonian normal geodesic equations preserve the support and the order of the distributional 1-form P .

A particularly simple example is when $P = a \otimes \delta_{x_0}$, with $a \in (\mathbb{R}^d)^*$ (that is, $P(X) = a(X(x_0))$). In that case, we have $h(\varphi, P) = \frac{1}{2}a \cdot a$, hence $\partial_\varphi h(\varphi, P) = 0$ and therefore P is constant along the geodesic flow, and $\partial_t \varphi(t, x) = e^{-\frac{|\varphi(t, x) - \varphi(t, x_0)|^2}{2\sigma}} a^T$. Note that, in this case, $\varphi(t, x_0) = x_0 + ta^T$ is a straight line with constant speed $a \cdot a$.

Gaussian kernels for sub-Riemannian distributions in \mathbb{R}^d . Let X_1, \dots, X_k be smooth pointwise linearly independent vector fields on \mathbb{R}^d , with $k \leq d$, which are bounded as well as all their derivatives. These vector fields generate a sub-Riemannian structure on \mathbb{R}^d (see [BR96, Mon02] for this classical construction), with horizontal curves $t \mapsto x(t)$ satisfying the differential equation $\dot{x}(t) = \sum_{j=1}^k u_j(t)X_j(x(t))$ for almost every t , with $u_j \in L^2(0, 1; \mathbb{R})$ for every $j = 1, \dots, k$. We consider the kernel $K(x, y)p = e(x, y) \sum_{j=1}^k p(X_j(y))X_j(x)$.

Let us first prove that K is the reproducing kernel of a Hilbert space of vector fields V . For any compactly supported 1-form P with distributional coefficients, an easy computation gives $K(x, \cdot)P = \sum_{j=1}^k (\int_{\mathbb{R}^d} e(x-y)P(y)(X_j(y)) dy) X_j(x)$, where $P(y)(X_j(y)) = P_1(y)X_j^1(y) + \dots + P_d(y)X_j^d(y)$. The vector $K(x, \cdot)P$ is well defined since the vector fields $y \mapsto e(x-y)X_j(y)$ are smooth and all their derivatives decrease exponentially at infinity, and P is a compactly supported 1-form with distributional coefficients. Then we have

$$P(K(\cdot, \cdot)P) = \sum_{j=1}^k \int_{\mathbb{R}^d} e(x-y)P(y)(X_j(y))P(x)(X_j(x)) dy dx.$$

In order to check that K is the reproducing kernel of a Hilbert space of vector fields, it is sufficient to check that $P(K(\cdot, \cdot)P) \geq 0$ for any such 1-form P , and $P(K(\cdot, \cdot)P) = 0$ if and only if the vector field $x \mapsto K(x, \cdot)P$ is identically equal to 0 (see [You10]). For such a P , we set $p_j(x) = P(x)(X_j(x))$, for $j = 1, \dots, k$. Then we have $K(x, \cdot)P = \sum_{j=1}^k \int_{\mathbb{R}^d} e(x-y)p_j(y) dy X_j(x)$, and therefore $P(K(\cdot, \cdot)P) = \sum_{j=1}^k \int_{\mathbb{R}^d} e(x-y)p_j(y)p_j(x) dy dx$. Since $(x, y) \mapsto e(x-y)$ is the heat kernel at time σ , it follows that $\int_{\mathbb{R}^d} e(x-y)T(y)T(x) dy dx \geq 0$ for any

distribution T on \mathbb{R}^d , with equality if and only if $T = 0$. Therefore, $P(K(\cdot, \cdot)P)$ is nonnegative, and is equal to 0 if and only if $p_1 = \dots = p_k = 0$, in which case $K(x, \cdot)P = 0$ for every x in \mathbb{R}^d . We have thus proved that K is the reproducing kernel of a Hilbert space of vector fields V .

Moreover, any element of V can be written as $x \mapsto u_1(x)X_1(x) + \dots + u_k(x)X_k(x)$ (and is horizontal with respect to the sub-Riemannian structure on \mathbb{R}^d induced by the X_i), where the u_j 's are analytic functions with all derivatives decreasing exponentially at infinity. In particular, the inclusions $V \subset H^s(\mathbb{R}^d, \mathbb{R}^d)$ are continuous for every integer s .

Remark 9.18. It could be more natural to replace $e(x - y)$ with the sub-Riemannian heat kernel associated with the sub-Laplacian $X_1^2 + \dots + X_k^2$, but since such kernels are much harder to compute, we will keep the Euclidean heat kernel.

For every $(\varphi, P) \in T^*\mathcal{D}^s(\mathbb{R}^d) = \mathcal{D}^s(\mathbb{R}^d) \times H^{-s}(\mathbb{R}^d, \mathbb{R}^d)$, the solution of $\partial_X H(\varphi, P, X) = 0$ is $X(\varphi, P)(x) = K(x, \varphi(\cdot))P = \sum_{j=1}^k (\int_{\mathbb{R}^d} e(x - \varphi(y))P(y)(X_j(\varphi(y))) dy) X_j(x)$. The normal Hamiltonian is given by

$$h(\varphi, P) = \frac{1}{2} \sum_{j=1}^k \int_{\mathbb{R}^d \times \mathbb{R}^d} e(\varphi(x) - \varphi(y))P(y)(X_j(\varphi(y)))P(x)(X_j(\varphi(x))) dy dx.$$

We infer that the normal geodesic equations are written as

$$\begin{aligned} \partial_t \varphi(t, x) &= \sum_{j=1}^k \left(\int_{\mathbb{R}^d} e(\varphi(t, x) - \varphi(t, y))P(t, y)(X_j(\varphi(t, y))) dy \right) X_j(\varphi(t, x)), \\ \partial_t P(t, x) &= \frac{1}{\sigma} \sum_{j=1}^k \int_{\mathbb{R}^d} e(\varphi(t, x) - \varphi(t, y))(\varphi(t, x) - \varphi(t, y))^T P(t, y)(X_j(\varphi(t, y)))P(t, x)(X_j(\varphi(t, x))) dy \\ &\quad - \sum_{j=1}^k \int_{\mathbb{R}^d} e(\varphi(t, x) - \varphi(t, y))P(t, y)(X_j(\varphi(t, y)))P(x)(dX_j(\varphi(t, x))) dy. \end{aligned}$$

Here, we have $P(x)(dX_j(\varphi(t, x))) = P_1(x)dX_j^1(\varphi(t, x)) + \dots + P_d(x)dX_j^d(\varphi(t, x))$. It is easy to check that this geodesic flow preserves the support of P and its regularity.

In the simple case where $P = a \otimes \delta_{x_0}$ with $a \in (\mathbb{R}^d)^*$, we have

$$X(\varphi, P)(x) = e(x - \varphi(x_0)) \sum_{j=1}^k a(X_j(\varphi(x_0)))X_j(x).$$

Since $e(0) = 1$, we get $h(\varphi, a \otimes \delta_{x_0}) = \frac{1}{2} \sum_{j=1}^k a(X_j(\varphi(x_0)))^2$. It is interesting to note that $h(\varphi, a \otimes \delta_{x_0}) = h^\Delta(\varphi(x_0), a)$, where h^Δ is the normal hamiltonian for the sub-Riemannian structure induced by the X_j 's on \mathbb{R}^d . Now we have $\partial_\varphi h(\varphi, a \otimes \delta_{x_0}) = \left(\sum_{r=1}^k a(X_r(\varphi(x_0)))a(dX_{r, \varphi(x_0)}) \right) \otimes \delta_{x_0}$. We see that the subbundle $\mathcal{D}^s(\mathbb{R}^d) \times (\mathbb{R}^{d*} \otimes \delta_{x_0}) \subset T^*\mathcal{D}^s(\mathbb{R}^d)$ is invariant under the Hamiltonian geodesic flow. Hence, if $t \mapsto (\varphi(t), P(t))$ is solution of the normal geodesic equations and $P(0) = a(0) \otimes \delta_{x_0}$, then there exists a curve $t \mapsto a(t) \in \mathbb{R}^d$ such that $P(t) = a(t) \otimes \delta_{x_0}$ for every t . Denoting by $x(t) = \varphi(t, x)$ for $x \in \mathbb{R}^d$ and by $x_0(t) = \varphi(t, x_0)$, we finally obtain

$$\dot{x}(t) = e(x(t) - x_0(t)) \sum_{j=1}^k a(t)(X_j(x_0(t)))X_j(x(t)), \quad \dot{a}(t) = - \sum_{j=1}^k a(X_j(x_0(t)))a(dX_j(x_0(t))).$$

In particular, if $x(0) = x_0$, we get

$$\dot{x}(t) = \sum_{j=1}^k a(X_j(x(t)))X_j(x(t)), \quad \dot{a}(t) = - \sum_{j=1}^k a(X_j(x(t)))a(dX_j(x(t))).$$

We recover the equations satisfied by a normal geodesic in \mathbb{R}^d for the sub-Riemannian structure induced by the X_j 's on \mathbb{R}^d , with initial momentum $a(0)$.

9.2 Shape spaces

9.2.1 Definition

Throughout the section, M is a smooth compact Riemannian manifold of dimension d . Let s_0 be the smallest integer such that $s_0 > d/2$. Shape spaces in M are spaces acted upon by the group of diffeomorphisms of M in a way that is compatible with its inverse limit Hilbert Lie group structure. The following definition is adapted from that of [ATTY13].

Definition 9.13. *Let \mathcal{S} be a Banach manifold modeled on a Banach space \mathcal{B} , equipped with a compatible complete distance $d_{\mathcal{S}}$. Assume that $\mathcal{D}^s(M)$, with $s = s_0 + \ell$ and $\ell \in \mathbb{N}^*$, acts continuously on M , according to the action*

$$\begin{aligned} \mathcal{D}^s(M) \times \mathcal{S} &\rightarrow \mathcal{S} \\ (\varphi, q) &\mapsto \varphi \circ q = R_q(\varphi) \end{aligned}$$

We say that \mathcal{S} is a shape space of order ℓ in M if the following three conditions are satisfied:

1. The action is semi-Lipschitz, that is, for every $q \in M$ there exists $\gamma_q > 0$ such that $d_{\mathcal{S}}(\varphi_1 \circ q, \varphi_2 \circ q) \leq \gamma_q d_{H^s(M,M)}(\varphi_1, \varphi_2)$ for all $(\varphi_1, \varphi_2) \in (\mathcal{D}^s(M))^2$.
2. For every $q \in \mathcal{S}$, the mapping $\varphi \mapsto \varphi \circ q$ is differentiable at e . This differential is denoted by ξ_q or by $\cdot \circ q$, and is called the infinitesimal action of $\Gamma^s(TM)$.
3. For every $k \in \mathbb{N}$, the mappings

$$\begin{aligned} R : \mathcal{S} \times \mathcal{D}^{s+k}(M) &\longrightarrow \mathcal{S} & \xi : \mathcal{S} \times \Gamma^{s+k}(TM) &\longrightarrow T\mathcal{S} \\ (q, \varphi) &\longmapsto \varphi \circ q & (q, X) &\longmapsto \xi_q(X) \end{aligned} \tag{9.12}$$

are of class \mathcal{C}^k .

Note that, according to this definition, we have the following immediate properties:

- the semi-Lipschitz property implies that the operator $\xi_q : \Gamma^s(TM) \rightarrow T_q\mathcal{S}$ is continuous;
- the mappings $q \mapsto \varphi \circ q$ and $q \mapsto \xi_q(X)$ are of class \mathcal{C}^k , for every φ in $\mathcal{D}^{s+k}(M)$ and bounded vector field X of class H^{s+k} ;
- the action by left composition of $\mathcal{D}^{s_0+\ell}(M)$ on itself makes it a shape space of order ℓ .

Definition 9.14. *We say that $q \in \mathcal{S}$ has a compact support if there exists a compact subset U of M such that, in the first condition of Definition 9.13, we have $d_{\mathcal{S}}(\varphi_1 \circ q, \varphi_2 \circ q) \leq \gamma_q d_{H^s(M,M)}(\varphi_1|_U, \varphi_2|_U)$.*

In other words, q has a compact support if $\varphi \circ q$ depends only on the restriction of φ to a compact subset of M .

Example 9.3. The most usual shape spaces are the manifolds of all differentiable (or even just topological) embeddings $q : S \rightarrow M$ of a compact Riemannian manifold S into M . In this case, $\mathcal{D}^{s_0+1}(M)$ acts on \mathcal{S} by composition. This example motivates the notation $\varphi \circ q$ chosen above to denote the action of φ on q .

Indeed, let $\ell \in \mathbb{N}^*$ and let S be a smooth compact Riemannian manifold. The space $\mathcal{S} = \mathcal{C}^\ell(S, M)$ of embeddings of S in M is a Banach manifold. The group of diffeomorphisms $\mathcal{D}^s(M)$, with $s = s_0 + \ell$, acts continuously on M by the left composition $(\varphi, q) \mapsto \varphi \circ q$, with infinitesimal action $(X, q) \mapsto X \circ q$. Then \mathcal{S} is a shape space (see [ATTY13]).

Another example is given by $\mathcal{S} = H^{r_0+\ell}(S, M)$, where r_0 is the smallest integer greater than $\dim(S)/2$. The restriction theorems for an element of $H^{s_0+\ell}(M, M)$ to a submanifold show that \mathcal{S} is also a shape space for the action by composition.

A particularly interesting case is obtained when taking $S = (s_1, \dots, s_n)$, $n \in \mathbb{N}$, a finite set (i.e., a zero-dimensional compact manifold). In that case, we obtain $\mathcal{S} = \mathcal{C}^\ell(S, M)$, which is identified with the space $\text{Lmk}_n(M) = \{(x_1, \dots, x_n) \in M^n \mid x_i \neq x_j \text{ if } i \neq j\}$, called the space of n landmarks in M . For $q = (x_1, \dots, x_n)$, the action of $\mathcal{D}^{s_0+1}(M)$ is given by $\varphi \circ q = (\varphi(x_1), \dots, \varphi(x_n))$. For a vector field X of class H^{s_0+1} on M , the infinitesimal action of X at q is given by $\xi_q(X) = (X(x_1), \dots, X(x_n))$. Spaces of landmarks are all shape spaces of order 1 (see [ATTY13] for more details).

All elements of all shape spaces described in this example have compact support.

Remark 9.19. Let \mathcal{S}_1 and \mathcal{S}_2 be shape spaces of respective orders ℓ_1 and ℓ_2 . Then the space $\mathcal{S}_1 \times \mathcal{S}_2$ is a shape space of order $\max(\ell_1, \ell_2)$, with the diagonal action $\varphi \circ (q_1, q_2) = (\varphi \circ q_1, \varphi \circ q_2)$.

Remark 9.20. Let \mathcal{S} be a shape space of order ℓ . Then $T\mathcal{S}$ is a shape space of order $\ell + 1$, with the action of $\mathcal{D}^{s_0+\ell_1+1}(M)$ on $T\mathcal{S}_1$ defined by $\varphi \circ (q, v) = (\varphi \circ q, \partial_q(\varphi \circ q)(v))$.

Curves on \mathcal{S} . We consider curves on \mathcal{S} induced by curves on $\mathcal{D}^s(M)$ as follows. For every shape $q_0 \in \mathcal{S}$, and for every $\varphi(\cdot) \in H^1(0, 1; \mathcal{D}^s(M))$, the curve of shapes $q(\cdot) = \varphi(\cdot) \circ q_0$ belongs to $H^1(0, 1; \mathcal{S})$ and is solution of the Cauchy problem

$$\dot{q}(t) = \xi_{q(t)}X(t) = X(t) \circ q(t), \quad q(0) = q_0, \quad (9.13)$$

where $X(\cdot) = \dot{\varphi}(\cdot) \circ \varphi(\cdot)^{-1} \in L^2(0, 1; \Gamma^s(M))$ is the logarithmic velocity of $\varphi(\cdot)$. We say that the curve $\varphi(\cdot) \in H^1(0, 1; \mathcal{D}^s(M))$ is a *lift* of the curve $q(\cdot) \in H^1(0, 1; \mathcal{S})$.

9.2.2 Sub-Riemannian structure on shape spaces

Let \mathcal{S} be a shape space of order $\ell \geq 1$ in M , and $s = s_0 + \ell$ as in Definition 9.13.

Let $\tilde{s} > s$, and let $(V, \langle \cdot, \cdot \rangle)$ be a Hilbert subspace of $\Gamma^{\tilde{s}}(TM) = T_e\mathcal{D}^{\tilde{s}}(M)$ with continuous inclusion. According to Section 9.1.2 (see in particular Definition 9.2), we consider the strong right-invariant sub-Riemannian structure induced by V on $\mathcal{D}^s(M)$.

This sub-Riemannian structure on $\mathcal{D}^s(M)$ induces itself a sub-Riemannian structure on \mathcal{S} , via the infinitesimal action ξ , as follows. For every $q \in \mathcal{S}$, we set $\mathcal{H}_q^S = \{X \circ q \mid X \in V\}$. The resulting distribution of subspaces $\mathcal{H}^S = \cup_{q \in \mathcal{S}} \mathcal{H}_q^S$ is the horizontal distribution induced by V on the shape space. The infinitesimal action $X \mapsto X \circ q = \xi_q(X)$ induces a Hilbert norm on \mathcal{H}_q^S by

$$\|w\|_{\mathcal{H}_q^S}^2 = \inf\{\langle X, X \rangle \mid w = X \circ q, X \in V\} \quad (9.14)$$

Indeed, if $X = X_0 + X_1$ with $X_0 \in \ker \xi_q$ and $X_1 \in (\ker \xi_q)^\perp$, then $\|w\|_{\mathcal{H}_q^S}^2 = \langle X_1, X_1 \rangle$, and therefore $\|\cdot\|_{\mathcal{H}_q^S}$ is a Hilbert norm on \mathcal{H}_q^S with inner product $\langle \cdot, \cdot \rangle_q$, and $\xi_q : (\ker \xi_q)^\perp \rightarrow \mathcal{H}_q^S$ is an isometry. With a slight abuse of notation, in what follows we will denote by $\xi_q^{-1} : \mathcal{H}_q^S \rightarrow (\ker \xi_q)^\perp$ the inverse of this isometry, which is actually the pseudo-inverse of ξ_q . Note that $\xi_q^{-1}\xi_q$ is the orthogonal projection onto $(\ker \xi_q)^\perp$.

We have thus shown that the subbundle \mathcal{H}^S of $T\mathcal{S}$, equipped with the fibered Hilbert metric (9.14), induces a sub-Riemannian structure on \mathcal{S} .

Remark 9.21. The continuity of $X \mapsto X \circ q$ implies that the topology on \mathcal{H}_q^S defined by the norm (9.14) is finer than the one induces from the inclusion $\mathcal{H}_q^S \hookrightarrow T_q\mathcal{S}$. The topologies coincide if and only if \mathcal{H}_q^S is closed.

Definition 9.15. An horizontal curve on \mathcal{S} is a curve $q(\cdot) \in H^1(0, 1; \mathcal{S})$ such that $\dot{q}(t) \in \mathcal{H}_{q(t)}^{\mathcal{S}}$ for almost every $t \in [0, 1]$. Equivalently, it is a curve such that $\dot{q}(t) = \xi_{q(t)} X(t) = X(t) \circ q(t)$ for almost every $t \in [0, 1]$ for some $X \in L^2(0, 1; V)$.

The lift of an horizontal curve $q(\cdot)$ on \mathcal{S} is an horizontal curve $\varphi(\cdot)$ on $\mathcal{D}^s(M)$, called horizontal lift of $q(\cdot)$, satisfying $q(\cdot) = \varphi(\cdot) \circ q(0)$.

Definition 9.16. The sub-Riemannian length $L^{\mathcal{S}}(q(\cdot))$ and the sub-Riemannian action $A^{\mathcal{S}}(q(\cdot))$ of an horizontal curve $q(\cdot) \in H^1(0, 1; \mathcal{S})$ are respectively defined by

$$L^{\mathcal{S}}(q(\cdot)) = \int_0^1 \|\dot{q}(t)\|_{\mathcal{H}_{q(t)}^{\mathcal{S}}} dt \quad \text{and} \quad A^{\mathcal{S}}(q(\cdot)) = \frac{1}{2} \int_0^1 \|\dot{q}(t)\|_{\mathcal{H}_{q(t)}^{\mathcal{S}}}^2 dt.$$

The sub-Riemannian distance $d_{SR}^{\mathcal{S}}$ between two elements q_0 and q_1 of \mathcal{S} is defined as the infimum of the length of horizontal curves steering q_0 to q_1 , with the agreement that $d_{SR}(q_0, q_1) = +\infty$ whenever there is no horizontal curve steering q_0 to q_1 .

An horizontal curve $q(\cdot) : [0, 1] \rightarrow \mathcal{S}$ is said to be minimizing if $d_{SR}^{\mathcal{S}}(q(0), q(1)) = L(q(\cdot))$.

For all $(q_0, q_1) \in \mathcal{S}^2$, we denote by $\Omega_{q_0}^{\mathcal{S}}$ the set of all horizontal curves $q(\cdot) \in H^1(0, 1; \mathcal{S})$ such that $q(0) = q_0$, and by $\Omega_{q_0, q_1}^{\mathcal{S}}$ is the set of all curves $q(\cdot) \in \Omega_{q_0}^{\mathcal{S}}$ such that $q(1) = q_1$.

Definition 9.17. A geodesic $q(\cdot)$ is an horizontal curve which is a critical point of the action mapping $A^{\mathcal{S}}$ restricted to $\Omega_{q(0), q(1)}^{\mathcal{S}}$.

We have the following result.

Lemma 9.10. For every $q_0 \in \mathcal{S}$, for every $q(\cdot) \in \Omega_{q_0}^{\mathcal{S}}$, for every horizontal lift $\varphi(\cdot)$ of $q(\cdot)$, we have $A^{\mathcal{S}}(q(\cdot)) \leq A(\varphi(\cdot))$. Moreover, for every $\varphi_0 \in \mathcal{D}^s(M)$ such that $\varphi_0 \circ q_0 = q_0$, there exists a unique horizontal lift $\varphi(\cdot) \in \Omega_{\varphi_0}$ of $q(\cdot)$ such that $A^{\mathcal{S}}(q(\cdot)) = A(\varphi(\cdot))$, and its logarithmic velocity is given by $X(\cdot) = \xi_{q(\cdot)}^{-1} \dot{q}(\cdot)$. It is called the minimizing lift of $q(\cdot)$ starting at φ_0 .

Proof. For almost every $t \in [0, 1]$, we set $X(t) = \xi_{q(t)}^{-1} \dot{q}(t) \in V$. This defines a measurable function. Since ξ_q^{-1} is an isometry onto its image, we have $\frac{1}{2} \int_0^1 \langle X(t), X(t) \rangle dt = \frac{1}{2} \int_0^1 \|\dot{q}(t)\|_{\mathcal{H}_{q(t)}^{\mathcal{S}}}^2 dt = A^{\mathcal{S}}(q(\cdot)) < \infty$, and therefore $X(\cdot) \in L^2(0, 1; V)$, and for every $\varphi_0 \in \mathcal{D}^s(M)$ there exists a unique horizontal curve $\varphi(\cdot) \in H^1(0, 1; \mathcal{D}^s(M))$ with logarithmic velocity $X(\cdot)$. If $\varphi_0 \circ q_0 = q_0$ then $\varphi(\cdot)$ is an horizontal lift of $q(\cdot)$ and satisfies $A^{\mathcal{S}}(q(\cdot)) = A(\varphi(\cdot))$.

Conversely, let $\varphi(\cdot)$ be an horizontal lift of $q(\cdot)$ with logarithmic velocity $X(\cdot)$. Then $\dot{q}(t) = X(t) \circ q(t)$ for almost every $t \in [0, 1]$, and $\|\dot{q}(t)\|_{\mathcal{H}_{q(t)}^{\mathcal{S}}} \leq \langle X(t), X(t) \rangle$, with equality if and only if $X(t) = \xi_{q(t)}^{-1} \dot{q}(t)$. Therefore, $A(q) \leq A(\varphi)$ with equality if and only if $X(t) = \xi_{q(t)}^{-1} \dot{q}(t)$ for almost every $t \in [0, 1]$. The lemma follows. \square

A consequence is that the minimizing lift of a minimizing geodesic on \mathcal{S} is a minimizing geodesic on $\mathcal{D}^s(M)$. The metric and geodesic completeness of \mathcal{S} for $d^{\mathcal{S}}$ follow (as in Theorem 9.1).

Theorem 9.6. The sub-Riemannian distance $d_{SR}^{\mathcal{S}}$ is indeed a distance (taking its values in $[0, +\infty]$), that is, $d_{SR}^{\mathcal{S}}(q_0, q_1) = 0$ implies $q_0 = q_1$.

Moreover, for every $q_0 \in \mathcal{S}$ with compact support, denoting by $\mathcal{S}_{q_0} = \{\varphi \circ q_0 \mid \varphi \in \mathcal{D}^s(M)\}$ be the orbit of q_0 , the metric space $(\mathcal{S}_{q_0}, d_{SR}^{\mathcal{S}})$ is complete and any two states $q, q' \in \mathcal{S}_{q_0}$ can be connected by a minimizing horizontal curve.

Proof. Let $(q_0, q_1) \in \mathcal{S}^2$ with $d_{SR}^{\mathcal{S}}(q_0, q_1) < \infty$, and let $(q^n(\cdot))_{n \in \mathbb{N}}$ be a minimizing sequence of horizontal curves in $H^1(0, 1; \mathcal{S})$ such that $q^n(0) = q_0$ and $q^n(1) = q_1$ for every n , with $A^{\mathcal{S}}(q^n(\cdot))$ converging to $2d_{SR}^{\mathcal{S}}(q_0, q_1)^2$. For every n , we define $X^n(\cdot) = \xi_{q^n(\cdot)}^{-1}(\dot{q}^n(\cdot)) \in$

$L^2(0, 1; V)$, and we define $\varphi^n(\cdot) = \varphi^{X^n}(\cdot)$. For every n , $\varphi^n(\cdot)$ is the minimizing lift of $q^n(\cdot)$ such that $\varphi^n(0) = e$. Up to some subsequence, we have $X^n(\cdot) \rightharpoonup \bar{X}(\cdot)$ in $L^2(0, 1; V)$. Let $\bar{\varphi}(\cdot) = \varphi^{\bar{X}}(\cdot)$. Since $\tilde{s} > s$, Lemma 9.2 implies that $d_{H^s(U, M)}(\varphi^n(1)|_U, \bar{\varphi}(1)|_U) \rightarrow 0$ as $n \rightarrow +\infty$ for every compact subset U of M . Since q_0 has a compact support, defining the horizontal curve $\bar{q}(\cdot) = \bar{\varphi}(\cdot) \circ q_0$ in \mathcal{S} , we get

$$d_{\mathcal{S}}(q_1, \bar{q}(1)) = d_{\mathcal{S}}(q^n(1), \bar{q}(1)) \leq \gamma d_{H^s(U, M)}(\varphi^n(1)|_U, \bar{\varphi}(1)|_U) \rightarrow 0.$$

Therefore, $d_{\mathcal{S}}(q_1, \bar{q}(1)) = 0$ and $\bar{q}(1) = q_1$. But

$$A^{\mathcal{S}}(\bar{q}(\cdot)) \leq A(\bar{X}(\cdot)) \leq \liminf_{n \rightarrow \infty} A(X^n(\cdot)) = \liminf_{n \rightarrow \infty} A^{\mathcal{S}}(q^n(\cdot)) = 2d_{SR}^{\mathcal{S}}(q_0, q_1).$$

The proof for the completeness of \mathcal{S}_{q_0} follows a similar line. \square

The momentum viewpoint As mentioned in Example 9.2, if V is a Hilbert space with continuous injection in $\Gamma^s(TM)$, then V is uniquely characterized by a positive semi-definite convolution kernel K , the reproducing kernel of V . K is a section of the vector bundle $\text{End}(T^*M, TM) \rightarrow M \times M$ such that for any continuous 1-form $p : M \rightarrow T^*M$,

$$\int_{M \times M} p_x(K(y, x)p_y) \, dy \, dx \geq 0.$$

More generally, this kernel also allows the conversion of a continuous linear form P on V into the corresponding vector field $X \in V$ that satisfies $P = \langle X, \cdot \rangle$: X is simply obtained by convolution of P with K

$$X(x) = P(K(\cdot, x)) = \int_M P(K(y, x)) \, dy.$$

For this reason, it is interesting to consider the problem of finding geodesics on a shape space \mathcal{S} from the dual viewpoint. Indeed, $(\mathcal{H}_q, \langle \cdot, \cdot \rangle_q)$ is a Hilbert space with continuous inclusion $i_q : \mathcal{H}_q \hookrightarrow T_q^*M$. Therefore, any continuous linear function (or momentum) $p \in T_q^*M$ restricts to a continuous linear form $i_q^*p \in \mathcal{H}_q^*$ which in turn corresponds to a unique horizontal vector v such that $i_q^*p = \langle v, \cdot \rangle_q$. In the following result, we show how to compute this vector v using the isometric operator $K^V : V^* \rightarrow V$ and the infinitesimal action ξ_q .

Lemma 9.11. *Let $q \in M$ and $p \in T_q^*M$. The momentum map of the action of $\mathcal{D}^s(TM)$ on \mathcal{S} is defined by $(q, p) \mapsto \xi_q^*p = p \diamond q \in \Gamma^s(TM)^*$. Then $X = K^V \xi_q^*p$ belongs to $(\ker \xi_q)^\perp$, and*

$$i_q^*p = \langle \xi_q X, \cdot \rangle_q = \langle K_q p, \cdot \rangle_q.$$

The operator in the right-hand side of this equality is defined as $K_q = \xi_q K^V \xi_q^ : T_q^*M \rightarrow \mathcal{H}_q$. It is symmetric (i.e. $p_1(K_q p_2) = p_2(K_q p_1)$), positive semi-definite ($p(K_q p) \geq 0$), and*

$$\langle X, X \rangle = \|K_q p\|_q^2 = p(K_q p).$$

The map $(q, p) \mapsto K_q p$ is of class $\mathcal{C}^{\tilde{s}-s}$.

*Finally, when \mathcal{H}_q is closed in T_q^*M , K_q is surjective.*

Recall that $s = s_0 + \ell$ and ℓ is the order of \mathcal{S} as a shape space.

Proof. Let $X = K^V \xi_q^*p$ and take $X_0 \in \ker(\xi_q)$. Then

$$\langle X, X_0 \rangle = \langle K_{\mathcal{H}}^e \xi_q^*p, X_0 \rangle = \xi_q^*p(X_0) = p(\xi_q X_0) = 0.$$

So X belongs to $\ker(\xi_q)^\perp$. But we know that ξ_q is an isometry when restricted to the orthogonal of its kernel, so for any $v \in \mathcal{H}_q$,

$$\langle K_q p, v \rangle_q = \langle \xi_q K^V \xi_q^*p, v \rangle_q = \langle K^V \xi_q^*p, \xi_q^{-1}v \rangle = p(\xi_q \xi_q^{-1}v) = p(v).$$

Hence $i_q^*p = \langle K_q p, \cdot \rangle_q$. The rest follows immediately. \square

This lemma gives a nice way to build horizontal vectors from momenta on \mathcal{S} .

Remark 9.22. The map $q \mapsto K_q$ is the cometric associated to the sub-Riemannian metric $\langle \cdot, \cdot \rangle_q$. In particular, as we will see in the next section, the reduced Hamiltonian on M is $h(q, p) = \frac{1}{2}p(K_qp)$.

Example 9.4. Take $\mathcal{S} = Lmk_d(n)$. An element q of \mathcal{S} is given by an n -tuple $q = (x_1, \dots, x_n)$ with $x_i \in M$, and the infinitesimal action of $X \in \Gamma^s(TM)$ is

$$X \circ q = \xi_q X = (X(x_1), \dots, X(x_n)).$$

A 1-form on \mathcal{S} is given by an n -tuple $p = (\alpha_1, \dots, \alpha_n)$ with $\alpha_i \in T_{x_i}^*M$, and

$$p(w) = \sum_{i=1}^n \alpha_i(w_i).$$

Now $p \diamond q(X) = p(X \circ q) = \sum_{i=1}^n \alpha_i(X(x_i))$. So

$$p \diamond q = \sum_{i=1}^n \alpha_i \otimes \delta_{x_i},$$

with δ_x the usual dirac mass. Therefore, if $(x, y) \mapsto K(y, x)$ is the reproducing kernel of the Hilbert space V , we get

$$K^V p \diamond q(x) = p \diamond q(K(\cdot, x)) = \sum_{i=1}^n K(x_i, x) \alpha_i$$

and

$$K_qp = \xi_q K^V p \diamond q = \left(\sum_{i=1}^n K(x_i, x_1) \alpha_i, \dots, \sum_{i=1}^n K(x_i, x_n) \alpha_i \right) \in T_{x_1}M \times \dots \times T_{x_n}M.$$

9.2.3 Normal sub-Riemannian geodesic equations on shape spaces

We assume that \mathcal{S} is a shape space in M of order ℓ , and that $\mathcal{D}^s(M)$, $s = s_0 + \ell$, is equipped with a right-invariant sub-Riemannian structure \mathcal{H} with continuous inclusion $V \hookrightarrow \Gamma^{\tilde{s}}(TM)$, $\tilde{s} \geq s + 1$.

For $q_0 \in \mathcal{S}$, define the end-point map $\text{end}_{q_0}^{\mathcal{S}} : L^2(0, 1; V) \rightarrow \mathcal{S}$ by $\text{end}_{q_0}^{\mathcal{S}}(X) = \text{end}_e(X) \circ q_0 = R_{q_0}(\text{end}_e(X)) = q(1)$, where end is the end-point map from e on $\mathcal{D}^s(M)$ as defined in Section 9.1.2. $\text{end}_{q_0}^{\mathcal{S}}$ is of class $\mathcal{C}^{\tilde{s}-s}$ as a composition of the $\mathcal{C}^{\tilde{s}-s}$ end_e map with the smooth map R_{q_0} .

The previous section shows that to find the minimizing geodesics between two points q_0, q_1 in M , we can just look for geodesics $t \mapsto \varphi(t)$ in $\mathcal{D}^s(M)$ starting at e and such that $\varphi(1) \circ q_0 = q_1$. This leads us to look for critical points of

$$A(X) = \frac{1}{2} \int_0^1 \langle X(t), X(t) \rangle dt$$

with respect to the constraints $\text{end}_{q_0}^{\mathcal{S}}(X) = \text{end}(X) \circ q_0 = q_1$.

Define the hamiltonian $H^{\mathcal{S}} : T^*\mathcal{S} \times V \rightarrow \mathbb{R}$ of the motion

$$H^{\mathcal{S}}(q, p, X) = p(X \circ q) - \frac{1}{2} \langle X, X \rangle = p(\xi_q X) - \frac{1}{2} \langle X, X \rangle.$$

Then we have the following lemma.

Lemma 9.12. *Let $q : [0, 1] \mapsto \mathcal{S}$ be an horizontal curve, and φ an horizontal lift with logarithmic speed $X = \dot{\varphi} \circ \varphi^{-1}$. Let $q_0 = q(0)$ and $q_1 = q(1)$. Let $p_1 \in T_{q_1}^* \mathcal{S}$. Then, for every $\delta X \in L^2(0, 1; V)$,*

$$dA_X(\delta X) = (d \operatorname{end}_{q_0}^{\mathcal{S}})^* p_1(\delta X) - \int_0^1 \partial_X H^{\mathcal{S}}(q(t), p(t), X(t))(\delta X(t)) dt,$$

where $p(t) \in T_{q(t)}^* M$ satisfies $p(1) = p_1$ and

$$(\dot{q}(t), \dot{p}(t)) = \nabla^\omega H^{\mathcal{S}}(q(t), p(t), X(t)). \quad (9.15)$$

Proof. First of all, $\nabla^\omega H^{\mathcal{S}}(q, p, X)$ is the symplectic gradient of $H^{\mathcal{S}}$ at (q, p) , with fixed parameter X , for the canonical symplectic form on T^*M . It should be pointed out that, since \mathcal{S} is modeled only on a Banach space \mathcal{B} , not a Hilbert space, ω is just a weak symplectic form and not every real map on T^*M has a symplectic gradient. However, the function $(q, p) \mapsto H^{\mathcal{S}}(q, p, X)$ does admit a symplectic gradient for fixed X . This is very simple to check: just consider a canonical coordinate system $(q, p) \in U \times \mathcal{B}^*$ around (q, p) . Then

$$\partial_{(q,p)} H_{q,p,X}^{\mathcal{S}}(\delta q, \delta p) = \delta p(\xi_q X) + \partial_q(p(\xi_q X))(\delta q) = \omega((\xi_q X, -\partial_q(\xi_q X)^* p), (\delta q, \delta p)),$$

so $\nabla^\omega H^{\mathcal{S}}(q, p, X) = (\xi_q X, -\partial_q(\xi_q X)^* p)$. In particular, the Cauchy problem (9.15) has linear coefficients in p , and so it does have a unique global solution.

The rest of the proof is precisely the same as that of Lemma 9.8. First of all, $p(t)$ exists and is uniquely defined for fixed p_1 , since $\nabla^\omega H^{\mathcal{S}}(q, p, X)$ is linear in p . Moreover, we have $d(\operatorname{end}_{q_0}^{\mathcal{S}})_X(\delta X) = \delta q(1)$, where, in a local trivialization,

$$\dot{\delta q}(t) = \partial_q(\xi_{q(t)} X(t))(\delta q(t)) + \xi_{q(t)}(\delta X(t)),$$

with $\delta q(0) = 0$. Then

$$\int_0^1 \partial_X H^{\mathcal{S}}(q(t), p(t), X(t))(\delta X(t)) dt = \int_0^1 p(t)(\xi_{q(t)}(\delta X(t))) dt - dA_X(\delta X).$$

But since $\xi_{q(t)}(\delta X(t)) = \dot{\delta q}(t) - \partial_q(\xi_{q(t)} X(t))(\delta q(t))$ we get

$$\begin{aligned} \int_0^1 p(t)(\xi_{q(t)}(\delta X(t))) dt &= \int_0^1 p(t)(\dot{\delta q}(t)) dt - \int_0^1 \partial_q(\xi_{q(t)} X(t))^* p(t)(\delta q(t)) dt \\ &= \int_0^1 p(t)(\dot{\delta q}(t)) + \dot{p}(t)(\delta q(t)) dt \\ &= p_1(\delta q(1)) = (d \operatorname{end}_{q_0}^{\mathcal{S}})_X^* p_1. \end{aligned}$$

□

Corollary 9.1. *Let $q : [0, 1] \rightarrow \mathcal{S}$ be an horizontal curve with end-points q_0 and q_1 , and X the logarithmic velocity of an horizontal lift φ of q . Then the following statements are equivalent*

1. *There exists $p_1 \neq 0 \in T_{q_1}^* \mathcal{S}$ such that $dA_X = (d \operatorname{end}_{q_0}^{\mathcal{S}})_X^* p_1$. In particular, $t \mapsto q(t)$ is a geodesic.*
2. *There exists an absolutely continuous map $t \mapsto p(t) \in T_{q(t)}^* \mathcal{S}$, with $p(1) = p_1$, such that for almost every $t \in [0, 1]$,*

$$\begin{aligned} 0 &= \partial_X H^{\mathcal{S}}(q(t), p(t), X(t)) = \xi_{q(t)}^* p(t) - \langle X(t), \cdot \rangle, \\ (\dot{q}(t), \dot{p}(t)) &= \nabla^\omega H^{\mathcal{S}}(q(t), p(t), X(t)) = (\xi_{q(t)} X(t), -\partial_q(\xi_{q(t)} X(t))^* p(t)). \end{aligned} \quad (9.16)$$

Such a curve is called a normal geodesic.

The first equation in (9.16) is satisfied if and only if

$$\xi_{q(t)}^* p(t) = \langle X(t), \cdot \rangle,$$

if and only if $X(t) = K^V \xi_{q(t)}^* p(t)$. Moreover, we know that in this case, $X = K^V \xi_q^* p$ always belongs to $(\ker \xi_q)^\perp$.

This means that if $\xi_{q(t)}^* p(t) = \langle X(t), \cdot \rangle$ for almost every t , then $t \mapsto X(t)$ is the logarithmic velocity of a *minimizing lift* of the horizontal curve q .

We now define the maximized hamiltonian $h^S : T^* \mathcal{S} \rightarrow \mathbb{R}$ by

$$h^S(q, p) = H^S(q, p, K^V \xi_q^* p) = \frac{1}{2} p(\xi_q K^V \xi_q^* p) = \frac{1}{2} p(K_q p) = \frac{1}{2} \|K_q p\|_q^2.$$

Remark 9.23. We have $h^S(q, p) = h(e, \xi_q^* p)$, where h is the reduced Hamiltonian on $T^* \mathcal{D}^s(M)$.

This is the same as the usual sub-Riemannian Hamiltonian defined in [Mon02]. It is of class at least $\mathcal{C}^{\bar{s}-s}$ and has a well-defined symplectic gradient of class $\mathcal{C}^{\bar{s}-s-1}$, given by

$$\nabla^\omega h^S_{q,p} = (\partial_p h^S(q, p), -\partial_q h^S(q, p)) = \left(K_q p, -\frac{1}{2} \partial_q p(K_q p) \right).$$

Theorem 9.7. *Let \mathcal{S} be a shape space of order ℓ in M , and \mathcal{H} be a strong right-invariant sub-Riemannian structure on $\mathcal{D}^s(M)$ such that $\xi : \mathcal{S} \times V \rightarrow T\mathcal{S}$ is of class \mathcal{C}^2 . Then, for every starting point $q_0 \in \mathcal{S}$ and initial momentum $p_0 \in T_{q_0}^* \mathcal{S}$, there is a unique normal geodesic $t \mapsto q(t)$ with corresponding momentum $t \mapsto p(t)$, defined on an interval $I \ni 0$, such that $(q(0), p(0)) = (q_0, p_0)$ and the curve $t \mapsto (q(t), p(t))$ satisfies the usual hamiltonian geodesic equation*

$$\begin{cases} \dot{q}(t) = \partial_p h^S(q(t), p(t)) & = K_{q(t)} p(t), \\ \dot{p}(t) = -\partial_q h^S(q(t), p(t)) & = -\partial_q (p(t)(K_{q(t)} p(t))). \end{cases}$$

In particular, $t \mapsto h^S(q(t), p(t))$ is constant and

$$A(q) = \frac{1}{2} \int_I h^S(q(t), p(t)) dt = \frac{\ell(I)}{2} h^S(q(0), p(0)).$$

Moreover, q is locally minimizing, i.e. for any $t_0 \in I$ and any $t_1 \in I$ close enough to t_0 is minimizing.

Remark 9.24. It should be noted that the same discussion as that of the end of Section 9.1.2, other geodesics may also exist. They may, again, come from the fact that ξ may be a "bad" parametrization of \mathcal{H}^S , or be "abnormal" geodesics (critical points of the end-point map), or exist for deeper reasons exclusive to infinite dimensional sub-Riemannian geometry.

However, just as is the case for $\mathcal{D}^s(M)$, critical points of functionals of the form

$$J(q) = A(q) + g(q(1))$$

are always normal geodesics. This is very useful in practical applications, where finding the exact minimizing geodesic between two shapes is impossible anyway.

Minimizing lifts of normal geodesics are normal geodesics It is clear that for a normal geodesic $t \mapsto (q(t), p(t))$ on a shape space \mathcal{S} , the horizontal curve $t \mapsto \varphi(t)$ in $\mathcal{D}^s(M)$ given by $\varphi(0) = e$ and logarithmic velocity $X(t) = K^V \xi_{q(t)}^* p(t)$ is the minimizing lift of $q(\cdot)$. Moreover, if $q(\cdot)$ is a minimizing geodesic, then so is $\varphi(\cdot)$ as a straightforward consequence of Lemma 9.10.

However, it is a well-known fact in dynamics [MR99] that $\varphi(\cdot)$ itself satisfies the Hamiltonian geodesic equations in $\mathcal{D}^s(M)$, with initial momentum $K^V \xi_{q(0)}^* p(0)$. More precisely, we have the

Proposition 3. *Take $(q_0, p_0) \in T^*\mathcal{S}$ and $P_0 = \xi_{q_0}^* p_0$. Then the Hamiltonian geodesic on $\mathcal{D}^s(M)$ starting at e with initial momentum P_0 is the minimizing lift of the Hamiltonian geodesic on \mathcal{S} starting at q_0 with initial momentum p_0 .*

In other words, the Hamiltonian flow on $T^\mathcal{D}^s(M)$ preserves the subbundle $R_{q_0}^* T^*\mathcal{S}$.*

Proof. Consider the mapping $t \mapsto \mu(t) = \xi_{q(t)}^* p(t)$, where $(q(t), p(t))$ is the Hamiltonian geodesic on \mathcal{S} starting at q_0 with initial momentum p_0 . Let $X(t) = K^V \mu(t)$, and $Y \in \Gamma^{s+1}(TM)$. Then

$$\begin{aligned} \dot{\mu}(Y) &= p\left(\partial_q(\xi_q(Y))(\xi_q(X)) - \partial_q(\xi_q(X))(\xi_q(Y))\right) = p([\xi(X), \xi(Y)](q)) \\ &= p(\xi_q([X, Y])) = ad_X^* \mu(Y). \end{aligned}$$

Therefore, $t \mapsto \mu(t)$ satisfies the necessary and sufficient condition for $X(t) = K^V \mu(t)$ to be the logarithmic velocity of a Hamiltonian geodesic $(\varphi(t), P(t))$ on $\mathcal{D}^s(M)$ with initial momentum $\mu(0) = P_0$. Since $X(t) = \xi_{q(t)}^* p(t)$, φ is also the minimizing lift of q starting at e . \square

9.2.4 Some examples of geodesic equations on classical shape spaces

Just as before we consider shape spaces in $M = \mathbb{R}^d$ to obtain simpler geodesic equations. We keep the notations from Section 9.1.5.

Spaces of continuous embeddings of compact manifolds and Gaussian kernels

Take the space Hilbert space $V \subset H^s(\mathbb{R}^d, \mathbb{R}^d)$ defined by the Gaussian kernel

$$K(x, y)p = e^{-\frac{|x-y|^2}{2\sigma}} p^T = e(x-y)p^T,$$

as in Section 9.1.5.

We let S be a compact Riemannian manifold, and $\mathcal{S} = \mathcal{C}^0(S, \mathbb{R}^d)$, a Banach manifold (and actually a Banach vector space). We have $T^*\mathcal{S} = \mathcal{S} \times \mathcal{C}^0(S, \mathbb{R}^d)^*$, with the elements of $\mathcal{C}^0(S, \mathbb{R}^d)^*$ being identified to maps $S \rightarrow (\mathbb{R}^d)^*$ with coefficients in the space $\mathcal{M}(S)$ of Radon measures on S , that is, elements of $\mathcal{M}(S) \otimes (\mathbb{R}^d)^*$.

Then, for $q : S \rightarrow \mathbb{R}^d$ continuous and $p = (p_1, \dots, p_d) \in \mathcal{M}(S) \otimes (\mathbb{R}^d)^*$,

$$(K^V \xi_q^* p)(x) = \int_S e(x - q(s)) dp^T(s).$$

Hence, we have $K_q p \in T_q \mathcal{S} = \mathcal{C}^0(S, \mathbb{R}^d)$ given by

$$(K_q p)(s) = \int_S e(q(s) - q(s')) dp^T(s').$$

The reduced Hamiltonian is therefore

$$h^{\mathcal{S}}(q, p) = \frac{1}{2} \int_{S \times S} e(q(s) - q(s')) \sum_{i=1}^d dp_i(s) dp_i(s') = \frac{1}{2} \int_{S \times S} e(q(s) - q(s')) [dp(s) \cdot dp(s')].$$

Using $e(x - y) = e(y - x)$ and $de_x(v) = -\frac{1}{\sigma}(x \cdot v)e(x)$, we get for any $\delta q \in T_q\mathcal{S}$,

$$\partial_q h^{\mathcal{S}}(q, p)(\delta q) = -\frac{1}{\sigma} \int_{S \times S} [(q(s) - q(s')) \cdot \delta q(s)] e(q(s) - q(s')) [dp(s) \cdot dp(s')].$$

In the end, the geodesic equations read

$$\begin{aligned} \dot{q}(t, s) &= \int_S e(q(t, s) - q(t, s')) dp^T(t, s'), \\ \dot{p}(t, s) &= \frac{1}{\sigma} \int_S e(q(t, s) - q(t, s')) (q(s) - q(s'))^T [dp(t, s') \cdot dp(t, s)]. \end{aligned}$$

In the special case of landmarks, we can write $\mathcal{S} = (\mathbb{R}^d)^n$, writing $q = (x_1, \dots, x_n)$ and $p = (p_1, \dots, p_n)(\mathbb{R}^{d*})^n \sim (\mathbb{R}^d)^n$, we get (omitting time for readability)

$$\begin{aligned} \dot{x}_i &= \sum_{j=1}^n e^{-\frac{|x_i - x_j|^2}{2\sigma}} p_j^T, \\ \dot{p}_i &= \frac{1}{\sigma} \sum_{j=1}^n e^{-\frac{|x_i - x_j|^2}{2\sigma}} (p_i \cdot p_j) (x_i - x_j)^T, \end{aligned} \quad i = 1, \dots, n.$$

Remark 9.25. It is worthwhile to note that in the neighbourhood of $q = (x_1, \dots, x_n)$ such that $i \neq j$ implies $x_i \neq x_j$, the structure here is actually Riemannian instead of simply sub-Riemannian.

An action of higher order In this paragraph, we keep our Gaussian kernel K as above and S our compact Riemannian manifold, but we take the shape space $\mathcal{S} = TC^0(S, \mathbb{R}^d) = C^0(S, \mathbb{R}^d) \times C^0(S, \mathbb{R}^d)$, a shape space of order 1 for the action

$$\varphi \circ (q, v)(s) = (\varphi(q(s)), d\varphi_{q(s)}(v(s))).$$

The three conditions for a shape space of order 1 are easily seen to be true. The infinitesimal action of a vector field X on \mathbb{R}^d is

$$\xi_{q,v}(X)(s) = (X(q(s)), dX_{q(s)}(v(s))).$$

Now a momentum on \mathcal{S} is a couple (p, l) of maps $S \rightarrow (\mathbb{R}^d)^*$ with coefficients in $\mathcal{M}(S)$, and

$$\xi_{q,v}^*(p, l)(X) = \int_S dp(s)(X(q(s))) + \int_S dl(s)(dX_{q(s)}v(s)).$$

Therefore,

$$K^V \xi_{q,v}^*(p, l)(x) = \int_S e(x - q(s)) dp^T(s) - \frac{1}{\sigma} \int_S e(x - q(s)) [v(s) \cdot (x - q(s))] dl^T(s).$$

Writing $\xi_{q,v} K^V \xi_{q,v}^*(p, l) = (w_1, w_2)$, and $e_{s,s'} = e(q(s) - q(s'))$, we get

$$\begin{aligned} w_1(s) &= \int_S e_{s,s'} dp^T(s') - \frac{1}{\sigma} \int_S e_{s,s'} [v(s') \cdot (q(s) - q(s'))] dl^T(s'), \\ w_2(s) &= -\frac{1}{\sigma} \int_S e_{s,s'} [v(s) \cdot (q(s) - q(s'))] dp^T(s') \\ &\quad + \frac{1}{\sigma^2} \int_S e_{s,s'} [v(s') \cdot (q(s) - q(s'))] [v(s) \cdot (q(s) - q(s'))] dl^T(s') \\ &\quad - \frac{1}{\sigma} \int_S e_{s,s'} [v(s') \cdot v(s)] dl^T(s'). \end{aligned}$$

The reduced hamiltonian is given by

$$\begin{aligned} h^S(q, v, p, l) &= \frac{1}{2} \int_{S \times S} e_{s,s'} [dp(s) \cdot dp(s')] - \frac{1}{2\sigma} \int_{S \times S} e_{s,s'} [v(s') \cdot v(s)] [dl(s) \cdot dl(s')] \\ &\quad - \frac{1}{2\sigma} \int_{S \times S} e_{s,s'} [(v(s)) \cdot (q(s) - q(s'))] [dl(s) \cdot dp(s') + dp(s) \cdot dl(s')] \\ &\quad + \frac{1}{2\sigma^2} \int_{S \times S} e_{s,s'} [v(s') \cdot (q(s) - q(s'))] [v(s) \cdot (q(s) - q(s'))] [dl(s) \cdot dl(s')]. \end{aligned}$$

From there, the geodesic equations are easily deduced. For example

$$\begin{aligned} \dot{l}(s)^T &= -\partial_v h^S(q, v, p, l)^T(s) = \frac{1}{\sigma} \int_S e_{s,s'} v(s') [l(s) \cdot dl(s')] \\ &\quad + \frac{1}{2\sigma} \int_S e_{s,s'} (q(s) - q(s')) [l(s) \cdot dp(s') + p(s) \cdot dl(s')] \\ &\quad - \frac{1}{\sigma^2} \int_S e_{s,s'} (q(s) - q(s')) [v(s') \cdot (q(s) - q(s'))] [l(s) \cdot dl(s')]. \end{aligned}$$

In the case of landmarks, we take $q = (x_1, \dots, x_n)$, $v = (v_1, \dots, v_n)$, $p = (p_1, \dots, p_n)$, and $l = (l_1, \dots, l_n)$. Denoting $x_{i,j} = x_i - x_j$, we get

$$\left. \begin{aligned} \dot{x}_i^T &= \sum_{j=1}^n \left(p_j - \frac{1}{\sigma} [v_j \cdot x_{i,j}] l_j \right) e(x_{i,j}), \\ \dot{v}_i^T &= \sum_{j=1}^n \left(-\frac{1}{\sigma} [v_i \cdot x_{i,j}] p_j - \frac{1}{\sigma^2} [v_i \cdot (x_{i,j})] [v_j \cdot x_{i,j}] l_j + \frac{1}{\sigma} [v_i \cdot v_j] l_j \right) e(x_{i,j}), \\ \dot{l}_i^T &= \frac{1}{\sigma} \sum_{j=1}^n [l_i \cdot l_j] e(x_{i,j}) v_j + \frac{1}{2\sigma} \sum_{j=1}^n [l_i \cdot p_j + l_i p_j] e(x_{i,j}) x_{i,j} \\ &\quad - \frac{1}{\sigma^2} \sum_{j=1}^n [v_j \cdot x_{i,j}] [l_i \cdot l_j] e(x_{i,j}) x_{i,j}. \end{aligned} \right\} \quad i = 1, \dots, n$$

The derivative of p_i is slightly more complex, as

$$\left. \begin{aligned} \dot{p}_i^T &= \frac{1}{\sigma} \sum_{j=1}^n \left(p_i \cdot p_j - \frac{1}{\sigma} [v_i \cdot x_{i,j}] [v_j \cdot x_{i,j}] [l_j \cdot l_j] - \frac{1}{\sigma} [v_i \cdot v_j] [l_i \cdot l_j] \right) e(x_{i,j}) x_{i,j} \\ &\quad - \frac{1}{\sigma^2} \sum_{j=1}^n ([v_i \cdot x_{i,j}] [l_i \cdot p_j] - [v_j \cdot x_{i,j}] [p_i \cdot l_j]) e(x_{i,j}) x_{i,j} \\ &\quad - \frac{1}{\sigma^3} \sum_{j=1}^n ([v_i \cdot x_{i,j}] [v_j \cdot x_{i,j}] [l_i \cdot l_j]) e(x_{i,j}) x_{i,j} \\ &\quad + \frac{1}{\sigma} \sum_{j=1}^n ([l_i \cdot p_j] v_i - [p_i \cdot l_j] v_j) e(x_{i,j}). \end{aligned} \right\} \quad i = 1, \dots, n.$$

Remark 9.26. Again, in the case of landmarks, there is an open and dense subset of \hat{S} on which the sub-Riemannian structure is actually Riemannian.

The case of sub-Riemannian Gaussian kernels Here, we compute the sub-Riemannian Hamiltonian geodesic equations for the shape space $\mathcal{S} = \mathcal{C}^0(S, \mathbb{R}^d)$, with S a compact Riemannian manifold, and with V defined by the reproducing kernel

$$K(x, y)p = e^{-\frac{|x-y|^2}{2\sigma}} \sum_{r=1}^k p(X_r(y)) X_r(x) = e(x, y) \sum_{r=1}^k p(X_r(y)) X_r(x),$$

for X_1, \dots, X_k smooth bounded vector fields on \mathbb{R}^d with bounded derivatives at every order, linearly independent at every point x of \mathbb{R}^d .

Now for $q \in \mathcal{S}$ and $p = (p_1, \dots, p_d) \in T_q^* \mathcal{S} = \mathcal{M}(S) \otimes (\mathbb{R}^d)^*$, we get

$$(K^V \xi_q^* p)(x) = \sum_{r=1}^k \int_S e(x - q(s)) dp(s)(X_r(q(s))) X_r(x),$$

where $dp(s)(X_r(q(s))) \in \mathcal{M}(S)$ is defined by $X_r^1(q(s)) dp_1(s) + \dots + X_r^d(q(s)) dp_d(s)$.

Then we get

$$K_q p(s) = K^V \xi_q^* p(q(s)) = \sum_{r=1}^k \int_S e(q(s) - q(s')) dp(s')(X_r(q(s')) X_r(q(s)),$$

and compute the reduced Hamiltonian

$$h^S(q, p) = \frac{1}{2} \sum_{r=1}^k \int_{S \times S} e(q(s) - q(s')) dp(s)(X_r(q(s))) dp(s')(X_r(q(s'))).$$

Then, reintroducing the notations $e_{s,s'} = e(q(s) - q(s'))$, we obtain

$$\begin{aligned} \partial_q h^S(q, p)(\delta q) &= -\frac{1}{\sigma} \sum_{r=1}^k \int_{S \times S} e_{s,s'} [\delta q(s) \cdot (q(s) - q(s'))] dp(s)(X_r(q(s))) dp(s')(X_r(q(s'))) \\ &\quad + \sum_{r=1}^k \int_{S \times S} e_{s,s'} dp(s)(dX_{r,q(s)}(\delta q(s))) dp(s')(X_r(q(s'))). \end{aligned}$$

In other words, as a Radon measure,

$$\partial_q h^S(q, p)(s) = \sum_{r=1}^k \int_S e_{s,s'} dp(s')(X_r(q(s'))) \left(p(s)(dX_{r,q(s)}) - \frac{1}{\sigma} p(s)(X_r(q(s)))(q(s) - q(s'))^T \right).$$

Here, $p(dX_{r,q})$ in $\mathcal{M}(S) \otimes (\mathbb{R}^d)^*$ is the $(\mathbb{R}^d)^*$ -valued Radon measure given by

$$dp(s)(dX_{r,q(s)}) = dp_1(s) \otimes dX_{r,q(s)}^1 + \dots + dp_d(s) \otimes dX_{r,q(s)}^d.$$

Therefore, the Hamiltonian geodesic equations are given by

$$\begin{aligned} \dot{q}(s) &= \sum_{r=1}^k \int_S e(q(s) - q(s')) dp(s')(X_r(q(s'))) X_r(q(s)), \\ \dot{p}(s) &= \sum_{r=1}^k \int_S e_{s,s'} dp(s')(X_r(q(s'))) \left(\frac{1}{\sigma} p(s)(X_r(q(s)))(q(s) - q(s'))^T - p(s)(dX_{r,q(s)}) \right). \end{aligned}$$

In the case of landmarks, $q = (x_1, \dots, x_n) \in (\mathbb{R}^d)^n$ and $p = (p_1, \dots, p_n) \in (\mathbb{R}^d)^{n*}$, and denoting $x_i - x_j = x_{ij}$, we obtain

$$\begin{aligned} \dot{x}_i &= \sum_{r=1}^k \sum_{j=1}^n e(x_{ij}) p_j(X_r(x_j)) X_r(x_i), \\ \dot{p}_i &= \sum_{r=1}^k \sum_{j=1}^n e(x_{ij}) p_j(X_r(x_j)) \left(\frac{1}{\sigma} p_i(X_r(x_i)) x_{ij}^T - p_i(dX_{r,x_i}) \right). \end{aligned} \tag{9.17}$$

Remark 9.27. An interesting fact is that $e(x_{ij}) \rightarrow \delta_{ij}$ as $\sigma \rightarrow 0$, where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. Moreover, $\frac{1}{\sigma}e(x_{ij})x_{ij} \rightarrow 0$ as $\sigma \rightarrow 0$ because $x_{ii} = 0$. Therefore, as σ goes to 0, the geodesic equations become

$$\begin{aligned}\dot{x}_i &= \sum_{r=1}^k p_i(X_r(x_i))X_r(x_i), \\ \dot{p}_i &= \sum_{r=1}^k -p_i(X_r(x_i))p_i(dX_{r,x_i}).\end{aligned}$$

For each couple $t \mapsto (x_i(t), p_i(t)) \in T^*\mathbb{R}^d$, we recognize the hamiltonian geodesic equation for the sub-Riemannian structure Δ induced on \mathbb{R}^d by X_1, \dots, X_k : the dynamic is that of a system of n points without interaction in (\mathbb{R}^d, Δ) .

So Equations (9.17) can be seen as a perturbation of such a dynamic by creating an interaction between the points x_i , which decreases exponentially with respect to the square of the Euclidean distances between the points.

9.3 Lifted Shape Spaces

In this section we introduce another geometry on a shape space $\hat{\mathcal{S}}$ for which there exists a smooth equivariant submersion π onto another shape space \mathcal{S} . This structure is obtained by lifting the sub-Riemannian structure on \mathcal{S} to $\hat{\mathcal{S}}$.

The geodesic equations obtained are more complex, but can be used to "keep track" of additional information on a shape space, or to emphasize the importance of certain points in the shape. Examples of applications of such structures have been given in [You12].

9.3.1 Lifted sub-Riemannian structure

We keep the notations from the previous sections.

Let us start with a simple case to clarify the concepts we wish to introduce. Let \mathcal{S} and \mathcal{S}' be two shape spaces of respective orders ℓ and ℓ' . Then $\mathcal{S} \times \mathcal{S}' =: \hat{\mathcal{S}}$ is a shape space of order $\hat{\ell} = \max(\ell, \ell')$ for the diagonal action of $\mathcal{D}^{s_0+\hat{\ell}}(M)$. Now take a right-invariant sub-Riemannian structure \mathcal{H} on $\mathcal{D}^{s_0+\hat{\ell}}(M)$ induced by a Hilbert space vector fields of class H^{s_0+k} , with $k \geq \hat{\ell}$. It also induces on \mathcal{S} (resp. \mathcal{S}' and $\hat{\mathcal{S}}$) the sub-Riemannian structure $\mathcal{H}^{\mathcal{S}}$ (resp. $\mathcal{H}^{\mathcal{S}'}$ and $\mathcal{H}^{\hat{\mathcal{S}}}$).

However, it is not the only such structure that one could equip on $\hat{\mathcal{S}}$. Indeed, let $\hat{q}_0 = (q_0, q'_0) \in \hat{\mathcal{S}}$ and define $q : [0, 1] \rightarrow \mathcal{S}$ an horizontal curve in \mathcal{S} with $q(0) = q_0$. Then there is a unique minimizing horizontal lift $t \mapsto \varphi(t)$ in $\mathcal{D}^s(M)$ of q such that $\varphi(0) = e$. Recall that its logarithmic velocity $X(t) = \dot{\varphi}(t) \circ \varphi^{-1}(t)$ is given by $X(t) = \xi_{q(t)}^{-1} \dot{q}(t)$.

This one-parameter family of diffeomorphisms, in turn, defines an horizontal curve

$$\hat{q}(t) = \varphi(t) \circ \hat{q}_0 = (\varphi(t) \circ q_0, \varphi(t) \circ q'_0)$$

in $\hat{\mathcal{S}}$, and

$$\dot{\hat{q}}(t) = X(t) \circ \hat{q}(t) = \xi_{\hat{q}(t)} X(t) = \xi_{\hat{q}(t)} \xi_{q(t)}^{-1} \dot{q}(t).$$

It is, in fact, the only curve starting at \hat{q}_0 in $\hat{\mathcal{S}}$ that projects on q through the canonical projection $\pi : (q, q') \mapsto q$ whose total action is equal to that of $q(\cdot)$. In other words, it is the unique minimizing lift of $q(\cdot)$ to $\hat{\mathcal{S}}$ starting at \hat{q}_0 .

Such minimizing lifts are exactly curves of class H^1 in $\hat{\mathcal{S}}$ with velocity belonging to

$$\xi_{\hat{q}} \left(\xi_{\pi(\hat{q})}^{-1} (\mathcal{H}^{\mathcal{S}}) \right),$$

which is a distribution of subspaces on $\hat{\mathcal{S}}$. We will denote

$$\mathcal{H}_\pi^{\hat{\mathcal{S}}} = \bigcup_{\hat{q} \in \hat{\mathcal{S}}} \xi_{\hat{q}} \left(\xi_{\pi(\hat{q})}^{-1}(\mathcal{H}^{\mathcal{S}}) \right).$$

This is a particular case of lifted sub-riemannian structures on shape spaces. We will now give a more general description. Simply assume that \mathcal{S} and $\hat{\mathcal{S}}$ are shape spaces in M of orders $\ell \leq \hat{\ell}$, and that we have a map $\pi : \hat{\mathcal{S}} \rightarrow \mathcal{S}$ that satisfies, for every $\varphi \in \mathcal{D}^{s_0 + \hat{\ell}}(M)$ and $\hat{q} \in \hat{\mathcal{S}}$,

$$\pi(\varphi \circ \hat{q}) = \varphi \circ \pi(\hat{q}).$$

We say that π is *equivariant*.

The composition rule for derivatives shows that for every $X \in \Gamma^{s_0 + \hat{\ell}}(TM)$, we have

$$d\pi(\xi_{\hat{q}} X) = d\pi(X \circ \hat{q}) = X \circ \pi(\hat{q}) = \xi_{\pi(\hat{q})} X.$$

This implies that $d\pi(\mathcal{H}^{\hat{\mathcal{S}}}) \subset \mathcal{H}^{\mathcal{S}}$ for any sub-Riemannian structure induced by a Hilbert space of vector fields $V \subset \Gamma^{\tilde{s}}(TM)$, $\tilde{s} > s_0 + \hat{\ell}$.

Let us give a few examples.

Example 9.5. Additional landmarks: Take $n \leq \hat{n} \in \mathbb{N}^*$, and $\mathcal{S} = Lmk_n(M)$ and $\hat{\mathcal{S}} = Lmk_{\hat{n}}(M)$. We define $\pi : Lmk_{\hat{n}}(M) \rightarrow Lmk_n(M)$ by

$$\pi(x_1, \dots, x_{\hat{n}}) = (x_1, \dots, x_n).$$

The mapping π is an equivariant submersion for the diagonal action of $\mathcal{D}^{s_0 + 1}(M)$: for every φ in $\mathcal{D}^{s_0 + 1}(M)$,

$$\pi(\varphi \circ (x_1, \dots, x_{\hat{n}})) = \pi(\varphi(x_1), \dots, \varphi(x_{\hat{n}})) = (\varphi(x_1), \dots, \varphi(x_n)) = \varphi \circ \pi(x_1, \dots, x_n).$$

Example 9.6. Embeddings of submanifolds: Let \hat{S} be a smooth compact Riemannian manifold, S a sub-manifold of \hat{S} without boundary, and let $\mathcal{S} = \mathcal{C}^0(S, M)$ and $\hat{\mathcal{S}} = \mathcal{C}^0(\hat{S}, M)$. Then the map $\pi : \hat{q} \mapsto \hat{q}|_S$ is an equivariant submersion.

Example 9.7. Action on the tangent space: Let \mathcal{S} be a shape space in M of order ℓ . Then $\hat{\mathcal{S}} = T\mathcal{S}$ is a shape space of order $\ell + 1$ for the action of $\mathcal{D}^{s_0 + \ell + 1}(M)$ given by the differential of the action on \mathcal{S}

$$\varphi \circ (q, v) = (\varphi \circ q, \partial_q(\varphi \circ q)(v)) \in T_{\varphi \circ q} \mathcal{S}.$$

Then the projection $\pi(q, v) = q$ is an equivariant submersion for these actions.

In the case of landmarks, i.e. $\mathcal{S} = Lmk_n(M)$, we can write $(q, v) = (x_1, \dots, x_n, v_1, \dots, v_n)$, with $v_i \in T_{x_i} M$, and

$$\varphi \circ (q, v) = (\varphi(x_1), \dots, \varphi(x_n), d\varphi_{x_1}(v_1), \dots, d\varphi_{x_n}(v_n)).$$

We now have the following lemma for an equivariant map $\pi : \hat{\mathcal{S}} \rightarrow \mathcal{S}$.

Lemma 9.13. *Let V be a Hilbert space of vector fields with continuous inclusion $V \subset \Gamma^{\tilde{s}}(TM)$. Assume that π is an horizontal submersion, that is, for every \hat{q} in $\hat{\mathcal{S}}$,*

$$d\pi(\mathcal{H}_{\hat{q}}^{\hat{\mathcal{S}}}) = \mathcal{H}_{\pi(\hat{q})}^{\mathcal{S}}.$$

Then the map π is also actually a sub-Riemannian submersion: for every $\hat{q} \in \hat{\mathcal{S}}$, the restriction of $d\pi_{\hat{q}}$ to the orthogonal of its kernel in the horizontal space $\mathcal{H}_{\hat{q}}^{\hat{\mathcal{S}}}$ is an isometry. Moreover, the inverse of this restriction, i.e. the pseudo-inverse of $d\pi_{\hat{q}}$, is given by

$$d\pi_{\hat{q}}^{-1} = \xi_{\hat{q}} \xi_{\pi(\hat{q})}^{-1}.$$

In particular, $\Pi_q = \xi_{\hat{q}} \xi_{\pi(\hat{q})}^{-1} d\pi_{\hat{q}}$ is the orthogonal projection $\mathcal{H}_{\hat{q}}^{\hat{\mathcal{S}}} \rightarrow (\ker(d\pi_{\hat{q}}) \cap \mathcal{H}_{\hat{q}}^{\hat{\mathcal{S}}})^\perp$.

Proof. Let $\hat{q} \in \hat{\mathcal{S}}$ and $q = \pi(\hat{q})$. Also take $w \in \mathcal{H}_q^{\mathcal{S}}$ and $\hat{w}_0 \in \mathcal{H}_{\hat{q}}^{\hat{\mathcal{S}}}$ such that $d\pi(\hat{w}_0) = 0$.

To prove that the pseudo-inverse of $d\pi_{\hat{q}}$ is $\xi_{\hat{q}}\xi_{\pi(\hat{q})}^{-1}$, we need to check the following two facts:

$$d\pi_{\hat{q}}\xi_{\hat{q}}\xi_q^{-1}(w) = w,$$

and

$$\xi_{\hat{q}}\xi_q^{-1}(w) \perp \hat{w}_0, \quad |\xi_{\hat{q}}\xi_q^{-1}(w)|_{\hat{q}} = |w|_q.$$

For the first one, recall that $d\pi_{\hat{q}}\xi_{\hat{q}} = \xi_{\pi(\hat{q})} = \xi_q$, so $d\pi_{\hat{q}}\xi_{\hat{q}}\xi_q^{-1}(w) = \xi_q\xi_q^{-1}(w) = w$.

For the second fact, let us first note that $X = \xi_q^{-1}(w)$ belongs to $\ker(\xi_{\hat{q}})^{\perp}$. Indeed, since $d\pi_{\hat{q}}\xi_{\hat{q}} = \xi_q$, we have $\ker(\xi_{\hat{q}})^{\perp} \subset \ker(\xi_q)^{\perp}$, and therefore $\xi_q^{-1}(w) \in \ker(\xi_q)^{\perp} \subset \ker(\xi_{\hat{q}})^{\perp}$. In particular,

$$X = \xi_{\hat{q}}^{-1}\xi_{\hat{q}}(X).$$

Let $X_0 = \xi_{\hat{q}}^{-1}(\hat{w}_0)$. We then have $\xi_{\hat{q}}(X_0) = \hat{w}_0$, and therefore

$$\xi_q(X_0) = d\pi_{\hat{q}}\xi_{\hat{q}}(X_0) = d\pi(\hat{w}_0) = 0.$$

Now, as both $\xi_{\hat{q}}^{-1}$ and ξ_q^{-1} are isometries onto their images, we finally obtain

$$\langle \xi_{\hat{q}}(X), \hat{w}_0 \rangle_{\hat{q}} = \langle X, X_0 \rangle = \langle w, \xi_q(X_0) \rangle_q = 0.$$

All that is left is to see that for $w \in \mathcal{H}_q^{\mathcal{S}}$,

$$\|w\|_q = \|d\pi_{\hat{q}}^{-1}w\|_{\hat{q}}.$$

This is a trivial consequence of the fact that $X = \xi_q^{-1}(w)$ belongs to $\ker(\xi_{\hat{q}})^{\perp}$, on which $\xi_{\hat{q}}$ is an isometry. \square

This lemma has a few trivial but important consequences, similar to Lemma 9.10.

Proposition 4. Sub-Riemannian lifts: *Let $\hat{q}_0 \in \hat{\mathcal{S}}$ and $q_0 = \pi(\hat{q}_0)$. Let $q : [0, 1] \rightarrow \mathcal{S}$ be an horizontal curve with $q(0) = q_0$. Then for any horizontal curve $t \mapsto \hat{q}(t)$ such that $\pi(\hat{q}(t)) = q(t)$ (such curves are called horizontal lifts of q to $\hat{\mathcal{S}}$),*

$$A(q) \leq A(\hat{q}).$$

Moreover, there is a unique such \hat{q} starting at \hat{q}_0 such that

$$A(q) = A(\hat{q}).$$

Its velocity is given by $\dot{\hat{q}}(t) = \xi_{\hat{q}(t)}\xi_{q(t)}^{-1}\dot{q}(t)$. We call it the minimizing lift of q to $\hat{\mathcal{S}}$ starting at \hat{q}_0 .

In particular, when q is a locally minimizing geodesic, so is its the minimizing lift.

We can now define the lifted sub-Riemannian structure on $\hat{\mathcal{S}}$ defined by π .

Definition 9.18. Lifted sub-Riemannian structure: *Let \mathcal{S} and $\hat{\mathcal{S}}$ be shape spaces in M of respective orders $\ell \leq \hat{\ell}$. Let \mathcal{H} be a right-invariant sub-Riemannian structure on $\mathcal{D}^{\mathfrak{so}+\hat{\ell}}(M)$ induced by vector fields of class $H^{\mathfrak{so}+k}$, $k \geq \hat{\ell}$. Assume there is an equivariant smooth submersion $\pi : \hat{\mathcal{S}} \rightarrow \mathcal{S}$.*

The lifted sub-Riemannian structure $\mathcal{H}_{\pi}^{\hat{\mathcal{S}}}$ on $\hat{\mathcal{S}}$ is the structure for which horizontal curves are exactly minimizing lifts of horizontal curves in \mathcal{S} . In other words, it is the distribution of subspaces on $\hat{\mathcal{S}}$ defined by

$$\mathcal{H}_{\pi}^{\hat{\mathcal{S}}} = \bigcup_{\hat{q} \in \hat{\mathcal{S}}} \xi_{\hat{q}} \left(\xi_{\pi(\hat{q})}^{-1}(\mathcal{H}^{\mathcal{S}}) \right),$$

with the Hilbert norm $\|\hat{w}\|_{\hat{q}} = \|d\pi_{\hat{q}}\hat{w}\|_{\pi(\hat{q})}$.

Lifted shapes can be used for studying how the deformation of a given shape induces the deformation of a "bigger" shape, i.e. one with more information.

We can then define horizontal curves for this new structure, the action A and length L of such curves (which coincide with the action and length for $\mathcal{H}^{\hat{S}}$), and the corresponding sub-Riemannian distance $d_{\pi}^{\hat{S}}$ on \hat{S} .

We have the following result, a straightforward consequence of Proposition 4.

Proposition 5. *Let $\hat{q} : [0, 1] \rightarrow \hat{S}$ be a curve of class H^1 , and define $q : t \mapsto \pi(\hat{q}(t))$ its projection onto \mathcal{S} . Then*

1. \hat{q} is horizontal for $\mathcal{H}_{\pi}^{\hat{S}}$ if and only if it is the minimizing lift of q . In this case, $A(\hat{q}) = A(q)$ and $L(\hat{q}) = L(q)$.
2. If q is a geodesic (resp. a minimizing geodesic), so is \hat{q} (for both sub-Riemannian structures).
3. As a consequence, π is a lipshitz-1 map: for every $\hat{q}_1, \hat{q}_2 \in \hat{M}$, $d_{\pi}^{\hat{S}}(\hat{q}_1, \hat{q}_2) \geq d^{\mathcal{S}}(\pi(\hat{q}_1), \pi(\hat{q}_2))$.

9.3.2 The end-point mapping under an additional assumption

Recall that we took V to have a continuous inclusion in $\Gamma^{s_0+\ell+1}(TM)$. We also require the following assumption, which will hold true for the rest of the section.

- The pseudo-inverse $d\pi^{-1} : \pi^*\mathcal{H}^{\mathcal{S}} \rightarrow \mathcal{H}_{\pi}^{\hat{S}}$ is of class \mathcal{C}^1 . In particular, so is the orthogonal projection tensor $\Pi_{\hat{q}} = d\pi_{\hat{q}}^{-1}d\pi_{\hat{q}}$ on $\mathcal{H}_{\pi}^{\hat{S}}$.

Remark 9.28. This assumption is automatically verified whenever the sub-Riemannian structure $\mathcal{H}^{\mathcal{S}}$ on the base space \mathcal{S} is a true fiber bundle of class \mathcal{C}^1 over \mathcal{S} . This is in particular the case when this sub-Riemannian structure is actually a Riemannian structure.

In finite dimension, the converse is also true, and the authors do not know wether that is the case in Banach shape spaces.

In this case, we can introduce another "infinitesimal action" $\zeta : \hat{S} \times V \rightarrow T\mathcal{S}$ of class \mathcal{C}^1 by

$$\zeta_{\hat{q}}(X) = \Pi_{\hat{q}}\xi_{\hat{q}}(X).$$

Horizontal curves are those curves $t \mapsto \hat{q}(t)$ of class H^1 for which there exists $X \in L^2(0, 1; V)$ such that for a.e. t ,

$$\dot{\hat{q}}(t) = \zeta_{\hat{q}(t)}(X(t)) = \Pi_{\hat{q}(t)}\xi_{\hat{q}(t)}(X(t))$$

Remark 9.29. There is no reason for $\hat{q}(t)$ to be equal to $\varphi^X(t) \circ \hat{q}_0$. In fact, this is almost never true.

We obviously get that $\mathcal{H}_{\pi}^{\hat{S}}$ is the image of ζ . Moreover, if $\hat{v} = \zeta_{\hat{q}}(Y)$, then

$$\|\hat{v}\|_{\hat{q}}^2 = \min_{Y \in V, \zeta_{\hat{q}}(Y) = \hat{v}} \langle Y, Y \rangle.$$

There exists a unique X such that this minimum is reached, given by $X = \zeta_{\hat{q}}^{-1}\dot{\hat{q}}(t)$, where $\zeta_{\hat{q}}^{-1}$ is the pseudo-inverse of ζ , which is, of course, equal to the restriction of $\xi_{\hat{q}}^{-1}$ to $\mathcal{H}_{\pi, \hat{q}}^{\hat{S}}$.

Therefore, for any curve $t \mapsto \hat{q}(t)$ in \hat{S} , horizontal for $\mathcal{H}_{\pi}^{\hat{S}}$, there is a unique time-dependent vector field $t \mapsto X(t) \in V$ with minimal action such that $\dot{\hat{q}}(t) = \zeta_{\hat{q}(t)}(X(t))$ almost everywhere.

Remark 9.30. For this particular time-dependent vector field X , we have

$$\xi_{\hat{q}(t)}(X(t)) = \xi_{\hat{q}(t)}\xi_{\hat{q}(t)}^{-1}(\dot{\hat{q}}(t)) = \dot{\hat{q}}(t) = \zeta_{\hat{q}(t)}(X(t)),$$

so we do have $\hat{q}(t) = \varphi^X(t) \circ \hat{q}(0)$ for every t .

Let $\hat{q}_0 \in \hat{\mathcal{S}}$. Since ζ is of class \mathcal{C}^1 , it is easy to see that so is the corresponding end-point mapping

$$\text{end}_{\hat{q}_0}^{\zeta} : L^2(0, 1; V) \rightarrow \hat{\mathcal{S}},$$

which to X associates $\hat{q}(1)$, where $\hat{q} : [0, 1] \rightarrow \hat{\mathcal{S}}$ is the unique solution to the Cauchy problem

$$\hat{q}(0) = \hat{q}_0, \quad \dot{\hat{q}} = \zeta_{\hat{q}}(X(t)).$$

Therefore, just as in Section 9.2.3, to find horizontal geodesics for the lifted structure between \hat{q}_0 and $\hat{q}_1 \in \hat{\mathcal{S}}$, we just need to find critical points of the action map $A : L^2(0, 1; V) \rightarrow \mathbb{R}$ with respect to the constraints $\text{end}_{\hat{q}_0}^{\zeta}(X) = \hat{q}_1$.

9.3.3 The lifted Hamiltonian geodesic equations

Introduce the lifted Hamiltonian $H^{\zeta} : T^*\hat{\mathcal{S}} \times V$ by

$$H^{\zeta}(\hat{q}, \hat{p}, X) = \hat{p}(\zeta_{\hat{q}}(X)) - \frac{1}{2}\|X\|^2.$$

The exact same proofs as in the previous sections give the following results.

Lemma 9.14. *Let $\hat{q} : [0, 1] \mapsto \hat{\mathcal{S}}$ be an horizontal curve for the lifted structure, and $X : [0, 1] \rightarrow V$ such that $\dot{\hat{q}} = \zeta_{\hat{q}}(X)$ almost everywhere. Let $\hat{q}_0 = \hat{q}(0)$ and $\hat{q}_1 = \hat{q}(1)$. Let $\hat{p}_1 \in T_{\hat{q}_1}^*\hat{\mathcal{S}}$. Then, for every $\delta X \in L^2(0, 1; V)$,*

$$dA_X(\delta X) = (d\text{end}_{\hat{q}_0}^{\zeta})^* \hat{p}_1(\delta X) - \int_0^1 \partial_X H^{\zeta}(\hat{q}(t), \hat{p}(t), X(t))(\delta X(t)) dt,$$

where $\hat{p}(t) \in T_{\hat{q}(t)}^*M$ satisfies $\hat{p}(1) = \hat{p}_1$ and

$$(\dot{\hat{q}}(t), \dot{\hat{p}}(t)) = \nabla^{\omega} H^{\zeta}(\hat{q}(t), \hat{p}(t), X(t)). \quad (9.18)$$

Corollary 9.2. *Let $\hat{q} : [0, 1] \rightarrow \mathcal{S}$ be an horizontal curve for the lifted structure with end-points \hat{q}_0 and \hat{q}_1 , and X as above. Then the following statements are equivalent*

1. *There exists $\hat{p}_1 \neq 0 \in T_{\hat{q}_1}^*\hat{\mathcal{S}}$ such that $dA_X = (d\text{end}_{\hat{q}_0}^{\zeta})^*_X \hat{p}_1$. In particular, $t \mapsto \hat{q}(t)$ is a geodesic.*
2. *There exists an absolutely continuous map $t \mapsto \hat{p}(t) \in T_{\hat{q}(t)}^*\hat{\mathcal{S}}$, with $\hat{p}(1) = \hat{p}_1$, such that for almost every $t \in [0, 1]$,*

$$\begin{aligned} 0 &= \partial_X H^{\zeta}(\hat{q}(t), \hat{p}(t), X(t)) = \zeta_{\hat{q}(t)}^* \hat{p}(t) - \langle X(t), \cdot \rangle, \\ (\dot{\hat{q}}(t), \dot{\hat{p}}(t)) &= \nabla^{\omega} H^{\zeta}(\hat{q}(t), \hat{p}(t), X(t)) = (\zeta_{\hat{q}(t)} X(t), -\partial_{\hat{q}}(\zeta_{\hat{q}(t)} X(t))^* \hat{p}(t)). \end{aligned} \quad (9.19)$$

Such a curve is called a normal geodesic.

The first equation in (9.19) is satisfied if and only if

$$\zeta_{\hat{q}(t)}^* \hat{p}(t) = \langle X(t), \cdot \rangle,$$

if and only if $X(t) = K^V \zeta_{\hat{q}(t)}^* \hat{p}(t)$.

As usual, we now take a look at the maximized hamiltonian $h^{\zeta} : T^*\hat{\mathcal{S}} \rightarrow \mathbb{R}$, defined by

$$h^{\zeta}(\hat{q}, \hat{p}) = H^{\zeta}(\hat{q}, \hat{p}, K^V \zeta_{\hat{q}(t)}^* \hat{p}(t)) = \frac{1}{2} \hat{p}(\zeta_{\hat{q}} K^V \zeta_{\hat{q}}^* \hat{p}) = \frac{1}{2} \hat{p}(d\pi_{\hat{q}}^{-1} K_q d\pi_{\hat{q}}^{-1*} \hat{p}) = \frac{1}{2} \|K_q d\pi_{\hat{q}}^{-1*} \hat{p}\|_q^2,$$

where $q = \pi(\hat{q})$.

It is interesting to note that in the case where $\mathcal{H}^{\mathcal{S}} = T\mathcal{S}$, i.e. if the sub-Riemannian structure on the smaller shape space is actually Riemannian, then $d\pi_{\hat{q}}^{-1*} \hat{p}$ is actually an element of $T_q^*\mathcal{S}$, and we obtain

$$h^{\zeta}(\hat{q}, \hat{p}) = h^{\mathcal{S}}(\pi(\hat{q}), d\pi_{\hat{q}}^{-1*} \hat{p}).$$

9.3.4 The momentum viewpoint and another form of the lifted geodesic equations

As we saw in Section 9.2, it can be difficult to compute $\xi_q^{-1}w$, and, therefore, $d\pi_{\hat{q}}^{-1}$. This means we need an easy way to build horizontal vectors. For this, we use the same method as in Section 9.2. Recall that we denote, for $\hat{q} \in \hat{\mathcal{S}}$ and $q = \pi(\hat{q})$,

$$K_{\hat{q}} = \xi_{\hat{q}} K^V \xi_{\hat{q}}^* : T_{\hat{q}}^* \hat{\mathcal{S}} \rightarrow \mathcal{H}_{\hat{q}}^{\hat{\mathcal{S}}}$$

and

$$K_q = \xi_q K^V \xi_q^* : T_q^* \mathcal{S} \rightarrow \mathcal{H}_q^{\mathcal{S}}.$$

Note that since $\xi_{\pi(\hat{q})} = d\pi_{\hat{q}} \xi_{\hat{q}}$, we have

$$K_{\pi(\hat{q})} = d\pi_{\hat{q}} K_{\hat{q}} d\pi_{\hat{q}}^*.$$

Lemma 9.15. *Let $\hat{q} \in \hat{\mathcal{S}}$, $q = \pi(\hat{q})$ and $p \in T_q^* \mathcal{S}$. Let $\hat{p} = d\pi_{\hat{q}}^* p$. Then*

$$K_{\hat{q}} \hat{p} = \xi_{\hat{q}} K^V \xi_{\hat{q}}^* p = d\pi_{\hat{q}}^{-1} K_q p \in \mathcal{H}_{\pi}^{\hat{\mathcal{S}}}.$$

In particular, $\text{Im}(K_{\hat{q}} d\pi_{\hat{q}}^*) \subset \mathcal{H}_{\pi}^{\hat{\mathcal{S}}}$ and

$$\|K_{\hat{q}} \hat{p}\|_{\hat{q}}^2 = \|K_q p\|^2 = p(K_q p).$$

Proof. We know that $\xi_q^{-1} K_q p = K^V \xi_q^* p$ and $d\pi_{\hat{q}} = \xi_{\hat{q}} \xi_q^{-1}$, so

$$d\pi_{\hat{q}}^{-1} K_q p = \xi_{\hat{q}} K^V \xi_{\hat{q}}^* p.$$

Since $d\pi_{\hat{q}} \xi_{\hat{q}} = \xi_q$, we have $\xi_q^* = \xi_{\hat{q}}^* d\pi_{\hat{q}}^*$, hence $\xi_q^* p = \xi_{\hat{q}}^* \hat{p}$ and

$$d\pi_{\hat{q}}^{-1} K_q p = \xi_{\hat{q}} K^V \xi_{\hat{q}}^* \hat{p} = K_{\hat{q}} \hat{p}.$$

Moreover,

$$\|K_{\hat{q}} \hat{p}\|_{\hat{q}}^2 = \|K^V \xi_{\hat{q}}^* \hat{p}\|^2 = \|K^V \xi_q^* p\|^2 = p(K_q p).$$

□

Example 9.8. Take $n \leq \hat{n} \in \mathbb{N}^*$, and $\mathcal{S} = \text{Lmk}_n(M)$ and $\hat{\mathcal{S}} = \text{Lmk}_{\hat{n}}(M)$. We define $\pi : \text{Lmk}_{\hat{n}}(M) \rightarrow \text{Lmk}_n(M)$, as above, by $\pi(x_1, \dots, x_{\hat{n}}) = (x_1, \dots, x_n)$. Now let V be a hilbert subspace of vector fields of class at least H^{s_0+1} , with associated kernel $K(x, y) : T_y^* M \rightarrow T_x M$.

Let $\hat{q} = (x_1, \dots, x_{\hat{n}})$ and $q = \pi(\hat{q})$. Now take some

$$p = (p_1, \dots, p_n) \in T_q^* \text{Lmk}_n(M) = T_{x_1}^* M \times \dots \times T_{x_n}^* M.$$

Then for $x \in M$,

$$X(x) = K^V \xi_q^* p(x) = \sum_{i=1}^n K(x, x_i) p_i,$$

so we get

$$d\pi^{-1} K_q p = \left(\sum_{i=1}^n K(x_1, x_i) p_i, \dots, \sum_{i=1}^n K(x_{\hat{n}}, x_i) p_i \right).$$

On the other hand, we have $\hat{p} = d\pi^* p = (p_1, \dots, p_n, 0, \dots, 0)$, so we also get

$$K_{\hat{q}} \hat{p} = \left(\sum_{i=1}^n K(x_1, x_i) p_i, \dots, \sum_{i=1}^n K(x_{\hat{n}}, x_i) p_i \right).$$

Example 9.9. Take $n \in \mathbb{N}^*$, $\mathcal{S} = Lmk_n(M)$ and $\hat{\mathcal{S}} = T Lmk_{\hat{n}}(M)$. We take $\pi : T Lmk_{\hat{n}}(M) \rightarrow Lmk_n(M)$ the usual projection. Now let V be a hilbert subspace of vector fields of class at least H^{s_0+2} , with associated kernel $K(x, y) : T_y^*M \rightarrow T_xM$.

Let $\hat{q} = (q, v) = (x_1, \dots, x_{\hat{n}}, v_1, \dots, v_n)$, where $q = (x_1, \dots, x_n)$ and $v_i \in T_{x_i}M$. Now take some $p = (p_1, \dots, p_n) \in T_q^*Lmk_n(M)$. Again, we have $\hat{p} = d\pi^*p = (p_1, \dots, p_n, 0, \dots, 0) \in T^*T Lmk_n(M)$, so we obtain

$$K_{\hat{q}}\hat{p} = (w_1, \dots, w_n, w'_1, \dots, w'_n),$$

with

$$w_i = \sum_{j=1}^n K(x_i, x_j)p_j \in T_{x_i}M$$

and

$$w'_i = \partial_{x_i} \left(\sum_{j=1}^n K(x_i, x_j)p_j \right) (v_i) \in T_{(x_i, v_i)}TM,$$

which is well-defined as the derivative at x_i of the map $K^V \xi_q^*p : M \rightarrow TM$ applied to v_i .

These examples show that the easiest way to find horizontal vectors (and, hence, horizontal curves) for the lifted structure is to consider those curves that are solution of the control system

$$\dot{q} = \xi_{\hat{q}} K^V \xi_{\pi(\hat{q})}^* u, \quad u \in T_{\pi(\hat{q})}^* \mathcal{S}.$$

In any local trivialization, this is a control system of class \mathcal{C}^1 with respect to \hat{q} even if the assumption from the previous sections that $d\pi^{-1}$ be of class \mathcal{C}^1 is not satisfied. In particular, it has a unique solution.

The action of such a curve is equal to

$$\frac{1}{2} \int_0^1 \frac{1}{2} u(t) (K_{\pi(\hat{q}(t))} u(t)) dt.$$

Then, one can introduce yet another Hamiltonian H , but this time defined on the vector bundle $T^*\hat{\mathcal{S}} \oplus_M \pi^*T^*\mathcal{S} \rightarrow \hat{\mathcal{S}}$, by

$$H(\hat{q}, \hat{p}, u) = \hat{p} \left(\xi_{\hat{q}} K^V \xi_{\pi(\hat{q})}^* u \right) - \frac{1}{2} u (K_{\pi(\hat{q})} u).$$

In this case, while $\partial_u H(\hat{q}, \hat{p}, u) : T_{\pi(\hat{q})}^* \mathcal{S} \rightarrow \mathbb{R}$ is well-defined as a derivative along fibers, the symplectic gradient

$$\nabla^{\hat{\omega}} H(\hat{q}, \hat{p}, u) \in TT^*\hat{\mathcal{S}}$$

is defined indepently of the trivialization of $\pi^*T^*\mathcal{S}$ if and only if $\partial_u H(\hat{q}, \hat{p}, u) = 0$, which is equivalent to

$$K_{\pi(\hat{q})} u = \xi_{\pi(\hat{q})} K^V \xi_{\hat{q}}^* p. \quad (9.20)$$

This implies in particular that $d\pi_{\hat{q}}^* u$ and p coincide on $\mathcal{H}_{\pi, \hat{q}}^{\hat{\mathcal{S}}}$.

Then

$$H(\hat{q}, \hat{p}, u) = \frac{1}{2} u (K_{\pi(\hat{q})} u) = h^{\mathcal{S}}(\pi(\hat{q}), u) = h_{\pi}^{\mathcal{S}}(\hat{q}, \hat{p}),$$

which does not depend on u as long as $K_{\pi(\hat{q})} u = \xi_{\pi(\hat{q})} K^V \xi_{\hat{q}}^* p$. Therefore, assuming that (9.20) can be solved in u , we can define the corresponding reduced Hamiltonian, which coincides with the reduced Hamiltonian $h_{\pi}^{\mathcal{S}}$ defined in the previous Section.

On the other hand, using the equalities $K_{\pi(\hat{q})} = d\pi_{\hat{q}}K_{\hat{q}}d\pi_{\hat{q}}^*$ and $\xi_{\pi(\hat{q})}^* = x_{\hat{q}}^*d\pi_{\hat{q}}^*$, we have, in a local trivialization,

$$\begin{aligned} \partial_{\hat{q}}H(\hat{q}, \hat{p}, u) &= \hat{p}((\partial_{\hat{q}}K_{\hat{q}})d\pi_{\hat{q}}^*u) + \hat{p}(K_{\hat{q}}(\partial_{\hat{q}}d\pi_{\hat{q}}^*)u) - \frac{1}{2}u(d\pi_{\hat{q}}(\partial_{\hat{q}}K_{\hat{q}})d\pi_{\hat{q}}^*u) - u(d\pi_{\hat{q}}K_{\hat{q}}(\partial_{\hat{q}}d\pi_{\hat{q}}^*)u) \\ &= (\hat{p} - \frac{1}{2}d\pi_{\hat{q}}^*u)((\partial_{\hat{q}}K_{\hat{q}})d\pi_{\hat{q}}^*u) + (\hat{p} - d\pi_{\hat{q}}^*u)(K_{\hat{q}}((\partial_{\hat{q}}d\pi_{\hat{q}}^*)u)). \end{aligned}$$

We deduce the following Hamiltonian geodesic equations

$$\begin{aligned} K_{\pi(\hat{q})}u &= \xi_{\pi(\hat{q})}K^V\xi_{\hat{q}}^*\hat{p}, \\ \dot{\hat{q}} &= \xi_{\hat{q}}K^V\xi_{\pi(\hat{q})}^*u, \\ \dot{\hat{p}} &= -(\hat{p} - \frac{1}{2}d\pi_{\hat{q}}^*u)((\partial_{\hat{q}}K_{\hat{q}})d\pi_{\hat{q}}^*u) - (\hat{p} - d\pi_{\hat{q}}^*u)(K_{\hat{q}}((\partial_{\hat{q}}d\pi_{\hat{q}}^*)u)). \end{aligned} \tag{9.21}$$

It is easy to check that when these equations admit a solution, this solution coincides with that of the geodesic equations (9.19) from the previous section. This is usually a better form for the geodesic equations, because there is no need to compute the derivative of ζ . The first form was needed because this control system using lifted momenta has "less" controls than the one we wanted to study, that is, we might have

$$\xi_{\hat{q}}K^V\xi_{\pi(\hat{q})}^*(T_{\pi(\hat{q})}^*\mathcal{S}) \subsetneq \mathcal{H}_{\pi}^{\hat{S}},$$

and therefore, we had to check that critical points of the action for the momentum point of view were indeed true critical points of the action.

9.3.5 Some examples of lifted geodesic equations

We consider some simple cases of lifted shape spaces and compute their geodesic equations. Note that, contrarily to examples from Section 9.2.4, even the landmark cases will be strictly sub-Riemannian.

Extra landmarks with the Gaussian kernel Let $n, m \in \mathbb{N}^*$, and consider $\mathcal{S} = Lmk_n(\mathbb{R}^d)$ and $\hat{\mathcal{S}} = Lmk_{n+m}(\mathbb{R}^d)$, with $\pi(x_1, \dots, x_{n+m}) = (x_1, \dots, x_n)$. Then $T^*\mathcal{S} = Lmk_n(\mathbb{R}^d) \times (\mathbb{R}^d)^{n*}$, and $T^*\hat{\mathcal{S}} = Lmk_{m+n}(\mathbb{R}^d) \times (\mathbb{R}^d)^{(m+n)*}$.

We assume that the reproducing kernel K of V is the diagonal Gaussian kernel, as introduced in Sections 9.1.5 and 9.2.4. Our goal is to compute the geodesic equations from (9.21) for this particular example. In other words, we first need to solve the linear system

$$K_{\pi(\hat{q})}u = \xi_{\pi(\hat{q})}K^V\xi_{\hat{q}}^*p.$$

Then, we shall compute $\partial\partial_{\hat{p}}H(\hat{q}, \hat{p}, u) = \xi_{\hat{q}}K^V\xi_{\pi(\hat{q})}^*u$ and $\partial_{\hat{q}}H(\hat{q}, \hat{p}, u) = (\hat{p} - \frac{1}{2}d\pi_{\hat{q}}^*u)((\partial_{\hat{q}}K_{\hat{q}})d\pi_{\hat{q}}^*u)$. Note that the term $(\hat{p} - d\pi_{\hat{q}}^*u)(K_{\hat{q}}((\partial_{\hat{q}}d\pi_{\hat{q}}^*)u))$ does not appear in $\partial_{\hat{q}}H(\hat{q}, \hat{p}, u)$ because $\partial_{\hat{q}}(d\pi_{\hat{q}}^*) = 0$.

Here, K_q can be identified to a $n \times n$ block matrix, with blocks of size $d \times d$, and the (i, j) -th bloc is given by $e(x_i - x_j)I_d$, with I_d the identity matrix. It is invertible (see[TY05]), and there exists $S = (S(x_i, x_j))_{i,j=1,\dots,n}$, a matrix of size $n \times n$ such that $\sum_{k=1}^n S(x_i, x_k)e(x_k - x_j) = 1$ if $i = j$ and 0 if $i \neq j$.

Then, for $\hat{q} = (x_1, \dots, x_{n+m})$, $u = (u_1, \dots, u_n)$ and $\hat{p} = (p_1, \dots, p_{n+m})$, we know that we have

$$K^V\xi_{\hat{q}}^*\hat{p}(x) = \sum_{j=1}^{n+m} e(x - x_j)p_j^T.$$

Hence, if $(v_1, \dots, v_n) = \xi_q K^V \xi_{\hat{q}}^* \hat{p}$, we obtain

$$v_j = \sum_{k=1}^{n+m} e(x_j - x_k) p_k^T, \quad j = 1, \dots, n.$$

So the equation $K_{\pi(\hat{q})} u = \xi_{\pi(\hat{q})} K^V \xi_{\hat{q}}^* p$ becomes

$$\sum_{k=1}^n e(x_j - x_k) u_k^T = \sum_{k=1}^{n+m} e(x_j - x_k) p_k^T, \quad i = 1, \dots, n.$$

This linear equation is solved in u by

$$u_i = \sum_{j=1}^n \sum_{k=1}^{n+m} S(x_i, x_j) e(x_j, x_k) p_k = p_i + \sum_{j=1}^n \sum_{k=n+1}^{n+m} S(x_i, x_j) e(x_j, x_k) p_k,$$

and therefore, if $\partial_{\hat{p}} H(\hat{q}, \hat{p}, u) = \xi_{\hat{q}} K^V \xi_{\hat{q}}^* u = (\hat{v}_1, \dots, \hat{v}_{n+m})$, we have

$$v_a = \sum_{i=1}^n e(x_a, x_i) u_i^T = \sum_{i=1}^n e(x_a, x_i) p_i^T + \sum_{i,j=1}^n \sum_{k=n+1}^{n+m} S(x_i, x_j) e(x_j, x_k) p_k^T.$$

Noting that $\partial_{\hat{q}} d\pi_{\hat{q}} = 0$, we must now compute

$$\partial_{\hat{q}} H(\hat{q}, \hat{p}, u) = (\hat{p} - \frac{1}{2} d\pi_{\hat{q}}^* u) ((\partial_{\hat{q}} K_{\hat{q}}) d\pi_{\hat{q}}^* u) + (\hat{p} - d\pi_{\hat{q}}^* u) = (\hat{p} - \frac{1}{2} d\pi_{\hat{q}}^* u) ((\partial_{\hat{q}} K_{\hat{q}}) d\pi_{\hat{q}}^* u).$$

If we denote $\hat{p}' = \hat{p} - \frac{1}{2} d\pi_{\hat{q}}^* u = (p'_1, \dots, p'_{n+m})$ and write $-\partial_{\hat{q}} H(\hat{q}, \hat{p}, u)^T = (\alpha_1, \dots, \alpha_{n+m})$, we have

$$\begin{aligned} \alpha_a &= \frac{1}{\sigma} \sum_{i=1}^n e(x_a - x_j) [p'_a \cdot u_i + u_a \cdot p'_i] (x_a - x_j)^T \\ &\quad + \frac{1}{\sigma} \sum_{i=n+1}^{n+m} e(x_a - x_i) [u_a \cdot p'_i] (x_a - x_i)^T, \quad a = 1, \dots, n, \\ \alpha_a &= \frac{1}{\sigma} \sum_{i=1}^n e(x_a - x_i) [p'_a \cdot u_i] (x_a - x_i)^T, \quad a = n+1, \dots, n+m. \end{aligned}$$

On the other hand, $p'_a = p_a - \frac{1}{2} u_a$, hence, we obtain

$$\begin{aligned} \alpha_a &= \frac{1}{\sigma} \sum_{i=1}^n e(x_a - x_i) [p_a \cdot u_i + u_a \cdot p_i - u_a \cdot u_i] (x_a - x_i)^T \\ &\quad + \frac{1}{\sigma} \sum_{i=n+1}^{n+m} e(x_a - x_i) [u_a \cdot p_i] (x_a - x_i)^T, \quad a = 1, \dots, n, \\ \alpha_a &= \frac{1}{\sigma} \sum_{i=1}^n e(x_a - x_i) [p_a \cdot u_i] (x_a - x_i)^T, \quad a = n+1, \dots, n+m. \end{aligned}$$

We can finally write the lifted geodesic equations on $\hat{\mathcal{S}}$

$$\begin{aligned}
u_a &= p_a + \sum_{i=1}^n \sum_{k=n+1}^{n+m} S(x_a, x_i) e(x_i - x_k) p_k, & a=1, \dots, n, \\
\dot{x}_a &= \sum_{i=1}^n e(x_a - x_i) u_i^T & a=1, \dots, n+m, \\
\dot{p}_a &= \frac{1}{\sigma} \sum_{i=1}^n e(x_a - x_i) [p_a \cdot u_i + u_a \cdot p_i - u_a \cdot u_i] (x_a - x_i)^T \\
&\quad + \frac{1}{\sigma} \sum_{i=n+1}^{n+m} e(x_a - x_i) [u_a \cdot p_i] (x_a - x_i)^T, & a=1, \dots, n, \\
\dot{p}_a &= \frac{1}{\sigma} \sum_{i=1}^n e(x_a - x_i) [p_a \cdot u_i] (x_a - x_i)^T, & a=n+1, \dots, n+m.
\end{aligned}$$

Tangent spaces of landmarks with the Gaussian kernel In this example, we keep our Gaussian kernel K as above and $\mathcal{S} = Lmk_n(\mathbb{R}^d)$ but we take $\hat{\mathcal{S}} = T\mathcal{S} = \mathcal{S} \times (\mathbb{R}^d)^n$ and $\pi(q, v) = q$. We denote $\hat{q} = (q, v) = (x_1, \dots, x_n, v_1, \dots, v_n)$. Again, we want to compute the geodesic equations (9.21), first by solving

$$K_{\pi(\hat{q})} u = \xi_{\pi(\hat{q})} K^V \xi_{\hat{q}}^* p,$$

then by computing $\partial \partial_{\hat{p}} H(\hat{q}, \hat{p}, u) = \xi_{\hat{q}} K^V \xi_{\pi(\hat{q})}^* u$ and $\partial_{\hat{q}} H(\hat{q}, \hat{p}, u) = (\hat{p} - \frac{1}{2} d\pi_{\hat{q}}^* u)((\partial_{\hat{q}} K_{\hat{q}}) d\pi_{\hat{q}}^* u)$. Again, the term $(\hat{p} - d\pi_{\hat{q}}^* u)(K_{\hat{q}}((\partial_{\hat{q}} d\pi_{\hat{q}}^*) u))$ does not appear in $\partial_{\hat{q}} H(\hat{q}, \hat{p}, u)$ because $\partial_{\hat{q}}(d\pi_{\hat{q}}^*) = 0$.

A momentum on $\hat{\mathcal{S}}$ is denoted $\hat{p} = (p, l) = (p_1, \dots, p_n, l_1, \dots, l_n)$, and we have

$$\xi_{q,v}^*(p, l)(X) = \sum_{j=1}^n p_j(X(x_j)) + l_j(dX_{x_j}(v_j)).$$

Therefore we get

$$K^V \xi_{q,v}^*(p, l)(x) = \sum_{j=1}^n e(x - x_j) p_j - \frac{1}{\sigma} \sum_{k=1}^n e(x - x_j) [(x - x_j) \cdot v_j] l_j^T.$$

Hence, the equation $K_{\pi(\hat{q})} u = \xi_{\pi(\hat{q})} K^V \xi_{\hat{q}}^* p$ becomes the linear problem

$$\sum_{k=1}^n e(x_j - x_k) u_k^T = \sum_{k=1}^n e(x_j - x_k) p_k^T - \frac{1}{\sigma} \sum_{k=1}^n e(x_j - x_k) [(x_j - x_k) \cdot v_k] l_k^T, \quad j = 1, \dots, n,$$

solved by

$$u_i = p_i - \frac{1}{\sigma} \sum_{j,k=1}^n S(x_i, x_j) e(x_j - x_k) [(x_j - x_k) \cdot v_k] l_k, \quad i = 1, \dots, n.$$

Then we compute $\xi_{q,v} K^V \xi_{\hat{q}}^* u = (w_1, \dots, w_n, w'_1, \dots, w'_n)$

$$\left. \begin{aligned}
w_a &= \sum_{i=1}^n e(x_a - x_i) u_i^T, \\
\left(= \sum_{i=1}^n e(x_a - x_i) p_i^T - \frac{1}{\sigma} \sum_{j,k=1}^n e(x_j - x_k) [(x_j - x_k) \cdot v_k] l_k^T \right) \\
w'_a &= -\frac{1}{\sigma} \sum_{i=1}^n e(x_a - x_i) [v_a \cdot (x_a - x_i)] u_i^T,
\end{aligned} \right\} \quad a = 1, \dots, n.$$

Now if we denote

$$\hat{p}' = \hat{p} - \frac{1}{2}d\pi_{\hat{q}}^*u = (p'_1, \dots, p'_n, l'_1, \dots, l'_n) = (p_1 - \frac{1}{2}u_1, \dots, p_n - \frac{1}{2}u_n, l_1, \dots, l_n),$$

we obtain

$$\hat{p}'(K_{\hat{q}}d\pi^*\hat{u}) = \sum_{i,j=1}^n e(x_i - x_j)[p'_i \cdot u_j] - \frac{1}{\sigma} \sum_{i,j=1}^n e(x_i - x_j)[v_i \cdot (x_i - x_j)][l_i \cdot u_j].$$

Therefore, for $-\partial_{\hat{q}}H(\hat{q}, \hat{p}', u) = (\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n)$, we get

$$\left. \begin{aligned} \alpha_a &= \frac{1}{\sigma} \sum_{j=1}^n e(x_a - x_j)[p'_a \cdot u_j + p'_j \cdot u_a](x_a - x_j)^T \\ &+ \frac{1}{\sigma} \sum_{j=1}^n e(x_a - x_j)([l_a \cdot u_j]v_a^T - [l_j \cdot u_a]v_j^T) \\ &- \frac{1}{\sigma^2} \sum_{j=1}^n e(x_a - x_j)[(x_a - x_j) \cdot ([l_a \cdot u_j]v_a - [l_j \cdot u_a]v_j)](x_a - x_j)^T, \\ \beta_a &= \frac{1}{\sigma} \sum_{j=1}^n e(x_a - x_j)[l_a \cdot u_j](x_a - x_j)^T, \end{aligned} \right\} a = 1, \dots, n.$$

Since $p'_i = p_i - \frac{1}{2}u_i$, $[p'_a \cdot u_j + p'_j \cdot u_a] = [p_a \cdot u_j + p_j \cdot u_a - u_j \cdot u_a]$ and we obtain the geodesic equations

$$\left. \begin{aligned} \dot{u}_a &= p_a - \frac{1}{\sigma} \sum_{j,k=1}^n S(x_a, x_j)e(x_j - x_k)[(x_j - x_k) \cdot v_k]l_k, \\ \dot{x}_a &= \sum_{i=1}^n e(x_a - x_i)u_i^T, \\ \dot{v}_a &= -\frac{1}{\sigma} \sum_{i=1}^n e(x_a - x_i)[v_a \cdot (x_a - x_i)]u_i^T \end{aligned} \right\} a = 1, \dots, n,$$

and

$$\left. \begin{aligned} \dot{p}_a &= \frac{1}{\sigma} \sum_{j=1}^n e(x_a - x_j)[p'_a \cdot u_j + p'_j \cdot u_a](x_a - x_j)^T \\ &+ \frac{1}{\sigma} \sum_{j=1}^n e(x_a - x_j)([l_a \cdot u_j]v_a^T - [l_j \cdot u_a]v_j^T) \\ &- \frac{1}{\sigma^2} \sum_{j=1}^n e(x_a - x_j)[(x_a - x_j) \cdot ([l_a \cdot u_j]v_a - [l_j \cdot u_a]v_j)](x_a - x_j)^T, \\ \dot{l}_a &= \frac{1}{\sigma} \sum_{j=1}^n e(x_a - x_j)[l_a \cdot u_j](x_a - x_j)^T, \end{aligned} \right\} a = 1, \dots, n.$$

Chapter 10

Shape deformation analysis from the optimal control viewpoint

The contents of this chapter were presented in Chapters 6 and 7.

Abstract

A crucial problem in shape deformation analysis is to determine a deformation of a given shape into another one, which is optimal for a certain cost. It has a number of applications in particular in medical imaging.

In this article we provide a new general approach to shape deformation analysis, within the framework of optimal control theory, in which a deformation is represented as the flow of diffeomorphisms generated by time-dependent vector fields. Using reproducing kernel Hilbert spaces of vector fields, the general shape deformation analysis problem is specified as an infinite-dimensional optimal control problem with state and control constraints. In this problem, the states are diffeomorphisms and the controls are vector fields, both of them being subject to some constraints. The functional to be minimized is the sum of a first term defined as geometric norm of the control (kinetic energy of the deformation) and of a data attachment term providing a geometric distance to the target shape.

This point of view has several advantages. First, it allows one to model general constrained shape analysis problems, which opens new issues in this field. Second, using an extension of the Pontryagin maximum principle, one can characterize the optimal solutions of the shape deformation problem in a very general way as the solutions of constrained geodesic equations. Finally, recasting general algorithms of optimal control into shape analysis yields new efficient numerical methods in shape deformation analysis. Overall, the optimal control point of view unifies and generalizes different theoretical and numerical approaches to shape deformation problems, and also allows us to design new approaches.

The optimal control problems that result from this construction are infinite dimensional and involve some constraints, and thus are nonstandard. In this article we also provide a rigorous and complete analysis of the infinite-dimensional shape space problem with constraints and of its finite-dimensional approximations.

Keywords: shape deformation analysis, optimal control, reproducing kernel Hilbert spaces, Pontryagin maximum principle, geodesic equations.

AMS classification: 58E99 49Q10 46E22 49J15 62H35 53C17 53C22 58D05

Introduction

The mathematical analysis of shapes has become a subject of growing interest in the past few decades, and has motivated the development of efficient image acquisition and segmentation methods, with applications to many domains, including computational anatomy and object recognition.

The general purpose of shape analysis is to compare two (or more) shapes in a way that takes into account their geometric properties. Two shapes can be very similar from a human's point of view, like a circle and an ellipse, but very different from a computer's automated perspective. In *Shape Deformation Analysis*, one optimizes a deformation mapping one shape onto the other and bases the analysis on its properties. This of course implies that a cost has been assigned to every possible deformation of a shape, the design of this cost function being a crucial step in the method. This approach has been used extensively in the analysis of anatomical organs from medical images (see [GM98]).

In this framework, a powerful and convenient approach represents deformations as flows of diffeomorphisms generated by time-dependent vector fields [DGM98, Tro95, Tro05]. Indeed, when considering the studied shapes as embedded in a real vector space \mathbb{R}^d , deformations of the whole space, like diffeomorphisms, induce deformations of the shape itself. The set of all possible deformations is then defined as the set of flows of time-dependent vector fields of a Hilbert space V , called space of "infinitesimal transformations", which is a subset of the space of all smooth bounded vector fields on \mathbb{R}^d .

This point of view has several interesting features, not the least of which being that the space of possible deformations is a well-defined subgroup of the group of diffeomorphisms, equipped with a structure similar to the one of a right-invariant sub-Riemannian metric [BR96, Mon02]. This framework has led to the development of a family of registration algorithms called *Large Deformation Diffeomorphic Metric Mapping* (LDDMM), in which the correspondence between two shapes comes from the minimization of an objective functional defined as a sum of two terms [AG04, BMTY05, JM00, MTY02, MTY06]. The first term takes into account the cost of the deformation, defined as the integral of the squared norm of the time-dependent vector field from which it arises. In a way, it is the total kinetic energy of the deformation. The second term is a data attachment penalizing the difference between the deformed shape and a target.

An appropriate class of Hilbert spaces of vector fields for V is the one of *reproducing kernel Hilbert spaces* (in short, RKHS) [Aro50], because they provide very simple solutions to the spline interpolation problem when the shape is given by a set of landmarks [TY11, You10], which is an important special case since it includes most practical situations after discretization. This framework allows one to use tools from Riemannian geometry [TY11], along with classical results from the theory of Lie groups equipped with right-invariant metrics [Arn66, Arn89, HMR98, MR99, You10]. These existing approaches provide an account for some of the geometric information in the shape, like singularities for example. However, they do not consider other intrinsic properties of the studied shape, which can also depend on the nature of the object represented by the shape. For example, for landmarks representing articulations of a robotic arm, the deformation can be searched so as to preserve the distance between certain landmarks. For cardiac motions, it may be relevant to consider deformations of the shape assuming that the movement only comes from a force applied only along the fiber structure of the muscle. In other words, it may be interesting to constrain the possible deformations (by considering non-holonomic constraints) in order to better fit the model.

In order to take into account such constraints in shape deformation problems, we propose to model these problems within the framework of *optimal control theory*, where the control system would model the evolution of the deformation and the control would be the time-dependent vector field (see preliminary ideas in [TY11]).

The purpose of this paper is to develop the point of view of *optimal control for shape deformation analysis* as comprehensively as possible. We will show the relevance of this framework, in particular because it can be used to model constrained shapes among many other applications.

Indeed, a lot of tools have been developed in control theory for solving optimal control problems with or without constraints. The well-known *Pontryagin maximum principle* (in short PMP, see [PBG64]) provides first-order conditions for optimality in the form of Hamiltonian extremal equations with a maximization condition permitting the computation of the optimal control. It has been generalized in many ways, and a large number of variants or improvements have been made over the past decades, with particular efforts in order to be able to address optimal control problems involving general state/control constraints (see the survey article [HSV95] and the many references therein). The analysis is, however, mainly done in finite dimension. Since shape analysis has a natural setting in infinite dimension (indeed, in 2D, the shape space is typically a space of smooth curves in \mathbb{R}^2), we need to derive an appropriate infinite-dimensional variant of the PMP for constrained problems. Such a variant is nontrivial and nonstandard, given that our constrained shape analysis problems generally involve an infinite number of equality constraints.

Such a PMP will allow us to derive in a rigorous geometric setting the (constrained) geodesic equations that must be satisfied by the optimal deformations.

Moreover, modeling shape deformation problems within the framework of optimal control theory can inherit from the many numerical methods in optimal control and thus lead to new algorithms in shape analysis.

The paper is organized as follows.

Section 10.1 is devoted to modeling shape deformation problems with optimal control. We first briefly describe, in Section 10.1.1, the framework of diffeomorphic deformations arising from the integration of time-dependent vector fields belonging to a given RKHS, and recall some properties of RKHS's of vector fields. In Section 10.1.2 we introduce the action of diffeomorphisms on a shape space, and we model and define the optimal control problem on diffeomorphisms which is at the heart of the present study, where the control system stands for the evolving deformation and the minimization runs over all possible time-dependent vector fields attached to a given RKHS and satisfying some constraints. We prove that, under weak assumptions, this problem is well posed and has at least one solution (Theorem 10.1). Since the RKHS is in general only known through its kernel, we then provide a kernel formulation of the optimal control problem and we analyze the equivalence between both problems. In Section 10.1.3 we investigate in our framework two important variants of shape spaces, which are lifted shapes and multi-shapes. Section 10.1.4 is devoted to the study of finite-dimensional approximations of the optimal control problem. Section 10.1.5 contains a proof of Theorem 10.1.

Section 10.2 is dedicated to the derivation of the constrained geodesic equations in shape spaces, that must be satisfied by optimal deformations. We first establish in Section 10.2.1 an infinite dimensional variant of the PMP which is adapted to our setting (Theorem 10.2). As an application, we derive in Section 10.2.2 the geodesic equations in shape spaces (Theorem 10.3), in a geometric setting, and show that they can be written as a Hamiltonian system.

In Section 10.3, we design some algorithms in order to solve the optimal control problem modeling the shape deformation problem. Problems without constraints are first analyzed in Section 10.3.1, and we recover some already known algorithms used in unconstrained shape spaces, however with a more general point of view. We are thus able to extend and generalize existing methods. Problems with constraints are investigated in Section 10.3.2 in view of solving constrained matching problems. We analyze in particular the augmented Lagrangian algorithm, and we also design a method based on shooting.

In Section 10.4 we provide numerical examples, investigating first a matching problem with constant total volume, and then a multishape matching problem.

10.1 Modelling shape deformation problems with optimal control

The following notations will be used throughout the paper. Let $d \in \mathbb{N}$ fixed. A vector $a \in \mathbb{R}^d$ can be as well viewed as a column matrix of length d . The Euclidean norm of a is denoted by $|a|$. The inner product $a \cdot b$ between two vectors $a, b \in \mathbb{R}^d$ can as well be written, with matrix notations, as $a^T b$, where a^T is the transpose of a . In particular one has $|a|^2 = a \cdot a = a^T a$.

Let X be a Banach space. The norm on X is denoted by $\|\cdot\|_X$, and the inner product by $(\cdot, \cdot)_X$ whenever X is a Hilbert space. The topological dual X^* of X is defined as the set of all linear continuous mappings $p : X \rightarrow \mathbb{R}$. Endowed with the usual dual norm $\|p\|_{X^*} = \sup\{p(x) \mid x \in X, \|x\|_X = 1\}$, it is a Banach space. For $p \in X^*$, the natural pairing between p and $w \in X$ is $p(w) = \langle p, w \rangle_{X^*, X}$, with the duality bracket. If $X = \mathbb{R}^n$ then p can be identified with a column vector through the equality $p(w) = p^T w$.

Let M be an open subset of X , and let Y be another Banach space. The Fréchet derivative of a map $f : M \rightarrow Y$ at a point $q \in M$ is written as df_q . When it is applied to a vector w , it is denoted by $df_q \cdot w$ or $df_q(w)$. When $Y = \mathbb{R}$, we may also write $\langle df_q, w \rangle_{X^*, X}$.

We denote by $W^{1,p}(0, 1; M)$ (resp. $H^1(0, 1; M)$) the usual Sobolev space of elements of $L^p(0, 1; M)$, with $1 \leq p \leq +\infty$ (resp., with $p = 2$) having a weak derivative in $L^p(0, 1; X)$. For $q_0 \in M$ we denote by $W_{q_0}^{1,p}(0, 1; M)$ (resp., by $H_{q_0}^1(0, 1; M)$) the space of all $q \in W^{1,p}(0, 1; M)$ (resp., $q \in H_{q_0}^1(0, 1; M)$) such that $q(0) = q_0$.

For every $\ell \in \mathbb{N}$, a mapping $\varphi : M \rightarrow M$ is called a \mathcal{C}^ℓ diffeomorphism if it is a bijective mapping of class \mathcal{C}^ℓ with an inverse of class \mathcal{C}^ℓ . The space of all such diffeomorphisms is denoted by $\text{Diff}^\ell(M)$. Note that $\text{Diff}^0(M)$ is the space of all homeomorphisms of M .

For every mapping $f : \mathbb{R}^d \rightarrow X$ of class \mathcal{C}^ℓ with compact support, we define the usual semi-norm

$$\|f\|_\ell = \sup \left\{ \left\| \frac{\partial^{\ell_1 + \dots + \ell_d} f(x)}{\partial x_1^{\ell_1} \dots \partial x_d^{\ell_d}} \right\|_X \mid x \in \mathbb{R}^d, (\ell_1, \dots, \ell_d) \in \mathbb{N}^d, \ell_1 + \dots + \ell_d \leq \ell \right\}.$$

We define the Banach space $\mathcal{C}_0^\ell(\mathbb{R}^d, \mathbb{R}^d)$ (endowed with the norm $\|\cdot\|_\ell$) as the completion of the space of vector fields of class \mathcal{C}^ℓ with compact support on \mathbb{R}^d with respect to the norm $\|\cdot\|_\ell$. In other words, $\mathcal{C}_0^\ell(\mathbb{R}^d, \mathbb{R}^d)$ is the space of vector fields of class \mathcal{C}^ℓ on \mathbb{R}^d whose derivatives of order less than or equal to ℓ converge to zero at infinity.

We define $\text{Diff}_0^\ell(\mathbb{R}^d)$ as the set of all diffeomorphisms of class \mathcal{C}^ℓ that converge to identity at infinity. Clearly, $\text{Diff}_0^\ell(\mathbb{R}^d)$ is the set of all $\varphi \in \text{Diff}^\ell(\mathbb{R}^d)$ such that $\varphi - \text{Id}_{\mathbb{R}^d} \in \mathcal{C}_0^\ell(\mathbb{R}^d, \mathbb{R}^d)$. It is a group for the composition law $(\varphi, \psi) \mapsto \varphi \circ \psi$.

Note that, if $\ell \geq 1$, then $\text{Diff}_0^\ell(\mathbb{R}^d)$ is an open subset of the affine Banach space $\text{Id}_{\mathbb{R}^d} + \mathcal{C}_0^\ell(\mathbb{R}^d, \mathbb{R}^d)$. This allows one to develop a differential calculus on $\text{Diff}_0^\ell(\mathbb{R}^d)$.

10.1.1 Preliminaries: deformations and RKHS of vector fields

Our approach to shape analysis is based on optimizing evolving deformations. A deformation is a one-parameter family of flows in \mathbb{R}^d generated by time-dependent vector fields on \mathbb{R}^d . Let us define this concept more rigorously.

Diffeomorphic deformations. Let $\ell \in \mathbb{N}^*$. Let

$$\begin{aligned} v : [0, 1] &\longrightarrow \mathcal{C}_0^\ell(\mathbb{R}^d, \mathbb{R}^d) \\ t &\longmapsto v(t) \end{aligned}$$

be a time-dependent vector field such that the real-valued function $t \mapsto \|v(t)\|_\ell$ is integrable. In other words, we consider an element v of the space $L^1(0, 1; \mathcal{C}_0^\ell(\mathbb{R}^d, \mathbb{R}^d))$.

According to the Cauchy-Lipshitz theorem, v generates a (unique) flow $\varphi : [0, 1] \rightarrow \text{Diff}_0^1(\mathbb{R}^d)$ (see, e.g., [AS04] or [Tré08, Chapter 11]), that is a one-parameter family of diffeomorphisms such that

$$\begin{aligned} \frac{\partial \varphi}{\partial t}(t, x) &= v(t, \varphi(t, x)), \\ \varphi(0, x) &= x, \end{aligned}$$

for almost every $t \in [0, 1]$ and every $x \in \mathbb{R}^d$. In other words, considering φ as a curve in the space $\text{Diff}_0^1(\mathbb{R}^d)$, the flow φ is the unique solution of

$$\begin{aligned} \dot{\varphi}(t) &= v(t) \circ \varphi(t), \\ \varphi(0) &= \text{Id}_{\mathbb{R}^d}. \end{aligned} \tag{10.1}$$

Such a flow φ is called a *deformation* of \mathbb{R}^d of class \mathcal{C}^ℓ .

Proposition 6. *The set of deformations of \mathbb{R}^d of class \mathcal{C}^ℓ coincides with the set*

$$\left\{ \varphi \in W^{1,1}(0, 1; \text{Diff}_0^\ell(\mathbb{R}^d)) \mid \varphi(0) = \text{Id}_{\mathbb{R}^d} \right\}.$$

In other words, the deformations of \mathbb{R}^d of class \mathcal{C}^ℓ are exactly the curves $t \mapsto \varphi(t)$ on $\text{Diff}_0^\ell(\mathbb{R}^d)$ that are integrable on $(0, 1)$ as well as their derivative, such that $\varphi(0) = \text{Id}_{\mathbb{R}^d}$.

Proof. Let us first prove that there exists a sequence of positive real numbers $(D_n)_{n \in \mathbb{N}}$ such that for every deformation φ of \mathbb{R}^d of class \mathcal{C}^ℓ , with $\ell \in \mathbb{N}^*$, induced by the time-dependent vector field $v \in L^1(0, 1; \mathcal{C}_0^\ell(\mathbb{R}^d, \mathbb{R}^d))$, one has

$$\sup_{t \in [0, 1]} \|\varphi(t) - \text{Id}_{\mathbb{R}^d}\|_i \leq D_i \exp\left(D_i \int_0^1 \|v(t)\|_i\right), \tag{10.2}$$

for every $i \in \{0, \dots, \ell\}$.

The case $i = 0$ is an immediate consequence of the integral formulation of (10.1). Combining the formula for computing derivatives of a composition of mappings with an induction argument shows that the derivatives of order i of $v \circ \varphi$ are polynomials in the derivatives of v and φ of order less than or equal to i . Moreover, these polynomials are of degree one with respect to the derivatives of v , and also of degree one with respect to the derivatives of φ of order i . Therefore we can write

$$\left| \frac{d}{dt} \partial_x^i \varphi(t, x) \right| \leq \|v(t)\|_i |\partial_x^i \varphi(t, x)| + \|v(t)\|_{i-1} P_i(|\partial_x^1 \varphi(t, x)|, \dots, |\partial_x^{i-1} \varphi(t, x)|), \tag{10.3}$$

where P_i is a polynomial independent of v and φ , and the norms of the derivatives of the $\partial_x^j \varphi(t, x)$ are computed in the space of \mathbb{R}^d -valued multilinear maps. The result then follows from Gronwall estimates and from an induction argument on i .

That any deformation of \mathbb{R}^d of class \mathcal{C}^ℓ is a curve of class $W^{1,1}$ in $\text{Diff}_0^\ell(\mathbb{R}^d)$ is then a direct consequence of (10.2) and (10.3).

Conversely, for every curve φ on $\text{Diff}_0^\ell(\mathbb{R}^d)$ of class $W^{1,1}$, we set $v(t) = \dot{\varphi}(t) \circ \varphi^{-1}(t)$, for every $t \in [0, 1]$. We have $\dot{\varphi}(t) = v(t) \circ \varphi(t)$ for almost every $t \in [0, 1]$, and hence it suffices to prove that $t \mapsto \|v(t)\|_\ell$ is integrable. The curve φ is continuous on $[0, 1]$ and therefore is bounded. This implies that $t \mapsto \varphi(t)^{-1}$ is bounded as well. The formula for computing derivatives of compositions of maps then shows that $\|v(t)\|_\ell$ is integrable whenever $t \mapsto \|\dot{\varphi}(t)\|_\ell$ is integrable, which completes the proof since φ is of class $W^{1,1}$. \square

Reproducing Kernel Hilbert Spaces of vector fields. Let us briefly recall the definition and a few properties of RKHS's (see [Aro50, TY11] for more details). Let $k \in \mathbb{N}^*$ be fixed.

Given a Hilbert space $(V, (\cdot, \cdot)_V)$, according to the Riesz representation theorem, the mapping $v \mapsto (v, \cdot)_V$ is a bijective isometry between V and V^* , whose inverse is denoted by K_V . Then for every $p \in V^*$ and every $v \in V$ one has $\langle p, v \rangle_{V^*, V} = (K_V p, v)_V$ and $\|p\|_{V^*}^2 = \|K_V p\|_V^2 = \langle p, K_V p \rangle_{V^*, V}$.

Definition 10.1. A *Reproducing Kernel Vector Space (RKHS) of vector fields of class \mathcal{C}_0^ℓ* is a Hilbert space $(V, (\cdot, \cdot)_V)$ of vector fields on \mathbb{R}^d , such that $V \subset \mathcal{C}_0^\ell(\mathbb{R}^d, \mathbb{R}^d)$ with continuous inclusion.

Let V be an RKHS of vector fields of class \mathcal{C}_0^ℓ . Then, for all $(b, y) \in (\mathbb{R}^d)^2$, by definition the linear form $b \otimes \delta_y$ on V , defined by $b \otimes \delta_y(v) = b^T v(y)$ for every $v \in V$, is continuous (actually this continuity property holds as well for every compactly supported vector-valued distribution of order at most ℓ on \mathbb{R}^d). By definition of K_V , there holds $\langle b \otimes \delta_y, v \rangle_{V^*, V} = (K_V(b \otimes \delta_y), v)_V$. The *reproducing kernel* K of V is then the mapping defined on $\mathbb{R}^d \times \mathbb{R}^d$, with values in the set of real square matrices of size d , defined by

$$K(x, y)b = K_V(b \otimes \delta_y)(x), \quad (10.4)$$

for all $(b, x, y) \in (\mathbb{R}^d)^3$. In other words, there holds $(K(\cdot, y)b, v)_V = b^T v(y)$, for all $(b, y) \in (\mathbb{R}^d)^2$ and every $v \in V$, and $K(\cdot, y)b = K_V(b \otimes \delta_y)$ is a vector field of class \mathcal{C}^ℓ in \mathbb{R}^d , element of V .

It is easy to see that $(K(\cdot, x)a, K(\cdot, y)b)_V = a^T K(x, y)b$, for all $(a, b, x, y) \in (\mathbb{R}^d)^4$, and hence that $K(x, y)^T = K(y, x)$ and that $K(x, x)$ is positive semi-definite under the assumption that no nontrivial linear combination $a_1^T v(x_1) + \dots + a_n^T v(x_n)$, with given distinct x_j 's can vanish for every $v \in V$. Finally, writing $K_V(a \otimes \delta_y)(x) = K(x, y)a = \int_{\mathbb{R}^d} K(x, s)a d\delta_y(s)$, we have

$$K_V p(x) = \int_{\mathbb{R}^d} K(x, y) dp(y), \quad (10.5)$$

for every compactly supported vector-valued distribution p on \mathbb{R}^d of order less than or equal to ℓ .¹

As explained in [Aro50, You10], one of the interests of such a structure is that we can define the kernel itself instead of defining the space V . Indeed a given kernel K yields a unique associated RKHS. It is usual to consider kernels of the form $K(x, y) = \gamma(|x - y|)Id_{\mathbb{R}^d}$ with $\gamma \in C^\infty(\mathbb{R})$. Such a kernel yields a metric that is invariant under rotation and translation. The most common model is when γ is a Gaussian function but other families of kernels can be used as well [TY05, You10].

10.1.2 From shape space problems to optimal control

We define a *shape space* in \mathbb{R}^d as an open subset M of a Banach space X on which the group of diffeomorphisms of \mathbb{R}^d acts in a certain way. The elements of M , called *states of the shape*, are denoted by q . They are usually subsets or immersed submanifolds of \mathbb{R}^d , with a typical definition of the shape space as the set $M = \text{Emb}^1(S, \mathbb{R}^d)$ of all embeddings of class \mathcal{C}^1 of a given Riemannian manifold S into \mathbb{R}^d . For example, if S is the unit circle then M is the set of all parametrized \mathcal{C}^1 simple closed curves in \mathbb{R}^d . In practical applications or in

1. Indeed, it suffices to note that

$$b^T K_V p(x) = (b \otimes \delta_x, K_V p)_{V^*, V} = (p, K_V b \otimes \delta_x) = \int_{\mathbb{R}^d} (K(y, x)b)^T dp(y) = b^T \int_{\mathbb{R}^d} K(x, y) dp(y).$$

numerical implementations, one has to consider finite-dimensional approximations, so that \mathcal{S} usually just consists of a finite set of points, and then M is a space of landmarks (see [TY05, You10] and see examples further).

Let us first explain how the group of diffeomorphisms acts on the shape space M , and then in which sense this action induces a control system on M .

The group structure of $\text{Diff}_0^\ell(\mathbb{R}^d)$. Let $\ell \in \mathbb{N}^*$. The set $\text{Diff}_0^\ell(\mathbb{R}^d)$ is an open subspace of the affine Banach space $\text{Id}_{\mathbb{R}^d} + \mathcal{C}_0^\ell(\mathbb{R}^d, \mathbb{R}^d)$ and also a group for the composition law. However, we can be more precise.

First of all, the mappings $(\varphi, \psi) \mapsto \varphi \circ \psi$ and $\varphi \mapsto \varphi^{-1}$ are continuous (this follows from the formula for the computation of the derivatives of compositions of mappings).

Moreover, for every $\psi \in \text{Diff}_0^\ell(\mathbb{R}^d)$, the right-multiplication mapping $\varphi \mapsto R_\psi(\varphi) = \varphi \circ \psi$ is Lipschitz and of class \mathcal{C}^1 , as the restriction of the continuous affine map $(\text{Id}_{\mathbb{R}^d} + v) \mapsto (\text{Id}_{\mathbb{R}^d} + v) \circ \psi$. Its derivative $(dR_\psi)_{\text{Id}_{\mathbb{R}^d}} : \mathcal{C}_0^\ell(\mathbb{R}^d, \mathbb{R}^d) \rightarrow \mathcal{C}_0^\ell(\mathbb{R}^d, \mathbb{R}^d)$ at $\text{Id}_{\mathbb{R}^d}$ is then given by $v \mapsto v \circ \psi$. Moreover, $(v, \psi) \mapsto v \circ \psi$ is easily seen to be continuous.

Finally, the mapping

$$\begin{aligned} \mathcal{C}_0^{\ell+1}(\mathbb{R}^d, \mathbb{R}^d) \times \text{Diff}_0^\ell(\mathbb{R}^d) &\rightarrow \mathcal{C}^\ell(\mathbb{R}^d, \mathbb{R}^d) \\ (v, \psi) &\mapsto v \circ \psi \end{aligned}$$

is of class \mathcal{C}^1 . Indeed we have $\|v \circ (\psi + \delta\psi) - v \circ \psi - dv_\psi \cdot \delta\psi\|_\ell = o(\|\delta\psi\|_\ell)$, for every $\delta\psi \in \mathcal{C}_0^\ell(\mathbb{R}^d, \mathbb{R}^d)$. Then, using the uniform continuity of any derivative $d^i v$ of order $i \leq \ell$, it follows that the mapping $\psi \mapsto dv_\psi$ is continuous. These properties are useful for the study of the Fréchet Lie group structure of $\text{Diff}_0^\infty(\mathbb{R}^d)$ [Omo74].

Group action on the shape space. In the sequel, we fix $\ell \in \mathbb{N}$, and we assume that the space $\text{Diff}_0^{\max(1, \ell)}(\mathbb{R}^d)$ acts continuously on M (recall that M is an open subset of a Banach space X) according to a mapping

$$\begin{aligned} \text{Diff}_0^{\max(1, \ell)}(\mathbb{R}^d) \times M &\rightarrow M \\ (\varphi, q) &\mapsto \varphi \cdot q, \end{aligned} \tag{10.6}$$

such that $\text{Id}_{\mathbb{R}^d} \cdot q = q$ and $\varphi \cdot (\psi \cdot q) = (\varphi \circ \psi) \cdot q$ for every $q \in M$ and all $(\varphi, \psi) \in (\text{Diff}_0^{\max(1, \ell)}(\mathbb{R}^d))^2$.

Definition 10.2. *M is a shape space of order $\ell \in \mathbb{N}$ if the action (10.6) is compatible with the properties of the group structure of $\text{Diff}_0^{\max(1, \ell)}(\mathbb{R}^d)$ described above, that is:*

- For every $q \in M$ fixed, the mapping $\varphi \mapsto \varphi \cdot q$ is Lipschitz with respect to the (weaker when $\ell = 0$) norm $\|\cdot\|_\ell$, i.e., there exists $\gamma > 0$ such that

$$\|\varphi_1 \cdot q - \varphi_2 \cdot q\|_X \leq \gamma \|\varphi_1 - \varphi_2\|_\ell \tag{10.7}$$

for all $(\varphi_1, \varphi_2) \in (\text{Diff}_0^{\max(1, \ell)}(\mathbb{R}^d))^2$.

- The mapping $\varphi \mapsto \varphi \cdot q$ is differentiable at $\text{Id}_{\mathbb{R}^d}$. This differential is denoted by ξ_q and is called the infinitesimal action of $\mathcal{C}_0^{\max(1, \ell)}(\mathbb{R}^d, \mathbb{R}^d)$. From (10.7) one has

$$\|\xi_q v\|_X \leq \gamma \|v\|_\ell,$$

for every $v \in \mathcal{C}_0^\ell(\mathbb{R}^d, \mathbb{R}^d)$, and if $\ell = 0$ then ξ_q has a unique continuous extension to the whole space $\mathcal{C}_0^0(\mathbb{R}^d, \mathbb{R}^d)$.

– The mapping

$$\begin{aligned} \xi : M \times \mathcal{C}_0^\ell(\mathbb{R}^d, \mathbb{R}^d) &\longrightarrow X \\ (q, v) &\longmapsto \xi_q v \end{aligned} \quad (10.8)$$

is continuous, and its restriction to $M \times \mathcal{C}_0^{\ell+1}(\mathbb{R}^d, \mathbb{R}^d)$ is of class \mathcal{C}^1 . In particular the mapping $q \mapsto \xi_q v$ is of class \mathcal{C}^1 , for every bounded vector field v of class $\mathcal{C}^{\ell+1}$.

Example 10.1. For $\ell \geq 1$, the action of $\text{Diff}_0^\ell(\mathbb{R}^d)$ on itself by left composition makes it a shape space of order ℓ in \mathbb{R}^d .

Example 10.2. Let $\ell \in \mathbb{N}$ and let S be a \mathcal{C}^ℓ smooth compact Riemannian manifold. Consider the space $M = X = \mathcal{C}^\ell(S, \mathbb{R}^d)$ equipped with its usual Banach norm. Then M is a shape space of order ℓ , where the action of $\text{Diff}_0^{\max(1, \ell)}(\mathbb{R}^d)$ is given by the composition $\varphi \cdot q = \varphi \circ q$. Indeed, it is continuous thanks to the rule for computing derivatives of a composition, and we also have

$$\|\varphi_1 \cdot q - \varphi_2 \cdot q\|_X \leq \gamma \|\varphi_1 - \varphi_2\|_\ell.$$

Moreover, given $q \in M$ and $v \in \mathcal{C}_0^\ell(\mathbb{R}^d, \mathbb{R}^d)$, $\xi_q v$ is the vector field along q given by $\xi_q(v) = v \circ q \in \mathcal{C}^\ell(M, \mathbb{R}^d)$. Finally, the formula for computing derivatives of a composition yields

$$\|v \circ (q + \delta q) - v \circ q - dv_q \cdot \delta q\|_X = o(\|\delta q\|_X),$$

for every $\delta q \in M$, and the last part of the definition follows. This framework describes most of shape spaces.

An interesting particular case of this general example is when $S = (s_1, \dots, s_n)$ is a finite set (zero-dimensional manifold), $X = (\mathbb{R}^d)^n$ and

$$M = \text{Lmk}_d(n) = \{(x_1, \dots, x_n) \in (\mathbb{R}^d)^n \mid x_i \neq x_j \text{ if } i \neq j\}$$

is a (so-called) space of n landmarks in \mathbb{R}^d . For $q = (x_1, \dots, x_n)$, the smooth action of order 0 is $\varphi \cdot q = (\varphi(x_1), \dots, \varphi(x_n))$. For $v \in \mathcal{C}_0^0(\mathbb{R}^2, \mathbb{R}^2)$, the infinitesimal action of v at q is given by $\xi_q(v) = (v(x_1), \dots, v(x_n))$.

Remarque 10.1. In most cases, and in all examples given throughout this paper, the mapping ξ restricted to $M \times \mathcal{C}_0^{\ell+k}(\mathbb{R}^d, \mathbb{R}^d)$ is of class \mathcal{C}^k , for every $k \in \mathbb{N}$.

Proposition 7. For every $q \in M$, the mapping $\varphi \mapsto \varphi \cdot q$ is of class \mathcal{C}^1 , and its differential at φ is given by $\xi_{\varphi \cdot q} dR_{\varphi^{-1}}$. In particular, given $q_0 \in M$ and given φ a deformation of \mathbb{R}^d of class $\mathcal{C}_0^{\max(1, \ell)}$, which is the flow of the time-dependent vector field v , the curve $t \mapsto q(t) = \varphi(t) \cdot q_0$ is of class $W^{1,1}$ and one has

$$\dot{q}(t) = \xi_{\varphi(t) \cdot q_0} \dot{\varphi}(t) \circ \varphi(t)^{-1} = \xi_{q(t)} v(t), \quad (10.9)$$

for almost every $t \in [0, 1]$.

Proof. Let $q_0 \in M$, fix $\varphi \in \text{Diff}_0^\ell(\mathbb{R}^d)$ and take $\delta\varphi \in T_\varphi \text{Diff}_0^\ell(\mathbb{R}^d) = \mathcal{C}_0^\ell(\mathbb{R}^d, \mathbb{R}^d)$. Then $\varphi + \delta\varphi \in \text{Diff}_0^\ell(\mathbb{R}^d)$ for $\|\delta\varphi\|_\ell$ small enough. We define $v = (dR_{\varphi^{-1}})_\varphi \delta\varphi = \delta\varphi \circ \varphi^{-1}$. We have

$$(\varphi + \delta\varphi) \cdot q = (\text{Id}_{\mathbb{R}^d} + v) \cdot (\varphi \cdot q) = \varphi \cdot q + \xi_{\varphi \cdot q} v + o(v) = \varphi \cdot q + \xi_{\varphi \cdot q} \dot{\varphi} \circ \varphi^{-1} + o(\delta\varphi),$$

and therefore the mapping $\varphi \mapsto \varphi \cdot q$ is differentiable at φ , with continuous differential $\xi_{\varphi \cdot q} dR_{\varphi^{-1}}$. The result follows. \square

The result of this proposition shows that the shape $q(t) = \varphi(t) \cdot q_0$ is evolving in time according to the differential equation (10.9), where v is the time-dependent vector field associated with the deformation φ .

At this step we make a crucial connection between shape space analysis and control theory, by adopting another point of view. The differential equation (10.9) can be seen as a control system on M , where the time-dependent vector field v is seen as a control. In conclusion, the group of diffeomorphisms acts on the shape space M , and this action induces a control system on M .

As said in the introduction, in shape analysis problems, the shapes are usually assumed to evolve in time according to the minimization of some objective functional [TY11]. With the control theory viewpoint developed above, this leads us to model the shape evolution as an optimal control problem settled on M , that we define hereafter.

Induced optimal control problem on the shape space. We assume that the action of $\text{Diff}_0^{\max(\ell, 1)}(\mathbb{R}^d)$ on M is smooth of order $\ell \in \mathbb{N}$. Let $(V, (\cdot, \cdot)_V)$ be an RKHS of vector fields of class \mathcal{C}_0^ℓ on \mathbb{R}^d . Let K denote its reproducing kernel (as defined in Section 10.1.1). Let Y be another Banach space. Most problems of shape analysis can be recast as follows.

Problem 1. Let $q_0 \in M$, and let $C : M \times V \rightarrow Y$ be a mapping such that $v \mapsto C(q, v) = C_q v$ is linear for every $q \in M$. Let $g : M \rightarrow \mathbb{R}$ be a function. We consider the problem of minimizing the functional

$$J_1(q, v) = \frac{1}{2} \int_0^1 \|v(t)\|_V^2 dt + g(q(1)) \quad (10.10)$$

over all $(q(\cdot), v(\cdot)) \in W_{q_0}^{1,1}(0, 1; M) \times L^2(0, 1; V)$ such that $\dot{q}(t) = \xi_{q(t)} v(t)$ and $C_{q(t)} v(t) = 0$ for almost every $t \in [0, 1]$.

In the problem above, q_0 stands for an initial shape, and C stands for continuous constraints. Recall that the infinitesimal action can be extended to the whole space $\mathcal{C}_0^\ell(\mathbb{R}^d, \mathbb{R}^d)$.

Note that if $t \mapsto v(t)$ is square-integrable then $t \mapsto \dot{q}(t)$ is square-integrable as well. Indeed this follows from the differential equation $\dot{q}(t) = \xi_{q(t)} v(t)$ and from Gronwall estimates. Therefore the minimization runs over the set of all $(q(\cdot), v(\cdot)) \in H_{q_0}^1(0, 1; M) \times L^2(0, 1; V)$.

Problem 1 is an infinite-dimensional optimal control problem settled on M , where the state $q(t)$ is a shape and the control $v(\cdot)$ is a time-dependent vector field. The constraints C can be of different kinds, as illustrated further. A particular but important case of constraints consists of *kinetic* constraints, i.e., constraints on the speed $\dot{q} = \xi_{q(t)} v$ of the state, which are of the form $C_{q(t)} \xi_{q(t)} v(t) = C_{q(t)} \dot{q}(t) = 0$. Pure state constraints, of the form $C(q(t)) = 0$ with a differentiable map $C : M \rightarrow Y$, are in particular equivalent to the kinetic constraints $dC_{q(t)} \cdot \xi_{q(t)} v(t) = 0$.

To the best of our knowledge, except very few studies (such as [You12]), only unconstrained problems have been studied so far (i.e., with $C = 0$). In contrast, the framework that we provide here is very general and permits to model and solve far more general constrained shape deformation problems.

Remarque 10.2. Assume V is an RKHS of class \mathcal{C}_0^1 , and let $v(\cdot) \in L^2(0, 1; V)$. Then v induces a unique deformation $t \mapsto \varphi(t)$ on \mathbb{R}^d , and the curve $t \mapsto q_v(t) = \varphi(t) \cdot q_0$ satisfies $q(0) = q_0$ and $\dot{q}_v(t) = \xi_{q_v(t)} v(t)$ for almost every $t \in [0, 1]$. As above, it follows from the Gronwall lemma that $q \in H_{q_0}^1(0, 1; M)$. Moreover, according to the Cauchy-Lipshitz theorem, if $\ell \geq 1$ then $q(\cdot)$ is the unique such element of $H_{q_0}^1(0, 1; M)$. Therefore, if $\ell \geq 1$ then Problem 1 is equivalent to the problem of minimizing the functional $v \mapsto J_1(v, q_v)$ over all $v \in L^2(0, 1; V)$ such that $C_{q_v(t)} v(t) = 0$ for almost every $t \in [0, 1]$.

Concerning the existence of an optimal solution of Problem 1, we need the following definition.

Definition 10.3. *A state q of a shape space M of order ℓ is said to have compact support if for some compact subset U of \mathbb{R}^d , for some $\gamma > 0$ and for all $(\varphi_1, \varphi_2) \in (\text{Diff}_0^{\max(\ell, 1)}(\mathbb{R}^d))^2$, we have*

$$\|\varphi_1 \cdot q - \varphi_2 \cdot q\| \leq \gamma \|(\varphi_1 - \varphi_2)|_U\|_\ell,$$

where $(\varphi_1 - \varphi_2)|_U$ denotes the restriction of $\varphi_1 - \varphi_2$ to U .

Except for $\text{Diff}_0^{\max(\ell, 1)}(\mathbb{R}^d)$ itself, every state of every shape space given so far in examples had compact support.

Theorem 10.1. *Assume that V is an RKHS of vector fields of class $\mathcal{C}^{\ell+1}$ on \mathbb{R}^d , that $q \mapsto C_q$ is continuous, and that g is bounded below and lower semi-continuous. If q_0 has compact support, then Problem 1 has at least one solution.*

In practice one does not usually have available a convenient, functional definition of the space V of vector fields. The RKHS V is in general only known through its kernel K , as already mentioned in Section 10.1.1 (and the kernel is often a Gaussian one). Hence Problem 1, formulated as such, is not easily tractable since one might not have a good knowledge (say, a parametrization) of the space V .

One can however derive, under a slight additional assumption, a different formulation of Problem 1 that may be more convenient and appropriate in view of practical issues. This is done in the next section, in which our aim is to obtain an optimal control problem only depending on the knowledge of the reproducing kernel K of the space V (and not directly on V itself), the solutions of which can be lifted back to the group of diffeomorphisms.

Kernel formulation of the optimal control problem. For a given $q \in M$, consider the transpose $\xi_q^* : X^* \rightarrow V^*$ of the continuous linear mapping $\xi_q : V \rightarrow X$. This means that for every $u \in X^*$ the element $\xi_q^*u \in V^*$ (sometimes called pullback) is defined by $\langle \xi_q^*u, v \rangle_{V^*, V} = \langle u, \xi_q(v) \rangle_{X^*, X}$, for every $v \in V$. Besides, by definition of K_V , there holds $\langle \xi_q^*u, v \rangle_{V^*, V} = \langle K_V \xi_q^*u, v \rangle_V$. The mapping $(q, u) \in M \times X^* \mapsto \xi_q^*u \in V^*$ is often called the *momentum map* in control theory [MR99].

We start our discussion with the following remark. As seen in Example 10.2, we observe that, in general, given $q \in M$ the mapping ξ_q is far from being injective (i.e., one-to-one). Its null space $\ker(\xi_q)$ can indeed be quite large, with many possible time-dependent vector fields v generating the same solution of $q(0) = q_0$ and $\dot{q}(t) = \xi_{q(t)}v(t)$ for almost every $t \in [0, 1]$.

A usual way to address this overdetermination consists of selecting, at every time t , a $v(t)$ that has minimal norm subject to $\xi_{q(t)}v(t) = \dot{q}(t)$ (resulting in a least-squares problem). This is the object of the following lemma.

Lemma 10.1. *Let $q \in M$. Assume that $\text{Range}(\xi_q) = \xi_q(V)$ is closed. Then, for every $v \in V$ there exists $u \in X^*$ such that $\xi_q v = \xi_q K_V \xi_q^* u$. Moreover, the element $K_V \xi_q^* u \in V$ is the one with minimal norm over all elements $v' \in V$ such that $\xi_q v' = \xi_q v$.*

Proof. Let \hat{v} denote the orthogonal projection of 0 on the space $A = \{v' : \xi_q v' = \xi_q v\}$, i.e., the element of A with minimal norm. Then \hat{v} is characterized by $\xi_q \hat{v} = \xi_q v$ and $\hat{v} \in \ker(\xi_q)^\perp$. Using the Banach closed-range theorem, we have $(\ker(\xi_q))^\perp = K_V \text{Range}(\xi_q^*)$, so that there exists $u \in X^*$ such that $\hat{v} = K_V \xi_q^* u$, and hence $\xi_q v = \xi_q K_V \xi_q^* u$. \square

Remarque 10.3. Note that the latter assertion in the proof does not require $\text{Range}(\xi_q)$ to be closed, since we always have $K_V(\text{Range}(\xi_q)) \subset (\ker(\xi_q))^\perp$.

Whether $\text{Range}(\xi_q) = \xi_q(V)$ is closed or not, this lemma and the previous discussion suggest replacing the control $v(t)$ in Problem 1 by $u(t) \in X^*$ such that $v(t) = K_V \xi_{q(t)}^* u(t)$. Plugging this expression into the system $\dot{q}(t) = \xi_{q(t)} v(t)$ leads to the new control system $\dot{q}(t) = K_{q(t)} u(t)$, where

$$K_q = \xi_q K_V \xi_q^*, \quad (10.11)$$

for every $q \in M$. The operator $K_q : X^* \rightarrow X$ is continuous and symmetric (i.e., $\langle u_2, K_q u_1 \rangle_{X^*, X} = \langle u_1, K_q u_2 \rangle_{X^*, X}$ for all $(u_1, u_2) \in (X^*)^2$), satisfies $\langle u, K_q u \rangle_{X^*, X} = \|K_V \xi_q^* u\|_V^2$ for every $u \in X$ and thus is positive semi-definite, and $(q, u) \mapsto K_q u$ is as regular as $(q, v) \mapsto \xi_q v$. Note that $K_q(X^*) = \xi_q(V)$ whenever $\xi_q(V)$ is closed.

This change of variable appears to be particularly relevant since the operator K_q is usually easy to compute from the reproducing kernel K_V of V , as shown in the following examples.

Example 10.3. Let $M = X = \mathcal{C}^0(S, \mathbb{R}^d)$ be the set of continuous mappings from a Riemannian manifold S to \mathbb{R}^d . The action of $\text{Diff}^0(\mathbb{R}^d)$ is smooth of order 0, with $\xi_q v = v \circ q$ (see Example 10.2). Let V be an RKHS of vector fields of class \mathcal{C}_0^1 on \mathbb{R}^d , with reproducing kernel K . Every $u \in X^*$ can be identified with a vector-valued Radon measure on S . Then

$$\langle \xi_q^* u, v \rangle_{V^*, V} = \langle u, v \circ q \rangle_{X^*, X} = \int_S v(q(s))^T du(s),$$

for every $q \in M$ and for every $v \in V$. In other words, one has $\xi_q^* u = \int_S du(s) \otimes \delta_{q(s)}$, and therefore, by definition of the kernel, we have $K_V \xi_q^* u = \int_S K_V(du(s) \otimes \delta_{q(s)}) = \int_S K(\cdot, q(s)) du(s)$. We finally infer that

$$K_q u(t) = \int_S K(q(t), q(s)) du(s).$$

Example 10.4. Let $X = (\mathbb{R}^d)^n$ and $M = \text{Lmk}_d(n)$ (as in Example 10.2). Then $\xi_q v = (v(x_1), \dots, v(x_n))$, and every $u = (u_1, \dots, u_n)$ is identified with a vector of X by $\langle u, w \rangle_{X^*, X} = \sum_{j=1}^n u_j^T w_j$. Therefore, we get $\xi_q^* u = \sum_{j=1}^n u_j \otimes \delta_{x_j}$, and $K_V \xi_q^* u = \sum_{j=1}^n K(x_j, \cdot) u_j$. It follows that

$$K_q u = \left(\sum_{j=1}^n K(x_1, x_j) u_j, \sum_{j=1}^n K(x_2, x_j) u_j, \dots, \sum_{j=1}^n K(x_n, x_j) u_j \right).$$

In other words, K_q can be identified with matrix of total size $nd \times nd$ and made of square block matrices of size d , with the block (i, j) given by $K(x_i, x_j)$.

Following the discussion above and the change of control variable $v(t) = K_V \xi_{q(t)}^* u(t)$, we are led to consider the following optimal control problem.

Problem 2. Let $q_0 \in M$, and let $C : M \times V \rightarrow Y$ be a mapping such that $v \mapsto C(q, v) = C_q v$ is linear for every $q \in M$. Let $g : M \rightarrow \mathbb{R}$ be a function. We consider the problem of minimizing the functional

$$J_2(q, u) = \frac{1}{2} \int_0^1 \langle u(t), K_{q(t)} u(t) \rangle_{X^*, X} dt + g(q(1)) = E(q(t)) + g(q(1)) \quad (10.12)$$

over all couples $(q(\cdot), u(\cdot))$, where $u : [0, 1] \rightarrow X^*$ is a measurable function and $q(\cdot) \in W_{q_0}^{1,1}(0, 1; M)$ are such that $\dot{q}(t) = K_{q(t)} u(t)$ and $C_{q(t)} K_V \xi_{q(t)}^* u(t) = 0$ for almost every $t \in [0, 1]$.

The precise relation between both problems is clarified in the following result.

Proposition 8. *Assume that $\ker(\xi_q) \subset \ker(C_q)$ and that $\text{Range}(\xi_q) = \xi_q(V)$ is closed, for every $q \in M$. Then Problems 1 and 2 are equivalent in the sense that $\inf J_1 = \inf J_2$ over their respective sets of constraints.*

Moreover, if $(\bar{q}(\cdot), \bar{u}(\cdot))$ is an optimal solution of Problem 2, then $(\bar{q}(\cdot), \bar{v}(\cdot))$ is an optimal solution of Problem 1, with $\bar{v}(\cdot) = K_V \xi_{\bar{q}(\cdot)}^ \bar{u}(\cdot)$ and $\bar{q}(\cdot)$ the corresponding curve defined by $\bar{q}(0) = q_0$ and $\dot{\bar{q}}(t) = K_{\bar{q}(t)} \bar{u}(t)$ for almost every $t \in [0, 1]$. Conversely, if $(\bar{q}(\cdot), \bar{v}(\cdot))$ is an optimal solution of Problem 1 then there exists a measurable function $\bar{u} : [0, 1] \rightarrow X^*$ such that $\bar{v}(\cdot) = K_V \xi_{\bar{q}(\cdot)}^* \bar{u}(\cdot)$, and $\bar{u}(\cdot)$ and $(\bar{q}(\cdot), \bar{u}(\cdot))$ is an optimal solution of Problem 2.*

Proof. First of all, if $J_2(q, u)$ is finite, then $v(\cdot)$ defined by $v(t) = K_V \xi_{q(t)} u(t)$ belongs to $L^2(0, 1; V)$ and therefore, using the differential equation $\dot{q}(t) = \xi_{q(t)} v(t)$ for almost every t and the Gronwall lemma, we infer that $q \in H_{q_0}^1(0, 1; M)$. The inequality $\inf J_1 \leq \inf J_2$ follows obviously.

Let us prove the converse. Let $\varepsilon > 0$ arbitrary, and let $v(\cdot) \in L^2(0, 1; V)$ and $q \in H_{q_0}^1(0, 1; M)$ be such that $J_1(q, v) \leq \inf J_1 + \varepsilon$, with $\dot{q}(t) = \xi_{q(t)} v(t)$ and $C_{q(t)} v(t) = 0$ for almost every $t \in [0, 1]$. We can write $v(t) = v_1(t) + v_2(t)$ with $v_1(t) \in \ker(\xi_{q(t)})$ and $v_2(t) \in (\ker(\xi_{q(t)}))^\perp = \text{Range}(K_V \xi_{q(t)}^*)$, for almost every $t \in [0, 1]$, with $v_1(\cdot)$ and $v_2(\cdot)$ measurable functions, and obviously one has $\int_0^T \|v_2(t)\|_V^2 dt \leq \int_0^T \|v(t)\|_V^2 dt$. Then, choosing $u(\cdot)$ such that $v_2(\cdot) = K_V \xi_{q(\cdot)}^* u(\cdot)$, it follows that $J_2(u) = J_1(v_2) \leq J_1(v) \leq \inf J_1 + \varepsilon$. Therefore $\inf J_2 \leq \inf J_1$. The rest is obvious. \square

Remark 10.4. Under the assumptions of Proposition 8 and of Theorem 10.1, Problem 2 has at least one solution $\bar{u}(\cdot)$, there holds $\min J_1 = \min J_2$, and the minimizers of Problems 1 and 2 are in one-to-one correspondance according to the above statement.

Remark 10.5. The assumption $\ker(\xi_q) \subset \ker(C_q)$ is satisfied in the important case where the constraints are kinetic, and is natural to be considered since it means that, in the problem of overdetermination in v , the constraints can be passed to the quotient (see Lemma 10.1). Actually for kinetic constraints we have the following interesting result (proved further, see Remark 10.17), completing the discussion on the equivalence between both problems.

Proposition 9. *Assume that V is an RKHS of vector fields of class at least $C_0^{\ell+1}$ on \mathbb{R}^d , that the constraints are kinetic, i.e., are of the form $C_q \xi_q v = 0$, and that the mapping $(q, w) \mapsto C_q w$ is of class C^1 . If $C_q \xi_q$ is surjective (onto) for every $q \in M$, then for every optimal solution \bar{v} of Problem 1 there exists a measurable function $\bar{u} : [0, 1] \rightarrow X^*$ such that $\bar{v} = K_V \xi_{\bar{q}}^* \bar{u}$, and \bar{u} is an optimal solution of Problem 2.*

Note that this result does not require the assumption that $\text{Range}(\xi_q) = \xi_q(V)$ be closed.

Remark 10.6. It may happen that Problems 1 and 2 do not coincide whenever $\text{Range}(\xi_q)$ is not closed. Actually, if the assumption that $\text{Range}(\xi_q)$ is closed is not satisfied then it may happen that the set of controls satisfying the constraints in Problem 2 be reduced to the zero control.

Let us provide a situation where this occurs. Let $v(q) \in \overline{\text{Range}(K_V \xi_q^*)} \setminus \text{Range}(K_V \xi_q^*)$ with $\|v(q)\|_V = 1$. In particular $v(q) \in (\ker(\xi_q))^\perp$. Assume that C_q is defined as the orthogonal projection onto $(\mathbb{R}v(q) \oplus \ker(\xi_q))^\perp = v(q)^\perp \cap (\ker(\xi_q))^\perp$. Then $\ker(C_q) = \mathbb{R}v(q) \oplus \ker(\xi_q)$. We claim that $\ker(C_q K_V \xi_q^*) = \{0\}$. Indeed, let $u \in X^*$ be such that $C_q K_V \xi_q^* u = 0$. Then on the one part $K_V \xi_q^* u \in \ker(C_q)$, and on the other part, $K_V \xi_q^* u \in \text{Range}(K_V \xi_q^*) \subset (\ker(\xi_q))^\perp$. Therefore $K_V \xi_q^* u \in \ker(C_q) \cap (\ker(\xi_q))^\perp = \mathbb{R}v(q)$, but since $v(q) \notin \text{Range}(K_V \xi_q^*)$, necessarily $u = 0$.

10.1.3 Further comments: lifted shape spaces and multishapes

In this section we provide one last way to study shape spaces and describe two interesting and important variants of shape spaces, namely lifted shape spaces and multishapes. We show that a slightly different optimal control problem can model the shape deformation problem in these spaces.

Lifted shape spaces. Lifted shapes can be used to keep track of additional parameters when studying the deformation of a shape. For example, when studying n landmarks (x_1, \dots, x_n) in \mathbb{R}^d , it can be interesting to keep track of how another point x is moved by the deformation.

Let M and \hat{M} be two shape spaces, open subsets of two Banach spaces X and \hat{X} respectively, on which the group of diffeomorphisms of \mathbb{R}^d acts smoothly with respective orders ℓ and $\hat{\ell}$. Let V be an RKHS of vector fields in \mathbb{R}^d of class $C_0^{\max(\ell, \hat{\ell})}$. We denote by ξ_q (respectively $\xi_{\hat{q}}$) the infinitesimal action of V on M (respectively \hat{M}). We assume that there exists a C^1 equivariant submersion $P : \hat{M} \rightarrow M$.

By equivariant, we mean that $P(\varphi \cdot \hat{q}) = \varphi \cdot P(\hat{q})$, for every diffeomorphism $\varphi \in \text{Diff}^{\max(\ell, \hat{\ell}), 1}$ and every $\hat{q} \in \hat{M}$. Note that this implies that $dP_{\hat{q}} \cdot \xi_{\hat{q}} = \xi_q$ and $\xi_{\hat{q}}^* dP_{\hat{q}}^* = \xi_q^*$.

For example, for $n < \hat{n}$, the projection $P : \text{Lmk}_d(\hat{n}) \rightarrow \text{Lmk}_d(n)$ defined by $P(x_1, \dots, x_{\hat{n}}) = (x_1, \dots, x_n)$ is a C^1 equivariant submersion. More generally, for a compact Riemannian manifold \hat{S} and a submanifold $S \subset \hat{S}$, the restriction mapping $P : \text{Emb}(\hat{S}, \mathbb{R}^d) \rightarrow \text{Emb}(S, \mathbb{R}^d)$ defined by $P(q) = q|_S$ is a C^1 equivariant submersion for the action by composition of $\text{Diff}^1(\mathbb{R}^d)$.

The constructions and results of Section 10.1.2 can be applied to this setting, and in particular the deformation evolution induces a control system on \hat{M} , as investigated previously.

Remarque 10.7. Let V be an RKHS of bounded vector fields of class $C_0^{\max(\ell, \hat{\ell})+1}$. Let g be a data attachment function on M and let C be a mapping of constraints. We set $\hat{g} = g \circ P$ and $\hat{C}_{\hat{q}} = C_{P(\hat{q})}$. Then a time-dependent vector field v in V is a solution of Problem 1 for M with constraints C and data attachment g if and only if it is also a solution of Problem 1 for \hat{M} with constraints \hat{C} and data attachment \hat{g} . This remark will be used for finite-dimensional approximations in Section 10.1.4.

One can however define a control system of a different form, by lifting the control applied on the smaller shape space M to the bigger shape space \hat{M} .

The method goes as follows. Let $q_0 \in M$ and $\hat{q}_0 \in P^{-1}(q_0)$. Consider a measurable map $u : [0, 1] \rightarrow X^*$ and the corresponding curve $q(\cdot)$ defined by $q(0) = q_0$ and $\dot{q}(t) = K_{q(t)} u(t)$ for almost every $t \in [0, 1]$, where $K_q = \xi_q K_V \xi_q^*$. This curve is the same as the one induced by the time-dependent vector field $v(\cdot) = K_V \xi_{q(\cdot)}^* u(\cdot)$. The deformation φ corresponding to the flow of v defines on \hat{M} a new curve $\hat{q}(t) = \varphi(t) \cdot \hat{q}_0$ with speed

$$\dot{\hat{q}}(t) = \xi_{\hat{q}} K_V \xi_{q(t)}^* u(t) = K_{\hat{q}(t)} dP_{\hat{q}(t)}^* u(t),$$

with $K_{\hat{q}} = \xi_{\hat{q}} K_V \xi_{\hat{q}}^*$. Note that $P(\hat{q}(t)) = q(t)$ for every $t \in [0, 1]$. We have thus obtained a new class of control problems.

Problem 3. Let $\hat{q}_0 \in \hat{M}$, and let $C : \hat{M} \times V \rightarrow Y$ be continuous and linear with respect to the second variable, with Y a Banach space. Let $g : \hat{M} \rightarrow \mathbb{R}$ be a real function on \hat{M} . We consider the problem of minimizing the functional

$$J_3(\hat{q}, u) = \frac{1}{2} \int_0^1 \langle u(t), K_{P(\hat{q}(t))} u(t) \rangle_{X^*, X} dt + g(\hat{q}(1))$$

over all $(\hat{q}(\cdot), u(\cdot))$, where $u : [0, 1] \rightarrow X^*$ is a measurable function and $\hat{q}(\cdot) \in W_{\hat{q}_0}^{1,1}(0, 1; \hat{M})$ are such that $\dot{\hat{q}}(t) = K_{\hat{q}(t)} dP_{\hat{q}(t)}^* u(t)$ and $C_{\hat{q}(t)} K_V \xi_{P(\hat{q}(t))}^* u(t) = 0$ for almost every $t \in [0, 1]$.

Note that, if g and C only depend on $P(\hat{q})$ then the solutions $u(t)$ of Problem 3 coincide with the ones of Problem 2 on M .

Problem 3 can be reformulated back into an optimal control problem on V and on \hat{M} , similar to Problem 1, by adding the constraints $D_{\hat{q}}v = 0$ where $D_{\hat{q}}v$ is the orthogonal projection of v on $\ker(\xi_{P(\hat{q})})$.

Some examples of lifted shape spaces can be found in [You12], where controls are used from a small number of landmarks to match a large number of landmarks, with additional state variables defining Gaussian volume elements. Another application of lifted shape spaces will be mentioned in Section 10.1.4, where they will be used to approximate infinite-dimensional shape spaces by finite-dimensional ones.

Multishapes. Shape analysis problems sometimes involve collections of shapes that must be studied together, each of them with specific properties associated with a different space of vector fields. These situations can be modeled as follows.

Consider some shape spaces M_1, \dots, M_k , open subsets of Banach spaces X_1, \dots, X_k , respectively, on which diffeomorphisms of \mathbb{R}^d acts smoothly on each shape space M_i with order ℓ_i . Let $k_i \geq 1$, and consider V_1, \dots, V_k , RKHS's of vector fields of \mathbb{R}^d respectively of class $\mathcal{C}_0^{\ell_i+k_i}$ with kernels K_1, \dots, K_k , as defined in Section 10.1.1. In such a model we thus get k control systems, of the form $\dot{q}_i(\cdot) = \xi_{i,q_i(\cdot)}v_i(\cdot)$, with the controls $v_i(\cdot) \in L^2(0, 1; V_i)$, $i = 1, \dots, k$. The shape space of a multi-shape is a space of the form $M = M_1 \times \dots \times M_k$. Let $q_0 = (q_{1,0}, \dots, q_{k,0}) \in M$. Similarly to the previous section we consider the problem of minimizing the functional

$$\sum_{i=1}^k \int_0^1 \|v_i(t)\|_{V_i}^2 dt + g(q_1(1), \dots, q_k(1)),$$

over all time-dependent vector fields $v_i(\cdot) \in L^2(0, 1; V_i)$, $i = 1, \dots, k$, and with $q_i(1) = \varphi_i(1) \cdot q_{i,0}$ where φ_i is the flow generated by v_i (note that, here, the problem is written without constraint).

As in Section 10.1.2, the kernel formulation of this optimal control problem consists of minimizing the functional

$$\frac{1}{2} \int_0^1 \sum_{i=1}^k \langle u_i(t), K_{q_i(t), i} u_i(t) \rangle_{X_i^*, X_i} dt + g(q(1)). \quad (10.13)$$

over all measurable functions $u(\cdot) = (u_1(\cdot), \dots, u_k(\cdot)) \in L^2(0, 1; X_1^* \times \dots \times X_k^*)$, where the curve $q(\cdot) = (q_1(\cdot), \dots, q_k(\cdot)) : [0, 1] \rightarrow M$ is the solution of $q(0) = q_0$ and

$$\dot{q}(t) = K_{q(t)} u(t) = \left(K_{1, q_1(t)} u_1(t), \dots, K_{k, q_k(t)} u_k(t) \right), \quad (10.14)$$

for almost every $t \in [0, 1]$, with $K_{i, q_i} = \xi_{i, q_i} K_{V_i} \xi_{i, q_i}^*$ for $i = 1, \dots, k$.

Obviously, without any further consideration, studying this space essentially amounts to studying each M_i separately, the only interaction possibly intervening from the final cost function g . More interesting problems arise however when the shapes can interact with each other, and are subject to consistency constraints. For example, assume that one studies a cat shape, decomposed into two parts for the body and the tail. Since these parts have very different properties, it makes sense to consider them *a priori* as two distinct shapes S_1 and S_2 , with shape spaces $M_1 = \mathcal{C}^0(S_1, \mathbb{R}^3)$ and $M_2 = \mathcal{C}^0(S_2, \mathbb{R}^3)$, each of them being associated with RKHS's V_1 and V_2 respectively. Then, in order to take account for the tail being attached to the cat's body, the contact point of the body and the tail of the cat must belong to both shapes and be equal. In other words, if $q_1 \in M_1$ represents the body and $q_2 \in M_2$ the tail,

then there must hold $q_1(s_1) = q_2(s_2)$ for some $s_1 \in S_1$ and $s_2 \in S_2$. This is a particular case of state constraints, i.e., constraints depending only on the state q of the trajectory.

Considering a more complicated example, assume that two (or more) shapes are embedded in a given background. Consider two states q_1 and q_2 in respective spaces $M_1 = \mathcal{C}^0(S_1, \mathbb{R}^d)$ and $M_2 = \mathcal{C}^0(S_2, \mathbb{R}^d)$ of \mathbb{R}^d . Assume that they represent the boundaries of two disjoint open subsets U_1 and U_2 of \mathbb{R}^d . We define a third space $M_3 = M_1 \times M_2$, whose elements are of the form $q_3 = (q_3^1, q_3^2)$. This shape space represents the boundary of the complement of $U_1 \cup U_2$ (this complement being the background). Each of these three shape spaces is acted upon by the diffeomorphisms of \mathbb{R}^d . Consider for every M_i an RKHS V_i of vector fields. The total shape space is then $M = M_1 \times M_2 \times M_3$, an element of which is given by $q = (q_1, q_2, q_3) = (q_1, q_2, q_3^1, q_3^2)$. Note that $\partial(U_1 \cup U_2) = \partial(U_1 \cup U_2)^c$. However, since (q_1, q_2) represents the left-hand side of this equality, and $q_3 = (q_3^1, q_3^2)$ the right-hand side, it only makes sense to impose the constraints $q_1 = q_3^1$ and $q_2 = q_3^2$. This model can be used for instance to study two different shapes that are required not to overlap during the deformation.

In this example, one can even go further: the background does not need to completely mimic the movements of the shapes. We can for example let the boundaries slide on each another. This imposes constraints on the speed of the shapes (and not just on the shapes themselves), of the form $C_q K_q u = 0$. See section 10.4 for additional details.

Multi-shapes are of great interest in computational anatomy and provide an important motivation to study shape deformation under constraints.

10.1.4 Finite dimensional approximation of optimal controls

The purpose of this section is to show that at least one solution of Problem 1 can be approximated by a sequence of solutions of a family of nested optimal control problems on finite-dimensional shape spaces with finite-dimensional constraints. We assume throughout that $\ell \geq 1$.

Let $(Y^n)_{n \in \mathbb{N}}$ be a sequence of Banach spaces and $(C^n)_{n \in \mathbb{N}}$ be a sequence of continuous mappings $C^n : M \times V \rightarrow Y^n$ that are linear and continuous with respect to the second variable. Let $(g^n)_{n \in \mathbb{N}}$ be a sequence of continuous functions on M , bounded from below with a constant independent of n . For every integer n , we consider the problem of minimizing the functional

$$J_1^n(v) = \frac{1}{2} \int_0^1 \|v(t)\|_V^2 dt + g^n(q(1)),$$

over all $v(\cdot) \in L^2(0, 1; V)$ such that $C_{q(t)}^n v(t) = 0$ for almost every $t \in [0, 1]$, where $q(\cdot) : [0, 1] \rightarrow M$ is the curve defined by $q(0) = q_0$ and $\dot{q}(t) = \xi_{q(t)} v(t)$ for almost every $t \in [0, 1]$. It follows from Theorem 10.1 that there exists an optimal solution $v^n(\cdot) \in L^2(0, 1; V)$. We denote by $q^n(\cdot)$ the corresponding curve.

Proposition 10. *Assume that V is an RKHS of vector fields of class $\mathcal{C}_0^{\ell+1}$ on \mathbb{R}^d and that the sequence $(\ker(C_q^n))_{n \in \mathbb{N}}$ is decreasing (in the sense of the inclusion) and satisfies*

$$\bigcap_{n \in \mathbb{N}} \ker(C_q^n) = \ker(C_q),$$

for every $q \in M$. Assume that g^n converges to g uniformly on every compact subset of M . Finally, assume that q_0 has compact support. Then the sequence $(v^n(\cdot))_{n \in \mathbb{N}}$ is bounded in $L^2(0, 1; V)$, and every cluster point of this sequence for the weak topology of $L^2(0, 1; V)$ is an optimal solution of Problem 1. More precisely, for every cluster point $\bar{v}(\cdot)$ of $(v^n(\cdot))_{n \in \mathbb{N}}$, there exists a subsequence such that $(v^{n_j}(\cdot))_{j \in \mathbb{N}}$ converges weakly to $\bar{v}(\cdot) \in L^2(0, 1; V)$, the sequence $(q^{n_j}(\cdot))_{j \in \mathbb{N}}$ of corresponding curves converges uniformly to $\bar{q}(\cdot)$, and $J_1^{n_j}(v^{n_j})$ converges to $\min J_1 = J_1(\bar{v})$ as j tends to $+\infty$, and $\bar{v}(\cdot)$ is a solution of Problem 1.

Proof. The sequence $(v^n(\cdot))_{n \in \mathbb{N}}$ is bounded in $L^2(0, 1; V)$ as a consequence of the fact that the functions g^n are uniformly bounded below. Let $\bar{v}(\cdot)$ be a cluster point of this sequence for the weak topology of $L^2(0, 1; V)$. Assume that $(v^{n_j}(\cdot))_{j \in \mathbb{N}}$ converges weakly to $\bar{v}(\cdot) \in L^2(0, 1; V)$. Denoting by $\bar{q}(\cdot)$ the curve corresponding to $\bar{v}(\cdot)$, the sequence $(q^{n_j}(\cdot))_{j \in \mathbb{N}}$ converges uniformly to $\bar{q}(\cdot)$ (see Lemma 10.2). Using the property of decreasing inclusion, we have $C_{q^{n_j}(\cdot)}^N v^{n_j}(\cdot) = 0$ for every integer N and every integer $j \geq N$. Using the same arguments as in the proof of Theorem 10.1 (see Section 10.1.5), it follows that $C_{\bar{q}(\cdot)} \bar{v}(\cdot) = 0$. Finally, since $\int_0^1 \|\bar{v}(t)\|_V^2 dt \leq \liminf \int_0^1 \|v^n(t)\|_V^2 dt$, and since g^n converges uniformly to g on every compact subset of M , it follows that $J_1(\bar{v}) \leq \liminf J_1^{n_j}(v^{n_j})$.

Since every $v \in \ker(C_q)$ belongs as well to $\ker(C_q^{n_j})$, it follows that $J_1^{n_j}(v^{n_j}) \leq J_1^{n_j}(v)$, for every time-dependent vector field $v(\cdot) \in L^2(0, 1; V)$ such that $C_{q(\cdot)} v(\cdot) = 0$, where $q(\cdot) : [0, 1] \rightarrow M$ is the curve corresponding to $v(\cdot)$. Since g^n converges uniformly to g , one has $J_1^{n_j}(v) \rightarrow J_1(v)$ as $n \rightarrow +\infty$. It follows that $\limsup J_1^{n_j}(v^{n_j}) \leq \min J_1$.

We have proved that $J_1(\bar{v}) \leq \liminf J_1^{n_j}(v^{n_j}) \leq \limsup J_1^{n_j}(v^{n_j}) \leq \min J_1$, and therefore $J_1(\bar{v}) = \min J_1$, that is, $\bar{v}(\cdot)$ is an optimal solution of Problem 1, and $J_1^{n_j}(v^{n_j})$ converges to $\min J_1 = J_1(\bar{v})$ as j tends to $+\infty$. \square

Application: approximation with finite dimensional shape spaces. Let S^1 be the unit circle of \mathbb{R}^2 , let $\ell \geq 1$ be an integer, $X = \mathcal{C}^\ell(S^1, \mathbb{R}^d)$ and let $M = \text{Emb}^\ell(S^1, \mathbb{R}^d)$ be the space of parametrized simple closed curves of class \mathcal{C}^ℓ on \mathbb{R}^d . We identify X with the space of all mappings $f \in \mathcal{C}^\ell([0, 1], \mathbb{R}^d)$ such that $f(0) = f(1)$, $f'(0) = f'(1)$, ..., $f^{(\ell)}(0) = f^{(\ell)}(1)$. The action of the group of diffeomorphisms of \mathbb{R}^d on M , defined by composition, is smooth of order ℓ (see Section 10.1.2). Let $q_1 \in M$ and $c > 0$ fixed. We define g by

$$g(q) = c \int_{S^1} |q(t) - q_1(t)|^2 dt,$$

for every $q \in M$. Consider pointwise kinetic constraints $C : M \times X \rightarrow \mathcal{C}^0(S^1, \mathbb{R}^m)$, defined by $(C_q \dot{q})(s) = F_{q(s)} \dot{q}(s)$ for every $s \in S^1$, with $F \in \mathcal{C}^0(\mathbb{R}^d, \mathcal{M}_{m,d}(\mathbb{R}))$, where $\mathcal{M}_{m,d}(\mathbb{R})$ is the set of real matrices of size $m \times d$.

Note that the multishapes constraints described in Section 10.1.3 are of this form.

Our objective is to approximate this optimal control problem with a sequence of optimal control problems specified on the finite dimensional shape spaces $\text{Lmk}_d(2^n)$.

For $n \in \mathbb{N}$, let ξ^n be the infinitesimal action of the group of diffeomorphisms on $\text{Lmk}_d(2^n)$, the elements of which are denoted by $q^n = (x_1^n, \dots, x_{2^n}^n)$. Define on the associated control problem the kinetic constraints $\tilde{C}_{q^n} \xi_{q^n}^n v = (F_{x_1^n} v(x_1^n), \dots, F_{x_{2^n}^n} v(x_{2^n}^n))$, and the data attachment function

$$\tilde{g}^n(q^n) = \frac{c}{2^n} \sum_{r=1}^{2^n} |x_r^n - q_1(2^{-n}r)|.$$

Let $v^n(\cdot)$ be an optimal control of Problem 1 for the above optimal control problem specified on $\text{Lmk}_d(2^n)$.

Proposition 11. *Every cluster point of the sequence $(v^n(\cdot))_{n \in \mathbb{N}}$ for the weak topology on $L^2(0, 1; V)$ is a solution of Problem 1 specified on $M = \text{Emb}^\ell(S^1, \mathbb{R}^d)$ with constraints and minimization functional respectively given by C and g defined above.*

Proof. Define the submersions $P^n : M \rightarrow \text{Lmk}_d(2^n)$ by

$$P^n(q) = (q(2^{-n}), \dots, q(2^{-n}r), \dots, q(1)).$$

Let $g^n = \tilde{g}^n \circ P^n$ and $C_q^n = \tilde{C}_{P^n(q)} dP_q$. In other words, g^n (resp. C^n) are the lifts of \tilde{g}^n (resp. \tilde{C}^n) from $\text{Lmk}_d(2^n)$ to M through P^n . Using Remark 10.7 on lifted shape spaces,

we infer that the optimal control $v_n(\cdot)$ of Problem 1 specified on the finite dimensional space $\text{Lmk}_d(2^n)$ with constraints \tilde{C}^n and data attachment \tilde{g}^n is also optimal for Problem 1 specified on the infinite dimensional set M with constraints C^n and data attachment g^n . Now, if $v \in \ker(C^n)$ for every integer n , then $F_{q(s)}v(q(s)) = 0$ for every $s = 2^{-n}r$ with $n \in \mathbb{N}$ and $r \in \{1, \dots, 2^n\}$. The set of such s is dense in $[0, 1]$, and $s \mapsto F_{q(s)}v(q(s))$ is continuous. Therefore $F_{q(s)}v(q(s)) = 0$ for every $s \in [0, 1]$, that is, $C_{q(\cdot)}\xi_{q(\cdot)}v(\cdot) = 0$. Since the converse is immediate, we get $\ker(C_{q(\cdot)}) = \bigcap_{n \in \mathbb{N}} \ker(C_{q(\cdot)}^n)$. Finally, since $q(\cdot)$ is a closed curve of class at least \mathcal{C}^ℓ with $\ell \geq 1$, it is easy to check that

$$g^n(q) = \frac{c}{2^n} \sum_{r=1}^{2^n} |q(2^{-n}r) - q_1(2^{-n}r)|$$

converges to g , uniformly on every compact subset of M . Therefore, Proposition 10 can be applied to the sequence $(v^n(\cdot))_{n \in \mathbb{N}}$, which completes the proof. \square

Remarque 10.8. The same argument works as well if we replace S^1 with any compact Riemannian manifold \mathcal{S} , and applies to the vertices of increasingly finer triangulations of \mathcal{S} .

10.1.5 Proof of Theorem 10.1

Let $(v_n(\cdot))_{n \in \mathbb{N}}$ be a sequence of $L^2(0, 1; V)$ such that $J_1(v_n)$ converges to its infimum. Let $(\varphi_n)_{n \in \mathbb{N}}$ be the corresponding sequence of deformations and let $(q_n(\cdot))_{n \in \mathbb{N}}$ be the sequence of corresponding curves (one has $q_n(t) = \varphi_n(t) \cdot q_0$ thanks to Remark 10.2). Since g is bounded below, it follows that the sequence $(v_n(\cdot))_{n \in \mathbb{N}}$ is bounded in $L^2(0, 1; V)$. The following lemma is well known (see [You10]), but we provide a proof for the sake of completeness.

Lemma 10.2. *There exist $\bar{v}(\cdot) \in L^2(0, 1; V)$, corresponding to the deformation $\bar{\varphi}$, and a sequence $(n_j)_{j \in \mathbb{N}}$ of integers such that $(v_{n_j}(\cdot))_{j \in \mathbb{N}}$ converges weakly to $\bar{v}(\cdot)$ and such that, for every compact subset U of \mathbb{R}^d ,*

$$\sup_{t \in [0, 1]} \|(\varphi_{n_j}(t, \cdot) - \bar{\varphi}(t, \cdot))|_U\|_\ell \xrightarrow{j \rightarrow +\infty} 0.$$

Proof of Lemma 10.2. Since the sequence $(v_n(\cdot))_{n \in \mathbb{N}}$ is bounded in the Hilbert space $L^2(0, 1; V)$, there exists a subsequence $(v_{n_j}(\cdot))_{j \in \mathbb{N}}$ converging weakly to some $\bar{v} \in L^2(0, 1; V)$. Besides, using (10.2) for $i = \ell + 1$ and the Ascoli theorem, we infer that for every compact subset U of \mathbb{R}^d , the sequence $(\varphi_{n_j})_{j \in \mathbb{N}}$ is contained in a compact subset of the space $\mathcal{C}^0([0, 1], \mathcal{C}_0^\ell(U, \mathbb{R}^d))$. Considering a compact exhaustion of \mathbb{R}^d and using a diagonal extraction argument, we can therefore extract a subsequence $(\varphi_{n_{j_k}})_{k \in \mathbb{N}}$ with limit $\bar{\varphi}$ such that, for any compact subset U of \mathbb{R}^d ,

$$\sup_{t \in [0, 1]} \|(\varphi_{n_{j_k}}(t, \cdot) - \bar{\varphi}(t, \cdot))|_U\|_\ell \xrightarrow{k \rightarrow +\infty} 0. \quad (10.15)$$

To complete the proof, it remains to prove that $\bar{\varphi}$ is the deformation induced by $\bar{v}(\cdot)$. On the first hand, we have $\lim_{k \rightarrow +\infty} \varphi_{n_{j_k}}(t, x) = \bar{\varphi}(t, x)$, for every $x \in \mathbb{R}^d$ and every $t \in [0, 1]$. On the second hand, one has, for every $k \in \mathbb{N}$,

$$\begin{aligned} \left| \varphi_{n_{j_k}}(t, x) - x - \int_0^t \bar{v} \circ \bar{\varphi}(s, x) ds \right| &= \left| \int_0^t \left(v_{n_{j_k}} \circ \varphi_{n_{j_k}}(s, x) - \bar{v} \circ \bar{\varphi}(s, x) \right) ds \right| \\ &\leq \left| \int_0^t v_{n_{j_k}} \circ \varphi_{n_{j_k}}(s, x) - v_{n_{j_k}} \circ \bar{\varphi}(s, x) ds \right| + \left| \int_0^t v_{n_{j_k}} \circ \bar{\varphi}(s, x) - \bar{v} \circ \bar{\varphi}(s, x) ds \right|. \end{aligned}$$

Set $C = \sup_{n \in \mathbb{N}} \int_0^1 \|v_n(t)\|_1 dt$, and define $\chi_{[0,t]} \delta_{\bar{\varphi}(\cdot, x)} \in V^*$ by $(\chi_{[0,t]} \delta_{\bar{\varphi}(\cdot, x)} | v)_{L^2(0,1;V)} = \int_0^t v \circ \bar{\varphi}(s, x) ds$. Then,

$$\begin{aligned} \left| \varphi_{n_{j_k}}(t, x) - x - \int_0^t \bar{v} \circ \bar{\varphi}(s, x) ds \right| &\leq C \sup_{s \in [0,t]} |\varphi_{n_{j_k}}(s, x) - \bar{\varphi}(s, x)| \\ &\quad + \left| (\chi_{[0,t]} \delta_{\bar{\varphi}(\cdot, x)} | v_{n_{j_k}} - \bar{v})_{L^2(0,1;V)} \right|, \end{aligned}$$

which converges to 0 as j tends to $+\infty$ thanks to (10.15) and to the weak convergence $v_{n_{j_k}}(\cdot)$ to $\bar{v}(\cdot)$. We thus conclude that $\bar{\varphi}(t, x) = x + \int_0^t \bar{v} \circ \bar{\varphi}(s, x) ds$, which completes the proof. \square

Setting $\bar{q}(t) = \bar{\varphi}(t) \cdot q_0$ for every $t \in [0, 1]$, one has $\dot{\bar{q}}(t) = \xi_{\bar{q}(t)} \bar{v}(t)$ for almost every $t \in [0, 1]$, and it follows from the above lemma and from the fact that q_0 has compact support that

$$\sup_{t \in [0,1]} \|\bar{q}(t) - q_{n_j}(t)\|_X \xrightarrow{j \rightarrow +\infty} 0.$$

The operator $v(\cdot) \mapsto C_{\bar{q}(\cdot)} v(\cdot)$ is linear and continuous on $L^2(0, 1; V)$, so it is also weakly continuous [BBI01]. We infer that the sequence $C_{q_{n_j}(\cdot)} v_{n_j}(\cdot)$ converges weakly to $C_{\bar{q}(\cdot)} \bar{v}(\cdot)$ in $L^2(0, 1; Y)$. Since $C_{q_{n_j}(\cdot)} v_{n_j}(\cdot) = 0$ for every $j \in \mathbb{N}$, it follows that $C_{\bar{q}(\cdot)} \bar{v}(\cdot) = 0$. In other words, the time-dependent vector field $\bar{v}(\cdot)$ satisfies the constraints.

It remains to prove that $\bar{v}(\cdot)$ is indeed optimal. From the weak convergence of the sequence $(v_{n_j}(\cdot))_{j \in \mathbb{N}}$ to $\bar{v}(\cdot)$ in $L^2(0, 1; V)$, we infer that

$$\int_0^1 \|\bar{v}(t)\|_V^2 dt \leq \liminf_{j \rightarrow +\infty} \int_0^1 \|v_{n_j}(t)\|_V^2 dt.$$

Besides, since g is lower continuous, $\liminf_{j \rightarrow +\infty} g(q_{n_j}(1)) \geq g(\bar{q}(1))$. Since $J_1(v_{n_j})$ converges to $\inf J_1$, it follows that $J_1(\bar{v}) = \inf J_1$.

10.2 Constrained geodesic equations in shape spaces

In this section, we derive first-order necessary conditions for optimality in Problem 1. We extend the well-known Pontryagin maximum principle (PMP) from optimal control theory to our infinite-dimensional framework, under the assumption that the constraints are surjective. This allows us to derive the constrained geodesic equations for shape spaces, and we show how they can be reduced to simple Hamiltonian dynamics on the cotangent space of the shape space.

10.2.1 First-order optimality conditions: PMP in shape spaces

We address the Pontryagin maximum principle in a slightly extended framework, considering a more general control system and a more general minimization functional than in Problem 1.

Let V be a Hilbert space, and let X and Y be Banach spaces. Let M be an open subset of X . Let $\xi : M \times V \rightarrow X$ and $C : M \times V \rightarrow Y$ be mappings of class \mathcal{C}^1 . Let $L : M \times V \rightarrow \mathbb{R}$ and $g : M \rightarrow \mathbb{R}$ be functions of class \mathcal{C}^1 . We assume that there exist continuous functions $\gamma_0 : \mathbb{R} \rightarrow \mathbb{R}$ and $\gamma_1 : X \rightarrow \mathbb{R}$ such that $\gamma_0(0) = 0$ and

$$\|dL_{q',v'} - dL_{q,v}\|_{X^* \times V^*} \leq \gamma_0(\|q' - q\|_X) + \gamma_1(q - q') \|v' - v\|_V, \quad (10.16)$$

for all $(q, q') \in M^2$ and all $(v, v') \in V^2$.

Let $q_0 \in M$. We consider the optimal control problem of minimizing the functional

$$J(q, v) = \int_0^1 L(q(t), v(t)) dt + g(q(1)) \tag{10.17}$$

over all $(q(\cdot), v(\cdot)) \in H_{q_0}^1(0, 1; M) \times L^2(0, 1; V)$ such that $\dot{q}(t) = \xi_{q(t)}v(t)$ and $C_{q(t)}v(t) = 0$ for almost every $t \in [0, 1]$. We define the *Hamiltonian* $H : M \times X^* \times V \times Y^* \rightarrow \mathbb{R}$ by

$$H(q, p, v, \lambda) = \langle p, \xi_q v \rangle_{X^*, X} - L(q, v) - \langle \lambda, C_q v \rangle_{Y^*, Y}. \tag{10.18}$$

It is a function of class \mathcal{C}^1 . Using the canonical injection $X \hookrightarrow X^{**}$, we have $\partial_p H = \xi_q v$.

Remarque 10.9. The estimate (10.16) on L is exactly what is required to ensure that the mapping $(q, v) \mapsto \int_0^1 L(q(t), v(t)) dt$ be well defined and Fréchet differentiable for every $(q, v) \in H_{q_0}^1(0, 1; M) \times L^2(0, 1; V)$. Indeed, the estimate $L(q(t), v(t)) \leq L(q(t), 0) + \gamma_1(q(t), q(t))\|v(t)\|_V^2$ implies the integrability property. The differentiability is an immediate consequence of the following estimate, obtained by combining (10.16) with the mean value theorem: for every $t \in [0, 1]$, and for some $s_t \in [0, 1]$, one has

$$\begin{aligned} & |L(q(t) + \delta q(t), v(t) + \delta v(t)) - L(q(t), v(t)) - dL_{q(t), v(t)}(\delta q(t), \delta v(t))| \\ & \leq \left(\gamma_0(\|\delta q(t)\|_X) + \gamma_1(s_t \delta q(t))\|\delta v(t)\| \right) (\|\delta q(t)\|_X + \|\delta v(t)\|_V). \end{aligned}$$

Theorem 10.2. *Assume that the linear operator $C_{q(t)} : V \rightarrow Y$ is surjective for every $q \in M$. Let $(q(\cdot), v(\cdot)) \in H_{q_0}^1(0, 1; X) \times L^2(0, 1; V)$ be an optimal solution of the above optimal control problem. Then there exist $p \in H^1(0, 1; X^*)$ and $\lambda \in L^2(0, 1; Y^*)$ such that $p(1) + dg_{q(1)} = 0$ and*

$$\begin{aligned} \dot{q}(t) &= \partial_p H(q(t), p(t), v(t), \lambda(t)), \\ \dot{p}(t) &= -\partial_q H(q(t), p(t), v(t), \lambda(t)), \\ \partial_v H(q(t), p(t), v(t), \lambda(t)) &= 0, \end{aligned} \tag{10.19}$$

for almost every $t \in [0, 1]$.

Remarque 10.10. This theorem is the extension of the usual PMP to our specific infinite dimensional setting. Any quadruple $q(\cdot), p(\cdot), v(\cdot), \lambda(\cdot)$ solution of the above equations is called an *extremal*. This is a "weak" maximum principle, in the sense that we derive the condition $\partial_v H$ along any extremal, instead of the stronger maximization condition

$$H(q(t), p(t), v(t), \lambda(t)) = \max_{w \in \ker(C_{q(t)})} H(q(t), p(t), w, \lambda(t))$$

for almost every $t \in [0, 1]$. Note however that, in the case of shape spaces, $v \mapsto H(q, p, v)$ is strictly concave and hence both conditions are equivalent.

Remarque 10.11. It is interesting to note that, if we set $V = H^\ell(\mathbb{R}^d, \mathbb{R}^d)$ with ℓ large enough, and if $L(q, v) = \frac{1}{2} \int_{\mathbb{R}^d} |v(x)|^2 dx$ and $C_q v = \operatorname{div} v$, then the extremal equations given in Theorem 10.2 coincide with the incompressible Euler equation. In other words, we recover the well-known fact that every divergence-free time-dependent vector field minimizing its L^2 norm (in t and x) must satisfy the incompressible Euler equation (see [Arn66]).

Remarque 10.12. Note that the surjectivity assumption is a strong one in infinite dimension. It is usually not satisfied in the case of shape spaces when Y is infinite dimensional. For instance, consider the shape spaces $M = \mathcal{C}^0(S, \mathbb{R}^d)$, with S a smooth compact Riemannian manifold. Let V be an RKHS of vector fields of class \mathcal{C}_0^1 , acting on M as described in Section 10.1.2. Let T be a submanifold of S of class \mathcal{C}^1 . Set $Y = \mathcal{C}^0(T, \mathbb{R}^d)$, and consider the kinetic constraints $C : M \times V \rightarrow Y$ defined by $C_q v = v \circ q|_T$. If q is differentiable along T , then no nondifferentiable map $f \in Y$ is in $\operatorname{Range}(C_q)$.

Remarque 10.13. It is possible to replace C_q with the orthogonal projection on $\ker(C_q)^\perp$, which is automatically surjective. However in this case the constraints become fiber-valued, and for the proof of our theorem to remain valid, one needs to assume that there exists a Hilbert space V_1 such that, for every $q_0 \in M$, there exists a neighborhood U of q_0 in M such that $\bigcup_{q \in U} \ker(C_q) \sim U \times V_1$.

Remarque 10.14. An important consequence of the surjectivity of C_q is that the norm on Y is equivalent to the Hilbert norm induced from V by C_q , for every $q \in M$. The operator $C_q K_V C_q^* : Y^* \rightarrow Y$ is the isometry associated with this norm. In particular Y must be reflexive, and hence $L^2(0, 1; Y^*) = L^2(0, 1; Y)^*$ (whereas we only have an inclusion for general Banach spaces).

Remarque 10.15. Theorem 10.2 withstands several generalizations. For instance it remains valid whenever we consider nonlinear constraints $C(q, v) = 0$ and a general Lagrangian L of class \mathcal{C}^1 without any estimates and a nonlinear control system $\dot{q} = \xi(q, v)$, provided that ξ is of class \mathcal{C}^1 and that $\partial_v C(q(t), v(t))$ is surjective for every $t \in [0, 1]$. However, this requires to consider $v \in L^\infty(0, 1; V)$ and makes the proof of the regularity of the Lagrange multipliers slightly more involved.

Before proving Theorem 10.2, it can be noted that many versions of the PMP can be found in the existing literature for infinite-dimensional optimal control problems – for instance, with dynamics consisting of partial differential equations, and however, most of the time, without constraint on the state. Versions of the PMP with state constraints can also be found in the literature (see the survey [HSV95]), most of the time in finite dimension, and, for the very few of them existing in infinite dimension, under the additional assumption that the constraints are of finite codimension. To the best of our knowledge, no version does exist that would cover our specific framework, concerning shape spaces, group actions, with an infinite number of constraints on the acting diffeomorphisms. The proof that we provide hereafter contains some subtleties such as Lemma 10.4, and hence Theorem 10.2 is a nontrivial extension of the usual PMP.

Proof of Theorem 10.2. We define the mapping $\Gamma : H_{q_0}^1(0, 1; M) \times L^2(0, 1; V) \rightarrow L^2(0, 1; X) \times L^2(0, 1; Y)$ by $\Gamma(q, v) = (\Gamma_1(q, v), \Gamma_2(q, v))$ with $\Gamma_1(q, v)(t) = \dot{q}(t) - \xi_{q(t)}v(t)$ and $\Gamma_2(q, v)(t) = C_{q(t)}v(t)$ for almost every $t \in [0, 1]$. The mapping Γ stands for the constraints imposed to the unknowns of the optimal control problem.

The functional $J : H_{q_0}^1(0, 1; M) \times L^2(0, 1; V) \rightarrow \mathbb{R}$ and the mapping Γ are of class \mathcal{C}^1 , and their respective differentials at some point (q, v) are given by

$$dJ_{(q,v)}(\delta q, \delta v) = \int_0^1 (\langle \partial_q L(q(t), v(t)), \delta q(t) \rangle_{X^*, X} + \langle \partial_v L(q(t), v(t)), \delta v(t) \rangle_{V^*, V}) dt + \langle dg_{q(1)}, \delta q(1) \rangle_{X^*, X},$$

for all $(\delta q, \delta v) \in H_0^1(0, 1; X) \times L^2(0, 1; V)$, and $d\Gamma_{(q,v)} = (d\Gamma_1(q,v), d\Gamma_2(q,v))$ with

$$\begin{aligned} (d\Gamma_1(q,v)(\delta q, \delta v))(t) &= \dot{\delta q}(t) - \partial_q \xi_{q(t)}v(t) \cdot \delta q(t) - \xi_{q(t)}\delta v(t), \\ (d\Gamma_2(q,v)(\delta q, \delta v))(t) &= \partial_q C_{q(t)}v(t) \cdot \delta q(t) + C_{q(t)}\delta v(t), \end{aligned}$$

for almost every $t \in [0, 1]$.

Lemma 10.3. *For every $(q, v) \in H_{q_0}^1(0, 1; M) \times L^2(0, 1; V)$, the linear continuous mapping $d\Gamma_{(q,v)} : H_0^1(0, 1; X) \times L^2(0, 1; V) \rightarrow L^2(0, 1; X) \times L^2(0, 1; Y)$ is surjective. Moreover the mapping $\partial_q \Gamma_1(q,v) : H_0^1(0, 1; M) \rightarrow L^2(0, 1; X)$ is an isomorphism.*

Proof. Let $(q, v) \in H_{q_0}^1(0, 1; M) \times L^2(0, 1; V)$ and $(a, b) \in L^2(0, 1; X) \times L^2(0, 1; Y)$. Let us prove that there exists $(\delta q, \delta v) \in H_0^1(0, 1; X) \times L^2(0, 1; V)$ such that

$$a(t) = \dot{\delta q}(t) - \partial_q \xi_{q(t)} v(t) \cdot \delta q(t) - \xi_{q(t)} \delta v(t), \quad (10.20)$$

$$b(t) = \partial_q C_{q(t)} v(t) \cdot \delta q(t) + C_{q(t)} \delta v(t), \quad (10.21)$$

for almost every $t \in [0, 1]$.

For every $\tilde{q} \in M$, $C_{\tilde{q}} : V \rightarrow Y$ is a surjective linear continuous mapping. It follows that $C_{\tilde{q}|(\ker(C_{\tilde{q}}))^\perp} : (\ker(C_{\tilde{q}}))^\perp \rightarrow Y$ is an isomorphism (note that $V = \ker(C_{\tilde{q}}) \oplus (\ker(C_{\tilde{q}}))^\perp$ since V is Hilbert). We set $A_{\tilde{q}} = (C_{\tilde{q}|(\ker(C_{\tilde{q}}))^\perp})^{-1} = K_V C_{\tilde{q}}^* (C_{\tilde{q}} K_V C_{\tilde{q}}^*)^{-1}$. Note that $\tilde{q} \mapsto A_{\tilde{q}}$ is of class \mathcal{C}^1 in a neighbourhood of $q([0, 1])$.

Assume for the moment that $\delta q(\cdot)$ is known. Then we choose $\delta v(\cdot)$ defined by $\delta v(t) = A_{q(t)} (b(t) - \partial_q C_{q(t)} v(t) \cdot \delta q(t))$ for almost every $t \in [0, 1]$, so that (10.21) is satisfied. Plugging this expression into (10.20) yields

$$\dot{\delta q}(t) - \partial_q \xi_{q(t)} v(t) \cdot \delta q(t) - \xi_{q(t)} A_{q(t)} (b(t) - \partial_q C_{q(t)} v(t) \cdot \delta q(t)) = a(t),$$

for almost every $t \in [0, 1]$. This is a well-posed linear differential equation with square-integrable coefficients in the Banach space X , which has a unique solution $\delta q \in H_0^1(0, 1; X)$ such that $\delta q(0) = 0$. This proves the statement.

Proving that the mapping $\partial_q \Gamma_{1(q,v)} : H_0^1(0, 1; M) \rightarrow L^2(0, 1; X)$, defined by $(\partial_q \Gamma_{1(q,v)} \cdot \delta q)(t) = \dot{\delta q}(t) - \partial_q \xi_{q(t)} v(t) \cdot \delta q(t)$ for almost every $t \in [0, 1]$, is an isomorphism follows the same argument, by Cauchy uniqueness. \square

Let $(q, v) \in H_{q_0}^1(0, 1; M) \times L^2(0, 1; V)$ be an optimal solution of the optimal control problem. In other words, (q, v) is a minimizer of the problem of minimizing the functional J over the set of constraints $\Gamma^{-1}(\{0\})$ (which is a \mathcal{C}^∞ manifold as a consequence of Lemma 10.3 and of the implicit function theorem). Since $d\Gamma_{(q,v)}$ is surjective (note that this fact is essential since we are in infinite dimension), it follows from [Kur76, Theorem 4.1] that there exists a nontrivial Lagrange multiplier $(p, \lambda) \in L^2(0, 1; X)^* \times L^2(0, 1; Y)^*$ such that $dJ_{(q,v)} + (d\Gamma_{(q,v)})^*(p, \lambda) = 0$. Moreover since Y is reflexive one has $L^2(0, 1; Y)^* = L^2([0, 1], Y^*)$, and hence we can identify λ with a square-integrable Y^* -valued measurable mapping, so that the Lagrange multipliers relation yields

$$\begin{aligned} 0 &= \langle dJ_{(q,v)} + (d\Gamma_{(q,v)})^*(p, \lambda), (\delta q, \delta v) \rangle \\ &= \langle p, \dot{\delta q} \rangle_{L^2(0,1;X)^*, L^2(0,1;X)} - \langle p, \partial_q \xi_{q(t)} v(t) \cdot \delta q \rangle_{L^2(0,1;X)^*, L^2(0,1;X)} - \langle p, \xi_{q(t)} \delta v \rangle_{L^2(0,1;X)^*, L^2(0,1;X)} \\ &\quad + \int_0^1 \left\langle \partial_q L(q(t), v(t)) + (\partial_q C_{q(t)} v(t))^* \lambda(t), \delta q(t) \right\rangle_{X^*, X} dt \\ &\quad + \int_0^1 \left\langle \partial_v L(q(t), v(t)) + C_{q(t)}^* \lambda(t), \delta v(t) \right\rangle_{V^*, V} dt + \langle dg_{q(1)}, \delta q(1) \rangle_{X^*, X}, \end{aligned} \quad (10.22)$$

for all $(p, \lambda) \in L^2(0, 1; X)^* \times L^2([0, 1], Y^*)$ and all $(\delta q, \delta v) \in H_0^1(0, 1; M) \times L^2(0, 1; V)$. Note that the space $L^2(0, 1; X)^*$ can be different from $L^2(0, 1; X^*)$ (unless X^* satisfies the Radon-Nikodym property, but there is no reason to consider such a Banach space X), and hence *a priori* p cannot be obviously identified with a square-integrable X^* -valued measurable mapping. Anyway, in the next lemma we show that this identification is possible, due to a hidden regularity property in (10.22).

Lemma 10.4. *We can identify p with an element of $L^2([0, 1], X^*)$, so that*

$$\langle p, r \rangle_{L^2(0,1;X^*), L^2(0,1;X)} = \int_0^1 \langle p(t), r(t) \rangle_{X^*, X} dt,$$

for every $r \in L^2(0, 1; X)$.

Proof of Lemma 10.4. For every $r \in L^2(0, 1; X)$, we define $\delta q \in H_0^1(0, 1; X)$ by $\delta q(s) = \int_0^s r(t) dt$ for every $s \in [0, 1]$ (Bochner integral in the Banach space X), so that $r = \dot{\delta q}$. Defining $\alpha \in L^2(0, 1; X)^*$ by

$$\begin{aligned} \langle \alpha, f \rangle_{L^2(0,1;X)^*, L^2(0,1;X)} &= \langle p, \partial_q \xi_q v \cdot f \rangle_{L^2(0,1;X)^*, L^2(0,1;X)} \\ &\quad - \int_0^1 \langle \partial_q L(q(t), v(t)) + (\partial_q C_{q(t)} v(t))^* \lambda(t), f(t) \rangle_{X^*, X} dt, \end{aligned} \quad (10.23)$$

for every $f \in L^2(0, 1; X)$, and taking $\delta v = 0$ in (10.22), we get

$$\langle p, r \rangle_{L^2(0,1;X)^*, L^2(0,1;X)} = \langle \alpha, \delta q \rangle_{L^2(0,1;X)^*, L^2(0,1;X)} - \langle dg_{q(1)}, \delta q(1) \rangle_{X^*, X}. \quad (10.24)$$

Let us express δq in another way with respect to r . By definition, one has $\delta q(s) = \int_0^s r(t) dt$ for every $s \in [0, 1]$, and this can be also written as $\delta q(s) = \int_0^1 \chi_{[t,1]}(s) r(t) dt$, with $\chi_{[t,1]}(s) = 1$ whenever $s \in [t, 1]$ and 0 otherwise. In other words, one has $\delta q = \int_0^1 \chi_{[t,1]} r(t) dt$ (Bochner integral). For every $t \in [0, 1]$, we define the operator $A_t : X \rightarrow L^2(0, 1; X)$ by $A_t x = \chi_{[t,1]} x$. It is clearly linear and continuous. Then, we have $\delta q = \int_0^1 A_t(r(t)) dt$, and therefore, using (10.24),

$$\langle p, r \rangle_{L^2(0,1;X)^*, L^2(0,1;X)} = \left\langle \alpha, \int_0^1 A_t(r(t)) dt \right\rangle_{L^2(0,1;X)^*, L^2(0,1;X)} - \langle dg_{q(1)}, \int_0^1 r(t) dt \rangle_{X^*, X}.$$

Now, interchanging the Bochner integrals and the linear forms, we infer that

$$\langle p, r \rangle_{L^2(0,1;X)^*, L^2(0,1;X)} = \int_0^1 \langle \alpha, A_t(r(t)) \rangle_{L^2(0,1;X)^*, L^2(0,1;X)} dt - \int_0^1 \langle dg_{q(1)}, r(t) \rangle_{X^*, X} dt,$$

and then, using the adjoint $A_t^* : L^2(0, 1; X)^* \rightarrow X^*$, we get

$$\langle p, r \rangle_{L^2(0,1;X)^*, L^2(0,1;X)} = \int_0^1 \langle A_t^* \alpha - dg_{q(1)}, r(t) \rangle_{X^*, X} dt.$$

Since this identity holds true for every $r \in L^2(0, 1; X)$, it follows that p can be identified with an element of $L^2(0, 1; X^*)$, still denoted by p , with $p(t) = A_t^* \alpha - dg_{q(1)}$ for almost every $t \in [0, 1]$. \square

Still using the notation introduced in the proof of Lemma 10.4, now that we know that $p \in L^2(0, 1; X^*)$, we infer from (10.23) that α can as well be identified with an element of $L^2([0, 1], X^*)$, with

$$\alpha(t) = (\partial_q \xi_{q(t)} v(t))^* p(t) - \partial_q L(q(t), v(t)) - (\partial_q C_{q(t)} v(t))^* \lambda(t),$$

for almost every $t \in [0, 1]$. Note that $\alpha(t) = \partial_q H(q(t), p(t), v(t), \lambda(t))$, where the Hamiltonian H is defined by (10.18).

Since $\alpha \in L^2(0, 1; X^*)$, we have, for every $x \in X$,

$$\begin{aligned} \langle A_t^* \alpha, x \rangle_{X^*, X} &= \langle \alpha, \chi_{[t,1]} x \rangle_{L^2(0,1;X^*), L^2(0,1;X)} = \int_0^1 \langle \alpha(s), \chi_{[t,1]}(s) x \rangle_{X^*, X} ds \\ &= \int_t^1 \langle \alpha(s), x \rangle_{X^*, X} ds = \left\langle \int_t^1 \alpha(s) ds, x \right\rangle_{X^*, X}, \end{aligned}$$

and therefore $A_t^* \alpha = \int_t^1 \alpha(s) ds$ for every $t \in [0, 1]$. It follows that $p(t) = A_t^* \alpha - dg_{q(1)} = \int_t^1 \alpha(s) ds - dg_{q(1)}$, and hence, that $p \in H^1(0, 1; X^*)$ and that $p(\cdot)$ satisfies the differential

equation $\dot{p}(t) = -\alpha(t) = -\partial_q H(q(t), p(t), v(t), \lambda(t))$ for almost every $t \in [0, 1]$ and $p(1) + dg_{q(1)} = 0$.

Finally, taking $\delta q = 0$ in (10.22) yields

$$\int_0^1 \langle \partial_v L(q(t), v(t)) + C_{q(t)}^* \lambda(t) - \xi_{q(t)}^* p(t), \delta v(t) \rangle_{V^*, V} dt = 0,$$

for every $\delta v \in L^2(0, 1; X)$, and hence $\partial_v L(q(t), v(t)) + C_{q(t)}^* \lambda(t) - \xi_{q(t)}^* p(t) = 0$ for almost every $t \in [0, 1]$, which exactly means that $\partial_v H(q(t), p(t), v(t), \lambda(t)) = 0$. The theorem is proved. \square

10.2.2 The geodesic equations in a shape space

We use the notations introduced in Section 10.1.2, and consider a shape space M of order $\ell \in \mathbb{N}$, V an RKHS of vector fields of class $\mathcal{C}_0^{\ell+1}$ on \mathbb{R}^d , and we set $L(q, v) = \frac{1}{2} \|v\|_V^2$. We assume that C and g are at least of class \mathcal{C}^1 and that $C_q : V \rightarrow Y$ is surjective for every $q \in M$.

In this context ξ is of class \mathcal{C}^1 . Let us apply Theorem 10.2. We have $\partial_v H(q, p, v, \lambda) = \xi_q^* p - C_q^* \lambda - (v, \cdot)_V$, and the condition $\partial_v H(q, p, v, \lambda) = 0$ is equivalent to $v = K_V(\xi_q^* p - C_q^* \lambda)$. Then we have $\partial_p H = \xi_q v = \xi_q K_V(\xi_q^* p - C_q^* \lambda) = K_q p - \xi_q K_V C_q^* \lambda$, where $K_q = \xi_q K_V \xi_q^*$ is defined by (10.11). Besides, $C_q v = C_q K_V \xi_q^* p - C_q K_V C_q^* \lambda = 0$ if and only if $C_q K_V C_q^* \lambda = C_q K_V \xi_q^* p$. Since C_q is surjective, it follows that $C_q K_V C_q^*$ is invertible, and hence $\lambda = \lambda_{q,p} = (C_q K_V C_q^*)^{-1} C_q K_V \xi_q^* p$. The mapping $(q, p) \mapsto \lambda_{q,p}$ defined as such is of class \mathcal{C}^1 and is linear in p . In particular, $v = v_{q,p} = K_V(\xi_q^* p - C_q^* \lambda_{q,p})$ is a function of class \mathcal{C}^1 of q and p and is linear in p . We have obtained the following result.

Theorem 10.3 (Geodesic equations in shape spaces). *Let $(q(\cdot), v(\cdot)) \in H_{q_0}^1(0, 1; M) \times L^2(0, 1; V)$ be a solution of Problem 1. There exists $p \in H^1(0, 1; X^*)$ such that*

$$v(t) = v_{q(t), p(t)} = K_V \left(\xi_{q(t)}^* p(t) - C_{q(t)}^* \lambda(t) \right),$$

for almost every $t \in [0, 1]$, and $p(\cdot)$ satisfies $p(1) + dg_{q(1)} = 0$ and the geodesic equations

$$\begin{aligned} \dot{q}(t) &= K_{q(t)} p(t) - \xi_{q(t)} K_V C_{q(t)}^* \lambda(t), \\ \dot{p}(t) &= -\partial_q \left\langle p(t), \xi_{q(t)} v(t) \right\rangle_{X^*, X} + \partial_q \left\langle \lambda(t), C_{q(t)} v(t) \right\rangle_{Y^*, Y}, \end{aligned} \quad (10.25)$$

for almost every $t \in [0, 1]$, with

$$\lambda(t) = \lambda_{q(t), p(t)} = (C_{q(t)} K_V C_{q(t)}^*)^{-1} C_{q(t)} K_V \xi_{q(t)}^* p(t).$$

Moreover the mapping $t \mapsto \frac{1}{2} \|v(t)\|^2$ is constant, and one has

$$J(v) = \frac{1}{2} \|v(0)\|_V^2 + g(q(1)).$$

Remarque 10.16. Defining the so-called reduced Hamiltonian $h : M \times X^* \rightarrow \mathbb{R}$ by

$$h(q, p) = H(q, p, v_{q,p}, \lambda_{q,p}),$$

we have a priori $\partial_q h = \partial_q H + \partial_v H(q, p, v_{q,p}, \lambda_{q,p}) \circ \partial_q (v_{q,p}) + \partial_\lambda H(q, p, v_{q,p}, \lambda_{q,p}) \circ \partial_q (\lambda_{q,p})$. But since $\lambda_{q,p}$ and $v_{q,p}$ are such that $\partial_v H(q, p, v_{q,p}, \lambda_{q,p}) = 0$ and $C_q v_{q,p} = \partial_\lambda H(q, p, v_{q,p}, \lambda_{q,p}) = 0$, it follows that $\partial_q h(q, p) = \partial_q H(q, p, v_{q,p}, \lambda_{q,p})$. Similarly, we have $\partial_p h(q, p) = \partial_p H(q, p, v_{q,p}, \lambda_{q,p})$. Therefore, in Theorem 10.3, the geodesics are the solutions of the Hamiltonian system

$$\dot{q}(t) = \partial_p h(q(t), p(t)), \quad \dot{p}(t) = -\partial_q h(q(t), p(t)).$$

Corollary 10.1. *Assume that the mappings C and ξ are of class \mathcal{C}^2 . Then h is of class \mathcal{C}^2 as well, and for every $(q_0, p_0) \in M \times X^*$, there exists $\varepsilon > 0$ and there exists a unique solution $(q, p) : [0, \varepsilon] \rightarrow M \times X^*$ of the geodesic equations (10.25) such that $(q(0), p(0)) = (q_0, p_0)$.*

Note that for most of shape spaces (at least, for all shape spaces given as examples in this paper), the mapping ξ is of class \mathcal{C}^2 whenever V is an RKHS of vector fields of class $\mathcal{C}^{\ell+2}$.

Example 10.5. A geodesic $(q(t), p(t)) = (x_1(t), \dots, x_n(t), p_1(t), \dots, p_n(t))$ on the landmark space $\text{Lmk}_d(n)$ must satisfy the equations

$$\dot{x}_i(t) = \sum_{j=1}^n K(x_i(t), x_j(t)) p_j(t), \quad \dot{p}_j(t) = -\frac{1}{2} \sum_{i=1}^n \partial_{x_i} \left(p_i(t)^T K(x_i(t), x_j(t)) p_j(t) \right),$$

where K is the kernel of V .

Remarque 10.17. Assume that the constraints are kinetic, i.e., are of the form $C_q \xi_q v = 0$ with $C_q : X \rightarrow Y$. Then $\dot{q}(t) = K_{q(t)}(p(t) - C_{q(t)}^* \lambda(t))$ and $\lambda_{q,p} = (C_q K_q C_q^*)^{-1} C_q K_q p$. Note that even if K_q is not invertible, the assumption that $C_q \xi_q$ is surjective means that $C_q K_q C_q^*$ is invertible. In particular, minimizing controls take the form $K_V \xi_q^* u$, with $u \in L^2(0, 1; X^*)$. This proves the contents of Proposition 9 (see Section 10.1.2).

Remarque 10.18. Let M (resp., M'), open subset of a Banach space X (resp., X'), be a shape space of order ℓ (resp. of order $\ell' \leq \ell$). Assume that there is a dense and continuous inclusion $X \hookrightarrow X'$, such that $M \hookrightarrow M'$ is equivariant (in particular we have $X'^* \hookrightarrow X^*$). For every $q_0 \in M \subset M'$, there are more geodesics emanating from q_0 on M than on M' (indeed it suffices to consider initial momenta $p_0 \in X^* \setminus X'^*$). These curves are not solutions of the geodesic equations on M' , and are an example of so-called *abnormal extremals* [AS04, Tré08]. Note that they are however not solutions of Problem 1 specified on X' , since Theorem 10.2 implies that such solutions are projections of geodesics having an initial momentum in X'^* .

An example where this situation is encountered is the following. Let S be a compact Riemannian manifold. Consider $X = \mathcal{C}^{\ell+1}(S, \mathbb{R}^d)$ and $X' = \mathcal{C}^{\ell}(S, \mathbb{R}^d)$, with actions defined in Definition 10.2. If $p(0)$ is a $\ell + 1$ -th order distribution, then the geodesic equations on X with initial momentum $p(0)$ yield an abnormal geodesic in X' .

Remarque 10.19. Following Remark 10.18, if the data attachment function $g : X \rightarrow \mathbb{R}^+$ can be smoothly extended to a larger space X' , then the initial momentum of any solution of Problem 1 is actually in $X'^* \subset X^*$.

For example, we set $g(q) = \int_S |q(s) - q_{\text{target}}(s)|^2 ds$, for some $q_{\text{target}} \in X$, with $X = \mathcal{C}^{\ell}(S, \mathbb{R}^d)$ and S a compact Riemannian manifold. Let $X' = L^2(S, \mathbb{R}^d)$. Then the momentum $p : [0, 1] \rightarrow X^*$ associated with a solution of Problem 1 (whose existence follows from Theorem 10.3) takes its values in $X'^* = L^2(S, \mathbb{R}^d)$.

The case of pure state constraints. Let us consider pure state constraints, i.e., constraints of the form $C(q) = 0$ with $C : M \rightarrow Y$. Recall that, if C is of class \mathcal{C}^1 , then they can be transformed into the mixed constraints $dC_q \cdot \dot{q} = dC_q \cdot \xi_q v = 0$. In this particular case, the geodesic equations take a slightly different form.

Proposition 12. *Assume that $C : M \rightarrow Y$ is of class \mathcal{C}^3 , and that $dC_q \cdot \xi_q : V \rightarrow Y$ is surjective for every $q \in M$. Consider Problem 1 with the pure state constraints $C(q) = 0$. If $v(\cdot) \in L^2(0, 1; V)$ is a solution of Problem 1, associated with the curve $q(\cdot) : [0, 1] \rightarrow M$, then there exists $\tilde{p} : [0, 1] \rightarrow X^*$ such that*

$$\dot{q}(t) = K_{q(t)} \tilde{p}(t), \quad \dot{\tilde{p}}(t) = -\frac{1}{2} \partial_q \langle \tilde{p}(t), K_{q(t)} \tilde{p}(t) \rangle_{X^*, X} + \partial_q \langle \tilde{\lambda}_{q(t), \tilde{p}(t)}, C(q(t)) \rangle_{Y^*, Y},$$

for almost every $t \in [0, 1]$, with

$$\tilde{\lambda}_{q,\tilde{p}} = (dC_q K_q dC_q^*)^{-1} \left(\frac{1}{2} dC_q \cdot K_q \partial_q \langle \tilde{p}, K_q \tilde{p} \rangle_{X^*,X} - d^2 C_q \cdot (K_q \tilde{p}, K_q \tilde{p}) - dC_q \cdot (\partial_q (K_q \tilde{p}) \cdot K_q \tilde{p}) \right).$$

Moreover one has $dC_{q_0} \cdot K_{q_0} \tilde{p}_0 = 0$ and $dg_{q(1)} + \tilde{p}(1) = dC_{q(1)}^* \nu$ for some $\nu \in Y^*$.

Note that all functions involved are of class \mathcal{C}^1 . Therefore, for a given initial condition $(q_0, \tilde{p}_0) \in T_{q_0}^* M$ with $dC_{q_0} \cdot K_{q_0} p_0 = 0$, there exists a unique geodesic emanating from q_0 with initial momentum \tilde{p}_0 .

Proof. The proof just consists in considering $\tilde{\lambda}(t) = \dot{\lambda}(t)$ and $\tilde{p} = p - dC_q^* \cdot \lambda$, where p and λ are given by Theorem 10.3, and then in differentiating $C(q(t)) = 0$ twice with respect to time. \square

Remarque 10.20. The mappings \tilde{p} and $\tilde{\lambda}$ can also be obtained independently of Theorem 10.3 as Lagrange multipliers for the mapping of constraints $\Gamma(q, v) = (\dot{q} - \xi_q v, C(q))$, using the same method as in the proof of Theorem 10.2, under the assumptions that $C : X \rightarrow Y$ is of class \mathcal{C}^2 and $dC_q \cdot \xi_q$ is surjective for every $q \in M$.

10.3 Algorithmic procedures

In this section we derive some algorithms in order to compute the solutions of the optimal control problem considered throughout. We first consider problems without constraint in Section 10.3.1, and then with constraints in Section 10.3.2.

10.3.1 Problems without constraints

Shape deformation analysis problems without constraint have already been studied in [BMTY05, DGM98, GTY, GTY06, JM00, MTY02, MTY06, Tro95] with slightly different methods, in different and specific contexts. With our previously developed general framework, we are now going to recover methods that are well known in numerical shape analysis, but with a more general point of view allowing us to generalize the existing approaches.

Gradient Descent. We adopt the notations, the framework and the assumptions used in Section 10.2.1, but without constraint. For $q_0 \in M$ fixed, we consider the optimal control problem of minimizing the functional J defined by (10.17) over all $(q(\cdot), v(\cdot)) \in H_{q_0}^1(0, 1; M) \times L^2(0, 1; V)$ such that $\dot{q}(t) = \xi_{q(t)} v(t)$ for almost every $t \in [0, 1]$. The Hamiltonian of the problem then does not involve the variable λ , and is the function $H : M \times X^* \times V \rightarrow \mathbb{R}$ defined by $H(q, p, v) = \langle p, \xi_q v \rangle_{X^*,X} - L(q, v)$.

We assume throughout that $g \in \mathcal{C}^1(X, \mathbb{R})$, that $\xi : M \times V \rightarrow X$ is of class \mathcal{C}^2 , is a linear mapping in v , and that $L \in \mathcal{C}^2(X \times V, \mathbb{R})$ satisfies the estimate (10.16).

As in the proof of Theorem 10.2, we define the mapping $\Gamma : H_{q_0}^1(0, 1; M) \times L^2(0, 1; V) \rightarrow L^2(0, 1; X)$ by $\Gamma(q, v)(t) = \dot{q}(t) - \xi_{q(t)} v(t)$ for almost every $t \in [0, 1]$. The objective is to minimize the functional J over the set $E = \Gamma^{-1}(\{0\})$.

According to Lemma 10.3, the mapping $\partial_q \Gamma_{(q,v)}$ is an isomorphism for all $(q, v) \in E$. Therefore, the implicit function theorem implies that E is the graph of the mapping $v \mapsto q_v$ which to a control $v \in L^2(0, 1; V)$ associates the curve $q_v \in H_{q_0}^1(0, 1; X)$ solution of $\dot{q}_v(t) = \xi_{q_v(t)} v(t)$ for almost every $t \in [0, 1]$ and $q_v(0) = q_0$. Moreover this mapping is, like Γ , of class \mathcal{C}^2 .

Then, as it was already explained in Remark 10.2 in the case where $L(q, v) = \frac{1}{2} \|v\|_V^2$, minimizing J over E is then equivalent to minimizing the functional $J_1(v) = J(q_v, v)$ over $L^2(0, 1; V)$.

Thanks to these preliminary remarks, the computation of the gradient of $J|_E$ then provides in turn a gradient descent algorithm.

Proposition 13. *The differential of J_1 is given by*

$$dJ_{1v} \cdot \delta v = - \int_0^1 \partial_v H(q_v(t), p(t), v(t)) \cdot \delta v(t) dt,$$

for every $\delta v \in L^2(0, 1; V)$, where $p \in H^1([0, 1], X^*)$ is the solution of $\dot{p}(t) = -\partial_q H(q_v(t), p(t), v(t))$ for almost every $t \in [0, 1]$ and $p(1) + dg_{q_v(1)} = 0$. In particular we have

$$\nabla J_1(v)(t) = -K_V \partial_v H(q_v(t), p(t), v(t)),$$

for almost every $t \in [0, 1]$.

Remarque 10.21. This result still holds true for Lagrangians L that do not satisfy (10.16), replacing $L^2(0, 1; V)$ with $L^\infty(0, 1; V)$. The gradient is computed with respect to the pre-Hilbert scalar product inherited from L^2 .

Proof. Let $v \in L^2(0, 1; V)$ be arbitrary. For every $\delta v \in L^2(0, 1; V)$, we have $dJ_{1v} \cdot \delta v = dJ_{(q_v, v)} \cdot (\delta q, \delta v)$, with $\delta q = dq_v \cdot \delta v$. Note that $(\delta q, \delta v) \in T_{(q_v, v)}E$ (the tangent space of the manifold E at (q_v, v)), since E is the graph of the mapping $v \rightarrow q_v$. Since $E = \Gamma^{-1}(\{0\})$, we have $\langle d\Gamma_{(q_v, v)}^* p, (\delta q, \delta v) \rangle = 0$ for every $p \in L^2(0, 1; X^*)$. Let us find some particular p such that $\langle dJ_{1(q_v, v)} + d\Gamma_{(q_v, v)}^* p, (\delta q, \delta v) \rangle$ only depends on δv .

Let $p \in H^1([0, 1], X^*)$ be the solution of $\dot{p}(t) = -\partial_q H(q_v(t), p(t), v(t))$ for almost every $t \in [0, 1]$ and $p(1) + dg_{q_v(1)} = 0$. Using the computations done in the proof of Theorem 10.2, we get

$$\begin{aligned} \langle dJ_{1(q_v, v)} + d\Gamma_{(q_v, v)}^* p, (\delta q, \delta v) \rangle &= \int_0^1 \left(\langle p(t), \dot{\delta q}(t) \rangle_{X^*, X} - \langle \partial_q H(q_v(t), p(t), v(t)), \delta q(t) \rangle_{X^*, X} \right. \\ &\quad \left. - \langle \partial_v H(q_v(t), p(t), v(t)), \delta v(t) \rangle_{V^*, V} \right) dt + dg_{q_v(1)} \cdot \delta q_v(1). \end{aligned}$$

Integrating by parts and using the relations $\delta q(0) = 0$ and $p(1) + dg_{q_v(1)} = 0$, we obtain

$$\langle dJ_{1(q_v, v)} + d\Gamma_{(q_v, v)}^* p, (\delta q, \delta v) \rangle = - \int_0^1 \partial_v H(q_v(t), p(t), v(t)) \cdot \delta v(t) dt$$

Since $\langle d\Gamma_{(q_v, v)}^* p, (\delta q, \delta v) \rangle = 0$, the proposition follows. \square

In the case of shape spaces, which form our main interest here, we have $L(q, v) = \frac{1}{2} \|v\|_V^2$, and then $\partial_v H(q, p, v) = \xi_q^* p - (v, \cdot)_V$ and $K_V \partial_v H(q, p, v) = K_V \xi_q^* p - v$. It follows that

$$\nabla J_1(v) = v - K_V \xi_q^* p.$$

In particular, if $v = K_V \xi_q^* u$ for some $u \in L^2([0, 1], X^*)$, then

$$v - h \nabla J_1(v) = K_v \xi_q^* (u - h(u - p)),$$

for every $h \in \mathbb{R}$. Therefore, applying a gradient descent algorithm does not change this form. It is then important to notice that this provides as well a gradient descent algorithm for solving Problem 2 (the kernel formulation of Problem 1) without constraint, and this in spite of the fact that $L^2([0, 1], X^*)$ is not necessarily a Hilbert space. In this case, if $p \in H^1([0, 1], X^*)$ is the solution of $\dot{p}(t) = -\partial_q H(q_v(t), p(t), v(t))$ for almost every $t \in [0, 1]$ and $p(1) + dg_{q_v(1)} = 0$, then $u - p$ is the gradient of the functional J_2 defined by (10.12) with respect to the symmetric nonnegative bilinear form $B_q(u_1, u_2) = \int_0^1 \langle u_1(t), K_{q(t)} u_2(t) \rangle_{X^*, X} dt$.

This gives a first algorithm to compute unconstrained minimizers in a shape space. We next provide a second method using the space of geodesics.

Gradient descent on geodesics: minimization through shooting. Since the tools are quite technical, in order to simplify the exposition we assume that the shape space M is finite dimensional, i.e., that $X = \mathbb{R}^n$ for some $n \in \mathbb{N}$. The dual bracket $\langle p, w \rangle_{X^*, X}$ is then identified with the canonical Euclidean product $p^T w$, and $K_q = \xi_q K_V \xi_q^* : X^* \rightarrow X$ is identified with a $n \times n$ positive semi-definite symmetric matrix. Theorem 10.3 and Corollary 10.1 (see Section 10.2.2) imply that the minimizers of the functional J_1 defined by (10.10) coincide with those of the functional

$$\hat{J}_1(p_0) = \frac{1}{2} p_0^T K_{q_0} p_0 + g(q(1)), \tag{10.26}$$

where $v = K_V \xi_q^* p$ and $(q(\cdot), p(\cdot))$ is the geodesic solution of the Hamiltonian system $\dot{q}(t) = \partial_p h(q(t), p(t))$, $\dot{p}(t) = -\partial_q h(q(t), p(t))$, for almost every $t \in [0, 1]$, with $(q(0), p(0)) = (q_0, p_0)$. Here, h is the reduced Hamiltonian (see Remark 10.16) and is given by $h(q, p) = \frac{1}{2} \langle p, \xi_q K_V \xi_q^* p \rangle_{X^*, X} = \frac{1}{2} \langle p, K_q p \rangle_{X^*, X}$. Therefore, computing a gradient of \hat{J}_1 for some appropriate bilinear symmetric nonnegative product provides in turn another algorithm for minimizing the functional J_1 . For example, if the inner product that we consider is the canonical one, then $\nabla \hat{J}_1(p_0) = K_{q_0} p_0 + \nabla(g \circ q(1))(p_0)$. The term $\nabla(g \circ q(1))(p_0)$ is computed thanks to the following well-known result.

Lemma 10.5. *Let $n \in \mathbb{N}$, let U be an open subset of \mathbb{R}^n , let $f : U \rightarrow \mathbb{R}^n$ be a complete smooth vector field on U , let G be the function of class \mathcal{C}^1 defined on U by $G(q_0) = g(q(1))$, where g is a function on U of class \mathcal{C}^1 and $q : [0, 1] \rightarrow \mathbb{R}^n$ is the solution of $\dot{q}(t) = f(q(t))$ for almost every $t \in [0, 1]$ and $q(0) = q_0$. Then $\nabla G(q_0) = Z(1)$ where $Z : [0, 1] \rightarrow \mathbb{R}^n$ is the solution of $\dot{Z}(t) = df_{q(1-t)}^T Z(t)$ for almost every $t \in [0, 1]$ and $Z(0) = \nabla g(q(1))$.*

In our case, we have $U = M \times \mathbb{R}^n$ and $f(q, p) = (\nabla_p h, -\nabla_q h) = (K_q p, -\frac{1}{2} \nabla_q (p^T K_q p))$. Note that we used the Euclidean gradient instead of the derivatives. This is still true thanks to the identification made between linear forms and vectors at the beginning of the section. We get $\nabla \hat{J}_1(p_0) = K_{q_0} p_0 + \alpha(1)$, where $Z(\cdot) = (z(\cdot), \alpha(\cdot))$ is the solution of $\dot{Z}(t) = df_{q(1-t), p(1-t)}^T Z(t)$ for almost every $t \in [0, 1]$ and $Z(0) = (z(0), \alpha(0)) = (\nabla g(q(1)), 0)$.

In numerical implementations, terms of the form $df^T w$, with f a vector field and w a vector, require a long computational time since every partial derivative of f has to be computed. In our context however, using the fact that the vector field $f(q, p)$ is Hamiltonian, the computations can be simplified in a substantial way. Indeed, using the commutation of partial derivatives, we get

$$df_{(q,p)}^T Z = \begin{pmatrix} \nabla_q (f(q, p), Z) \\ \nabla_p (f(q, p), Z) \end{pmatrix} = \begin{pmatrix} \nabla_q (\nabla_p h^T z - \nabla_q h^T \alpha) \\ \nabla_p (\nabla_p h^T z - \nabla_q h^T \alpha) \end{pmatrix} = \begin{pmatrix} \partial_p (\nabla_q h) \cdot z - \partial_q (\nabla_q h) \cdot \alpha \\ \partial_p (\nabla_p h) \cdot z - \partial_q (\nabla_p h) \cdot \alpha \end{pmatrix}.$$

Replacing h with its expression, we get

$$df_{(q,p)}^T Z = \begin{pmatrix} \nabla_q (p^T K_q z) - \frac{1}{2} \partial_q (\nabla_q (p^T K_q p)) \cdot \alpha \\ K_q z - \partial_q (K_q p) \cdot \alpha \end{pmatrix}.$$

Therefore, instead of computing $d(\nabla_q (p^T K_q p))^T \alpha$, which requires the computation of all partial derivatives of $\nabla_q (p^T K_q p)$, it is required to compute only one of them, namely the one with respect to α . Let us sum up the result in the following proposition.

Proposition 14. *We have $\nabla \hat{J}_1(p_0) = K_{q_0} p_0 + \alpha(1)$, where $(z(\cdot), \alpha(\cdot))$ is the solution of*

$$\begin{aligned} \dot{z}(t) &= \nabla_q (p(1-t)^T K_{q(1-t)} z(t)) - \frac{1}{2} \partial_q (\nabla_q (p(1-t)^T K_{q(1-t)} p(1-t))) \cdot \alpha(t), \\ \dot{\alpha}(t) &= K_{q(1-t)} z(t) - \partial_q (K_{q(1-t)} p(1-t)) \cdot \alpha(t) \end{aligned}$$

with $(z(0), \alpha(0)) = (\nabla g(q(1)), 0)$, and $(q(t), p(t))$ satisfies the geodesic equations $\dot{q}(t) = K_{q(t)}p(t)$ and $\dot{p}(t) = -\frac{1}{2}\nabla_q(p(t)^T K_{q(t)}p(t))$ for almost every $t \in [0, 1]$, with $q(0) = q_0$ and $p(0) = p_0$.

A gradient descent algorithm can then be used in order to minimize \hat{J}_1 and thus J_1 .

10.3.2 Problems with constraints

In this section, we derive several different methods devoted to solve numerically constrained optimal control problems on shape spaces. To avoid using overly technical notation in their whole generality, we restrict ourselves to the finite-dimensional case. The methods can however be easily adapted to infinite-dimensional shape spaces. We use the notation, the framework and the assumptions of Section 10.1.2.

Let $X = \mathbb{R}^n$ and let M be an open subset of X . For every $q \in M$, we identify K_q with a $n \times n$ symmetric positive semi-definite real matrix. Throughout the section, we focus on kinetic constraints and we assume that we are in the conditions of Proposition 9, so that these constraints take the form $C_q K_q u = 0$. Note that, according to Proposition 8, in this case Problems 1 and 2 are equivalent. Hence in this section we focus on Problem 2, and thanks to the identifications above the functional J_2 defined by (10.12) can be written as

$$J_2(u) = \frac{1}{2} \int_0^1 u(t)^T K_{q(t)} u(t) dt + g(q(1)).$$

Note (and recall) that pure state constraints, of the form $C(q) = 0$, are treated as well since, as already mentioned, they are equivalent to the kinetic constraints $dC_q \cdot K_q u = 0$.

The augmented Lagrangian method. This method consists of minimizing iteratively unconstrained functionals in which the constraints have been penalized. Although pure state constraints are equivalent to kinetic constraints, in this approach they can also be treated directly. The method goes as follows. In the optimal control problem under consideration, we denote by $\lambda : [0, 1] \rightarrow \mathbb{R}^k$ the Lagrange multiplier associated with the kinetic constraints $C_q K_q u = 0$ (its existence is ensured by Theorem 10.2). We define the augmented cost function

$$J_A(u, \lambda_1, \lambda_2, \mu) = \int_0^1 L_A(q(t), u(t), \lambda(t), \mu) dt + g(q(1)),$$

where L_A , called augmented Lagrangian, is defined by

$$L_A(q, u, \lambda, \mu) = L(q, u) - \lambda^T C_q K_q u + \frac{1}{2\mu} |C_q K_q u|^2,$$

with, here, $L(q, u) = \frac{1}{2} u^T K_q u$. Let $q_0 \in M$ fixed. Choose an initial control u_0 (for example, $u_0 = 0$), an initial function λ_0 (for example, $\lambda_0 = 0$), and an initial constant $\mu_0 > 0$. At step ℓ , assume that we have obtained a control u_ℓ generating the curve q_ℓ , a function $\lambda_\ell : [0, 1] \rightarrow \mathbb{R}^k$, and a constant $\mu_\ell > 0$. The iteration $\ell \rightarrow \ell + 1$ is defined as follows. First, minimizing the unconstrained functional $u \mapsto J_A(u, \lambda_\ell, \mu_\ell)$ over $L^2(0, 1; \mathbb{R}^n)$ yields a new control $u_{\ell+1}$, generating the curve $q_{\ell+1}$ (see further in this section for an appropriate minimization method). Second, λ is updated according to

$$\lambda_{\ell+1} = \lambda_\ell - \frac{1}{\mu_\ell} C_{q_{\ell+1}} K_{q_{\ell+1}} u_{\ell+1}.$$

Finally, we choose $\mu_{\ell+1} \in (0, \mu_\ell]$ (many variants are possible in order to update this penalization parameter, as is well-known in numerical optimization).

Under some appropriate assumptions, as long as μ_ℓ is smaller than some constant $\beta > 0$, u_ℓ converges to a control u^* which is a constrained extremum of J_2 . Note that it is not required to assume that μ_ℓ converge to 0. More precisely we infer from [IK08, Chapter 3] the following convergence result.

Proposition 15 (Convergence of the augmented Lagrangian method). *Assume that all involved mappings are least of class C^2 and that $C_q K_q$ is surjective for every $q \in M$. Let u^* be an optimal solution of Problem 2 and let q^* be its associated curve. Let λ^* be the Lagrange multiplier (given by Theorem 10.2) associated with the constraints. We assume that there exist $c > 0$ and $\mu > 0$ such that*

$$(\partial_u^2 J_A)_{(u^*, \lambda^*, \mu)} \cdot (\delta u, \delta u) \geq c \|\delta u\|_{L^2(0,1;\mathbb{R}^n)}^2, \quad (10.27)$$

for every $\delta u \in L^2(0, 1; \mathbb{R}^n)$. Then there exists a neighborhood of u^* in $L^2(0, 1; \mathbb{R}^n)$ such that, for every initial control u_0 in this neighborhood, the sequence $(u_\ell)_{\ell \in \mathbb{N}}$ built according to the above algorithm converges to u^* , and the sequence $(\lambda_\ell)_{\ell \in \mathbb{N}}$ converges to λ^* , as ℓ tends to $+\infty$.

Remark 10.22. Assumption (10.27) may be hard to check for shape spaces. As is well-known in optimal control theory, this coercivity assumption of the bilinear form $(\partial_u^2 J_A)_{(u^*, \lambda^*, \mu)}$ is actually equivalent to the nonexistence of *conjugate points* of the optimal curve q^* on $[0, 1]$ (see [BCT07, BFT06] for this theory and algorithms of computation). In practice, computing conjugate points is a priori easy since it just consists of testing the vanishing of some determinants; however in our context the dimension n is expected to be large and then the computation may become difficult numerically.

Remark 10.23. Pure state constraints $C(q) = 0$ can either be treated in the above context by replacing $C_q K_q$ with $dC_q \cdot K_q$, or can as well be treated directly by replacing $C_q K_q$ with $C(q)$ in the algorithm above.

Any of the methods described in Section 10.3.1 can be used in order to minimize the functional J_A with respect to u . For completeness let us compute the gradient in u of J_A at the point (u, λ, μ) .

Lemma 10.6. *There holds*

$$\nabla_u J_A(u, \lambda, \mu) = u + C_q^T \lambda + \frac{1}{\mu} C_q^T C_q K_q u - p,$$

where $p(\cdot)$ is the solution of

$$\begin{aligned} \dot{p}(t) &= -\partial_q(p(t)^T K_{q(t)} u(t)) + \partial_q L_A(q(t), u(t), \lambda(t), \mu) \\ &= \partial_q \left(\left(\frac{u(t)}{2} - p(t) \right)^T K_{q(t)} u(t) \right) + \lambda(t)^T \partial_q C_{q(t)} u(t) + \frac{1}{2\mu} \partial_q \left(u(t)^T K_{q(t)} C_{q(t)}^T C_{q(t)} K_{q(t)} u(t) \right) \end{aligned}$$

for almost every $t \in [0, 1]$ and $p(1) + dg_{q(1)} = 0$.

Remark 10.24. For pure state constraints $C(q) = 0$, there simply holds $\nabla_u J_A = u - p$ and the differential equation in p is

$$\dot{p}(t) = \partial_q \left(\left(\frac{u(t)}{2} - p(t) \right)^T K_{q(t)} u(t) \right) + \lambda(t)^T dC_{q(t)} + \frac{1}{\mu} C(q(t))^T dC_{q(t)}.$$

Proof of Lemma 10.6. We use Proposition 13 with $L(q, u, t) = L_A(q, u, \lambda(t), \mu)$, with λ and μ fixed (it is indeed easy to check that this proposition still holds true when the Lagrangian also depends smoothly on t). The differential of J_A with respect to u is then given by

$$d(J_A)_u \delta u = \int_0^1 \left(\partial_u L_A(q(t), u(t), \lambda(t), \mu) - p(t)^T K_{q(t)} \right) \delta u(t) dt, \quad (10.28)$$

where $p(\cdot)$ is the solution of $\dot{p}(t) = -\partial_q(p(t)^T K_{q(t)} u(t)) + \partial_q L_A(q(t), u(t), \lambda(t), \mu)$ for almost every $t \in [0, 1]$ and $p(1) + dg_{q(1)} = 0$. To get the result, it then suffices to identify the differential $d(J_A)_u$ with the gradient $\nabla J_A(u)$ with respect to the inner product on $L^2(0, 1; R^n)$ given by $(u_1, u_2)_{L^2(0, 1; R^n)} = \int_0^1 u_1(t)^T K_{q(t)} u_2(t) dt$. \square

The advantage of the augmented Lagrangian method is that, at every step, each gradient is “easy” to compute (at least as easy as in the unconstrained case). The problem is that, as in any penalization method, a lot of iterations are in general required in order to get a good approximation of the optimal solution, satisfying approximately the constraints with enough accuracy.

The next method we propose tackles the constraints without penalization.

Constrained minimization through shooting. We adapt the usual shooting method used in optimal control (see, e.g., [Tré08]) to our context. For $(q, p) \in M \times X^* = M \times \mathbb{R}^n$, we define $\lambda_{q,p}$ as in Theorem 10.3 by

$$\lambda_{q,p} = (C_q K_q C_q)^{-1} C_q K_q p.$$

We also denote $\pi_q p = p - C_q^T \lambda_{q,p}$. In particular, $v_{q,p} = K_q \pi_q p$.

Remarque 10.25. A quick computation shows that $\pi_q p$ is the orthogonal projection of p onto $\ker(C_q)$ for the inner product induced by K_q .

According to Theorem 10.3, Corollary 10.1 and Remark 10.17 (see Section 10.2.2), the minimizers of J_2 have to be sought among the geodesics $(q(\cdot), p(\cdot))$ solutions of (10.25), and moreover $p(0)$ is a minimizer of the functional

$$\hat{J}_2(p_0) = \frac{1}{2} \|v_{q_0, p_0}\|_V^2 + g(q(1)) = \frac{1}{2} p_0^T \pi_{q_0} K_{q_0} \pi_{q_0} p_0 + g(q(1)).$$

The geodesic equations (10.25) now take the form

$$\begin{aligned} \dot{q}(t) &= \partial_p h(q(t), p(t)) = K_{q(t)} \pi_{q(t)} p(t), \\ \dot{p}(t) &= -\partial_q h(q(t), p(t)) = -\frac{1}{2} p(t)^T \pi_{q(t)}^T (\nabla_q K_{q(t)}) \pi_{q(t)} p(t) + \lambda_{q(t), p(t)}^T (\nabla_q C_{q(t)}) K_{q(t)} \pi_{q(t)} p(t), \end{aligned}$$

where h is the reduced Hamiltonian (see Remark 10.16). It follows from Proposition 14 that $\nabla \hat{J}_2(p_0) = \pi_{q_0}^T K_{q_0} \pi_{q_0} p_0 + \alpha(1)$, where $Z(\cdot) = (z(\cdot), \alpha(\cdot))$ is the solution of

$$\begin{aligned} \dot{z}(t) &= \partial_p (\nabla_q h(q(1-t), p(1-t))) \cdot z(t) - \partial_q (\nabla_q h(q(1-t), p(1-t))) \cdot \alpha(t) \\ \dot{\alpha}(t) &= \partial_p (\nabla_p h(q(1-t), p(1-t))) \cdot z(t) - \partial_q (\nabla_p h(q(1-t), p(1-t))) \cdot \alpha(t) \end{aligned}$$

with $(z(0), \alpha(0)) = (\nabla g(q(1)), 0)$. Replacing $\nabla_q h$ and $\nabla_p h$ by their expression, we get

$$\begin{aligned} \dot{z}(t) &= p(1-t)^T \pi_{q(1-t)}^T (\nabla_q K_{q(1-t)}) \pi_{q(1-t)} z(t) - \lambda_{q(1-t), p(1-t)}^T (\nabla_q C_{q(1-t)}) K_{q(1-t)} \pi_{q(1-t)} z(t) \\ &\quad - \partial_q \left(\frac{1}{2} p(1-t)^T \pi_{q(1-t)}^T (\nabla_q K_{q(1-t)}) \pi_{q(1-t)} p(1-t) \right. \\ &\quad \left. - \lambda_{q(1-t), p(1-t)}^T (\nabla_q C_{q(1-t)}) K_{q(1-t)} \pi_{q(1-t)} p(1-t) \right) \cdot \alpha(t) \\ \dot{\alpha}(t) &= K_{q(1-t)} \pi_{q(1-t)} z(t) - \partial_q (K_{q(1-t)} \pi_{q(1-t)} p(1-t)) \cdot \alpha(t) \end{aligned}$$

In practice, the derivatives appearing in these equations can be efficiently approximated using finite differences.

This algorithm of constrained minimization through shooting has several advantages compared with the previous augmented Lagrangian method. The first is that, thanks to the geodesic reduction, the functional \hat{J}_2 is defined on a finite-dimensional space (at least whenever the shape space itself is finite dimensional) and hence $\nabla \hat{J}_2(p_0)$ is computed on a finite-dimensional space, whereas in the augmented Lagrangian method $\nabla_u J_A$ was computed on the infinite-dimensional space $L^2(0, 1; \mathbb{R}^n)$.

A second advantage is that, since we are dealing with constrained geodesics, all resulting curves satisfy the constraints with a good numerical accuracy, whereas in the augmented Lagrangian method a large number of iterations was necessary for the constraints to be satisfied with an acceptable numerical accuracy.

This substantial gain is however counterbalanced by the computation of $\lambda_{q(t), p(t)}$, which requires solving of a linear equation at every time $t \in [0, 1]$ along the curve (indeed, recall that $\lambda_{q,p} = (C_q K_q C_q^*)^{-1} C_q K_q p$). The difficulty here is not just that this step is time-consuming, but rather the fact that the linear system may be ill-conditioned, which indicates that this step may require a more careful treatment. One possible way to overcome this difficulty is to solve this system with methods inspired from quasi-Newton algorithms. This requires however a particular study that is beyond the scope of the present article (see [Arg13] for results and algorithms).

10.4 Numerical examples

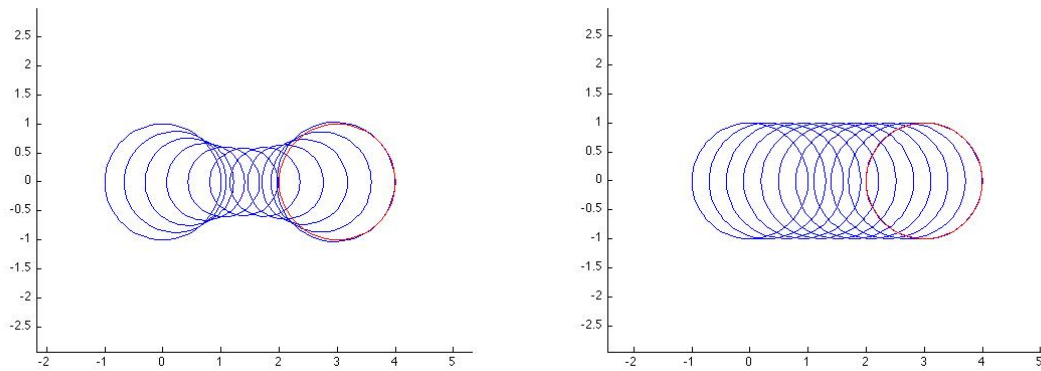
10.4.1 Matching with constant total volume

In this first example we consider a very simple constraint, namely, a constant total volume. Consider $S = S^{d-1}$, the unit sphere in \mathbb{R}^d , and let $M = \text{Emb}^1(S, \mathbb{R}^d)$ be the shape space, acted upon with order 1 by $\text{Diff}(\mathbb{R}^d)$. Consider as in [TY11] the RKHS V of smooth vector fields given by the Gaussian kernel K with positive scale σ defined by $K(x, y) = e^{-\frac{|x-y|^2}{\sigma^2}} I_d$. An embedding $q \in M$ of the sphere is the boundary of an open subset $U(q)$ with total volume given by $\text{Vol}(U(q)) = \int_S q^* \omega$, where ω is a $(d-1)$ -form such that $d\omega = dx_1 \wedge \cdots \wedge dx_d$. Let q_0 be an initial point and let q_1 be a target such that $\text{Vol}(U(q_0)) = \text{Vol}(U(q_1))$. We impose as a constraint to the deformation $q(\cdot)$ to be of constant total volume, that is, $\text{Vol}(U(q(t))) = \text{Vol}(U(q_0))$. The data attachment function is defined by $g(q) = d(q, q_1)^2$, with d a distance between submanifolds (see [VG05] for examples of such distances).

For the numerical implementation, we take $d = 2$ (thus $S = S^1$) and M is a space of curves, which is discretized as landmarks $q = (x_1, \dots, x_n) \in \text{Lmk}_2(n)$. The volume of a curve is approximately equal to the volume of the polygon $P(q)$ with vertices x_i , given by $\text{Vol}(P(q)) = \frac{1}{2}(x_1 y_2 - y_2 x_1 + \cdots + x_n y_1 - y_n x_1)$.

If one does not take into account a constant volume constraint, a minimizing flow matching a circle on a translated circle usually tends to shrink it along the way (see Figure 10.1(a)). If the volume is required to remain constant then the circle looks more like it were translated towards the target, though the diffeomorphism itself does not look like a translation (see Figure 10.1(b)).

The implementation of the shooting method developed in Section 10.3.2 leads to the diffeomorphism represented on Figure 10.2.



(a) Matching trajectories without constraints. (b) Matching trajectories with constant volume.

Figure 10.1: Matching trajectories.

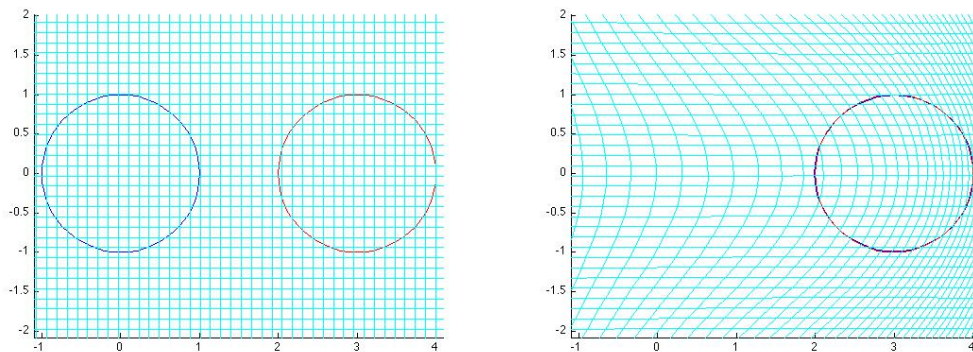


Figure 10.2: Constant volume experiment. On the left: initial condition (in blue) and target (in red). On the right: matching.

10.4.2 Multishape matching

We consider the multishape problem described in Section 10.1.3. We define the shape spaces M_1, \dots, M_k , by $M_j = \text{Emb}(S_j, \mathbb{R}^d)$ for every $j \in \{1, \dots, k-1\}$ and $M_k = M_1 \times \dots \times M_{k-1}$ (background space). For every $j \in \{1, \dots, k\}$ we consider a reproducing kernel K_i and a reduced operator $K_{q,j}$ for every $q \in M_j$.

In the following numerical simulations, each q_j is a curve in \mathbb{R}^2 , so that $S_1 = \dots = S_k = S^1$, the unit circle. The function g appearing in the functional (10.13) is defined by

$$g(q_1, \dots, q_k) = \sum_{j=1}^{k-1} \left(d(q_j, q^{(j)})^2 + d(q_k^j, q^{(j)})^2 \right),$$

where $q_k = (q_k^1, \dots, q_k^{k-1})$ and $q^{(1)}, \dots, q^{(k-1)}$ are given target curves. The distance d is a distance between curves (see [VG05] for examples of such distances). We consider two types of compatibility constraints between homologous curves q_j and q_k^j : either the identity (or stitched) constraint $q_j = q_k^j$, or the identity up to reparametrization (or sliding) constraint $q_k^j = q_j \circ f$ for some (time-dependent) diffeomorphism f of S^1 . Note that, since the curves have the same initial condition, the latter constraint is equivalent to imposing that $\dot{q}_k^j - \dot{q}_j \circ f$

is tangent to q_k^j , which can also be written as

$$(v_k(t, q_k^j) - v_j(t, q_k^j)) \cdot \nu_k^j = 0,$$

where $v_j = K_j \xi_{q_j}^* u_j$ and ν_k^j is normal to q_k^j . In the numerical implementation, the curves are discretized into polygonal lines, and the discretization of the control system (10.14) and of the minimization functional (10.13) is done by reduction to landmark space, as described in section 10.1.4. The discretization of the constraint in the identity case is straightforward. For the identity up to reparametrization (or sliding) constraint, the discretization is slightly more complicated and can be done in two ways. A first way is to add a new state variable ν_k^j which evolves while remaining normal to q_k^j , according to

$$\dot{\nu}_k^j = -dv_k^j(q_k^j)^T \nu_k^j,$$

which can be written as a function of the control u_k and of the derivatives of K_k (this is an example of lifted state space, as discussed in Section 10.1.3). A second way, which is computationally simpler and that we use in our experiments, avoids introducing a new state variable and uses finite-difference approximations. For every $j = 1, \dots, k-1$, and every line segment $\ell = [z_\ell^-, z_\ell^+]$ in q_k^j (represented as a polygonal line), we simply use the constraint $\nu_\ell \cdot (v_j(\ell) - v_k(\ell)) = 0$, where ν_ℓ is the unit vector perpendicular of $z_\ell^+ - z_\ell^-$ and $v_j(\ell) = \frac{1}{2}(v_j(z_\ell^-) + v_j(z_\ell^+))$. Note that the vertices z_ℓ^- and z_ℓ^+ are already part of the state variables that are obtained after discretizing the background boundaries q_k .

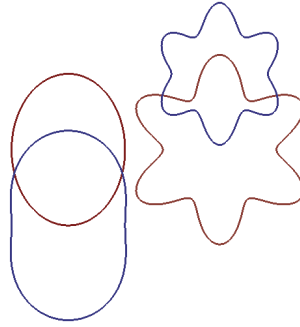


Figure 10.3: Multishape Experiment: Initial (blue) and target (red) sets of curves.

With these choices, the discretized functional and its associated gradient for the augmented Lagrangian method are obtained with a rather straightforward – albeit lengthy – computation. In Figure 10.3, we provide an example comparing the two constraints. In this example, we take $k = 2$ and use the same radial kernel $K_1 = K_2$ for the two shapes, letting $K_1(x, y) = \gamma(|x - y|/\sigma_1)$, with

$$\gamma(t) = (1 + t + 2t^2/5 + t^3/15)e^{-t}.$$

The background kernel is $K_3(x, y) = \gamma(|x - y|/\sigma_3)$, with $\sigma_1 = 1$ and $\sigma_3 = 0.1$. The desired transformation, as depicted in Figure 10.3, moves a curve with elliptical shape upwards, and a flower-shaped curve downwards, each curve being, in addition, subject to a small deformation. The compared curves have a diameter of order 1.

The solutions obtained using the stitched and sliding constraints are provided in Figures 10.4(a) and 10.4(b), in which we have also drawn a deformed grid representing the diffeomorphisms induced by the vector fields v_1 , v_2 and v_3 in their relevant regions. The consequence of the difference between the kernel widths inside and outside the curves on the regularity of the deformation is obvious in both experiments. One can also note differences in the deformation inside between the stitched and sliding cases, the second case being more regular thanks to the relaxed continuity constraints at the boundaries. Finally, we mention the fact that the numerical method that we illustrate here can be easily generalized to triangulated surfaces instead of polygonal lines.

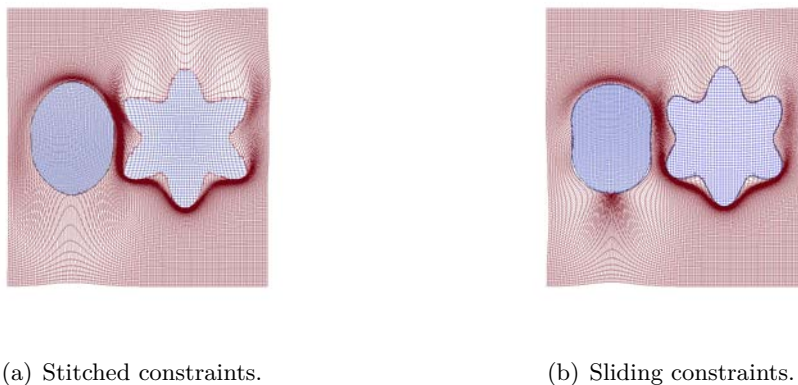


Figure 10.4: Multishape Experiment.

10.5 Conclusion and open problems

The purpose of this paper was to develop a very general framework for the analysis of shape deformations, along with practical methods to find an optimal deformation in that framework. The point of view of control theory gives powerful tools to attain this goal. This allows in particular for the treatment of constrained deformation, which had not been done before.

Now that a concrete setting has been fixed, many new developments can be expected. First of all, the minimization algorithms in the case of constrained shapes are quite slow for a very high number of constraints. Moreover, we did not study any of the geometric aspects of shape deformation spaces. For example, we only briefly mentioned the infinite-dimensional sub-Riemannian structure that the RKHS induces on both groups of diffeomorphism and shape spaces. Sub-Riemannian geometry in infinite dimension and codimension is still a very open subject with very few results. The sub-Riemannian geometry in this paper is particularly difficult and interesting to study because the horizontal spaces may not be closed in the ambient space. A key difference with finite dimension is that some geodesics might exist that are neither normal nor abnormal.

More general control problems can be designed. One can, for example, use a second-order approach, with a control system taking the form

$$\dot{q} = \partial_p H(q, p), \quad \dot{p} = -\partial_q H(q, p) + f(q, u)$$

in which the original state is lifted to the cotangent space ($q \rightarrow (q, p)$), and the new control is u . These models have been introduced for shapes in [TV12], with $f(q, u) = u$, providing a way to interpolate smoothly between multiple shapes. We are currently exploring applications of

this approach to model muscle-like motions, with external forces constrained to being collinear to the fibers.

Another glaring issue comes from the assumption of the surjectivity of the constraints in Theorem 10.2. Indeed, in most practical cases, such as multishapes in Section 10.1.3, this assumption fails. In finite dimension, problems can occur when the rank of C_q changes with q . They are usually solved by taking higher-order derivatives of C_q on the sets on which it is not maximal. This does not seem easily possible with Banach spaces. Another problem comes from the incompatibility of topologies between the Hilbert space V and the Banach space Y in which the constraints are valued: C_q may not have a closed range, in which case we could find "missing" Lagrange multipliers. If $C_q(V)$ were constant, this could be solved simply by restricting Y to $C_q(V)$ equipped with the Hilbert topology induced by C_q and V , but since it is not constant, this might be impossible. It would be very interesting, both for control theory in general and for shape deformation analysis in particular, to find a way to address this problem.

Chapter 11

Symmetric rank-one update

The contents of this chapter were presented in Section 7.4.

Abstract

The symmetric rank-one update method is well known in optimization for its applications in quasi-Newton algorithms. In particular, Conn, Gould, and Toint proved in 1991 that the matrix sequence resulting from this method approximates the Hessian of the minimized function under a suitable linear-independence assumption. Extending their idea, we prove that symmetric rank-one updates can be used to approximate any sequence of symmetric invertible matrices, which has applications to more general problems, such as the computation of constrained geodesics in shape analysis imaging problems. We also provide numerical simulations for the method and some of these applications.

11.1 Introduction

Let d be an integer and $f : \mathbb{R}^d \rightarrow \mathbb{R}$ a \mathcal{C}^2 function. We consider the problem of minimizing f over \mathbb{R}^d . A well-known efficient algorithm to numerically solve this minimization problem is Newton's method: starting at some point x_0 , it considers the sequence

$$x_{k+1} = x_k - h_k H(f)_{x_k}^{-1} \nabla f(x_k),$$

with ∇f the gradient of f , $H(f)$ its Hessian, and $h_k > 0$ some appropriate step length.

However, very often the Hessian of f is too computationally demanding difficult to compute, leading to the introduction of so-called quasi-Newton methods. The methods define a sequence

$$x_{k+1} = x_k - h_k B_k^{-1} \nabla f(x_k),$$

where (B_k) is a sequence of symmetric matrices such that

$$B_{k+1}(x_{k+1} - x_k) = \nabla f(x_{k+1}) - \nabla f(x_k). \quad (11.1)$$

Indeed, since

$$\nabla f(x_{k+1}) - \nabla f(x_k) = \left(\int_0^1 H(f)_{x_k + t(x_{k+1} - x_k)} dt \right) (x_{k+1} - x_k) \simeq H(f)_{x_k} (x_{k+1} - x_k),$$

we get

$$B_{k+1}(x_{k+1} - x_k) \simeq H(f)_{x_k} (x_{k+1} - x_k).$$

It is reasonable to expect B_k to be close to $H(f)_{x_k}$ in the direction $s_k = x_{k+1} - x_k$ (see [DM77, DS96, GMW81]).

There are many ways to build a matrix sequence (B_k) satisfying (11.1). However, it was proved in [CGT91] and [Sch04] that some of these methods let B_k approximate $H(f)_{x_k}$ in all directions instead of just one, i.e.,

$$\|B_k - H(f)_{x_k}\| \xrightarrow[k \rightarrow \infty]{} 0$$

which implies

$$\|B_k - H(f)_{x_*}\| \xrightarrow[k \rightarrow \infty]{} 0,$$

with the additional assumption of the uniform linear independence of the sequence

$$s_k = x_{k+1} - x_k,$$

a notion that will be recalled later. In [CGT91] for example, this is proved for the update of B_k by

$$y_k = \nabla f(x_{k+1}) - \nabla f(x_k) = A_k s_k, \quad r_k = y_k - B_k s_k, \quad B_{k+1} = B_k + \frac{r_k r_k^T}{r_k^T s_k}, \quad (11.2)$$

with

$$A_k = \int_0^1 H(f)_{x_k + t(x_{k+1} - x_k)} dt.$$

In this paper, our aim is to generalize the approach in [CGT91] by defining the above symmetric rank-one algorithm for any sequence of symmetric matrices (A_k) and vectors (s_k) , and derive a convergence result, opening a wider range of applications.

For instance, if a sequence A_k converges to an invertible matrix A_* , then we can use the above algorithm to approximate the inverse A_*^{-1} of the limit A_* . Indeed, let (e_0, \dots, e_{d-1}) be the canonical vector basis of \mathbb{R}^d . We define the sequence (s_k) in \mathbb{R}^d by

$$s_k := A_k e_{k \bmod d}, \quad y_k = A_k^{-1} s_k = e_{k \bmod d}. \quad (11.3)$$

This sequence is uniformly linearly independent, hence the sequence B_k defined by (11.2) will converge to A_*^{-1} . The rate of convergence depends on the dimension d and on the rate of convergence of A_k , but B_k is much easier to compute than A_k^{-1} .

One of the motivating applications is the computation of geodesics constrained to embedded submanifolds of Riemannian spaces. Indeed, to obtain a geodesic between two fixed points of a submanifold, we need to find a converging sequence of maps $t \mapsto \lambda_k(t)$ given implicitly by an equation of the form

$$A_k(t) \lambda_k(t) = c_k(t),$$

where $A_k(t)$ is a convergent sequence of symmetric, positive definite matrices of high dimension (see [ATTY13]). The λ_k are the Lagrange multipliers induced by the equations of the submanifold. It may be very computationally demanding to solve such a linear system for every time t and every step k . Instead, we can take

$$\lambda_k(t) = B_k(t) c_k(t),$$

with $B_k(t)$ obtained by applying the symmetric rank-one algorithm described in the previous paragraph. This is particularly useful in Shape Spaces, where the studied manifolds have a very high dimension and a very complex metric.¹

1. The present article was actually motivated by such a problem appearing in shape analysis, investigated in [ATTY13].

This paper is structured as follows. We give the general framework in Section 11.2, then state the main result after recalling two equivalent definitions of the uniform linear independence of a sequence of vectors in Section 11.3. Then, Section 11.4 is dedicated to intermediate results that will, along with notions developed in Section 11.5, lead to the proof of our theorem. Section 11.6 presents numerical simulations for approximating random matrices and their inverse. Finally, in Section 11.7, we describe the shape deformation problem for which the algorithm was introduced, and apply the symmetric rank-one update method to some simple examples, comparing it to the classical method described in [ATY13].

11.2 Notations and symmetric rank-one algorithm

Consider a sequence $(A_k)_{k \in \mathbb{N}}$ of real square symmetric $d \times d$ matrices. Assume that this sequence converges to some matrix A_* , i.e.,

$$\|A_k - A_*\| \xrightarrow[k \rightarrow \infty]{} 0,$$

where $\|\cdot\|$ is the operator norm on the space $M_d(\mathbb{R})$ of $d \times d$ matrices induced by the canonical Euclidean norm $|\cdot|$ on \mathbb{R}^d . Then define

$$\eta_{k,l} = \sup_{k \leq i \leq l} \|A_i - A_k\|, \quad \text{and} \quad \eta_{k,*} = \sup_{i \geq k} \|A_i - A_k\|$$

for all $k \leq l \in \mathbb{N}$. Note that

$$\forall k \leq l \in \mathbb{N}, \quad \eta_{k,l} \leq \eta_{k,*} \quad \text{and} \quad \eta_{k,*} \rightarrow 0 \quad \text{as} \quad k \rightarrow \infty.$$

Now let $(s_k)_{k \in \mathbb{N}}$ be a sequence of vectors in \mathbb{R}^d .

The objective is to find a somewhat simple sequence $(B_k)_{k \in \mathbb{N}}$ of symmetric matrices such that $B_k \rightarrow A_*$, using only s_k and $y_k = A_k s_k$.

We use the symmetric rank-one update method from [CGT91]. Start with $B_0 = I_d$, the $d \times d$ identity matrix. Define for $k \in \mathbb{N}$

$$y_k = A_k s_k, \quad r_k = (A_k - B_k) s_k = y_k - B_k s_k,$$

and compute

$$B_{k+1} = B_k + \frac{r_k r_k^T}{r_k^T s_k}.$$

It is of course required that $r_k^T s_k \neq 0$ for every k .

11.3 Main Result

For every k , we have

$$B_{k+1} s_k = B_k s_k + r_k = B_k s_k + y_k - B_k s_k = y_k,$$

so

$$A_k s_k = B_{k+1} s_k.$$

The main idea is that if $A_k, A_{k+1}, \dots, A_{k+m}$ are not too far from each other (which is the case for k large enough), we expect $B_{k+m} s_{k+i}$ to be close to $A_{k+m} s_{k+i}$ for $i \leq m$. Then, if we can extract from every finite subsequence (s_k, \dots, s_{k+m}) a vector basis of \mathbb{R}^d , we will obtain the desired convergence.

For a more precise statement, we next define the notion of uniform linear independence. The most intuitive and geometric definition is the following.

Definition 11.1. Take a sequence $s = (s_k)_{k \in \mathbb{N}}$ of vectors in \mathbb{R}^d , $d \in \mathbb{N}^*$, and let $m \geq d$ be an integer. Then s is said to be m -uniformly linearly independent if there exists some constant $\alpha > 0$, such that for all $k \in \mathbb{N}$, there are d integers $k + 1 \leq k_1 < \dots < k_d \leq k + m$ such that

$$|\det(s_{k_1}, \dots, s_{k_d})| \geq \alpha |s_{k_1}| \dots |s_{k_d}|.$$

In other words, from every finite segment of (s_k) of length m , we can extract a linear basis s_{k_1}, \dots, s_{k_d} that will, once normalized, form a parallelepiped that does not become flat as k goes to infinity.

Remark: A sufficient condition for $s = (s_k)_{k \in \mathbb{N}}$ to be m -uniformly linearly independent is the following. If the sequence of subsets $(\{s_{k+1}, \dots, s_{k+m}\})_{k \in \mathbb{N}}$ converges to a subset $\{s_{*,1}, \dots, s_{*,m'}\}$, with $m' \leq m$ a positive integer, that generates \mathbb{R}^d , then, for some integer k_0 large enough, $(s_{k_0+k})_{k \in \mathbb{N}}$ is m -uniformly linearly independent. This is an obvious consequence of the continuity of the determinant.

Another definition was given in [CGT91] after [OR00] as follows.

Definition 11.2. A sequence $s = (s_k)_{k \in \mathbb{N}}$ of vectors in \mathbb{R}^d , d a positive integer, is said to be (m, β) -uniformly linearly independent, where $d \leq m \in \mathbb{N}$ and $\beta > 0$, if for all $k \in \mathbb{N}$, there are d integers $k + 1 \leq k_1 < \dots < k_d \leq k + m$ such that

$$\left| \lambda \left(\frac{s_{k_1}}{|s_{k_1}|}, \dots, \frac{s_{k_d}}{|s_{k_d}|} \right) \right| \geq \beta,$$

where $\lambda(M)$ is the singular value of the square matrix M of smallest magnitude.

Remark: A sequence $s = (s_k)$ in \mathbb{R}^d is (m, β) -uniformly linearly independent for some $m \geq d$ and $\beta > 0$ if and only if it is m -uniformly linearly independent in the sense of Definition 11.1. Indeed, let $v_1, \dots, v_d \in \mathbb{R}^d$, and denote $V = \begin{pmatrix} v_1 & & v_d \\ |v_1| & & |v_d| \end{pmatrix}$. If $|\lambda(V)| \geq \beta > 0$, then $\det(V) \geq \beta^d$, which proves the first part of the equivalence. On the other hand, we know that the eigenvalue of V with largest modulus has modulus less than $\sqrt{d} \max_{i=1, \dots, d} \frac{|s_{k_i}|}{|s_{k_i}|} = \sqrt{d}$. Now, assume that $\det(V) \geq \alpha > 0$. Then $|\lambda(V)| \geq \frac{\alpha}{d^{\frac{d-1}{2}}}$, ensuring the second part of the equivalence.

Theorem 11.1. Let (A_k) , (s_k) , (y_k) , (r_k) and (B_k) be defined as in Section 11.2, with (A_k) having a limit A_* . Assume that there exists a constant $c > 0$ such that for every integer k ,

$$|r_k^T s_k| \geq c |r_k| |s_k|.$$

Then, for every $\beta > 0$ such that (s_k) is (m, β) -uniformly linearly independent in the sense of Definition 11.2, we have for every $k \in \mathbb{N}$ the quantitative estimates

$$\|B_{k+m} - A_*\| \leq \left(1 + \left(\frac{2+c}{c} \right)^{m+1} \right) \frac{\sqrt{d}}{\beta} \eta_{k,*}. \quad (11.4)$$

The next two sections are dedicated to the proof of this theorem.

11.4 First estimates

In this section, we give upper bounds on

$$\left| (B_{k+m} - A_k) \frac{s_k}{|s_k|} \right|,$$

and deduce estimates on

$$\left| \frac{(B_{k+m} - A_*)x}{|x|} \right|$$

for a particular set of $x \in \mathbb{R}^d$.

Proposition 16. *Let $(A_k)_{k \in \mathbb{N}}$ be a sequence of real symmetric matrices in $M_d(\mathbb{R})$, $d \in \mathbb{N}$, and (s_k) be any sequence in \mathbb{R}^d . Define y_k , B_k and r_k as above. Assume that there exists a constant $0 < c \leq 1$ and for every $k \in \mathbb{N}$,*

$$r_k^T s_k \geq c|r_k||s_k|.$$

Then, for every $l \geq k + 1$,

$$|(A_k - B_l)s_k| \leq \left(\frac{2+c}{c} \right)^{l-k-1} \eta_{k,l-1}|s_k|.$$

Proof: We prove this inequality by induction on l , with $k \in \mathbb{N}$ fixed. For $l = k + 1$, we know that $B_{k+1}s_k = A_k s_k = y_k$, hence

$$|(A_k - B_{k+1})s_k| = 0.$$

We will use the notation

$$IH(l) := \left(\frac{2+c}{c} \right)^{l-k-1} \eta_{k,l-1}|s_k|,$$

where IH stands for Induction Hypothesis. Now, assume the result to be true for some $l \geq k + 1$, i.e.,

$$|(A_k - B_l)s_k| \leq \left(\frac{2+c}{c} \right)^{l-k-1} \eta_{k,l-1}|s_k| = IH(l). \quad (11.5)$$

Let us prove that

$$|(A_k - B_{l+1})s_k| \leq \left(\frac{2+c}{c} \right)^{l-k} \eta_{k,l}|s_k| = IH(l+1).$$

Note that

$$\begin{aligned} |(A_k - B_{l+1})s_k| &= |A_k s_k - (B_l + \frac{r_l r_l^T}{r_l^T s_l})s_k| \\ &= |A_k s_k - B_l s_k - \frac{r_l r_l^T s_k}{r_l^T s_l}| \\ &\leq |(A_k - B_l)s_k| + \frac{|r_l||r_l^T s_k|}{c|r_l||s_l|} \\ &\leq IH(l) + \frac{|r_l^T s_k|}{c|s_l|}. \end{aligned} \quad (11.6)$$

Let us find a bound for $\frac{|r_l^T s_k|}{c|s_l|}$, the second term of the right-hand side. First we have

$$\begin{aligned} |r_l^T s_k| &= |y_l^T s_k - s_l^T B_l s_k| \\ &\leq |y_l^T s_k - s_l^T y_k| + |s_l^T (y_k - B_l s_k)| \\ &= |y_l^T s_k - s_l^T y_k| + |s_l^T (A_k - B_l)s_k| \\ &\leq |y_l^T s_k - s_l^T y_k| + |s_l| IH(l). \end{aligned}$$

However, since A_l is symmetric and $y_l = A_l s_l$,

$$|y_l^T s_k - s_l^T y_k| = |s_l^T (A_l - A_k)s_k| \leq \eta_{k,l}|s_l||s_k|,$$

from which we deduce

$$|r_l^T s_k| \leq \eta_{k,l} |s_l| |s_k| + IH(l) |s_l|.$$

Using (11.6), we get

$$\begin{aligned} |(A_k - B_{l+1})s_k| &\leq IH(l) + \frac{|r_l^T s_k|}{c|s_l|} \\ &\leq IH(l) + \frac{1}{c} \eta_{k,l} |s_k| + \frac{1}{c} IH(l) \\ &= \left(1 + \frac{1}{c}\right) IH(l) + \frac{1}{c} \eta_{k,l} |s_k| \\ &= \frac{1+c}{c} \left(\frac{2+c}{c}\right)^{l-k-1} \eta_{k,l-1} |s_k| + \frac{1}{c} \eta_{k,l} |s_k| \\ &\leq \left(\frac{2+c}{c}\right)^{l-k} \eta_{k,l} |s_k| = HI(l+1), \end{aligned}$$

where the last inequality comes from the simple fact that $\eta_{k,l-1} \leq \eta_{k,l}$. \square

This proposition shows that if A_k, A_{k+1}, \dots, A_l are close to one another (i.e., if $\eta_{k,l}$ is small), then $B_l s_k$ stays quantifiably close to $A_k s_k$.

Now, note that $\|A_* - A_k\| \leq \eta_{k,*}$, and $\eta_{k,*}$ decreases to 0 as k goes to infinity. Keeping the same assumptions, we obtain the following result.

Corollary 11.1. *Take $m, k \in \mathbb{N}$, and let $x \in \mathbb{R}^d$ be in the span of s_k, \dots, s_{k+m} . If*

$$\frac{x}{|x|} = \sum_{i=0}^m \lambda_i \frac{s_{k+i}}{|s_{k+i}|}, \quad \lambda_0, \dots, \lambda_m \in \mathbb{R},$$

then

$$\frac{|B_{k+m+1}x - A_*x|}{|x|} \leq \eta_{k,*} \left(1 + \left(\frac{2+c}{c}\right)^m\right) \sum_{i=0}^m |\lambda_i|.$$

Proof: First, it follows from Lemma 16 that

$$\begin{aligned} \frac{|B_{k+m+1}x - A_*x|}{|x|} &\leq \sum_{i=0}^m \frac{|\lambda_i|}{|s_{k+i}|} |B_{k+m+1}s_{k+i} - A_*s_{k+i}| \\ &\leq \sum_{i=0}^m \frac{|\lambda_i|}{|s_{k+i}|} \left(|B_{k+m+1}s_{k+i} - A_{k+i}s_{k+i}| + |A_*s_{k+m} - A_{k+i}s_{k+i}| \right) \\ &\leq \sum_{i=0}^m |\lambda_i| \left(\left(\frac{2+c}{c}\right)^i \eta_{k,k+m} + \eta_{k,*} \right). \end{aligned}$$

Define

$$C(m) = \left(1 + \left(\frac{2+c}{c}\right)^m\right)$$

and, using $\eta_{k,k+m} \leq \eta_{k,*}$, we get

$$\frac{|B_{k+m+1}x - A_*x|}{|x|} \leq \eta_{k,*} C(m) \sum_{i=0}^m |\lambda_i|.$$

The result follows. \square

In particular, if we can let k go to infinity while keeping $\sum_{i=0}^m |\lambda_i|$ bounded, then we obtain $B_{k+m}x \rightarrow A_*x$. Thus, if we can do it for all $x \in \mathbb{R}^d$, we will have proved that $B_k \rightarrow A_*$.

Thus, if we prove that every normalized vector $x \in \mathbb{R}^d$ is a uniformly bounded linear combination of s_k, \dots, s_{k+m} as k goes to infinity, we have proved our theorem. In the next section of this paper, we will define a third notion of uniform linear independence of a sequence directly related with this property and prove that it is equivalent to the previous definitions.

11.5 Uniform m -span of a sequence and applications

In order to investigate the subspace on which $B_k \rightarrow A_*$, we need a notion that is more general than uniform linear independence, that of a uniform span of a sequence of vectors.

Definition 11.3. Let $s = (s_k)_{k \geq 0}$ be a sequence in \mathbb{R}^d , and let $m \in \mathbb{N}$. We say that a vector x in \mathbb{R}^d is uniformly in the m -span of s if for some fixed $\gamma_x > 0$,

$$\forall k \in \mathbb{N}, \quad \exists \lambda_0, \dots, \lambda_m \in \mathbb{R} \quad \frac{x}{|x|} = \sum_{i=0}^m \lambda_i \frac{s_{k+i}}{|s_{k+i}|} \quad \text{and} \quad \sum_{i=0}^m |\lambda_i| \leq \gamma_x. \quad (11.7)$$

We denote by $US_m(s)$ the set of all such vectors.

$US_m(s)$ is a vector sub-space of \mathbb{R}^n . Moreover, there exists a constant $\gamma > 0$ such that Property (11.7) holds for all $x \in US_m(s)$ with $\gamma_x = \gamma$, i.e.,

$$\exists \gamma > 0, \quad \forall k \in \mathbb{N}, x \in US_m(s), \quad \exists \lambda_0, \dots, \lambda_m \in \mathbb{R}, \quad \frac{x}{|x|} = \sum_{i=0}^m \lambda_i \frac{s_{k+i}}{|s_{k+i}|} \quad \text{and} \quad \sum_{i=0}^m |\lambda_i| \leq \gamma. \quad (11.8)$$

To prove the existence of γ in (11.8), it suffices to consider an orthonormal basis $(x_i)_i$ of $US_m(s)$, associated with some constants $(\gamma_{x_i})_{1 \leq i \leq d}$, in Property (11.7). Then we can just take $\gamma = \gamma_{x_1} + \dots + \gamma_{x_d}$.

Remark: There holds $US_m(s) \subset \bigcap_{k=0}^{\infty} \text{span}(s_k, \dots, s_{k+m})$.

Example: Define the sequence $s = (s_k)$ by

$$s_k = \begin{cases} e_{k[d]} & \text{when } k \bmod d \neq d-1 \bmod d, \\ e_0 + \frac{1}{k} e_{d-1} & \text{when } k \bmod d = d-1 \bmod d, \end{cases}$$

Then

$$US_m(s) = \begin{cases} \{0\} & \text{if } 0 \leq m \leq n-2 \\ \text{span}(e_0, \dots, e_{d-2}) & \text{otherwise.} \end{cases}$$

Using this definition, a simple application of Corollary 11.1 gives the following result.

Proposition 17. Let (A_k) , (s_k) , (y_k) , (r_k) and (B_k) be defined as in Section 11.2, assuming that (A_k) has a limit A_* and that $|r_k^T s_k| \geq c|r_k||s_k|$ for some fixed constant $c > 0$. Then, for every $m \in \mathbb{N}$

$$\sup_{x \in US_m(\gamma)} \frac{|B_{k+m+1}x - A_*x|}{|x|} \leq C(m)\gamma\eta_{k,*}, \quad (11.9)$$

where γ is taken from (11.8) and

$$C(m) = \left(1 + \left(\frac{2+c}{c} \right)^m \right).$$

Finally the main result is a consequence of this proposition and of the following lemma.

Lemma 11.1. *Let $s = (s_k)_{k \geq 0}$ be a sequence in \mathbb{R}^d , and let $m \in \mathbb{N}$. Then s is (m, β) -uniformly linearly independent if and only if $US_m(s) = \mathbb{R}^d$. Moreover, we can take $\gamma = \frac{\sqrt{d}}{\beta}$ in (11.8).*

Proof: Let v_1, \dots, v_d be linearly independent elements of \mathbb{R}^d and define the invertible matrix

$$V = \left(\frac{v_1}{|v_1|}, \dots, \frac{v_d}{|v_d|} \right).$$

Let $\Lambda = (\lambda_1, \dots, \lambda_d)^T \in \mathbb{R}^d$, be such that $x = V\Lambda$ has $|x| = 1$. Then

$$\sum_{i=1}^d |\lambda_i| \leq \sqrt{d}|\Lambda| = \sqrt{d}|V^{-1}x| \leq \frac{\sqrt{d}}{|\lambda(V)|}.$$

This proves that if a sequence $s = (s_k)$ in \mathbb{R}^d is (m, β) -uniformly linearly independent, then $US_m(s) = \mathbb{R}^d$ and we can take $\gamma_m(s) = \frac{\sqrt{d}}{\beta}$.

On the other hand, take a unit vector $x \in \mathbb{R}^d$ such that

$$V^{-1T}V^{-1}x = \frac{1}{\lambda(V)^2}x.$$

Then, denoting $(\lambda_1, \dots, \lambda_d)^T = \Lambda = V^{-1}x$,

$$\frac{1}{|\lambda(V)|} = |\lambda(V)| \frac{1}{|\lambda(V)|^2} = |\lambda(V)| |V^{-1T}V^{-1}x| = |\lambda(V)| |V^{-1T}\Lambda| \leq |\Lambda| \leq \sum_{i=1}^d |\lambda_i|,$$

which proves the converse. \square

Our main result is proved.

11.6 Numerical simulations

In this section, after running numerical simulations of our algorithm on random symmetric matrices, we check that the sequence of inverses of a sequence of matrices can indeed be approximated.

All simulations were done using Matlab on a standard desktop computer.

11.6.1 Approximation of a sequence of matrices

Here we test the algorithm on random symmetric matrices with coefficients generated by a normalized Gaussian law.

Let $d \in \mathbb{N}^*$ and define a square symmetric matrix $A_* = \frac{1}{2}(M + M^T)$, where the entries of the $d \times d$ matrix M were chosen at random using the normalized Gaussian law. Then, fix $0 < \lambda < 1$, and define the sequence $(A_k)_{k \in \mathbb{N}}$ of symmetric matrices obtained by perturbing A_* as follows

$$A_k = A_* + \frac{\lambda^k}{2}(M_k + M_k^T),$$

where M_k is a matrix with random coefficients taken uniformly in $[-1, 1]$. This gives $\|M_k\| \leq d$. Obviously, $A_k \rightarrow A_*$ linearly as $k \rightarrow \infty$. More precisely, we have

$$\|A_k - A_*\| \leq d\lambda^k,$$

$\varepsilon(1, 10, k, \lambda)$	k=10	k=20	k=50	k=100
$\lambda = 0.9$	6.5e4	2.3e5	960	5.0
$\lambda = 0.5$	6.5e4	63	5.9e-8	0
$\lambda = 0.1$	6.5e4	6.4e-6	0	0

Table 11.1: Estimates from Theorem 11.1

	k=10		k=20		k=50		k=100	
	mean(δ_k)	max(δ_k)	mean(δ_k)	max(δ_k)	mean(δ_k)	max(δ_k)	mean(δ_k)	max(δ_k)
$\lambda = 0.9$	14	4700	7.5	3800	0.18	8.0	1e-3	0.1
$\lambda = 0.5$	2.7	1700	6.4e-4	3.e-2	7.4e-13	5.3 e-11	0	0
$\lambda = 0.1$	0.068	7.0	8.3e-12	2.1e-9	0	0	0	0

Table 11.2: Simulation results for symmetric rank-one update

so $\eta_{k,*} \leq d\lambda^k$.

Remark: While the Gaussian law is better suited to generate random real numbers, we wanted to have clear bounds on the norm of the perturbations $\frac{\lambda^k}{2}(M_k + M_k^T)$. This is why we only took coefficients with absolute value less than one for each M_k .

We define the $(d, 1)$ -uniformly linearly independent sequence of unit vectors $(s_k)_{k \in \mathbb{N}}$ by the formula

$$s_k = e_{k \bmod d}, \quad k \in \mathbb{N},$$

where (e_0, \dots, e_{d-1}) is the canonical basis of \mathbb{R}^d . Then we apply the symmetric rank-one update to obtain a sequence of symmetric matrices $(B_k)_{k \in \mathbb{N}}$, starting with $B_0 = I_d$. If we assume that there exists $c > 0$ such that $|r_k^T s_k| \geq c|r_k||s_k|$ for every $k \in \mathbb{N}$ then we can apply Theorem 11.1 with $m = d, \beta = 1$, and obtain

$$\|B_{k+d} - A_*\| \leq \varepsilon(c, d, k, \lambda)$$

where

$$\varepsilon(d, k, \lambda, c) = \left(1 + \left(\frac{1+c}{c}\right)^{d+1}\right) d^{3/2} \lambda^k.$$

Remark: In the algorithm, the term $r_k^T s_k = e_{k \bmod d}^T (A_k - B_k) e_{k \bmod d}$ is just the $k \bmod d$ -th diagonal term of $(A_k - B_k)$. It is difficult to give an *a priori* estimate on the term c in Theorem 11.1. For example, if the diagonal terms of the A_k are equal to one, since $B_0 = I_d, r_k^T s_k = 0$ for every k and B_k will never be updated.

Table 11.1 computes the best (i.e., the smallest) upper bounds $\varepsilon(c, d, k, \lambda)$ possible for $d = 10$ for different values of k and λ . They are obtained by taking $c = 1$. This will let us compare the rates of convergence of our simulations with the best possible estimates obtainable by Theorem 11.1. A zero corresponds to a numerical value smaller than the machine epsilon 2.2e-16.

We computed the final distance $\delta_k = \|B_k - A_*\|$ between B_k and A_* for $d = 10$, various values of λ , and various numbers of steps k . We repeated the simulation 1000 times for each value of λ and k , each time with new random values for both A_* and every $M_k, k \in \mathbb{N}$. Table 11.2 gives the mean value and the maximum value of δ_k over these 1000 simulations for each number of steps and each λ .

Comparing the two tables, we see that the numerical convergence rate is even better than the best possible one given by Theorem 11.1. These simulations strongly support the theoretical results.

	k=10		k=20		k=50		k=100	
	mean(δ'_k)	max(δ'_k)	mean(δ'_k)	max(δ'_k)	mean(δ'_k)	max(δ'_k)	mean(δ'_k)	max(δ'_k)
$\lambda = 0.9$	65	2.4e4	20	4500	35	2.4e4	0.22	91
$\lambda = 0.5$	33	1.3e4	2.5	1800	2.8e-8	2.8 e-5	2.8e-9	2.6e-6
$\lambda = 0.1$	14	6700	4.8e-9	1.9e-6	3.7e-11	2.8e-8	2.2e-12	1.4e-10

Table 11.3: Simulation results for symmetric rank-one update on inverses

11.6.2 Approximation of a sequence of inverses

As mentioned in the introduction, our algorithm lets us compute the inverse A_*^{-1} of the limit A_* provided A_* is invertible.

Indeed, consider the following sequences for the symmetric rank-one algorithm

$$s_k := A_k e_{k \bmod d}, \quad y_k = A_k^{-1} s_k = e_{k \bmod d}. \quad (11.10)$$

In other words, s_k is the $k \bmod d$ -th column of A_k . Then the sequence $(s_k)_{k \in \mathbb{N}}$ is (d, β) -linearly independent for some $\beta > 0$ starting at some k_0 large enough since, as k goes to infinity, the finite set $\{s_k, \dots, s_{k+d-1}\} = \{A_k e_{k \bmod d}, \dots, A_{k+d-1} e_{k+d-1 \bmod d}\}$ will converge to the generating set $\{A_* e_0, \dots, A_* e_{d-1}\}$, and the sequence $(s_{k_0+k})_{k \in \mathbb{N}}$ is therefore d -uniformly linearly independent for some k_0 big enough. The smallest singular value of the matrix

$$\left(\frac{s_k}{|s_k|}, \dots, \frac{s_{k+d-1}}{|s_{k+d-1}|} \right) \quad (11.11)$$

will converge to that of

$$\left(\frac{A_* e_0}{|A_* e_0|}, \dots, \frac{A_* e_{d-1}}{|A_* e_{d-1}|} \right),$$

which depends only on A_* , which gives an insight on the correct value of β . The value of c in Theorem 11.1, however, cannot be guessed here either.

Take the sequence $(B_k)_{k \in \mathbb{N}}$ obtained by applying the symmetric rank-one update method, with starting point $B_0 = I_d$ and using the sequence $(s_k)_{k \in \mathbb{N}}$ defined by (11.11). Assuming that there exists $c > 0$ such that $|r_k^T s_k| \geq c |r_k| |s_k|$ for every $k \in \mathbb{N}$, this sequence converges to A_*^{-1} by Theorem 11.1. The rate of convergence depends on the dimension d and the rate of convergence of A_k . Note that B_k is much easier to compute than A_k^{-1} . Indeed, the complexity for the computation of the inverse of a $d \times d$ matrix is greater than the $O(d^2)$ complexity required in each symmetric rank-one update. This can be useful when solving approximately converging sequences of linear equations, as we will show in the next section.

In our numerical simulation, we computed the distance δ'_k between B_k and A_*^{-1} for different values of k and λ . We used the same sequence (A_k) with random coefficients as in the previous section, with (A_k) converging linearly to a random (but invertible) symmetric matrix A_* with rate $0 < \lambda < 1$. For each number of steps and each value of λ considered, we repeated this simulation 1000 times for different A_* and different random matrices A_k . Table 11.3 gives the mean value and maximum value of δ'_k over these 1000 simulations.

We see that $\delta'_k = \|B_k - A_*^{-1}\|$ does decrease to zero, but with a slower rate than that of the approximation of A_* itself given in Table 11.2. Moreover, the maximal value is significantly larger than the mean value for this distance. A reasonable explanation for both discrepancies is that A_* can be ill-conditioned when generated in such a random way. This can cause them to be almost singular, which would have two consequences. First, the rate of convergence of $(A_k^{-1})_{k \in \mathbb{N}}$ to A_*^{-1} is slower than that of $(A_k)_{k \in \mathbb{N}}$ to A_* . Therefore, the $\eta_{k,*}$ from Theorem 11.1

	k=10		k=20		k=50		k=100	
	mean(δ'_k)	max(δ'_k)	mean(δ'_k)	max(δ'_k)	mean(δ'_k)	max(δ'_k)	mean(δ'_k)	max(δ'_k)
$\lambda = 0.9$	23	2300	11	2200	0.25	25	2.3e-3	1.2
$\lambda = 0.5$	2.8	220	1.2e-3	0.16	1.3e-12	4.8e-10	1.5e-15	8.7e-12
$\lambda = 0.1$	0.36	38	7.4e-12	7e-9	1.1e-14	4.8e-12	8.2e-15	1.4e-12

Table 11.4: Simulation results for inverses of matrices with singular values in $[0.5, 1.5]$

	k=10	k=20	k=50	k=100
$\lambda = 0.9$	83	56	47	12
$\lambda = 0.5$	85	29	1.0e-4	1.2 e-7
$\lambda = 0.1$	54	2.1e-4	3.7e-7	1.0e-9

Table 11.5: Simulation results for inverses of matrices y_k randomly generated

is larger than in the case described in Section 11.6.1. Second, the sequence $s_k = A_k e_{k \bmod d}$ is less uniformly linearly independent.

To test this hypothesis, we tried the simulation again but this time we forced A_* to have a good conditioning. This will make the sequence of matrices A_k uniformly well-conditioned. This kind of sequence can appear in certain numerical simulations of PDEs, in cases where the A_k are discretized versions of a positive-definite self-adjoint operator.

For the simulation, we took A^* so that its singular values all belong to $[0.5, 1.5]$. For this, we used $A_* = O^T D O$, where D is a diagonal matrix of size 10×10 with diagonal coefficients randomly generated using the uniform law on $[0.5, 1.5]$, and O was obtained by orthonormalizing the columns of a random matrix Z , whose coefficients were generated using a Gaussian law. Leaving the rest of the process unchanged, we performed 1000 simulations for the same values of λ and k as those from Table 11.3 and obtained Table 11.4.

As expected, the numbers on Table 11.4 show that the sequence $(B_k)_{k \in \mathbb{N}}$ converges to A_*^{-1} much faster than in the case of a purely random A_* . In fact, the convergence is almost as good as the one shown by Table 11.2 in the previous section. This confirms that the method is more effective with sequences of matrices that are well scaled.

We also did an extra simulation in the case of a purely random A_* : since the sequence $(s_k)_{k \in \mathbb{N}}$ in (11.10) has no reason to be particularly good (i.e., uniformly linearly independent with a nice constant), we applied our algorithm this time by taking a new sequence for $(y_k)_{k \in \mathbb{N}}$, letting each y_k be a random vector with coefficients taken along a normal Gaussian law at each step. We still set $s_k = A_k y_k$. This is done in the hopes that, on average, the sequence s_k could be more uniformly linearly independent. We computed the mean values for $\delta'_k = \|B_k - A_*^{-1}\|$ over 1000 repetition of this simulation. They are given in Table 11.5.

This experiment shows that taking a random sequence of vectors $(y_k)_{k \in \mathbb{N}}$ is not as effective as taking the y_k periodically equal to the canonical basis of \mathbb{R}^d .

For large values of d , as k goes to infinity, this method gives us an approximation of the whole sequence $(A_k^{-1})_{k \in \mathbb{N}}$ and is faster than computing the inverse of A_k at every step. Indeed, the complexity of one rank-one update is only in $O(d^2)$.

Remark: This method does not allow for better computations of inverse of a badly scaled matrix A by setting $A_k = A$ for every k . A quick Matlab simulation showed that the command `inv(A)` gives more precise results.

11.7 An application: optimal deformations of constrained shapes

The main problem of shape deformation analysis is to find an optimal deformation from one shape to another. From the numerical point of view, a shape is usually a collection of distinct points $q^T = (x_1^T, \dots, x_n^T)$ where n is a positive integer and each x_i is an element of \mathbb{R}^d . They are usually a discretization of the boundary of a certain domain in \mathbb{R}^d . The space of such shapes is called the space $Lmk_d(n)$ of n landmarks in \mathbb{R}^d , i.e.,

$$Lmk_d(n) = \{q = (x_1^T, \dots, x_n^T)^T \in \mathbb{R}^{nd}, i \neq j \Rightarrow x_i \neq x_j\}.$$

A deformation of an initial shape q_0 , with $q_0^T = (x_{1,0}^T, \dots, x_{n,0}^T)$, is a curve $t \mapsto q(t)$ with $q(0) = q_0$, of Sobolev class $W^{1,2}$, that is, an absolutely continuous curve with square-integrable speed.

11.7.1 Large deformation diffeomorphic metric mapping.

The so-called LDDMM (Large Deformation Diffeomorphic Metric Mapping) setting for studying deformations of landmarks is used in computational anatomy [BMTY05, DGM98, Tro05].

We start by considering a Reproducing Kernel Hilbert Space V of smooth vector fields on \mathbb{R}^d , that is, a subspace V of $\mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R}^d)$ equipped with a Hilbert product $\langle \cdot, \cdot \rangle_V$ such that the inclusion $V \hookrightarrow \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R}^d)$ is continuous. For such a space, there exists a matrix-valued *reproducing kernel* $K : \mathbb{R}^d \times \mathbb{R}^d \rightarrow M_d(\mathbb{R}^d)$ such that, for any $x, v \in \mathbb{R}^d$, and every $X \in V$,

$$\langle K(\cdot, x)v, X \rangle_V = v^T X(x).$$

Such a space V is completely determined by its reproducing kernel K . The kernel we use in this example is given by

$$K(x, y) = e^{-\frac{|x-y|^2}{2\sigma}} I_d,$$

for some $\sigma > 0$.

Then, one considers deformations $t \mapsto q(t) \in Lmk_d(n)$ of the form

$$q(t)^T = (\varphi^X(t) \cdot q_0)^T = (\varphi^X(t, x_{1,0})^T, \dots, \varphi^X(t, x_{n,0})^T),$$

where $(\varphi^X(t))_{t \in [0,1]}$ is a family of diffeomorphisms of \mathbb{R}^d , flow of a time-dependent vector field $t \mapsto X(t, \cdot) \in V$ on \mathbb{R}^d such that $t \mapsto \|X(t, \cdot)\|_V$ is square-integrable. The optimal deformation $t \in [0, 1] \mapsto \varphi^X(t) \cdot q_0$ from a starting shape q_0 to a target shape q_1 is the one such that the total energy $\frac{1}{2} \int_0^1 \langle X(t), X(t) \rangle_V dt$ is minimal. The reason of using such a setting is that it actually provides an optimal deformation of the full space \mathbb{R}^d thanks to the flow φ^X .

In particular, it preserves the fact that $q(0)$ is the discretization of the boundary of a certain domain U of \mathbb{R}^d : $q(t)$ will then be a discretization of the boundary of the deformed domain $U(t) = \varphi(t, U)$.

Since it is extremely hard to determine this minimum, one usually considers the penalized functional

$$\hat{J}(X) = \frac{1}{2} \int_0^1 \langle X(t), X(t) \rangle_V dt + g(\varphi^X(1) \cdot q_0),$$

where $g : \mathbb{R}^d \rightarrow \mathbb{R}$ is a smooth data attachment term, minimal at q_1 . A classical result [You10], consequence of the solution to the spline interpolation problem for vector fields in V , is that minimizing \hat{J} is equivalent to minimizing the functional

$$J(u) = \frac{1}{2} \int_0^1 \sum_{i,j=1}^n e^{-\frac{|x_i(t) - x_j(t)|^2}{2\sigma}} u_i(t)^T u_j(t) dt + g(q(1)), \quad u_i \in L^2(0, 1; \mathbb{R}^d), \quad i = 1, \dots, n,$$

where $q(t) = (x_1(t), \dots, x_n(t))$ satisfies the control system

$$q(0) = q_0, \quad \dot{x}_i(t) = \sum_{j=1}^n e^{-\frac{|x_i(t)-x_j(t)|^2}{2\sigma}} u_j(t) \quad \text{a.e. } t \in [0, 1], \quad i = 1, \dots, n.$$

We can retrieve the corresponding flow φ^X by integrating the vector field

$$X(t, x) = \sum_{j=1}^n e^{-\frac{|x-x_j(t)|^2}{2\sigma}} u_j(t).$$

We can write this differential equation as $\dot{q}(t) = K_{q(t)}^\sigma u(t)$, where K_q^σ is the bloc matrix of size $nd \times nd$ consisting of blocks of size $d \times d$, with the (i, j) -th block is equal to $e^{-\frac{|x_i(t)-x_j(t)|^2}{2\sigma}} I_d$, and $u = (u_1^T, \dots, u_n^T)^T \in (\mathbb{R}^d)^n$. This is a symmetric, positive-definite matrix for every q in $Lmk_d(n)$, and we also have

$$\sum_{i,j=1}^n e^{-\frac{|x_i-x_j|^2}{2\sigma}} u_i^T u_j = u^T K_q^\sigma u.$$

In this form, this an optimal control problem in finite dimension. It has an optimal solution which satisfies certain Hamiltonian equations, and can be solved numerically (see [ATTY13, Tro05, You10]).

11.7.2 Shapes with constraints.

The shape deformation problem which motivated the symmetric rank-one update described in this paper is an extension of the one described in the previous section, aimed at studying several shapes simultaneously. Let n^1, n^2 be two positive integers. Assume that we have two different starting shapes q_0^1 and q_0^2 , belonging to different landmarks spaces $Lmk_d(n^1)$ and $Lmk_d(n^2)$, each one being a discretization of a different domain U^1 and U^2 with $U^1 \cap U^2 = \emptyset$.

Usually, one would just consider the total shape to be the union of those two shapes, and deform it using a single diffeomorphism. However, the objects we want to model should be considered as two independent shapes, as in the case of images of different parts of the brain.

This is why, instead, we would like to model a deformation of q_0^1 and q_0^2 such that they evolve independently from one another (each one being deformed by a different diffeomorphism), but are immersed in a deformable background (deformed by a third diffeomorphism), whose boundary coincides with the union of the boundaries of U^1 and U^2 : one obtains a discretization q^3 of this boundary by concatenation $(q_0^3)^T = ((q_0^1)^T, (q_0^2)^T) \in Lmk_d(n^3)$, where $n^3 = n^1 + n^2$. The total shape $q_0^T = ((q_0^1)^T, (q_0^2)^T, (q_0^3)^T)$ belongs to the space $Lmk_d(n^1) \times Lmk_d(n^2) \times Lmk_d(n^3)$.

The main reason for considering constrained shape deformation is to model such a system [ATTY13]. Indeed, to model a deformation $q(t)^T = (q^1(t)^T, q^2(t)^T, q^3(t)^T)$ of the total shape, we can use the control system

$$q(0) = q_0, \quad \dot{q}^1(t) = K_{q^1(t)}^{\sigma^1} u^1(t), \quad \dot{q}^2(t) = K_{q^2(t)}^{\sigma^2} u^2(t), \quad \dot{q}^3(t) = K_{q^3(t)}^{\sigma^3} u^3(t), \quad \text{a.e. } t \in [0, 1],$$

where $u^i \in L^2(0, 1; \mathbb{R}^{dn^i})$, $i = 1, 2, 3$. This control system can be written $\dot{q}(t) = K_{q(t)} u(t)$, where

$$K_q = \begin{pmatrix} K_{q^1}^{\sigma^1} & 0 & 0 \\ 0 & K_{q^2}^{\sigma^2} & 0 \\ 0 & 0 & K_{q^3}^{\sigma^3} \end{pmatrix}.$$

Note that the matrix K_q is symmetric and positive definite for every $q \in Lmk_d(n^1) \times Lmk_d(n^2) \times Lmk_d(n^3)$. The functional we want to minimize is therefore given by

$$J(u) = \frac{1}{2} \int_0^1 u(t)^T K_{q(t)} u(t) dt + g(q(1)).$$

K_q is a symmetric positive definite matrix. However, one also needs to preserve the condition $\dot{q}^3(t)^T = ((q^1(t))^T, (q^2(t))^T)$ for every $t \in [0, 1]$ (that is, the boundary of the deformed background coincides with the boundaries of the deformed domains). This means that we should impose on our control system the constraints $(\dot{q}^3)^T = ((\dot{q}^1)^T, (\dot{q}^2)^T)$, i.e.,

$$K_{\dot{q}^3}^{\sigma^3} u^3(t) = ((K_{q^1}^{\sigma^1} u^1(t))^T, (K_{q^2}^{\sigma^2} u^2(t))^T)^T \quad \text{a.e. } t \in [0, 1].$$

These linear constraints can be as written $CK_{q(t)}u(t) = 0$, with

$$C = \begin{pmatrix} I_{n^3} & -I_{n^3} \end{pmatrix}.$$

In conclusion, we wish to find a minimum of $J(u) = \frac{1}{2} \int_0^1 u(t)^T K_{q(t)} u(t) dt + g(q(1))$ over all square-integrable functions $u : [0, 1] \rightarrow (\mathbb{R}^d)^{n^1+n^2+n^3}$ such that $q(0) = q_0$ and, for almost every t in $[0, 1]$, $\dot{q}(t) = K_{q(t)}u(t)$ and $CK_{q(t)}u(t) = 0$.

Now, define for $q \in Lmk_d(n^1) \times Lmk_d(n^2) \times Lmk_d(n^1+n^2)$ the symmetric positive definite $n^3 \times n^3$ matrix $A_q = CK_qC^T$.

Then, according to an appropriate version of the Pontryagin Maximum Principle ([ATTY13, Tré08]), if u is optimal for our constrained minimization problem, and q is the curve such that $q(0) = q_0$ and $\dot{q} = K_q u$, then there exists an absolutely continuous curve $p : [0, 1] \in (\mathbb{R}^d)^{n^1+n^2+n^3}$ with square-integrable speed such that $p(1) + \nabla g_{q(1)} = 0$, and, for almost every t in $[0, 1]$,

$$\begin{cases} u(t) = \left(p(t) - C^T A_{q(t)}^{-1} CK_{q(t)} p(t) \right), \\ \dot{q}(t) = K_{q(t)} \left(p(t) - C^T A_{q(t)}^{-1} CK_{q(t)} p(t) \right), \\ \dot{p}(t) = -\frac{1}{2} \left(p(t) - C^T A_{q(t)}^{-1} CK_{q(t)} p(t) \right)^T \nabla_q K_{q(t)} \left(p(t) - C^T A_{q(t)}^{-1} CK_{q(t)} p(t) \right). \end{cases} \quad (11.12)$$

Here, we used the notation $a^T \nabla_q K_q b$ for the gradient of the map $q \mapsto a^T K_q b$ at q , for $a, b \in (\mathbb{R}^d)^{n^1+n^2+n^3}$ fixed. Since this is a differential equation with smooth coefficient, it has a unique solution for fixed (q_0, p_0)

Remark: The system (11.12) actually consists of the Hamiltonian geodesic equations on the submanifold defined by $C = 0$ for the Riemannian metric whose cometric tensor is K_q .

Moreover, in this case, $t \mapsto u(t)^T K_{q(t)} u(t)$ is constant. Hence, the minimization of J reduces to the minimization of

$$\tilde{J}(p_0) = \frac{1}{2} \left(p_0 - C^T A_{q_0}^{-1} CK_{q_0} p_0 \right)^T K_{q_0} \left(p_0 - C^T A_{q_0}^{-1} CK_{q_0} p_0 \right) + g(q(1))$$

with respect to the initial momentum $p_0 = p(0)$. Note that this reduction is fundamental in our approach.

The computation the gradient of \tilde{J} requires solving an adjoint equation with coefficients depending on the derivatives of the right-hand side of (11.12). This is described in more details in [ATTY13].

All operations involved in the computation of this gradient have a number of multiplications have complexity in $O((n^1 + n^2)^2)$ at each time step, for some constant $m > 0$, with the distinct exception of the computation of the inverse of A_q (or, at least, the resolution of linear equations of the form $A_q a = b$ which appear both in (11.12) and several times in the associated adjoint equation), whose complexity is higher.

11.7.3 Implementation of the symmetric rank-one update.

One of the most time-consuming aspects of this method is the computation, at each time step, of the inverse of A_q . Therefore, we can try to speed things up by applying a symmetric rank-one update as follows.

Let $f_0, \dots, f_{d(n^1+n^2)-1}$ be the canonical vector basis of $\mathbb{R}^{dn^1+dn^2}$. For any $k \in \mathbb{N}$, define

$$y_k = e_{k \bmod d(n^1+n^2)}.$$

We start with the initial momentum $p_0 = 0$, and let $B_0(t) = A_{q_0}^{-1}$ for each $t \in [0, 1]$. The computation of $A_{q_0}^{-1}$ is necessary to compute the different gradients anyway, so this does not add any extra time, and gives a better initial conditioning of the matrix. Then, assuming we have constructed an initial momentum p_k and a family of matrices $B_k(t)$, $t \in [0, 1]$, we use (11.12) to compute a trajectory $x_k(t)$, replacing A_q^{-1} by $B_k(t)$. Finally, at each time t , we define

$$\begin{aligned} s_k(t) &= A_{q_k(t)} y_k, \\ r_k(t) &= B_k(t) s_k(t) - y_k, \\ B_{k+1}(t) &= B_k(t) + \frac{r_k(t) r_k^T(t)}{r_k^T(t) s_k(t)}. \end{aligned}$$

We can then compute the gradient of \tilde{J} with an adjoint equation, where any directional derivative

$$\partial_q(A_{q_k(t)} a)^{-1}(b) = -A_{q_k(t)}^{-1} \partial_q(A_{q_k(t)})(b) A_{q_k(t)}^{-1} a, \quad a \in \mathbb{R}^{n^3}, \quad b \in (\mathbb{R}^d)^{n^1+n^2+n^3}$$

is replaced with $-B_k(t) \partial_q A_{q_k(t)}(v) B_k(t) a$. This lets us perform the minimization of \tilde{J} using gradient descent or a regular quasi-Newton algorithm.

As long as the algorithm gives a converging sequence of initial momenta p_k , the trajectories $q_k(t)$ will also converge to a trajectory $q_*(t)$, making each $A_{q_k(t)}$, with $t \in [0, 1]$ fixed, a converging sequence, with invertible limit $A_*(t)$. Therefore, each $B_k(t)$, for $t \in [0, 1]$ fixed, converges to $A_*(t)$ as $k \rightarrow \infty$. In other words, as $k \rightarrow \infty$, we are indeed computing the true gradient of \tilde{J} .

11.7.4 Numerical simulations

We consider the shapes q_0^1 and q_0^2 as an equidistant discretization of $n^1 = n^2 = n$ points of circles of radius 1 in \mathbb{R}^2 , centered at $(-1, 0)$ (resp. $(2, 0)$), and we try to match them with circles of radius 1 (resp. 0.9) centered at $(-0.75, 0)$ (resp. $(1.5, 0)$). This is actually a difficult matching to perform without the context of constrained shape deformation, because deforming two separate shapes into targets that are so close to each other is very difficult using a single diffeomorphism. Constrained multiple shapes provide an appropriate framework for finding such a matching.

The following values are taken for the constants: $\sigma^1 = 0.5$, $\sigma^2 = 0.3$ and $\sigma^3 = 0.1$. We used 10 time steps, with step length $\Delta t = 0.1$.

The first thing that we compared is the time necessary to accomplish one gradient step thanks to the adjoint equations to (11.12) as described in [ATTY13]. Then we measured the time required to complete the algorithm with the same stopping condition on the gradient algorithm. Finally, we compare how much the constraints are satisfied in the final deformation obtained thanks to the gradient algorithm. This is done by computing the total value of the lack of satisfaction of the constraints

$$\gamma(u(\cdot)) = \sqrt{\int_0^1 |CK_{q(t)} u(t)|^2 dt}.$$

Number of points	n=30		n=60	
Method used	Regular method	Rank-one update	Regular method	Rank-One update
Time for one gradient	3.3	1	9.8	2.6
Total time	63	23	69	23
$\gamma(u(\cdot))$	1.e-15	0.66	1.e-15	1.1

Table 11.6: Comparison of the time necessary to perform a gradient descent algorithm and satisfaction of constraints

We obtain Table 11.6. It gives a comparison between the regular method of computing the exact solutions of any linear system $A_q a = b$ appearing in the adjoint equations, and the one using the symmetric rank-one update method described in the previous paragraph.

As expected, the symmetric rank-one algorithm is much faster. However, since the regular method computes the exact solutions for satisfying the constraints, those are satisfied with great precision. This is not the case when using the rank-one update method where we obtain $\gamma(u(\cdot)) = 1.1$ in the case $n = 60$. However, this value is actually quite small, since we are computing an Euclidean norm in $\mathbb{R}^{n^1+n^2} = \mathbb{R}^{120}$. Indeed, Figure 11.1 is a picture of the final deformation obtained by the rank-one update.

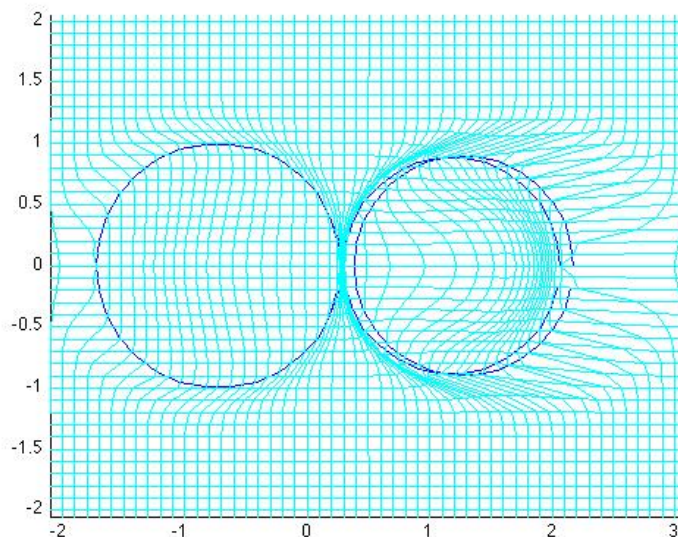


Figure 11.1: Multiple shape experiment: matching circles with the rank-one update method.

In Figure 11.1, there are four circles (two of which coincide), one for each of the two shapes, and two for the background. Since we are in the framework of LDDMM, the deformations are induced by diffeomorphisms, showcased by their action on a grid. In fact, there are three diffeomorphisms: one for the deformation of the first shape (circle on the left), shown by the grid deformation in the left disk, one for the deformation of the second shape (circle on the right), represented by the grid deformation on the disk on the right-hand side, and the diffeomorphism acting on the background. This one is represented by the deformation it induces on the rest of the grid.

The constraints forced the two circles on the left-hand side to coincide. The difference between the two circles on the right shows that the constraints are only approximately satisfied. Note that while visible, the difference is rather small. We could force the constraints to

be better satisfied by increasing the number of gradient steps. In fact, we obtained numbers as small as $1e-4$ by taking a great many number of steps (around 2000 or so). However, in this case, it is faster to use the regular method.

11.8 Conclusion

The symmetric rank-one update used in quasi-Newton methods for minimizing real functions can be generalized to a more abstract framework. The algorithm this paper obtained can be used to approximate sequences of symmetric matrices. This opens several possible applications, one of which is the computation of the sequence of inverses of a sequence of invertible matrices. This can be applied to compute constrained optimal controls for which the computation of the inverse of the constraints is too computationally demanding. This is particularly useful when tackling problems of shape deformations, where the number of constraints can be quite large. We are hopeful that other applications of the symmetric rank-one update method can be given.

However, some limitations remain. In particular, the rate of convergence may be low, since, in higher dimensions, the term $\left(\frac{1+c}{c}\right)^{m+1}$ in the main theorem will be quite large, because m is at least equal to the number of columns in the matrices of the sequence and is therefore large in high dimensions. Moreover, just as for the classical quasi-Newton algorithms, there is no clear way to find a lower bound for c , even in the more simple cases.

Bibliography

- [ABC⁺10] A. A. Agrachev, U. Boscain, G. Charlot, R. Ghezzi, and M. Sigalotti. Two-dimensional almost-Riemannian structures with tangency points. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 27(3):793–807, 2010.
- [AC09] A. A. Agrachev and M. Caponigro. Controllability on the group of diffeomorphisms. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 26(6):2503–2509, 2009.
- [AG04] B. Avants and J. C. Gee. Geodesic estimation for large deformation anatomical shape averaging and interpolation. *Neuroimage*, 23:S139–S150, 2004.
- [Agr09] A. A. Agrachëv. Any sub-Riemannian metric has points of smoothness. *Dokl. Akad. Nauk*, 424(3):295–298, 2009.
- [Arg13] S. Arguillère. Approximating sequences of symmetric matrices with a generalized quasi-newton algorithm. *Preprint*, 2013.
- [Arn66] V. Arnold. Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l’hydrodynamique des fluides parfaits. *Ann. Inst. Fourier (Grenoble)*, 16(fasc. 1):319–361, 1966.
- [Arn89] V. Arnold. *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1989.
- [Aro50] N. Aronszajn. Theory of reproducing kernels. *Trans. Amer. Math. Soc.*, 68:337–404, 1950.
- [AS04] A. A. Agrachev and Y. L. Sachkov. *Control theory from the geometric viewpoint*, volume 87 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2004. Control Theory and Optimization, II.
- [AT] S. Arguillère and E. Trélat. Sub-riemannian geometry on groups of diffeomorphisms and applications to shape spaces.
- [ATTY13] S. Arguillère, E. Trélat, A. Trouvé, and L. Younes. Control theory and shape analysis. *Preprint*, arxiv.org/abs/1401.0661, 2013.
- [BBHM13] M. Bauer, M. Bruveris, P. Harms, and P. W. Michor. Geodesic distance for right invariant sobolev metrics of fractional order on the diffeomorphism group. *Annals of Global Analysis and Geometry*, 44(1):5–21, 2013.
- [BBI01] D. Burago, Y. Burago, and S. Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.
- [BCF⁺13] C. Bertoglio, D. Chapelle, M. A. Fernández, J.-F. Gerbeau, and P. Moireau. State observers of a vascular fluid-structure interaction model through measurements in the solid. *Comput. Methods Appl. Mech. Engrg.*, 256:149–168, 2013.
- [BCT07] B. Bonnard, J.-B. Caillaud, and E. Trélat. Second order optimality conditions in the smooth case and applications in optimal control. *ESAIM Control Optim. Calc. Var.*, 13(2):207–236, 2007.

- [BFT06] B. Bonnard, L. Faubourg, and E. Trélat. *Mécanique céleste et contrôle des véhicules spatiaux*, volume 51 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 2006.
- [BHM11] M. Bauer, P. Harms, and P. W. Michor. Sobolev metrics on shape space of surfaces. *J. Geom. Mech.*, 3(4):389–438, 2011.
- [BHM12] M. Bauer, P. Harms, and P. W. Michor. Sobolev metrics on shape space, II: Weighted Sobolev metrics and almost local metrics. *J. Geom. Mech.*, 4(4):365–383, 2012.
- [Bis84] Jean-Michel Bismut. *Large deviations and the Malliavin calculus*, volume 45 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, 1984.
- [BMTY05] M. F. Beg, M. I. Miller, A. Trounev, and L. Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *Int. J. Comput. Vis.*, 61(2):139–157, 2005.
- [Bou83] R. D. Bourgin. *Geometric aspects of convex sets with the Radon-Nikodým property*, volume 993 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1983.
- [BR96] A. Bellaïche and J.-J. Risler, editors. *Sub-Riemannian geometry*, volume 144 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 1996.
- [BV13] M Bruveris and F.-X. Vialard. On completeness of groups of diffeomorphisms. *Preprint, arxiv.org/abs/1403.2089*, 2013.
- [CDPT07] L. Capogna, D. Danielli, S. D. Pauls, and J. T. Tyson. *An introduction to the Heisenberg group and the sub-Riemannian isoperimetric problem*, volume 259 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 2007.
- [CGT91] A. R. Conn, N. I. Gould, and P. L. Toint. Convergence of quasi-Newton matrices generated by the symmetric rank one update. *Mathematical Programming*, 50(2):177–195, 1991.
- [CJT06] Y. Chitour, F. Jean, and E. Trélat. Genericity results for singular curves. *J. Differential Geom.*, 73(1):45–73, 2006.
- [CJT08] Y. Chitour, F. Jean, and E. Trélat. Singular trajectories of control-affine systems. *SIAM J. Control Optim.*, 47(2):1078–1095, 2008.
- [Cla09] B. Clarke. *The completion of the manifold of Riemannian metrics with respect to its L^2 metric*. PhD thesis, University of Leipzig, arxiv.org/pdf/0904.0177v1.pdf, 2009.
- [Day40] M. M. Day. The spaces L^p with $0 < p < 1$. *Bull. Amer. Math. Soc.*, 46:816–823, 1940.
- [DGM98] P. Dupuis, U. Grenander, and M. I. Miller. Variational problems on flows of diffeomorphisms for image matching. *Quart. Appl. Math.*, 56(3):587–600, 1998.
- [DM77] J. E. Dennis, Jr. and J. J. Moré. Quasi-Newton methods, motivation and theory. *SIAM Rev.*, 19(1):46–89, 1977.
- [DM98] I. L. Dryden and K. V. Mardia. *Statistical shape analysis*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Ltd., Chichester, 1998.
- [DS80] P.I. Dubnikov and S.N. Samborskii. Controllability criterion for systems in a banach space (generalization of chow’s theorem). *Ukrainian Mathematical Journal*, 32(5):429–432, 1980.
- [DS96] J. E. Dennis, Jr. and R. B. Schnabel. *Numerical methods for unconstrained optimization and nonlinear equations*, volume 16 of *Classics in Applied Mathematics*.

- Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1996. Corrected reprint of the 1983 original.
- [EM70] D. G. Ebin and J. Marsden. Groups of diffeomorphisms and the motion of an incompressible fluid. *Ann. of Math. (2)*, 92:102–163, 1970.
- [ES96] J. Eichhorn and R. Schmid. Form preserving diffeomorphisms on open manifolds. *Ann. Global Anal. Geom.*, 14(2):147–176, 1996.
- [Eva10] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.
- [FR10] A. Figalli and L. Rifford. Mass transportation on sub-Riemannian manifolds. *Geom. Funct. Anal.*, 20(1):124–159, 2010.
- [GM98] U. Grenander and M. I. Miller. Computational anatomy: an emerging discipline. *Quart. Appl. Math.*, 56(4):617–694, 1998. Current and future challenges in the applications of mathematics (Providence, RI, 1997).
- [GMV12] E. Grong, I. Markina, and A. Vasil'ev. Sub-riemannian geometry on infinite-dimensional manifolds. *Preprint, arxiv.org/abs/1201.2251*, 2012.
- [GMW81] P. E. Gill, W. R. Murray, and M. H. Wright. *Practical optimization*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1981.
- [GTY] J. Glaunès, A. Trouvé, and L. Younes. Diffeomorphic matching of distributions: a new approach for unlabelled point-sets and sub-manifolds matching. In *CVPR 2004*, pages 712–718.
- [GTY06] J. Glaunès, A. Trouvé, and L. Younes. Modeling planar shape variation via Hamiltonian flows of curves. In *Statistics and analysis of shapes*, Model. Simul. Sci. Eng. Technol., pages 335–361. Birkhäuser Boston, Boston, MA, 2006.
- [HMR98] D. D. Holm, J. Marsden, and T. S. Ratiu. Euler-poincaré models of ideal fluids with nonlinear dispersion. *Phys. Rev. Lett.*, 80(19):4173–4176, 1998.
- [HSV95] R. F. Hartl, S. P. Sethi, and R. G. Vickson. A survey of the maximum principles for optimal control problems with state constraints. *SIAM Review*, 37(2):181–218, 1995.
- [IK08] K. Ito and K. Kunisch. *Lagrange multiplier approach to variational problems and applications*, volume 15 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.
- [JM00] S. C. Joshi and M. I. Miller. Landmark matching via large deformation diffeomorphisms. *IEEE Transcript Image Processing*, 9(8):1357–1370, 2000.
- [KBCL99] D. G. Kendall, D. Barden, T. K. Carne, and H. Le. *Shape and shape theory*. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester, 1999.
- [Ken84] D. G. Kendall. Shape manifolds, Procrustean metrics, and complex projective spaces. *Bull. London Math. Soc.*, 16(2):81–121, 1984.
- [KM97] A. Kriegl and P. W. Michor. *The convenient setting of global analysis*, volume 53 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1997.
- [Kur76] S. Kurcyusz. On the existence and non-existence Lagrange multipliers in Banach spaces. *J. Optimization Theory Appl.*, 20(1):81–110, 1976.
- [KW09] B. Khesin and R. Wendt. *The geometry of infinite-dimensional groups*, volume 51 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series*

- of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer-Verlag, Berlin, 2009.
- [LS95] W. Liu and H. J. Sussman. Shortest paths for sub-Riemannian metrics on rank-two distributions. *Mem. Amer. Math. Soc.*, 118(564):x+104, 1995.
- [LY95] X. J. Li and J. M. Yong. *Optimal control theory for infinite-dimensional systems*. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, 1995.
- [Mac05] K. C. H. Mackenzie. *General theory of Lie groupoids and Lie algebroids*, volume 213 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2005.
- [MM07] P. W. Michor and D. Mumford. An overview of the Riemannian metrics on spaces of curves using the Hamiltonian approach. *Appl. Comput. Harmon. Anal.*, 23(1):74–113, 2007.
- [Mon02] R. Montgomery. *A tour of subriemannian geometries, their geodesics and applications*, volume 91 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2002.
- [MP10] G. Misiólek and S. C. Preston. Fredholm properties of Riemannian exponential maps on diffeomorphism groups. *Invent. Math.*, 179(1):191–227, 2010.
- [MPAM06] F. Monroy-Pérez and A. Anzaldo-Meneses. The step-2 nilpotent $(n, n(n+1)/2)$ sub-Riemannian geometry. *J. Dyn. Control Syst.*, 12(2):185–216, 2006.
- [MR99] J. E. Marsden and T. S. Ratiu. *Introduction to mechanics and symmetry*, volume 17 of *Texts in Applied Mathematics*. Springer-Verlag, New York, second edition, 1999. A basic exposition of classical mechanical systems.
- [MTY02] M. I. Miller, A. Trouvé, and L. Younes. On the metrics and euler-lagrange equations of computational anatomy. *Annu. Rev. Biomed. Eng.*, 4:375–405, 2002.
- [MTY06] M. I. Miller, A. Trouvé, and L. Younes. Geodesic shooting for computational anatomy. *J. Math. Imaging Vision*, 24(2):209–228, 2006.
- [Omo74] H. Omori. *Infinite dimensional Lie transformation groups*. Lecture Notes in Mathematics, Vol. 427. Springer-Verlag, Berlin, 1974.
- [OR00] J. M. Ortega and W. C. Rheinboldt. *Iterative solution of nonlinear equations in several variables*, volume 30 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000. Reprint of the 1970 original.
- [PBGM64] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. *The mathematical theory of optimal processes*. Translated by D. E. Brown. A Pergamon Press Book. The Macmillan Co., New York, 1964.
- [RT05] L. Rifford and E. Trélat. Morse-Sard type results in sub-Riemannian geometry. *Math. Ann.*, 332(1):145–159, 2005.
- [Sch74] G. Schuller. On the order of convergence of certain quasi-Newton methods. *Numer. Math.*, 23:181–192, 1974.
- [Sch04] R. Schmid. Infinite dimensional Lie groups with applications to mathematical physics. *J. Geom. Symmetry Phys.*, 1:54–120, 2004.
- [Sma96] C. G. Small. *The statistical theory of shape*. Springer Series in Statistics. Springer-Verlag, New York, 1996.
- [Sus73] H. J. Sussmann. Orbits of families of vector fields and integrability of distributions. *Trans. Amer. Math. Soc.*, 180:171–188, 1973.

- [Tré08] E. Trélat. *Contrôle Optimal: Théorie et Applications*. 2008.
- [Tro95] A. Trouvé. Action de groupe de dimension infinie et reconnaissance de formes. *C. R. Acad. Sci. Paris Sér. I Math.*, 321(8):1031–1034, 1995.
- [Tro05] A. Trouvé. Diffeomorphism groups and pattern matching in image analysis. *International Journal of Computational Vision*, 37(1):17, 2005.
- [TV12] A. Trouvé and F.-X. Vialard. Shape splines and stochastic shape evolutions: a second order point of view. *Quart. Appl. Math.*, 70(2):219–251, 2012.
- [TY05] A. Trouvé and L. Younes. Local geometry of deformable templates. *SIAM J. Math. Anal.*, 37(1):17–59 (electronic), 2005.
- [TY11] A. Trouvé and L. Younes. Shape spaces. In O. Scherzer, editor, *Handbook of Mathematical Methods in Imaging*, pages 1309–1362. Springer New York, 2011.
- [VG05] M. Vaillant and J. Glaunès. Surface matching via currents. In G. Christensen and M. Sonka, editors, *Information Processing in Medical Imaging*, volume 3565 of *Lecture Notes in Computer Science*, pages 1–5. Springer Berlin, Heidelberg, 2005.
- [You10] L. Younes. *Shapes and diffeomorphisms*, volume 171 of *Applied Mathematical Sciences*. Springer-Verlag, Berlin, 2010.
- [You12] L. Younes. Constrained Diffeomorphic Shape Evolution. *Found. Comput. Math.*, 12(3):295–325, 2012.