

Hamiltonian Approach to Shape Spaces in a Diffeomorphic Framework : From the Discontinuous Image Matching Problem to a Stochastic Growth Model

François-Xavier Vialard

▶ **To cite this version:**

François-Xavier Vialard. Hamiltonian Approach to Shape Spaces in a Diffeomorphic Framework : From the Discontinuous Image Matching Problem to a Stochastic Growth Model. Mathematics [math]. École normale supérieure de Cachan - ENS Cachan, 2009. English. NNT : . tel-00400379

HAL Id: tel-00400379

https://theses.hal.science/tel-00400379

Submitted on 30 Jun 2009

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.

N° ENSC-2009

THÈSE DE DOCTORAT
DE L'ÉCOLE NORMALE SUPÉRIEURE DE CACHAN

Présentée par
Monsieur François-Xavier VIALARD

pour obtenir le grade de

DOCTEUR DE L'ÉCOLE NORMALE SUPÉRIEURE DE CACHAN

Domaine : **MATHEMATIQUES**

Sujet de la thèse :

Hamiltonian Approach to Shape Spaces in a Diffeomorphic Framework : From the Discontinuous Image Matching Problem to a Stochastic Growth Model

Rapporteurs :

Antonin Chambolle	Directeur de recherches CNRS, Ecole Polytechnique
David Mumford	Professeur, Brown University
Ofer Zeitouni	Professeur, University of Minnesota et Weizmann Institute of Science

Thèse présentée et soutenue à Cachan le 7 Mai 2009 devant le jury composé de :

Antonin Chambolle	Directeur de recherches CNRS, Ecole Polytechnique (Rapporteur)
Darryl Holm	Professeur, Imperial College
Renaud Keriven	Professeur, Ecole Nationale des Ponts et Chaussées
Yves Meyer	Professeur émérite, Ecole Normale Supérieure de Cachan
Alain Trouvé	Professeur, Ecole Normale Supérieure de Cachan, (Directeur de thèse)
Laurent Younes	Professeur, Johns Hopkins University

Centre de Mathématiques et de Leurs Applications
(ENS CACHAN/CNRS/UMR 8536)
61, avenue du Président Wilson, 94235 CACHAN CEDEX (France)"

REMERCIEMENTS

Je remercie très chaleureusement mon directeur de thèse Alain Trouvé pour m'avoir encadré durant ce travail et pour m'avoir apporté beaucoup d'un point de vue scientifique. Je tiens à souligner la grande qualité, à mes yeux, des échanges mathématiques que nous avons pu entretenir tout au long de ces trois années. Ceux-ci ont donné naissance aux principaux résultats de cette thèse dans une atmosphère libre, très propice à la créativité. D'autre part sa disponibilité et son soutien constant m'ont été indispensables pour aller au bout de ce travail.
Je remercie de la même manière Laurent Younes, qui m'acceuillit par deux fois au CIS de John Hopkins University. Ces séjours furent l'occasion de quelques discussions très fructueuses et me furent profitables sur d'autres plans que seulement scientifiques.

Antonin Chambolle, David Mumford et Ofer Zeitouni ont accepté d'être rapporteurs de ma thèse et sa forme finale leur doit beaucoup. Je les remercie pour leur relecture attentive et de leur critiques qui ont ouvert de riches discussions mathématiques. Leurs remarques ont suscité une amélioration sensible du manuscrit, notamment celles d'Antonin Chambolle sur la partie technique d'analyse fine. Je remercie également David Mumford pour ses nombreux et pertinents conseils.
Darryl Holm, Yves Meyer et Renaud Keriven ont accepté de faire partie de mon jury de thèse. Je leur en suis très reconnaissant. Leurs différentes sensibilités scientifiques représentent aussi la diversité des sujets abordés dans cette thèse. Cette diversité fut une source importante de motivation et de satisfaction personnelle.

Les chercheurs du laboratoire de l'ENS Cachan (CMLA) ont toujours été disposés à répondre à mes questions (Yves Meyer, Nicolas Vayatis, Jean-François Aujol, Jean-Michel Morel, Laurent Desvillettes...) et je les remercie pour cette disponibilité. Plus généralement, je remercie tous les mathématiciens qui m'ont reçu pour discuter ou qui ont répondu par mails à mes questions. Le CMLA m'a donné l'occasion de rencontrer et d'avoir le plaisir de collaborer avec Filippo Santambrogio. Nous avons pu ensemble établir les résultats de la partie d'analyse fine dans le cadre mathématique adéquat. Je le remercie de l'intérêt qu'il a porté à mes questions mathématiques.
Bernard Chalmond m'a permis de terminer mon travail de thèse dans de très bonne conditions, je lui en suis très redevable. Le CMLA a aussi été un cadre agréable de travail et je pense notamment aux trois secrétaires Véronique, Micheline et Virginie.

Durant ces trois dernières années, les conférences annuelles organisées par Laurent Younes, David Mumford et autres ont été un élément important de motivation et d'intégration dans ce cercle scientifique. Je remercie très chaleureusement Darryl Holm de m'offrir la possibilité de continuer mon travail à l'Imperial College au sein d'un projet scientifique très ambitieux et excitant.

Je remercie Jean-François Legall sans lequel je n'aurais peut-être pas pu m'engager dans ce travail de recherche. Sur un autre plan, je remercie aussi Claude Martini, Pierre Cohort, José Da Fonseca, Frédéric Patras et tous les autres sans qui ma période de thèse n'aurait pas été aussi intéressante.

Mes derniers remerciements vont à ma famille et très spécialement à ma mère, qui a toujours tenu à ce que j'aille au bout de ma thèse. Je lui dédie ce travail. Enfin, je remercie très affectueusement Florence qui m'a soutenu depuis le début de ma thèse et sans qui cette période de thèse aurait été infiniment plus terne.

Résumé

Ce travail de thèse se situe dans le contexte de l'appariement d'images par difféomorphismes qui a été récemment développé dans le but d'applications à l'anatomie computationnelle et l'imagerie médicale. D'un point de vue mathématique, on utilise l'action de groupe de difféomorphismes de l'espace euclidien pour décrire la variabilité des formes biologiques.
Le cas des images discontinues n'était compris que partiellement. La première contribution de ce travail est de traiter complètement le cas des images discontinues en considérant comme modèle d'image discontinues l'espace des fonctions à variations bornées. On apporte des outils techniques pour traiter les discontinuités dans le cadre d'appariemment par difféomorphismes. Ces résultats sont appliqués à la formulation Hamiltonienne des géodésiques dans le cadre d'un nouveau modèle qui incorpore l'action d'un difféomorphisme sur les niveaux de grille de l'image pour prendre en compte un changement d'intensité. La seconde application permet d'étendre la théorie des métamorphoses développée par A.Trouvé et L.Younes aux fonctions discontinues. Il apparait que la géométrie de ces espaces est plus compliquée que pour des fonctions lisses.

La seconde partie de cette thèse aborde des aspects plus probabilistes du domaine. On étudie une perturbation stochastique du système Hamiltonien pour le cas de particules (ou landmarks). D'un point de vue physique, on peut interpréter cette perturbation comme des forces aléatoires agissant sur les particules. Il est donc naturel de considérer ce modèle comme un premier modèle de croissance de forme ou au moins d'évolutions aléatoires de formes.
On montre que les solutions n'explosent pas en temps fini presque sûrement et on étend ce modèle stochastique en dimension infinie sur un espace de Hilbert bien choisi (en quelque sorte un espace de Besov ou Sobolev sur une base de Haar). En dimension infinie la propriété précédente reste vraie et on obtient un important (aussi d'un point de vue numérique) résultat de convergence du cas des particules vers le cas de dimension infinie. Le cadre ainsi développé est suffisamment général pour être adaptable dans de nombreuses situations de modélisation.

Summary

This thesis takes place in the context of image matching within the framework of large deformation diffeomorphisms. With important applications to medical imaging and computational anatomy, this approach uses the action of diffeomorphisms groups in order to classify images. One of the first issue to deal with is to compute the distance between objects on which can act the group of diffeomorphisms.
The case of discontinuous images was very partially understood. The first part of this thesis is devoted to fully tackle the case of discontinuous images in any dimension. Namely the images are assumed to be functions of bounded variations. We have provided technical tools to deal with discontinuous images within the diffeomorphism framework. The first application developed is a Hamiltonian formulation of the geodesic equations for a new model including a change of contrast in the images which is represented by an action of a diffeomorphism on the values of the level lines of the image. The second one is an extension of the metamorphosis framework developed by A.Trouvé and L.Younes to SBV functions, which points out that the geometry of such spaces is much more complicated than the one with smooth functions.

The second part of the thesis takes place in the probabilistic side of the field. Taking advantage of the Hamiltonian formulation, we aim to study stochastic perturbations of the geodesic equations. From a physical point of view, the perturbation we consider

affects the forces on the particles and not the speed of the particles. In some sense, this model could be an interesting dynamical model for growth of shapes or at least random evolution of shapes.
We have proven that the solutions of the system in the case of landmarks do not blow up and that there exists a SDE in infinite dimension on a suitable Hilbert space (actually a sort of Besov or Sobolev space on Haar basis) which extends the landmark case. In infinite dimension, the solutions are also defined for all times. Moreover an important convergence result of the landmark case to the infinite dimensional case is proven. Finally, let us precise that the structure of the noise is general enough to account for correlation between points of the curve in the noise.

Contents

1 Introduction **7**
- 1.1 The space of deformations . . . 11
  - 1.1.1 The group of diffeomorphisms . . . 11
  - 1.1.2 Admissible vector fields spaces . . . 14
- 1.2 Geodesic equations on diffeomorphism groups . . . 17
  - 1.2.1 EPDiff . . . 17
  - 1.2.2 Matching problems . . . 19
- 1.3 The Hamiltonian formulation . . . 24
  - 1.3.1 The deterministic case . . . 24
  - 1.3.2 A Stochastic perturbation of the Hamiltonian system . . . 27

2 Discontinuities and matching **29**
- 2.1 Introduction . . . 29
- 2.2 A simple approach: Lipschitz piecewise functions . . . 30
- 2.3 The case of SBV and BV functions . . . 38
  - 2.3.1 Statement of the results and notations . . . 39
  - 2.3.2 Density lemmas . . . 42
  - 2.3.3 SBV and continuous functions in any dimension . . . 46
- 2.4 The one-dimensional case . . . 51
- 2.5 Reduction to the one-dimensional case . . . 52
  - 2.5.1 Extension to other penalty terms . . . 55
- 2.6 Toward the application to geodesic equations . . . 57

3 Geodesic equations **58**
- 3.1 Introduction . . . 58
- 3.2 A new model . . . 59
  - 3.2.1 Geodesic equations . . . 59
  - 3.2.2 Geodesic flow . . . 64
  - 3.2.3 Comments and conclusion . . . 69
- 3.3 Applications to metamorphosis . . . 70
  - 3.3.1 Introduction to metamorphosis . . . 70
  - 3.3.2 Geodesic equations and existence result . . . 72
- 3.4 Simple examples . . . 75

4 Hamiltonian formulation **80**
- 4.1 Heuristic derivation of the Hamiltonian formulation . . . 80
- 4.2 The case of discontinuous images . . . 83
  - 4.2.1 Weak formulation . . . 84
  - 4.2.2 Uniqueness of the weak solutions . . . 86
  - 4.2.3 On the existence of weak solutions . . . 88
- 4.3 Conclusion . . . 89

5 Stochastic growth model **90**
5.1 Introduction . . . 90
5.2 The landmark case . . . 93
5.3 On the limit of the stochastic model . . . 95
5.3.1 On the deterministic case of curves in the plane . . . 96
5.3.2 Toward the stochastic extension . . . 96
5.4 The spaces P and Q . . . 98
5.5 Cylindrical Brownian motion and stochastic integral . . . 109
5.5.1 Cylindrical Brownian motion in $L^2(S^1, \mathbb{R})$ . . . 110
5.5.2 Cylindrical Brownian motion in $L^2(T_m, \mathbb{R})$ . . . 111
5.5.3 Cylindrical Brownian motion in $H = L^2(T_m, \mathbb{R}^d)$ . . . 112
5.5.4 Stochastic integral . . . 112
5.6 Solutions to the SDE on $P \times Q$ . . . 113
5.6.1 A trace Lemma . . . 117
5.7 Approximations Lemmas . . . 118
5.8 Applications and numerical simulations . . . 121
5.9 Possible extensions and conclusion . . . 127

6 Conclusion **130**
6.1 On the deterministic part . . . 132
6.2 On the stochastic growth model . . . 135
6.2.1 Spline approach on shape spaces . . . 135
6.2.2 The stochastic model . . . 136
6.2.3 Statistical issues . . . 138

A BV and SBV functions **140**

B Other Lemmas and additional remarks **144**
B.1 Hilbert spaces F_s and H_s . . . 144
B.2 Some additional Lemmas . . . 145

Chapter 1

Introduction

Contents

1.1 The space of deformations 11
1.2 Geodesic equations on diffeomorphism groups 17
1.3 The Hamiltonian formulation 24

Motivated by applications to computational anatomy and computer vision, the study of objects with a visual structure (shapes, images, landmarks, ...) under the action of a transformation group has been actively developed over the past several years. The aim of this endeavor is to classify data from their representation and to recognize features on these data. From the medical imaging point of view, it means for instance automatic detection of pathologies through the use of statistical tests. As the collection of images of organs is rapidly increasing, it opens the opportunity to provide practitioners with tools to help making of diagnosis. In the case of registration (i.e. matching) between two medical images, the bijection condition is relevant for the practitioners: the shape of the organ should be in some sense a one to one deformation of a template shape. This condition necessarily implies a not so obvious model of the deformation and may require some mathematical foundations.

In more mathematical words, one general approach is to act on the objects via a group action (the mathematical notion of action). At this level of generality, the underlying idea is to put the modeling effort on the group action and not on the object themselves. For example, in the case of 3D medical images, the group could be a diffeomorphism group (group of bijective smooth applications) of the ambient space. In image analysis, this framework was first introduced by the work of U. Grenander ([GM98],[Gre93]), and the field known as the large deformation diffeomorphic matching approach has been widely studied since this first preliminary work especially by A.Trouvé, L.Younes and M.Miller ([Tro95], [MTY06],[MY01]). The question of building a distance on space of shapes had been a well addressed subject in these works, by using the projection of right invariant distance of the group on the shape space. In fact the distance needs to be strong enough in order to build a useful Riemannian framework ([MM05b]). Answers to this question and the development of sound foundations from a mathematical point of view

have been increasing since the late 90's and it involved ideas from fluid flow motion and hydrodynamical notions ([Arn78]). In order to perform a statistical analysis on a group of observed shapes, at least the distance computed on any couple in this observed set of shapes should be meaningful. To go further, this model of shape representation should also encode local informations of data and is undoubtedly a rich framework to use statistical tools. Yet, a consistent statistical approach is not as well established as the geometrical aspects of this field, although the seminal ideas of large deformation by diffeomorphisms were already motivated by this question. One attempt to address this problem is carefully developed in the thesis [All07] in the case of small deformations. In the case of landmarks, a well-known approach was initiated by the work of D.G. Kendall known as 'Statistical Shape Analysis' ([KBCL99],[DM98]). D.G. Kendall also tackled the problem of modeling shape diffusion ([Ken77]) in the case of landmarks and his approach is currently developed ([BDG07]). In their work the trajectories of the landmarks follows stochastic processes such as Brownian motion or Ornstein-Uhlenbeck process, yet not constrained by a diffeomorphic condition.
The *large deformation diffeomorphisms* approach has seen in the past five years tremendous developments as for the biomedical applications, [GQMY08, WBR$^+$07, DPTA08] to name a few. And some efficient algorithms have been developed using the different structures of the geodesic equations such as [MM08, Cot08]. On the statistics aspects, very recent works have been carried out successfully in [MMTY08, AKTA06, AKT07]. Our work, focused on the mathematical foundations, has three main contributions for the geometric, analytic and probabilistic aspects of the field.

Brief outline of the contributions of our thesis:

Motivated by the derivation of the Hamiltonian equations of image matching in the *large deformation diffeomorphisms* framework, we needed to work out the structure of the momentum (arising in the EPDiff equation) which was given by the differentiation of a functional (the penalty term). It was not known how to differentiate the functional in the case of discontinuous images in gray-level (i.e. discontinuous functions).

- From a variational point of view, we determined the effect of discontinuities in the matching of images in the large diffeomorphisms setting. We solved with the Semi-Differentiation Theorem 15 the problem of the differentiation of the functional in the case of BV functions in any dimension for a large range of penalty terms.

We do not expect this theorem to be extended much further. However, our approach based on density arguments and decomposition methods raised interesting questions about decomposition of BV functions. For example, the generalization of the decomposition of a BV function into a sum of a continuous function and a SBV function would lead to a more general Semi-Differentiation Theorem. Once the differentiation is performed, the

structure of the momentum is known and as expected it has a singular part due to the jump set of the template image. This structure enables to derive the Hamiltonian equation in this case.

- We introduced a new model in Section 3.2 which takes into account a group action on the gray-level of the image, while keeping an action by diffeomorphisms on the support of the image. The geodesic equations for this model are given in theorem 20 on the space of SBV functions . We derived in Chapter 4 weak Hamiltonian equations for which we proved existence and uniqueness results.

This first model opens the way for more realistic models and for the development of a multilayer approach, which will be pursued elsewhere. Another consequence of our semi-differentiation result is the following application to the metamorphosis framework introduced in [TY05].

- We worked out the geodesic equations in theorem 22 for metamorphosis in the case of discontinuous images, namely SBV. On the way, we showed in Subsection 3.4 that this extension implies complications in the geometry of this infinite dimensional Riemannian manifold.

On the probabilistic side, we took advantage of the Hamiltonian formulation of the equations to propose a second order stochastic differential equation (SDE) to model time evolutions of shapes. This approach is very natural in order to turn deterministic evolution equations into SDE and this approach gives a simple model which deserves to be studied.

- We introduced a stochastic growth model for shapes within the *large deformation diffeomorphisms* framework. Having in mind statistical analysis of time dependent biological data, this model would be an interesting framework to apply. For example, we can hope to design quantitative methods for early detection of abnormal growth of an organ as well as to obtain probabilistic atlases of biological growths of organs.

From a mathematical point of view, we first worked out this model in the finite dimensional case of landmarks and then we extended it to shape spaces by finding a proper Hilbert space in which we studied the equations.

- To study the equations in infinite dimension, we introduced a new Hilbert space of functions F_s in Section 5.4 very close to a Sobolev space on a Haar basis. This space (available in any dimension) enjoys important properties (17,19,20,25,27,28) that are easy to prove and are really needed in our approach. These include stability by product as well as stability by composition with smooth functions. The Hilbert space also contains smooth functions which can represent shapes and its dual contains almost surely a cylindrical Brownian motion on L^2. Moreover, the Hilbert space admits numerical

approximations on finite dimensional spaces that can be reached by projecting the SDE on landmarks spaces and we proved in Proposition 33 that the solutions converge to the solution of the SDE on $F_s \times F_{-s}$.

It should be emphasized that the SDE can have a very general variance term and the framework can be extended to deal with other sources of noise. In what follows, we present in more details the organization of the thesis.

Plan of the thesis:

The first part of this thesis is devoted to the understanding of the matching equations between two images in gray level which can be discontinuous (namely SBV). It requires some technical work on the space of bounded variation functions and the Chapter 2 is devoted to the presentation of these results. The applications of this work are detailed in the Chapter 3 for two different cases:

- The first one details a new model for the matching between two images introducing a contrast transformation (a diffeomorphism on the co-domain of the image), and a geometric transformation (a diffeomorphism on the image domain) is as usual taken into account.
- The second application deals with metamorphosis and the aim is to enlight the results of the article on metamorphosis [TY05] with these new analysis tools.

For each case, we derive the geodesic equations from a variational approach and we detail the structure of the initial momentum. The last part of Chapter 3 gives some toy examples for the two models and in particular, we show that the introduction of discontinuities in the metamorphosis set-up bring some additional complications: some interesting results on geometrical aspects of [TY05] are false in this context.
The Chapter 4 gives a weak Hamiltonian formulation of the matching equations in the case of piecewise Lipschitz functions derived in Chapter 3. This formulation is not only important from the numerical point of view but it will also be the foundation for the stochastic model of growth we present in Chapter 5. The Hamiltonian formulation gives also some perspectives on how to introduce a multilayer approach within the large deformation by diffeomorphisms framework.
In the Chapter 5 of this thesis, we study a stochastic perturbation of the Hamiltonian equations in the case of landmarks by introducing a random force. We then extend it to a shape diffusion model. This model (though an a priori model) could be applied as biological growth model in order to make early diagnosis.

After a brief introduction to the mathematical background of the field, we will give a comprehensive introduction to our work and our different approaches.

1.1 The space of deformations

This section gives a short presentation of the diffeomorphism group we will consider. It is based on the corresponding Chapter in [You08], which gives detailed and comprehensive proofs. The first part presents the construction of the group of diffeomorphisms from a space of vector fields and the second part gives an overview of the Reproducing Kernel Hilbert Spaces (RKHS) in order to construct suitable vector fields.

1.1.1 The group of diffeomorphisms

Let $p \geq 1$ a positive integer, U be an open set of $\mathbb{R}^d$ and $V \subset \mathcal{X}_0^p(U, \mathbb{R}^d)$ the space of C^p vector fields vanishing on the boundary of U and at infinity with V a Hilbert space such that there exists a constant C_V for which,

$$|v|_{p,\infty} \leq C_V |v|_V \,. \tag{1.1}$$

The RKHS V will be said admissible if $p \geq 1$ and if we want higher regularity we will write $p-$admissible. From the variational point of view, we will essentially need the Hilbert space $L^2([0,1], V)$ but we introduce also $L^1([0,1], V)$. With the Cauchy-Schwarz inequality, $|v|_{L^1} \leq |v|_{L^2}$. As a consequence, all the controls with the L^1 norm will be valid for the L^2 norm. All these inequalities are also true if the $L^p([0,1], V)$ norm is replaced by the $L^p([0,1], \chi_0^p)$ norm. We now present a theorem which is a summary of the results which can be found in [You08] (chapter 12).

Theorem 1. *If $v \in L^1([0,1], V)$ then its flow is defined: if $t \in [0,1]$*

$$\begin{cases} \partial_t \phi_{0,t}^v(x) = v_t(\phi_{0,t}(x)) \\ \phi_{0,0}^v(x) = \mathrm{Id} \,. \end{cases} \tag{1.2}$$

For all time $s, t \in [0,1]$, $\phi_{s,t}^v := \phi_{0,t}^v \circ [\phi_{0,s}^v]^{-1}$ is a C^p diffeomorphism.
The application

$$(t, x) \in [0,1]^2 \times U \mapsto \phi_{0,t}^v(x) \in U \tag{1.3}$$

is Lipschitz in both variables, more precisely:

$$\begin{aligned} &\mathrm{Lip}(\phi_{0,t}^v) \leq e^{c_V |v|_{L^1}} \,, \\ &|\phi_{0,t}^v - \phi_{0,s}^v|_\infty \leq \int_s^t c_V |v_r|_V \, dr \,. \end{aligned} \tag{1.4}$$

The derivatives of $\phi_{s,t}^v$ verify the integral equation for $k \leq p$:

$$d_x^k \phi_{s,t}^v(x) = \mathrm{Id} + \int_s^t d^k [v_r \circ \phi_{s,t}^v(x)] \, dr \,, \tag{1.5}$$

and there exists constants C, C' independent on v such that for all $s, t \in [0,1]$,

$$|\phi_{s,t}^v|_{p,\infty} \leq C e^{C' c_V |v|_{L^1}} \,. \tag{1.6}$$

These results are different from the classical ODE result for vector fields since the control on the vector field is done via the L^1 norm. The main tool of the proof is the Gronwall lemma. Another important point is the dependence on the vector field v. Especially, we will need the weak convergence property to prove the existence of a deformation of minimal cost.

Theorem 2. *Let $u, v \in L^1([0,1], V)$ be two time dependent vector fields, then*

$$|\phi^u_{0,t} - \phi^v_{0,t}|_\infty \leq c_V |u - v|_{L^1} e^{c_V |v|_{L^1}} . \tag{1.7}$$

If $p \geq 2$, a bound on the derivative is given by

$$|d\phi^u_{0,t} - d\phi^v_{0,t}|_\infty \leq C'' |u - v|_{L^1} e^{c_V |u|_{L^1}} , \tag{1.8}$$

with C'' depending continuously on $|u|_{L^1}$ and $|v|_{L^1}$.

As the proof of inequality (1.8) is not explicitly given in [You08] or [Gla05], we give the details in appendix, proposition 44. Finally, we present the fundamental result which enables to prove the existence of a minimizer for a large range of functionals. This result states the uniform convergence of the flow on every compact under the weak convergence of the vector fields. We also add a uniform convergence of the derivatives for the strong convergence.

Theorem 3. *If $v_n \rightarrow_{L^2([0,1],V)} v$, then $d^k \phi^{v_n}_{t,s} \rightarrow d^k \phi^v_{t,s}$ for the uniform convergence on every compact uniformly in $t, s \in [0,1]$ for $k \leq p$.*
If $u_n \rightharpoonup_{L^2([0,1],V)} u$ (weak convergence), then $\phi^{u_n}_{t,s} \rightarrow \phi^u_{t,s}$ for the uniform convergence on every compact uniformly in $t, s \in [0,1]$. If $K \subset \mathbb{R}^d$ a compact set, then

$$\sup_{(t,s,x) \in [0,1]^2 \times K} |\phi^{u_n}_{t,s}(x) - \phi^u_{t,s}(x)| \rightarrow_{n \mapsto +\infty} 0 .$$

More than this result can be found in [Gla05] since if $p \geq 2$ the convergence of the first derivative is proved to be uniform on every compact set. (The weak convergence on $L^2([0,1], V)$ can be replaced by the weak convergence on $L^1([0,1], V)$ as it is stated in [You08].)

The set of diffeomorphisms G_V is defined as

$$G_V := \{\phi^v_{0,1} \,|\, v \in L^2([0,1], V)\} . \tag{1.9}$$

This set is naturally a group (the inverse of the diffeomorphism generated by v_t can be found as the flow of v_{1-t} and a time rescaling makes the product well defined). A right invariant distance can be defined on G_V:

Theorem 4. *Introducing $d(\psi_1, \psi_2) := \inf\{|v|_{L^2} \,|\, \psi_1 = \phi^v_{0,1} \circ \psi_2\}$, (G_V, d) is a complete metric space and d is a right invariant distance*

Note that replacing $L^2([0,1],V)$ by $L^1([0,1],V)$ leads to the same group G_V (reparameterization argument) and the distance (replacing the L^2 norm by the L^1 norm) is the same. The important point is that this distance is right invariant: for any $\phi \in G_V$,

$$d(\psi_1 \circ \phi, \psi_2 \circ \phi) = d(\psi_1, \psi_2)\,.$$

This construction enables to obtain the existence of a minimizing vector field v between two diffeomorphisms in G_V. The main argument of the proof of the theorem below is the weak compactness of bounded balls in Hilbert spaces and the use of Theorem 3. This is the key framework for the *large deformation diffeomorphisms* approach.

Theorem 5. *Let $\psi_1, \psi_2 \in G_V$ be two diffeomorphisms, then there exists $v \in L^2([0,1],V)$ such that $\psi_1 = \psi_2 \circ \phi^v_{0,1}$, and $|v|_{L^2} = d(\psi_1,\psi_2)$.*

Such a minimizer is also a minimizer towards a reparameterization of the time variable and it shows that the quantity v_t is constant in time.
The last point to present is the differentiation of the flow with respect to a variation of the vector field. Let us define the adjoint for ϕ a diffeomorphism of $\mathbb{R}^n$ and v a vector field on $\mathbb{R}^n$ by $Ad_\phi v = (d\phi\, v) \circ \phi^{-1}$.

Theorem 6. *Let $(u_t, v_t) \in L^2([0,1],V)$ be two time dependent vector fields, and denote by $\phi^\varepsilon_{0,t}$ the flow generated by the vector field $u_t + \varepsilon v_t$, then we have:*

$$\partial_{\varepsilon|_{\varepsilon=0}} \phi^\varepsilon_{0,1}(x) = \int_0^1 [d\phi_{t,1}]_{\phi_{0,t}(x)} v_t(\phi_{0,t}(x))dt = d\phi_{0,1}(\int_0^1 Ad_{\phi_{t,0}}(v_t)\,dt)\,.$$

We can rewrite the formula to derive the expression used for the geodesic equations with the equation

$$\partial_{\varepsilon|_{\varepsilon=0}} \phi_{0,1} \circ (\phi^\varepsilon_{0,1})^{-1} = -\int_0^1 Ad_{\phi_{t,1}}(v_t)\,dt\,. \tag{1.10}$$

We give here a short proof of the perturbation of the flow of a time dependent vector field with respect to the vector field. We assume in the proposition that the vector fields are smooth but it can be proven with weaker assumptions on the regularity of vector fields. (See [You08], for a detailed proof.)

Proof. Let us introduce the notation $A_t \in \mathbb{R}^n$ defined by: $A_t(\phi_{0,t}(x)) = \partial_{\varepsilon|_{\varepsilon=0}} \phi^\varepsilon_{0,t}(x)$. Differentiating this expression with respect to the time variable:

$$\frac{d}{dt} A_t(\phi_{0,t}(x)) = du_t(A_t(\phi_{0,t}(x))) + v_t(\phi_{0,t}(x)),$$

Remark that the expression above can be written as, with $\mathcal{L}$ the Lie derivative(for a definition of the Lie derivative, the reader can refer to [Sak96]):

$$\mathcal{L}_{u_t} A_t = \frac{d}{du}_{|u=0} [d\phi_{0,t+u}]^{-1}(A_{t+u}(\phi_{t+u}(x))) = [d\phi_{0,t}]^{-1}(v_t(\phi_{0,t}(x))).$$

By integration in time, we obtain the result. □

This construction is rather general and it is strongly dependent on the choice of the admissible space of vector fields V. The next part is devoted to the construction of such admissible spaces.

1.1.2 Admissible vector fields spaces

We have required V to be a Hilbert space of C^1 vector fields with the admissibility condition $|v|_{1,\infty} \leq |v|_V$. It is a particular case from the general situation when V is a Hilbert space of functions defined on a domain U for which the pointwise evaluation is a bounded operator: if $x \in U$, there exists a constant M such that defining

$$\delta_x : v \in V \mapsto v(x) \in \mathbb{R}^d \,, \tag{1.11}$$

we have $|v(x)| \leq M\,|v|_V$.
The vector valued case does not differ from the scalar case for most of the important points of this theory. Such a space is called a Reproducing Kernel Hilbert Space. As we do not need further generalizations of RKHS, we restrict ourselves to Hilbert spaces of functions for which the evaluation at any point is a bounded linear operator.
The canonical isomorphism between V and its dual V^* will be denoted:

$$K : V^* \mapsto V \,, \tag{1.12}$$

and $L : V \mapsto V^*$ its inverse.
Hence the Riesz theorem, for each $x \in U$ and $p \in T_x^* = \mathbb{R}^d$ a vector (T_x is the tangent space at point x and T_x^* is the cotangent space at point x. Of course in this euclidean case, all the (co)tangent spaces at a point can be identified to $\mathbb{R}^d$), there exists an element of V denoted $k(.,x)p$ such that

$$\langle k(.,x)p, v\rangle_V = \langle v(x), p\rangle \,. \tag{1.13}$$

An equivalent notation will be $K\delta_x^p$ with δ_x^p the linear form involved on the right hand side of (1.13). Then it defines a function of two variables k called the kernel:

$$\begin{aligned} &k : (y,x) \in U \times U \mapsto k(y,x) \in L(T_x^*, T_y) \,, \\ &k(y,x) : p \in T_x^* = k(y,x)p \in T_y \,. \end{aligned}$$

The first properties are:

- $k(x,y) = k(y,x)^*$
- k is a positive matrix in the following sense: developing the norm for $(x_i)_{i\in[1,n]}$ being n distinct points and $(p_i)_{i\in[1,n]}$ being n vectors in $\mathbb{R}^d$,

$$|\sum_{i=1}^{n} k(.,x_i)p_i|_V^2 \geq 0 \;\Leftrightarrow\; \sum_{i,j} p_j k(x_j,x_i)p_i \geq 0 \,. \tag{1.14}$$

An important theorem (Moore-Aronszajn) shows the correspondence between RKHS and positive matrices.

Theorem 7. *For any positive matrix k on $U \times U$ there exists a uniquely determined Hilbert space V of functions defined on U with the reproducing kernel k.*

The proof of Theorem 7 can be found in [Aro50]. If $(v_i)_{i\in\mathbb{N}}$ is an total orthonormal basis of the Hilbert space V, then we have using the definition and the decomposition $k(y,x) = \sum_{i\in\mathbb{N}} v_i(y).v_i(x)$ with the notation

$$v(y).v(x) : p \in T_x^* \mapsto \langle v(x), p\rangle v(y) \in T_y\,.$$

Before giving some examples, we present some regularity results we will need in Chapters 3 and 5: what is the relation between the regularity of the kernel and the regularity of the injection of V? For the continuity, we have the equivalence

Theorem 8. *Let V a RKHS of vector fields on U, then there is an equivalence between*

- $V \subset C^0(U, \mathbb{R}^d)$,
- *k is continuous on $U \times U$.*

Proof. If k is continuous then we apply the Cauchy-Schwarz inequality with $f \in V$, $x, y \in U$ and $p \in \mathbb{R}^d$,

$$\begin{aligned}|\langle f(x), p\rangle - \langle f(y), p\rangle| &= |\langle k(.,x)p - k(.,y)p, f\rangle_V| \\ &\leq |pk(x,x)p + pk(y,y) - 2pk(x,y)p|^{\frac{1}{2}}|f|_V\,, \quad (1.15)\end{aligned}$$

to obtain that f is continuous.
Assume that $V \subset C^0(U, \mathbb{R}^d)$, then $k(.,x)p$ are continuous functions. Using an orthonormal basis $(v_i)_{i\in\mathbb{N}}$, we have

$$k(x,x) = \sum_{i\in\mathbb{N}} v_i(x).v_i(x), \quad (1.16)$$

with v_i continuous functions. With the equation (1.16), $pk(x,x)p$ is the limit of an increasing sequence of continuous functions. By Dini's theorem, the convergence is uniform and the limit is continuous. Hence, k is continuous on the diagonal of $U \times U$. Now using inequality (1.15), applied to $f = k(.,y)p$ we can conclude by triangular inequality:

$$|k(x,y) - k(x',y')| \leq |k(x,y) - k(x',y)| + |k(x',y) - k(x',y')|\,,$$

Using inequality (1.15), the result is obtained. □

We showed this simple proof in order to get the reader used to this tool. For higher regularity results, we have the following assertion:

Theorem 9. *If k has continuous derivatives,*

$$\frac{\partial^{m+n}k(x,y)}{\partial_m x \partial_n y}\, m, n \in [1, r]\,,$$

then $V \subset C^r(U, \mathbb{R}^d)$.

A proof can be found in the book [Sai97].
Before showing some examples and ways to build such kernels, we define positive definite kernels: the case of equality in (1.14) implies that $\sum_{i\in[1,n]} k(.,x_i)p_i = 0$, i.e. for all $v \in V$, $\sum_{i\in[1,n]}\langle v(x_i), p_i\rangle = 0$. From now on, we will assume that this can not be true unless $p_i = 0\,\forall i \in [1,n]$. We say that the kernel is positive definite. Roughly speaking, this assumption means somehow that the space V is large enough (for example, V can not be of finite dimension).
The important consequence is that if $\mathbf{x} := (x_1,\ldots,x_n)$ for n distinct points and $\mathbf{p} = (p_1,\ldots,p_n)$, the operator $k(\mathbf{x}) : T^*_{x_1}\times\ldots\times T^*_{x_n} \mapsto T_{x_1}\times\ldots\times T_{x_n}$ defined by $k(\mathbf{x}) = [k(x_i,\mathbf{x})\mathbf{p}]_{i\in[1,n]}$ is an isomorphism.

Proposition 1. *Let k be a positive definite kernel. If $\mathbf{x}$ is a group of n distinct points in $\mathbb{R}^d$ and $\mathbf{u} \in [\mathbb{R}^d]^n$ a vector, there exists a unique vector field $v \in V$ such that*

$$v(x_i) = u_i\,\forall i \in [1,n], \tag{1.17}$$

and which minimizes the norm on V among the vector fields in V verifying (1.17)*. Moreover,*

$$v(.) = \sum_{i\in[0,n]} k(.,x_i)p_i,$$

such that $k(\mathbf{x})\mathbf{p} = \mathbf{u}$.

Proof. First remark that the affine subspace W of vector fields verifying (1.17) is nonempty since the kernel is positive definite (see the remark above). If u verifies (1.17), then

$$W = u + W_0\,,$$

with $W_0 \doteq \mathrm{Vect}\{w \,|\, \langle k(.,x_i)p_i, w\rangle = 0\,; p_i \in \mathbb{R}^d\,,\ i \in [1,n]\}$. If u is a minimizer of the interpolation problem, then u is orthogonal to W_0.
As a consequence,

$$u \in \mathrm{Vect}\{k(.,x_i)p_i \,|\, p_i \in \mathbb{R}^d\,,\ i \in [1,n]\}$$

which is the orthogonal of W_0.
Now, the minimizer is unique since $k(\mathbf{x})$ is an isomorphism. □

The first example of RKHS are Sobolev spaces $H^m(U,\mathbb{R}^d)$ with m large enough to use Morrey's theorem. As we will deal with kernels in the geodesic equations, it is convenient to introduce RKHS via their kernel. On $\mathbb{R}^d$ kernels that are invariant by translation can be written as $k(x-y)$. If the dimension $d = 1$, a large range of positive kernels are known: let μ a positive finite Borel measure on $\mathbb{R}$ and let k be its Fourier transform defined by

$$k(x) = \int_{\mathbb{R}} e^{itx} d\mu(x)\,,$$

then $k(x,y) = k(x-y)$ is a positive kernel.
Bochner's theorem states the converse: every positive kernel is the Fourier transform

of a positive finite Borel measure. Now, to obtain positive kernel in dimension d, the simplest way is to introduce $k(|x-y|)\mathrm{Id}$ as candidate. The condition on k to provide a positive kernel in any dimension d is given by the Schoenberg theorem in complement to Bochner's theorem:

Theorem 10 (Schoenberg). *Suppose that $f : \mathbb{R}_+ \mapsto \mathbb{R}_+$ is a continuous function, then the two following assertions are equivalent:*

- *the function $x \in \mathbb{R}^n \mapsto f(|x|)$ is a positive kernel,*
- *the function $t \in \mathbb{R}_+ \mapsto f(\sqrt{t})$ is the Laplace transform of a positive finite Borel measure on $\mathbb{R}_+$.*

A kernel that verifies this assumption is the well-known Gaussian kernel $k(x,y) = e^{-\frac{|x-y|^2}{\sigma^2}}$ with σ a scaling positive parameter. Another smooth kernel is the Cauchy kernel $k(x,y) = \frac{1}{1+\frac{|x-y|^2}{\sigma^2}}$. The exponential kernel $k(x,y) = e^{-\frac{|x-y|}{\sigma}}$ is another choice but singular at point 0. These three kernels are positive definite. The smoothness of the kernel will be an important point in the Chapter 3 since it will ensure the existence in all time of the geodesic flow. It is not the case for the exponential kernel, because of its singularity at the origin (the Riemannian manifold of two landmarks in one dimension is not complete).
More detailed discussions on RKHS are developed in [Sai97] and [Aro50]. To focus on vector fields RKHS, further issues (especially the case of kernels which are affine at infinity) are presented in [You08] and [Gla05].

1.2 Geodesic equations on diffeomorphism groups

With a right-invariant metric, many of the finite dimensional properties of the geodesics can be formally extended to infinite dimensional setting as emphasized in [Arn78]. We chose to present a formal derivation of the Euler equations, the Euler geodesic equation and the conservation of momentum (see (1.21) and (1.22)). Then we will properly derive the geodesic equations in two matching problems: the landmark case and the case of smooth images.

1.2.1 EPDiff

A comprehensive presentation of Euler-Poincaré equations is given in the books [Hol08b, Hol08a]. Once the group of transformations is well defined, one can deal with the minimization of functionals of this type:

$$D(\mathrm{Id}, \phi)^2 + H(\phi)\,, \tag{1.18}$$

with H a penalty term. We do not discuss conditions on the penalty term, yet in Chapter 3 continuity w.r.t. the pointwise convergence will be sufficient. A variational approach

to minimize (1.18) leads to a geodesic on the group of diffeomorphisms G_V. Since the metric is right invariant, the shortest path in $L^2([0,1],V)$ satisfies the Euler equations on Lie group of infinite dimension. The key point is the conservation of the momentum: the initial momentum is transported through the action of the flow generated by the geodesic. In this section, we derive formally the geodesic equations on the diffeomorphism group from the point of view of Euler-Poincaré principle (see [Hol08b]). This is essentially based on [MTY06] and [Viz08]. Formally means that we do not discuss the well-posedness of the involved quantities.

In our case the Lie algebra is V, an admissible space of vector fields. Assume that $v \in L^2([0,1],V)$ is a minimizer of $\int_0^1 |u_t|_V^2\, dt$ (the existence was discussed briefly in Subsection 1.1.1) with the end-point condition $\phi_{0,1} = \psi$. Let us introduce $\phi^\varepsilon_{0,t}$ a smooth variation of the flow with ε a real parameter.

We write,

$$\partial_t \phi^\varepsilon_{0,t} = v^\varepsilon_t \circ \phi^\varepsilon_{0,t}\,,$$
$$\partial_\varepsilon \phi^\varepsilon_{0,t} = w^\varepsilon_t \circ \phi^\varepsilon_{0,t}\,.$$

Using the Schwarz identity,

$$\partial_{\varepsilon,t}\phi^\varepsilon_{0,t} = \partial_{t,\varepsilon}\phi^\varepsilon_{0,t}\,,$$

we have

$$\partial_\varepsilon v^\varepsilon_t = \partial_t w^\varepsilon_t + dw^\varepsilon_t(v^\varepsilon_t) - dv^\varepsilon_t(w^\varepsilon_t).$$

With the notation for u, v two smooth vector fields,

$$\mathrm{ad}_u(v) = [u,v]\,,$$

we get:

$$\partial_\varepsilon v^\varepsilon_t = \partial_t w^\varepsilon_t - \mathrm{ad}_{v^\varepsilon_t} w^\varepsilon_t\,.$$

The differentiation of the energy gives at $\varepsilon = 0$,

$$\int_0^1 (Lv_t, \partial_\varepsilon v_t)\, dt = \int_0^1 (Lv_t, \partial_t w^\varepsilon_t - \mathrm{ad}_{v^\varepsilon_t} w^\varepsilon_t)\, dt = 0\,. \tag{1.19}$$

By integration by part and since $w_0 = w_1 = 0$, we get:

$$\int_0^1 (\frac{d}{dt} Lv_t + \mathrm{ad}^*_{v_t} Lv_t, w_t)\, dt = 0\,, \tag{1.20}$$

with ad^* the adjoint of ad. For all w_t coming from variations of $\phi^\varepsilon_{0,t}$ wich gives at least all continuous path w in V such that $w_0 = w_1 = 0$, which is sufficient to conclude,

$$\frac{d}{dt} Lv_t + \mathrm{ad}^*_{v_t} Lv_t = 0\,. \tag{1.21}$$

Since $\mathrm{ad}_{v_t} w = \frac{d}{ds}_{|s=0} \mathrm{Ad}_{\phi_{t,t+s}} v$, with $\phi_{t,s}$ the flow of u_t and $\mathrm{Ad}_\psi v := d\psi(v \circ \psi^{-1})$ for ψ a diffeomorphism, we obtain the first Euler theorem also known as the conservation of the momentum:

$$\frac{d}{dt}[\mathrm{Ad}^*_{\phi_{0,t}} Lv_t] = 0\,. \tag{1.22}$$

1.2.2 Matching problems

We will now derive the geodesic equations for landmarks matching which is the simplest case to illustrate the conservation of momentum. This is a well-known case and the content of this paragraph can be found in [You08] or [Mic08]. Introduced in [JM00] the landmark matching equations have a Hamiltonian formulation which is well described in [ATY05] or [VMTY04].

We are interested in the matching of two groups of n points in $\mathbb{R}^d$, $\mathbf{x} = (x_1, \dots, x_n)$ as initial data and $\mathbf{y} = (y_1, \dots, y_n)$. Let us consider the matching functional,

$$\mathcal{J}(u) = \int_0^1 |u_t|_V^2 \, dt + d(\phi_{0,1}.x, y)^2 \,, \tag{1.23}$$

where d is the euclidean distance on $\mathbb{R}^{nd}$ and $\phi_{0,1}$ the flow associated to the vector field $u \in L^2([0,1], V)$. The notation $\phi.x$ stands for the action of ϕ on the group of points $\mathbf{x}$ which is defined by $\phi.x := (\phi(x_1), \dots, \phi(x_n))$. Note that we do not require $\mathbf{x}$ to be a group of distinct points in $\mathbb{R}^d$ (neither for $\mathbf{y}$).

The existence of a minimizer follows the proof of Theorem 20 which actually contains the landmark case. Hence, the geodesic equations are still valid in this context. Remark that a minimizer $u \in L^2([0,1], V)$ for the functional (1.23) is also a minimizer of the exact matching problem:

$$\min\{\mathcal{E}(v) = \int_0^1 |v_t|_V^2 \, dt \,|\, \phi_{0,1}^v.x = \phi_{0,1}^u.x\}. \tag{1.24}$$

More generally if $\mathbf{x}$ and $\mathbf{y}$ are two groups of n points, what is the orbit of $\mathbf{x}$ under the action of the group of deformations? Clearly to answer this question, we can assume that $\mathbf{x}$ is a group of n distinct points.

Proposition 2. *The action of the group is transitive among the group of n distinct points if $n > 1$. For $n = 1$, two groups of points are in the same orbit if and only if the permutation σ ordering $\mathbf{x}$ in increasing order: $x_{\sigma(1)} < x_{\sigma(2)} < \ldots < x_{\sigma(n)}$ is the permutation ordering $\mathbf{y}$ in increasing order.*

Proof. If $\mathbf{x}$ is a group of n distinct points, namely $x_i \neq x_j$ for $i \neq j$, then by injectivity property of the diffeomorphism $\phi_{0,1}$ we have $\phi(x_i) \neq \phi(x_j)$. For the one dimension case, $\phi_{0,1}$ is a diffeomorphism in the connected component of the identity and thus is an increasing diffeomorphism and it preserves order.

Now assume that there exists a C^1 path $c : [0,1] \mapsto [\mathbb{R}^d]^n$ such that $c(0) = \mathbf{x}$ and $c(1) = \mathbf{y}$ and $c^i(t) \neq c^j(t)$ for $t \in [0,1]$ and $i, j \in [1, n]$. This is the case in dimension 1 if the order is respected and the dimension is greater than two or equal, this is the case for any couple of group of n distinct points.

We claim that there exists a vector field $v \in L^2([0,1], V)$ such that $\phi_t.x = c(t)$. Just consider the time dependent vector field v defined by $v_t(.) = k(., c(t))p(t)$ with $p(t)$ the vector of $[\mathbb{R}^d]^n$ uniquely determined by $k(c(t), c(t))p(t) = \frac{d}{dt}c(t)$ since $k(c(t), c(t))$ is

invertible. The map $p(t)$ is clearly C^1 (inverse mapping theorem) and we have by the definition:

$$\frac{d}{dt}c^i(t) = v_t(c^i(t))\, i \in [0,n]\,, \tag{1.25}$$

which implies that $\phi_{0,t}.x = c(t)$ for all $t \in [0,1]$. □

Remark 1. *The vector field introduced in the above proof is the unique minimizer for the norm on $L^2([0,1],V)$ of the vector fields v verifying* (1.25)*.*

We aim to study a minimizing path for the inexact matching (1.23).

Proposition 3. *A minimizer v of the inexact matching* (1.23) *verifies the following geodesic equation,*

$$v_t(.) = -\sum_{i=1}^{n} k(., \phi_{0,t}(x_i))[d\phi_{t,0}]^*_{\phi_{0,t}(x_i)}(p_i)\,, \tag{1.26}$$

with $[p_i]_{i\in[1,n]}$ a vector defined on the cotangent space of $[q_i(0)]_{i\in[1,n]}$ called the initial momentum.

Proof. Differentiating the functional (1.23), we obtain for a minimizer u_t:

$$\int_0^1 \langle v_t, u_t\rangle + \sum_{i=1}^{n} \langle (d\phi^*_{0,1}(\phi_{0,1}(x_i) - y_i)), \int_0^1 [\mathrm{Ad}_{\phi_{t,0}} u_t](x_i)\, dt\rangle = 0\,.$$

the second term can be rewritten with the kernel,

$$\int_0^1 \sum \langle k(., \phi_{0,t}(x_i))[d\phi_{t,0}]^*_{|\phi_{0,t}(x_i)}(p_i), u_t\rangle_V\, dt\,,$$

with $p_i = d\phi^*_{0,1}(\phi_{0,1}(x_i) - y_i))$. Then, it implies (1.26) since this equality is verified for all $u \in L^2([0,1],V)$. □

We observe that the momentum which corresponds to the Euler equation is

$$Lv_t = \sum_{i=1}^{n} K\delta^{p_i(t)}_{\phi_{0,t}(x_i)}$$

with $p_i(t) = [d\phi_{t,0}]^*_{\phi_{0,t}(x_i)}(p_i)$. The differentiation of the matching term gives in this case the structure of the momentum, which is a sum of Dirac operators. More generally, the differentiation of the matching term will give some information about the structure of the momentum.

Let us address the case of smooth images. In the case of images, the action is defined by:

$$I \mapsto I \circ \phi_{1,0},$$

with I an image which will be considered as a function defined on a domain $U \subset \mathbb{R}^d$. The inverse taken in the composition makes the action geometric: for example, if χ_S is the indicator function of a domain S, then $\chi_S \circ \phi_{1,0} = \chi_{\phi_{0,1}(S)}$. The matching problem

between two images I_0 the template and I_{targ} the target image can be written as the minimization of:

$$\mathcal{J}(v) = \int_0^1 |v_t|_V^2 \, dt + \int_U H(I_0 \circ \phi_{1,0}, I_{targ}) d\mu \,, \tag{1.27}$$

with μ the Lebesgue measure on U. The function H can be the euclidean distance or any other cost function.

If I_0 is smooth enough, namely C^1 and $I_{targ} \in L^\infty(U)$, it is easy to differentiate the matching term in (1.27). For example if $H(x,y) = |x-y|^2$, we obtain:

$$\int_U 2(I_0 \circ \phi_{1,0} - I_{targ}) \langle \nabla I_0, \partial_\varepsilon \phi_{0,1}^\varepsilon \rangle d\mu \,.$$

Using the Theorem 6, we have with $I_t := I_0 \circ \phi_{t,0}$ and for a variation $v_t + \varepsilon u_t$ of the vector field:

$$\partial_\varepsilon \int_U (I_1 \circ (\phi_{0,1} \circ \phi_{1,0}^\varepsilon) - I_{targ})^2 \, d\mu = - \int_U (I_1 - I_{targ}) \langle \nabla I_1, \int_0^1 \mathrm{Ad}_{\phi_{t,1}}(u_t) \rangle \, d\mu \,.$$

In order to use the kernel, remark that by a change of variable, if $f, g \in L^2(U, \mathbb{R}^d)$ are two vector fields and ψ a diffeomorphism with its first derivative bounded (to give a sense to the following equation),

$$\int_U \langle f, \mathrm{Ad}_\psi g \rangle \, d\mu = \int_U \langle d\psi^* f \circ \psi, g \rangle \mathrm{Jac}(\psi) d\mu \,. \tag{1.28}$$

We have also,

$$\nabla(I \circ \psi) = d\phi^* (\nabla I \circ \phi) \,.$$

Thus we can write, with $\Delta = (I_0 \circ \phi_{1,0} - I_{targ})$

$$\int_U 2(I_0 \circ \phi_{1,0} - I_{targ}) \langle \nabla I_0, \mathrm{Ad}_{\phi_{t,1}}(u_t) \rangle d\mu = \int_U p_t \langle \nabla I_t, u_t \rangle_V \,, \tag{1.29}$$

with $p_t = \mathrm{Jac}(\phi_{t,1}) \Delta \circ \phi_{t,1}$. Hence we end up with the proposition:

Proposition 4. *If I_0 is C^1 and $I_{targ} \in L^\infty(U)$, a minimizer of the functional* (1.27) *with $H(x,y) = |x-y|^2$ verifies the geodesic equation*

$$v_t(.) = \int_U k(., x) \nabla I_t(x) p_t(x) d\mu(x) \,, \tag{1.30}$$

with $p_t = \mathrm{Jac}(\phi_{t,1}) \Delta \circ \phi_{t,1}$ and $\Delta = (I_0 \circ \phi_{1,0} - I_{targ})$.

In this case, the momentum is a normal vector field to the level sets of I_t. This normality constraint is a general fact as addressed in [MTY06], however the structure of this momentum (here a function densely defined) comes from the matching term, which has a regularizing effect. Let us consider the exact matching of two images which lie in the same orbit: I and $I \circ \phi_{1,0}$. We assume in addition that I has a strict global maximum at $x_0 \in U$. Consider a solution v_t of the landmark matching between x_0 and $x_1 \in U$ with $x_0 \neq x_1$. Then, it is a solution of the exact matching between the two images since the two maximums shall match. In this case, this momentum can not be represented by a

dense function, yet the normality constraint is still verified. It appears to be a degenerated case of the inexact matching problem (1.27).

One problem partially understood was the case of discontinuous images. For example, assume that $I_0 = \chi_S$ the characteristic function of a domain S, then the structure of the momentum was totally understood in the 2 dimensional case with a smooth boundary, namely piecewise C^1 (see [GTL06]). The structure of the momentum is a normal vector field along the boundary of S. Having in mind the coarea formula, this is a first step to extend the result to functions of bounded variation (BV functions). The differentiation of the matching term is done in Chapter 2 for a wide range of cost functions H and for BV functions (see appendix A). Not surprisingly, we obtain a condition on I_{targ} which at least is assumed to be also a BV function. However, this framework is general enough to cover a lot cases of application, especially the 3D case for the matching of two shapes.
Another application is developed thanks to the differentiation lemma proved in Chapter 2 for the metamorphosis framework exposed in [TY05]. In this context, the theory was developed in the case of H^1 functions. We will extend it to SBV functions thanks to the their attempt to take as much as possible into account the case of discontinuities.
We give a short presentation of the semi-differentiation lemma for the reader who wants to avoid this technical Chapter 2. Consider for example two Lipschitz open sets U and V. One may want to deform one of these open sets while the second remains unchanged (figure 1.1). The basic case is the following:

$$J_t = \int_V \chi_U \circ \phi_t^{-1} dx = \mu(V \cap \phi_t(U)),$$

with μ the Lebesgue measure. For instance, our semi-differentiation lemma answers to the differentiation of J_t and we obtain a Stokes formula in which, roughly speaking, one takes only into account the deformation viewed in V.

Proposition 5. *We have,*

$$\partial_{t|t=0^+} \mu(V \cap \phi_t(U)) = \int_{\partial U} \langle X, n \rangle \tilde{\mathbf{1}}_V(X) d\mathcal{H}^{n-1},$$

with $\tilde{\mathbf{1}}_V(X)(y) = \lim_{\epsilon \mapsto 0^+} \mathbf{1}_{\bar{V}}(y + \epsilon X)$*, if the limit exists and* 0 *elsewhere. And we denote by* n *the outer unit normal to* ∂U*.*

This proposition is a corollary of the differentiation Theorem 15 in its full generality; under some smoothness assumptions on H, we have, with $Im(U) = BV(U) \cap L^\infty(U)$, also denoted by B.

Theorem 11. *Let* H *be a locally Lipschitz function from* $\mathbb{R}^2$ *to* $\mathbb{R}_+$*,* $(f, g) \in \mathrm{B} \times \mathrm{B}$ *be two images,* X *be a Lipschitz time dependent vector field* C^1 *in space and* ϕ_t *be its associated flow. We define the functional*

$$J_t(f, g) = \int_U H(f \circ \phi_t^{-1}(x), g(x)) dx\,, \tag{1.31}$$

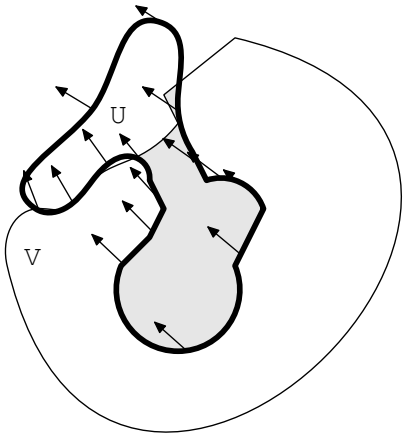


Figure 1.1: Evolution of the area between two Lipschitz curves. (The arrows represent the value of X on the boundary of U)

then, under the additional assumption that H is C^1 in the first variable (the derivative w.r.t. to such a variable being denoted by $\nabla_1 H$), we have

$$\partial_{t=0^+} J_t = \int \langle \partial_1 H(f(x), g_{X_0}(x)), -X_0 \rangle dx \,, \tag{1.32}$$

where $\partial_1 H(f, g_{X_0})$ is a part of the BV derivative of $H(f, g)$, defined by

$$\partial_1 H(f(x), l) = \nabla_1 H(f(x), l)(\nabla f(x) + D^c f(x)) + j_H(x)\mathcal{H}^{n-1} \llcorner J_f,$$

with $j_H(x) = (H(f^+(x), l) - H(f^-(x), l))\, \nu_f(x)$.

We give a sketch of the proof which can be found with the application of this Theorem in [VS09]. The proof follows three reductions.

• *First reduction*

It is sufficient to prove the result for autonomous vector fields (which do not depend on the time variable). Comparing the penalty term for the constant vector field X_0 and the initial one, we get the result with the estimation in $o(t)$ of the distance between the flows.

• *Second reduction*

It is sufficient to prove the result for $F(x, y) = xy$. As $Im(U)$ is an algebra since the functions are bounded, the formula is easily true for polynomial functions. Then, approximating $\nabla_1 F$ on the first coordinate by a polynomial function, we get the result with the following control if f is C^1

$$\begin{aligned} |J_t(f, g) - J_0(f, g)| &\leq \mathrm{Lip}_1(F) \int_U |f \circ \phi_t^{-1} - f| dx \\ &\leq \mathrm{Lip}_1(F) ||X||_\infty \int_0^t \int_U |\nabla (f \circ \phi_s^{-1})| dx\, ds \,. \end{aligned}$$

Here, $\mathrm{Lip}_1(F)$ is the Lipschitz constant of F on $\{(f(x), g(y))|(x, y) \in U^2\}$. The inequality is also valid for $f \in BV$ by approximation. By a change of variable, we get the result. Remark the sup norm of the vector field on a small time interval $[0, t_0]$ in the bound, so that the contribution of the equilibrium points ($X^{-1}(\{0\})$) to the differentiation result is null. In the case of the product xy, we have the following estimation: for $t \leq t_0$, if $g \in Im(U)$ and $f \in BV(U)$ (also true if $g \in L^\infty(U)$), we have

$$|J_t(f, g) - J_0(f, g)| \leq Ct||g||_\infty||f||_{BV}.$$

We emphasize the continuity of this result with respect to the sup norm for g and the BV norm for f. By a change of variable, we have the same result switching the role f and g.

• *Third reduction*

Remark first that if the result is true for $F(x, y) = xy$ then it is true if g is an uniform limit in $Im(U)$. We claim that it is sufficient to prove the result for the one dimensional case.

As we pointed it out before, we only need to focus on points x such that $X(x) \neq 0$. By the flow-box theorem, we obtain through a change of variables (the functions are extended on $\mathbb{R}^n$ with the value 0), $J_t = \int_{\mathbb{R}^n} f \circ \psi(x - tv)\, g \circ \psi\, \mathrm{Jac}(\psi)dx$. Remark that we have to deal with the Jacobian of ψ, we need to assume that it is continuous in order to apply the result we prove in one dimension. This is allowed thanks to our assumptions on V. Hence $g \circ \psi\, \mathrm{Jac}(\psi)$ lies in the closure of $Im(U)$ under the uniform norm. We then use the Theorem (3.108 in [AFP00]) which exhaustingly explains the behavior of the one-dimensional restrictions of a BV function and the dominated convergence theorem to conclude. In one dimension, the result is a straightforward verification.

Yet this proof requires a regularity on the vector field higher than the one we can obtain through the initial way we followed. First proving the result for Lipschitz piecewise functions, we can reach less regularity for the vector field and then conclude by (strong) density of Lipschitz piecewise functions in SBV functions (result proved in 2). This space of functions will be used in the Chapter 4 for the Hamiltonian formulation.

1.3 The Hamiltonian formulation

1.3.1 The deterministic case

The Euler equation on the conservation of momentum (1.22) says that the initial momentum is transported by the action of the flow. It is well-known that in the landmark case, the equations can be transformed in a Hamiltonian system:

Proposition 6. *The equation* (1.26) *can be rewritten as*

$$v_t(.) = k(., \mathbf{x}_t)\mathbf{p}_t\,,$$

and we have,

$$\begin{cases} \dot{\mathbf{p}}_t = -\mathbf{p}_t \partial_1 k(\mathbf{x}_t, \mathbf{x}_t)\mathbf{p}_t\,, \\ \dot{\mathbf{x}}_t = k(\mathbf{x}_t, \mathbf{x}_t)\mathbf{p}_t\,, \end{cases} \tag{1.33}$$

which are the Hamiltonian equations for the landmark matching: with

$$H(p,q) = \frac{1}{2}\mathbf{p}k(\mathbf{x})\mathbf{p},$$

the equations (1.33) *can be rewritten*

$$\begin{cases} \dot{\mathbf{p}}_t = -\partial_{\mathbf{x}} H\,, \\ \dot{\mathbf{x}}_t = \partial_{\mathbf{p}} H\,. \end{cases}$$

Proof. With the proposition 3, recall that we have

$$v_t(.) = \sum_{i=1}^{n} k(., \phi_{0,t}(x_i))[d\phi_{t,0}]^*_{\phi_{0,t}(x_i)}(p_i)\,.$$

Then, with the differentiation of $[\mathbf{p}_t]_i = [d\phi_{t,0}]^*_{\phi_{0,t}(x_i)}(p_i)$, :

$$\begin{aligned} &\partial_t d\phi_{0,t} = dv_t.d\phi_{0,t}\,, \\ &\partial_t [d\phi_{t,0}]^*_{\phi_{0,t}} = \partial_t d\phi_{0,t}^{-1*} = -dv_t^*.d\phi_{0,t}^{-1*}\,, \end{aligned}$$

and $dv_t^*(.) = \sum_{i=1}^n \partial_1 k(., \phi_{0,t}(x_i))[\mathbf{p}_t]_i$ gives the result. □

These Hamiltonian equations can be obtained from the optimal control point of view (see [ATY05]). We will pay more attention to this tool in Chapter 4. From these equations and from the proof of proposition 2, we can prove that there exists a trajectory which minimizes the exact landmark matching problem (1.24) which verifies the Hamiltonian equations (1.33). This Hamiltonian formulation is the equation of geodesics on the Riemannian manifold of landmarks (see [Mic08],[Gla05]).
Why to develop the Hamiltonian formulation of the geodesic equations?
First of all the Hamiltonian formulation changes the parameterization of the problem. One needs to encode the class of objects on which acts the group and the momentum associated to this object which has the same complexity in some cases of interests (Hilbert spaces): in the landmark case, the problem is reduced to an ODE in finite dimension. On the numerical side, this reduction of dimension (if so) should decrease the computational cost. It enables also to design numerical algorithms using shooting methods. For example, a Newton method on the initial momentum is developed in [ATY05]. This approach was also used in [MM08].

As for the landmark case, we can try to obtain Hamiltonian equations in the case of images. Thanks to proposition 4, we can differentiate for $f \in C_0^\infty(U)$ a test function,

$$\frac{d}{dt} < p_t, f >_{L^2} = - < p_t, \langle \nabla f, v_t \rangle >_{L^2}\ .$$

As a result, we get the Hamiltonian equations which have a weak sense (at least for the momentum equation):

$$\begin{cases} \dot{I}_t = \langle \nabla I_t, v_t \rangle = \partial_p H(p,q)\,, \\ \dot{p}_t = < p_t, \langle \nabla ., v_t \rangle >_{L^2} = -\partial_I H(p,q)\,, \end{cases} \tag{1.34}$$

where H is defined by,

$$H(I,p) = \frac{1}{2}\int_{U\times U} p_t(x)\nabla I(x)k(x,y)p_t(y)\nabla I(y)\,dxdy\,.$$

In [GTL06] the Hamiltonian formulation for the case of generalized closed curves (by generalized we mean element in $L^2(S_1,\mathbb{R}^2)$) is proposed. The structure of the equation is very closed to the landmark case: in Chapter 5, we explicitely show that the landmark case is a particular case of generalized closed curves. The momentum is proved to be a normal vector field to the (smooth) curve if the matching term is geometric (i.e. invariant to the reparameterization). The introduction to Chapter 4 will adopt the control theory point of view to propose a heuristic approach to derive the Hamiltonian equations in general. Then, we will detail the Hamiltonian formulation for the model introduced in Chapter 3 on discontinuous images, which introduces (in addition to the geometric deformation) a deformation of the contrast term through the left composition with a diffeomorphism:

$$I \mapsto \eta_{0,1}\circ I\circ\phi_{1,0},$$

where η is the flow generated by a vector field $(s_t)\in L^2([0,1],S)$ with S an admissible space of vector fields on $[0,1]$ for example. The interval $[0,1]$ stands for the gray-level set. Then the geodesic equations come from the minimization of the functional:

$$\mathcal{J}(v_t,s_t) = \frac{\lambda}{2}\int_0^1 \|v_t\|_V^2 dt + \frac{\beta}{2}\int_0^1 \|s_t\|_S^2 dt + \int_M H(\eta_{0,1}\circ I_0\circ\phi_{1,0}(x), I_{targ}(x))dx, \tag{1.35}$$

with β,λ two scaling parameters.
The derivation of the geodesic equations is proved in Chapter 3 on SBV functions. Once the structure of the initial momentum is known, we can address the problem of the Hamiltonian formulation: a weak formulation is showed in 4.2.1 for piecewise Lipschitz images which are introduced in 2.2. As for the case of curves in 2D, we obtain a strong formulation of the equations for a cartoon image in any dimension (3D for instance).

Apart from numerical advantages, the Hamiltonian system encodes the evolution in the initial momentum and it could be tempting to perform statistical study on this representation (as developed in [VMTY04]). From a template, a collection of objects is represented as a momentum, which lies in a linear space. At the first order, it linearizes the spaces of objects and this representation leads to more meaningful statistical quantities. A statistical model for large deformations is proposed in [All07] as the generalization of the work in the small deformation setting. The model on the initial momentum is a Gaussian distribution and it is a first step to modelize the empirical distribution of shapes. Yet the geometry is still encoded in the cotangent space and this should be reflected in the statistical model, since the geometry of the Riemannian manifold of landmarks is now better understood ([Mic08, MM05a]).
Finally, this formulation models physical evolution (as N points vortices model in hydrodynamics [New01]) and it can lead to other models of evolution. The last part of this

thesis is devoted to the study of a stochastic growth model which is a perturbation of this Hamiltonian system.

1.3.2 A Stochastic perturbation of the Hamiltonian system

We can interpret the geodesic evolution of a shape as a real time evolution. From a mechanical point of view, the landmarks are particles for which the position variable has one more regularity order than the momentum. Perturbing the momentum equation with a noise term is relevant from physical considerations point of view, since it means adding a random force to the system. Then, the position variable will evolve smoothly if the momentum is continuous. Also perturbing the momentum equation by a continuous noise will give a continuous evolution of the momentum. Let us write the perturbed equations in the case of landmarks:

$$\dot{p}_t = -\partial_q H + \varepsilon dB_t \,,$$
$$\dot{q}_t = \partial_p H \,,$$

hence, p_t being continuous the evolution of q_t is C^1 in time. Here, the noise is the Brownian motion since it is the prototype of continuous noise.
In addition the diffeomorphic property will be conserved by this model since it will lead to a random path of diffeomorphisms given by the flow of the vector field:

$$v_t(x) = k(x, q_t)p_t \,.$$

Thus this stochastic model could lead to an interesting dynamic model for the growth of biological shapes, having in mind such applications as early diagnosis in the medical field. On the mathematical side, the study these perturbed geodesics can produce interesting results since the most probable path leading to a point is not the geodesic one. However before addressing these questions or the possible applications of the model, we will cope with two fundamental questions which are crucial

- Do the solutions of the stochastic model for landmarks blow up in finite time?
- Does there exist an extension of the stochastic model for shapes (or is the model consistent when the number of landmarks increases)?

We will provide positive answers to both questions. The chosen method for the first point is the control of the perturbed Hamiltonian via the Ito formula and then to follow the proof of existence in all time for the deterministic case. For the extension to shapes as a limit of the stochastic system on landmarks, we chose to find a framework which contains the landmark case as well as continuous shapes. To this end, we will consider landmarks as piecewise constant functions defined for example on $L^2(S_1, \mathbb{R})$ and we will introduce a sort of Sobolev space on the Haar basis to pass to the limit. The space F_s on which we work is closed to a Besov space but easier to work with in our context. We will obtain

an almost sure convergence of the solutions of the stochastic system when increasing the number of landmarks to the continuous case.
Perturbing the momentum equation by a Brownian term leads to a consistent growth model with smooth trajectories. For concrete applications, this model needs to be enhanced: the white noise could be correlated by a term depending on the position of the landmarks. Remark that in this formulation the noise is Lagrangian, i.e. the noise is supported by each particle. We could be interested by a geometrical noise which is not included in the generalization we will develop. Also the way of dealing with the noise is closely related to the geometry of the space (in this case the space of landmarks), see for instance [Lar01]). The important point is that the model seems to be robust enough to deal with other different noise terms such as Poisson processes or others.

Last, the conclusion will further perspectives and consequences of our work on the deterministic side as well as on the probabilistic side.

Chapter 2

Discontinuities and matching: a semi differentiation lemma

Contents

2.1 Introduction 29
2.2 A simple approach: Lipschitz piecewise functions 30
2.3 The case of SBV and BV functions 38
2.4 The one-dimensional case 51
2.5 Reduction to the one-dimensional case 52
2.6 Toward the application to geodesic equations 57

The reader shall refer to the paragraph 1.2.2 for a summary and a sketch of the proof of the main result of this Chapter. We just recall that this Chapter is devoted to differentiate the attachment term of this functional

$$\mathcal{J}(v) = \int_0^1 |v_t|_V^2 \, dt + \int_U |I_0 \circ \phi_{1,0} - I_{targ}|^2 d\mu \,, \tag{2.1}$$

with respect to a variation of the diffeomorphism.

2.1 Introduction

This Chapter aims to develop the material we need to deal with discontinuities within the diffeomorphic approach. The large deformation via diffeomorphisms was first carried out on smooth templates. The only major discontinuous case that was completely understood was the case of a binary image in 2D with a piecewise smooth set of discontinuities. However, this case is rich enough to hopefully extend it to more general situations. Of course, a large literature is at hand to deal with discontinuities, especially with the space of functions of bounded variations (BV functions). The use of such spaces as BV or SBV functions (Special Bounded Variations, BV functions such that the Cantor part of the distributional derivative vanishes) is now widespread in the mathematical imaging community. Instead of directly trying to extend the approach to these spaces, we preferred to extend the simple case of smooth close curves to a set of functions which contains enough functions to be interesting. We ended up with the set of piecewise Lipschitz functions,

which are afterwards a subset of SBV functions: a function is piecewise Lipschitz if there exists a partition of the definition domain in Lipschitz domains for which the restriction of the function on each Lipschitz domain is Lipschitz. Not only this set was big enough for concrete applications but it also enables the possibility to introduce a multilayer approach, which will be partially addressed in the Chapter on the Hamiltonian formulation of the geodesic equations.

Back to the Euler-Poincaré equation, the problem as mentioned previously was to understand the structure of the momentum in the discontinuous case. In the case of a binary image on a piecewise smooth domain in 2D, the momentum was encoded in a normal vector field to the boundary. Hence we expected that the initial momentum for the Lipschitz piecewise functions would have a dense part (as for smooth functions) and a singular part which should be a normal vector field to the boundary of the Lipschitz domains of the partition. Thanks to simple analytical tools, this result was obtained and it opened the way to extend a little further.

Although it was not really needed for current applications, we then wanted to extend our results to the case of BV functions, which was done thanks to the collaboration with Filippo Santambrogio. To tackle this extension, two ways were developed: first we extend our previous result to SBV functions by proving that the closure of Lipschitz functions in BV functions is precisely the SBV functions space; second we modified our approach through the use of change of coordinates to trivialize the flow of the vector field in order to treat the problem in one dimension. This second approach needs a stronger assumption on the vector field regularity than the one needed for the first approach. In both ways we could not avoid to prove the result for smoother functions and then obtain the general case with a density argument.

This Chapter is divided into two main parts.

In the first sections, we will present the semi-differentiation lemma for the case of Lipschitz piecewise smooth functions and then the extension to the case of SBV and BV functions (which do not have the same generality level in these two cases). We will also address the problem of differentiating other penalty terms than the L^2 norm. We obtain a large generalization of our work which is sufficiently satisfying to be presented. However our result does not contain the L^1 norm even if we expect the result to be true for this penalty term.

In the last section we present an application which extends the result of [TY05] for metamorphosis to the case of SBV functions.

2.2 A simple approach: Lipschitz piecewise functions

Definition 1. *A Lipschitz open domain is an open subset U of $\mathbb{R}^n$ connected, bounded and nonempty such that for every $x \in \partial U$ there exist*

- *an open ball centered in x $B(x, \epsilon)$ with $\epsilon > 0$,*

- *an affine orthonormal basis* $\mathcal{B} = (e_1(x), \dots, e_n(x))$ *of* $\mathbb{R}^n$ *in* x,
- *a Lipschitz function* w *defined on* $\mathrm{Vect}(e_1(x), \dots, e_{n-1}(x)) = \mathbb{R}^{n-1}$,

such that if $z \in \mathbb{R}^n$ *is described by its coordinates in* $\mathcal{B}$, $(z_1, \dots, z_n)$ *then*

$$U \cap B(x, \epsilon) = \{z \in B(x, \epsilon) | z_n > w(z - z_n e_n)\}.$$

The next definition is introduced to get the existence of partitions of a Lipschitz domain in Lipschitz domains, since this definition allows points of the boundary to belong or not to the domain.

Definition 2. *A Lipschitz domain is a subset* D *of* $\mathbb{R}^n$ *if there exists an open domain* U *such that*

$$U \subset D \subset \bar{U}.$$

On the definition of a Lipschitz domain U: we use here a definition of Lipschitz domains which can be found in Chapter 2 of [DZ01], among other characterizations. Note that the set of Lipschitz domain is not stable under bi-Lipschitz homeomorphism (homeomorphism such that it and its inverse are Lipschitz) as stated in the previous reference.
We will denote by $\mathrm{Im}(M)$ the set of piecewise Lipschitz functions defined as follows,

Definition 3. *If* $f \in Im(M)$*, there exists a partition of* M *in Lipschitz domains* $(U_i)_{i\in[0,k]}$ *for an integer* $k \geq 0$*, and the restriction* $f_{|U_i}$ *is Lipschitz.*

Remark 2. *The extension theorem of Lipschitz function in* $\mathbb{R}^n$ *enables to consider that on each* U_i, $f_{|U_i}$ *is the restriction of a Lipschitz function defined on* $\mathbb{R}^n$.

Example 1. *The most simple example is a piecewise constant function,* $f = \sum_{i=1}^{k} a_i \mathbf{1}_{U_i}$ *with* $a_i \in \mathbb{R}$.

Remark 3. *A partition satisfying the definition 3 is not uniquely defined and in fact it does not seem possible to reach a uniqueness condition with an additional assumption such as a partition of minimal length. Obviously such a minimal partition will contain the jump set of the function.*
Moreover, let us consider the equivalence relation on the domain of definition of the piecewise Lipschitz function f *defined by* $x \sim y$ *if there exists a connected open set on which the restriction of* f *is continuous. Then it gives a partition in open sets of the domain, but this partition is not a Lipschitz one.*

In what follows, to simplify the notations a Lipschitz vector field on M **will be assumed to be a time dependent vector field on** M **which is Lipschitz in time and space.**
To tackle the differentiation of the penalty term in the functional 2.1, we can develop the L^2 difference:

$$\int_U |I_0 \circ \phi_{1,0} - I_{targ}|^2 d\mu = \int_U (I_0 \circ \phi_{1,0})^2 - 2 I_{targ} I_0 \circ \phi_{1,0} + I_{targ}^2 \, d\mu \,.$$

The last term does not depend on $\phi_{1,0}$ and as a consequence has a null contribution in the differentiation. If I_0^2 is a function of bounded variation (BV function), the first term $(I_0 \circ \phi_{1,0})^2 = I_0^2 \circ \phi_{1,0}$ in the integral can be differentiated from the definition of a BV function. The last term, which involves the two functions (which can both have discontinuities) is much more difficult to differentiate and as far as we know, this situation was not treated in the literature. This is why our interest in this Chapter will be focused on the differentiation of $J_t = \int_U I_{targ} I_0 \circ \phi_{1,0} d\mu$ with respect to a variation of the flow.

Lemma 1. *Let U, V be two bounded Lipschitz domains of $\mathbf{R}^n$. Let X be a Lipschitz vector field on $\mathbb{R}^n$ and ϕ_t be the associated flow. Finally, let g and f be Lipschitz real functions on $\mathbf{R}^n$. Consider the following quantity depending on t,*

$$J_t = \int_{\phi_t(U)} f \circ \phi_t^{-1} g \mathbf{1}_V d\mu,$$

where μ is the Lebesgue measure, then

$$\partial_{t|t=0^+} J_t = \int_U -\langle \nabla f, X \rangle g \mathbf{1}_V d\mu + \int_{\partial U} \langle X, n \rangle f g \tilde{\mathbf{1}}_V(X) d\mathcal{H}^{n-1}. \tag{2.2}$$

with $\tilde{\mathbf{1}}_V(X)(y) = \lim_{\epsilon \hookrightarrow 0^+} \mathbf{1}_{\bar{V}}(y + \epsilon X)$, if the limit exists, 0 elsewhere. And n denotes the outer unit normal on ∂U.

The proof of the lemma will follow two steps. The first step is to prove the lemma when the two involved open sets are C^1 and the second step will extend the result to Lipschitz domains. In the following, X will be a time dependent vector field on $\mathbb{R}^n$ Lipschitz in both space and time variables.
Our first proposition is a little more general than the final proposition since we will assume g only continuous. This hypothesis will be needed when we will make use of a change of variable. With the previous notations, we have

Proposition 7. *Let U, V be C^1 domains of $\mathbf{R}^n$, X be a Lipschitz vector field and ϕ_t be the associated flow. Let f and g be respectively Lipschitz and continuous real functions on $\mathbf{R}^n$. Consider the following quantity depending on t,*

$$J_t = \int_{\phi_t(U)} f \circ \phi_t^{-1} g \mathbf{1}_V d\mu,$$

where μ is the Lebesgue measure, then

$$\partial_{t|t=0^+} J_t = \int_U -\langle \nabla f, X \rangle g \mathbf{1}_V d\mu + \int_{\partial U} \langle X, n \rangle f g \tilde{\mathbf{1}}_V(X) d\mathcal{H}^{n-1}. \tag{2.3}$$

To prove this proposition, we need two simple lemmas. The first one somehow tells that there exists a convenient change of coordinates to represent the boundary of the first open set as a function defined on the boundary of the second open set.

Lemma 2. *If $x \in \partial U \cap \partial V$ with U, V two C^1 open sets in $\mathbb{R}^n$, then there exists ψ a C^1 diffeomorphism of $\mathbb{R}^n$, W a neighborhood of x and k a C^1 function defined on $\mathbb{R}^{n-1}$ such that*

1. $\psi(x) = 0$,
2. $\psi(U \cap W) = \{(z,t) \in \mathbb{R}^{n-1} \times \mathbb{R} | t > 0\} \cap B(0,1)$,
3. $\psi(V \cap W) = \{(r,s) \in \mathbb{R}^{n-1} \times \mathbb{R} | s > k(r)\} \cap B(0,1)$ *or*
   $\psi(V \cap W) = \{(r,s) \in \mathbb{R}^{n-1} \times \mathbb{R} | s < k(r)\} \cap B(0,1)$.

Proof. As this is a local result, we will not detail how to find precisely the neighborhood W. As U is a C^1 open set, there exists a C^1 diffeomorphism ϕ for which the two first conditions are satisfied. For the third condition, remark that $\psi(V)$ is still a C^1 open domain. As a consequence, the normal vector v to the tangent plane H at 0 to $\psi(V)$ is well defined. The boundary $\partial\psi(V)$ can be described as a C^1 function f defined on H. We denote by $(e_1, \ldots, e_n)$ the canonical orthonormal basis of $\mathbb{R}^n$. Two cases appear, the first one we deal with is the generic one.
If $\langle v, e_n \rangle \neq 0$, let $A \in L(\mathbb{R}^n)$ be the linear isomorphism defined by: $A(v) = e_n$ and $(e_1, \ldots, e_{n-1}) = Ker(A - Id)$. See figure 2.1. Denote by p_1 the orthogonal projection on H and p_2 the orthogonal projection on $(e_1, \ldots, e_{n-1})$. A consequence of the hypothesis $\langle v, e_n \rangle \neq 0$ is that

$$p_{2|H} : H \mapsto Ker(A - Id)$$

is an isomorphism. We denote by $z = p_1(z) + \lambda(z)e_n$ and $z = p_2(z) + \alpha(z)n$ the two decompositions of z with $\lambda, \alpha \in (\mathbb{R}^n)'$, so we obtain: $A(z) = p_1(p_2(z)) + (\lambda(p_2(z)) + \alpha(z))e_n$. Then with the change of variable $x = p_2(z)$, we get

$$A \circ \psi(\partial V) = \{(x, \lambda(p_2^{-1}(x)) + f \circ p_2^{-1}(x)) | x \in p_2(H \cap W)\}.$$

$f \circ p_2^{-1}$ is clearly C^1 and we obtain the lemma in this case with $A \circ \phi$.

If $\langle n, e_n \rangle = 0$, we can choose another system of coordinates for which we fall in the first case, since our previous argument remains valid for every hyperplane H for which the boundary can be described as a C^1 function. As this condition is clearly open, the lemma is proven.

□

The second lemma states the evolution of the function describing the boundary of an open set under the flow of a vector field.

Lemma 3. *Let ϕ_t be the flow of the vector field Lipschitz X on $\mathbb{R}^n$ (with $\|X\|$ bounded on $\mathbb{R}^n$). Let $V = \{(x,z) \in \mathbb{R}^{n-1} \times \mathbb{R} | z > w(x)\}$ be an open set with w a C^1 function, and we introduce $w_t(x) = \inf\{z \in bar\mathbb{R} | (x,z) \in \phi_t(V)\}$.*
Then, there exists $\varepsilon > 0$ such that w_t is defined in $\mathbb{R}$ for $t \in]-\varepsilon, \varepsilon[$ and is C^1 in both variables. The partial derivative is

$$\partial_{t|t=0} w_t(x) = -\langle \nabla w(x), p_1(X(x, w(x))) \rangle + p_2(X(x, w(x))),$$

with p_1 and p_2 orthogonal projections respectively on $\mathbb{R}^{n-1} \times \{0\}$ and $\{0\}^{n-1} \times \mathbb{R}$.

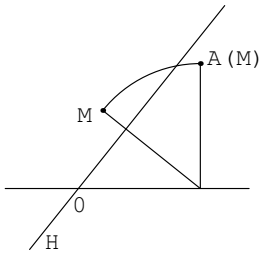


Figure 2.1: The linear transformation A

Proof. The proof is the straightforward application of the implicit function theorem to the function

$$F(x,t) = p_1(\phi_t(x, w(x))) - x_0,$$

which is clearly C^1. Note that $\partial_x F(x_0, 0) = Id_{|\mathbb{R}^{n-1}}$, so we obtain for each $x_0 \in \mathbb{R}^{n-1}$ a C^1 function $x_0 : t \mapsto x_0(t)$ such that the equation $F(x,t) = 0 \Leftrightarrow x = x_0(t)$ on a neighborhood of $(x_0, 0)$. We get by implicit function theorem the first derivative of $x_0(t)$:

$$\partial_{t|t=0} x_t \quad = \quad -p_1(X_0(x_0, w(x_0))).$$

We deduce also,

$$w_t(x) = p_2(\phi_t(x_t, w(x_t))), \tag{2.4}$$

and that $t \mapsto w_t(x_0)$ is a C^1 function for every x_0. We also get by differentiation of the equation (2.4),

$$\partial_{t_{|t=0}} w_t(x_0) = -\langle \nabla w(x_0), p_1(X_0(x_0, w(x_0)))\rangle + p_2(X(x_0, w(x_0))).$$

□

Remark 4. *In the first lemma, one can suppose that V is a Lipschitz open set. However our proof can not be extended if we assume the open sets to be Lipschitz domains and the change of coordinates to be a Lipschitz homeomorphism. As there exist bi-Lipschitz homeomorphisms that do not preserve the Lipschitz regularity of domains, the result even seems to be false in this case.*
The second lemma can be extended to Lipschitz regularity for the function w. The generalization of the proof could use an implicit function theorem with Lipschitz regularity which can be found in [PS03] and some additional details. The final result is the almost everywhere differentiability of w_t.

We can now present the proof of the semi-differentiation lemma 1.

Proof. Remark that the formula is additive in f, then by a partition of unity we only need to deal with a neighborhood of the boundary of U. Let W be a neighborhood of $x \in \partial U$, if $\partial V \cap W = \emptyset$ then the desired formula is simply the Stokes formula. Hence, we only need to focus on the intersection set of the two boundaries which is compact. We will use a change of variable to prove the result. Observe that the formula is invariant under the change of variable with a C^1 diffeomorphism. It comes from the equality:

$$\int (f1_U) \circ \phi_t^{-1}(g1_V)dx = \int (f1_U) \circ \psi \circ \tilde{\phi}_t^{-1}(g1_V) \circ \psi \mathrm{Jac}(\psi)dx, \qquad (2.5)$$

with $\tilde{\phi}_t = \psi \circ \phi_t \circ \psi^{-1}$. Remark that the Jacobian is only continuous in equation (2.5), which explains the weaker assumption on g in proposition 7.

Thanks to lemma 2, through a change of variables with a C^1 diffeomorphism of a neighborhood W' of a point $x_0 \in \partial U \cap \partial V$ we can assume that $U \cap W' = \{(x, y) \in \mathbb{R}^{n-1} \times \mathbb{R} | y > w(x)\}$ with w a C^1 function defined on $\mathbb{R}^{n-1}$, and $V \cap W' = \{(x, y) \in \mathbb{R}^{n-1} \times \mathbb{R} | y > 0\}$. (The symmetric case $y < 0$ can be treated following precisely the same proof.) We also assume thanks to the partition of unity that f and g have compact support in W'. Finally, we can write

$$J_t = \int_{\mathbb{R}^{n-1}} \int_{[w_t(x)]_+}^{+\infty} f \circ \phi_t^{-1}(x)g(x)dx,$$

with $[w_t(x)]_+$ the positive part of $w_t(x)$ and w_t is the deformation of the boundary ∂U under the flow ϕ_t, introduced in lemma 3. Alternatively, we can define w_t with $w_t(x) = \inf\{z | (x, z) \in \phi_t(V)\}$. By the implicit function theorem, we obtain easily as detailed in lemma 3:

$$\partial_{t|t=0} w_t(x) = -\langle \nabla w(x), p_1(X(x, w(x)))\rangle + p_2(X(x, w(x))),$$

with p_1, p_2 respectively the orthogonal projections on $\mathbb{R}^{n-1}$ and $\mathbb{R}$. Then, composed with the Lipschitz function $x \mapsto [x]_+$, a.e. on $\mathbb{R}^n$:

$$\partial_{t|t=0_+} [w_t]_+(x) = -\langle \nabla w(x), p_1(\tilde{X}(x, w(x)))\rangle + p_2(\tilde{X}(x, w(x))),$$

with $\tilde{X} = \tilde{\mathbf{1}}_V(X)X$. To see this point, observe that this is obvious if $w(x) > 0$ and if $w(x) < 0$. On the set $w^{-1}(\{0\})$ we have (as a direct consequence of the co-area theorem but it can also be proved directly as it is detailed in appendix lemma 21) $\nabla w = 0$ a.e.

Now, we can differentiate under the integral to get

$$\partial_{t=0_+} J_t = \int_{\mathbb{R}^{n-1}} \int_{[w_t(x)]_+}^{+\infty} -\langle \nabla f, X\rangle dx + \int_{\mathbb{R}^{n-1}} f(w(x))g(w(x)) \left(\langle \nabla w(x), p_1(X(x, w(x)))\rangle - p_2(X(x, w(x)))\right) dx\,.$$

The second term of the expression can be rewritten in the more compact form,

$$\int_{\mathbb{R}^{n-1}} f(w(x))g(w(x)) \left(\langle \nabla w(x), p_1(X(x, w(x)))\rangle - p_2(X(x, w(x)))\right) dx = \int_{\partial U} \langle X, n\rangle f g \tilde{\mathbf{1}}_V(X) d\mathcal{H}^{n-1}\,,$$

which proves the result.

□

We will extend the result to Lipschitz domains, we will use the following approximation result which is a straightforward corollary of the approximation of a Lipschitz function by a C^1 function.

Proposition 8. *Let U be a bounded Lipschitz domain, for any $\epsilon > 0$ there exists V a C^1 domain such that, $S = U \setminus \bar{V} \cup V \setminus \bar{U}$, is a rectifiable open set verifying:*

$$\mu(S) < \epsilon \tag{2.6}$$

$$\mathcal{H}^{n-1}(\partial S) < \epsilon. \tag{2.7}$$

Back to the semi-differentiation lemma, we will need stronger regularity assumptions on g since we will use a symmetry argument.

Proof of the semi-differentiation lemma 1
The first remark is that we can bound the derivative of J_t,

$$\limsup_{t\to 0} |\frac{1}{t}(J_t(U) - J_0(U))| \leq (\mu(U)\text{Lip}(f)|X|_\infty + C\mathcal{H}^{n-1}(\partial U)|f|_\infty)|g|_\infty\,, \tag{2.8}$$

with C a constant defined later.
To prove the inequality (2.8), we proceed as follows

$$|\frac{J_t - J_0}{t}| \leq |\frac{J_t - J_t^1}{t}| + |\frac{J_t^1 - J_0}{t}|,$$

with $J_t^1 = \int_U f \circ \phi_t^{-1} g 1_V dx$. We first write:

$$|\frac{J_t^1 - J_0}{t}| \leq \int_U \frac{1}{|t|}|f \circ \phi_t^{-1} - f||g|_\infty dx \leq \int_U \text{Lip}(f)\frac{|\phi_t^{-1}(x) - x|}{|t|}|g|_\infty dx\,.$$

With $\lim_{t\to 0} |\frac{\phi_t^{-1}(x)-x}{t}| = |X(x)|$, we get

$$\lim_{t\mapsto 0} |\frac{J_t^1 - J_0}{t}| \leq \int_U \text{Lip}(f)|g|_\infty |X|_\infty dx\,.$$

Then we write

$$|\frac{J_t - J_t^1}{t}| \leq \frac{1}{|t|}\int_{\mathbb{R}^n} |1_{\phi_t(U)} - 1_U||f|_\infty |g|_\infty dx.$$

Since the domain U is Lipschitz we have for t small enough such that $\text{Lip}(\phi_t) \leq 2$ and $|X_t| < M$,

$$\mu(\Delta(U, \phi_t(U)) \leq t \max(2, M)^n \mathcal{H}^{n-1}(\partial U),$$

with $\Delta(U, V) = U \setminus V \cup V \setminus U$. This result is obtained by introducing for $s_0 > 0$ small enough,

$$\Psi : (t, x) \in [-s_0, s_0] \times \partial U \mapsto \phi_t(x) \in \mathbb{R}^n\,.$$

Remark that $\Delta(U, \phi_{s_0}(U)) \subset \Psi([0, s_0] \times \partial U)$ (thanks to a connexity argument). We have that Ψ is Lipschitz and $\text{Lip}(\Psi) \leq \max(2, M)$. Since $\mathcal{H}^n([0, s_0] \times \partial U) = s_0 \mathcal{H}^{n-1}(\partial U)$, we have

$$\mathcal{H}^n(\Psi([0, s_0] \times \partial U)) \leq s_0 \max(2, M)^n \mathcal{H}^{n-1}(\partial U)\,.$$

And we then obtain the inequality (2.8) since:

$$\frac{1}{|t|} \int_{\mathbb{R}^n} |1_{\phi_t(U)} - 1_U||f|_\infty |g|_\infty dx \leq C\mathcal{H}^{n-1}(\partial U)|f|_\infty |g|_\infty\,,$$

with $C = \max(2, M)^n$.

This remark enables to pass the formula to a domain U of Lipschitz regularity. Let us write $J_t(U) = \int_{\phi_t(U)} f \circ \phi_t^{-1} g 1_V dx$ and we also define by extension of the formula in the C^1 case:

$$\partial_{t=0_+} J_t(U) \doteq \int_U - < \nabla f, X > g \mathbf{1}_V d\mu + \int_{\partial U} < X, n > f g \tilde{\mathbf{1}}_V(X) d\mathcal{H}^{n-1}.$$

Let U_ε be a C^1 open set approximating U according to proposition 8. We have:

$$J_t(U) - J_t(U_\varepsilon) = J_t(U \setminus U_\varepsilon) - J_t(U_\varepsilon \setminus U)\,,$$

hence we get by applying the inequality (2.8) for the last of the following inequalities

$$\begin{aligned}
&\limsup_{t \to 0} \frac{1}{|t|} |J_t(U) - J_t(U_\varepsilon) - J_0(U) - J_0(U_\varepsilon)| = \\
&\limsup_{t \to 0} \frac{1}{|t|} |J_t(U \setminus U_\varepsilon) - J_0(U \setminus U_\varepsilon) - (J_t(U_\varepsilon \setminus U) - J_0(U_\varepsilon \setminus U))| \\
&\leq \limsup_{t \to 0} \frac{1}{|t|} |J_t(U \setminus U_\varepsilon) - J_0(U \setminus U_\varepsilon)| + \limsup_{t \to 0} \frac{1}{|t|} |J_t(U_\varepsilon \setminus U) - J_0(U_\varepsilon \setminus U)| \\
&\leq \varepsilon(\text{Lip}(f)|X|_\infty + C|f|_\infty)|g|_\infty\,.
\end{aligned}$$

It implies that

$$|\limsup_{t \to 0_+} \frac{J_t(U) - J_0(U)}{t} \; - \; \partial_{s=0_+} J_s(U_\varepsilon)| \quad \leq \quad \varepsilon(\text{Lip}(f)|X|_\infty \; + \; C|f|_\infty)|g|_\infty\,.$$

In addition we have directly from the definition of $\partial_{s=0_+} J_s(U)$ that:

$$|\partial_{s=0_+} J_s(U) - \partial_{s=0_+} J_s(U_\varepsilon)| \leq \varepsilon(\text{Lip}(f)|X|_\infty + C|f|_\infty)|g|_\infty\,,$$

and by triangular inequality,

$$\begin{aligned}
|\limsup_{t \to 0_+} \frac{J_t(U) - J_0(U)}{t} - \partial_{s=0_+} J(U)| &\leq |\partial_{t=0_+} J(U) - \partial_{t=0_+} J(U_\varepsilon)| \\
&+ \varepsilon(\text{Lip}(f) + |f|_\infty)|g|_\infty |X|_\infty \leq 2\varepsilon(\text{Lip}(f)|X|_\infty + C|f|_\infty)|g|_\infty\,.
\end{aligned}$$

Hence, we get the result for U of Lipschitz regularity.

It remains to pass the formula to V Lipschitz. Again with the approximation of proposition 8, there exists V_ε such that $\mu(\bar{V} \setminus V_\varepsilon) + \mu(\bar{V_\varepsilon} \setminus V) < \varepsilon$ and $\mathcal{H}^{n-1}(\partial(\bar{V} \setminus V_\varepsilon)) + \mathcal{H}^{n-1}(\partial(\bar{V_\varepsilon} \setminus V)) < \varepsilon$. Now redefining some previous notation:

$$J_t(V) \doteq \int_{\phi_t(U)} f \circ \phi_t^{-1} g 1_V \, dx\,.$$

By the same argument than the one developed above, we just need to ensure that the formula pass to the limit. Thanks to a change of variable with $y = \phi_t(x)$, we swap the role of V and U. The only difference is that there is a divergence term in the formula but as X is assumed to be a Lipschitz vector field, this term is bounded by the Lipschitz constant of X. After the change of variable, we have

$$J_t = \int_{\phi_t^{-1}(V)} g \circ \phi_t f 1_U \mathrm{Jac}(\phi_t) dx \,,$$

hence we see that we need to control $\lim_{t\to 0} |\frac{\mathrm{Jac}(\phi_t)-1}{t}| = |\mathrm{div}(X)| \leq n\mathrm{Lip}(X)$. Finally we get the same kind of inequality 2.8 but for V. With the same arguments, we get the result. □

The final result on $\mathrm{Im}(M)$ is the following,

Theorem 12. *Let $(f, g) \in Im(M)^2$ be two images, X be a Lipschitz vector field on $\mathbb{R}^n$ and ϕ_t be the associated flow. Let us define:*

$$J_t = \int_M f \circ \phi_t^{-1}(x) g(x) d\mu(x),$$

then the semi-differentiation of J_t is

$$\partial_{t|t=0+} J_t = \int_M -\langle \nabla f, X\rangle g dx - \int (f_+ - f_-)\tilde{g}\langle \nu_f, X\rangle d\mathcal{H}^{n-1}. \tag{2.9}$$

with $\tilde{g}_X(x) := \lim_{t\mapsto 0^+} g(\phi_t(x))$ if the limit exists and if not, $\tilde{g}_X(x) = 0$.

Proof: Writing f as $f = \sum_{i=1}^k f \mathbf{1}_{x\in U_i}$ where $(U_i)_{i=1,\dots,k}$ is the partition in domains associated to f, and using the same expression for g, by linearity of integration, we fall in previous case. □

We can rewrite the equation (2.9) in a more compact form:

$$\partial_{t|t=0+} J_t = -\int_M \langle Df, X\rangle \tilde{g},$$

with Df the notation for the derivative for SBV function and $\tilde{g}$ is the function defined above. Remark that $\mu\ a.e.\ \tilde{g} = g$, these two functions differ on J_f. The second notation is the proper one to extend our results to the SBV case.

2.3 The case of SBV and BV functions

We aim to generalize this lemma in two directions. First the goal is to extend the formula to SBV or BV functions. Second we want to give a version of this lemma in order to be able to differentiate other terms than the square of the L^2 norm for the penalty term in the functional (2.1). Sometimes, we will work with a C^1 vector field instead of a Lipschitz vector field. This is not restrictive for current applications. On the other hand, Lipschitz assumptions are usual assumptions to define the flow ϕ_t. For the applications to geodesic equations the hypothesis could be slightly weaker, as discussed in Subsection 2.6.

2.3.1 Statement of the results and notations

We need some basic properties of BV functions. If g is a BV function, the precise representative of g is defined $\mathcal{H}^{n-1}$ a.e. We denote by (g^+, g^-, ν) the precise representative of g (g^+ and g^- being the upper and lower value at each point and ν the normal vector to the jump set, denoted by J_g, pointing in the direction of the upper value). To make the notations shorter we introduce the algebra B composed of bounded BV functions on $\mathbb{R}^n$ with compact support. If we denote $BV_c(\mathbb{R}^n)$ the subset of functions of compact support in $BV(\mathbb{R}^n)$, then $\mathrm{B} = BV_c(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$. In the definition below, a vector field is an mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$ without any further assumption.

Definition 4. *If X is a vector field on $\mathbb{R}^n$ and g a BV function, we define g_X by $g_X(x) = g(x)$ if $x \notin J_g$. On J_g, we define $\mathcal{H}^{n-1}$ a.e.*

- $g_X(x) = g^+(x)$ *if* $\langle \nu(x), X(x) \rangle > 0$,
- $g_X(x) = g^-(x)$ *if* $\langle \nu(x), X(x) \rangle < 0$,
- *else* $\langle \nu(x), X(x) \rangle = 0$ *and* $g_X(x) = \frac{g^-(x)+g^+(x)}{2}$.

Hence, g_X lies in $BV(\mathbb{R}^n) \times L^1(J_g; \mathcal{H}^{n-1})$.

Remark 5. *In order to make use of change of variables formulas, the action by a diffeomorphism ψ is given by*

$$(g \circ \psi)_X \circ \psi^{-1} = g_{d\psi(X \circ \psi^{-1})}.$$

Remark that if X is $\mathcal{H}^{n-1}$ measurable (for example for continuous vector fields), then g_X is also measurable. In this article this will always be the case, thanks to the Lipschitz or even C^1 assumptions we will use.

The main result of the paper is the following, which we will refer to as "semi-differentiation result". Here $\mathcal{S}_1$ and $\mathcal{S}_2$ represent functional spaces that will be detailed in the different extensions of the statement (they could be Lip^p, BV, SBV, B...).

To stress its generality, we give its statement in the time-dependent case. Yet, in the whole paper we will only deal with the autonomous case, but a remark will show how to extend the results to time-dependent vector fields. We will consider time dependent vector fields $X(t,x) : \mathbb{R} \times \mathbb{R}^n \mapsto \mathbb{R}^n$ which are continuous in both variables and Lipschitz in x (with, for simplicity a Lipschitz constant which does not depend on t). Mostly we will use the notation $X_t(x) = X(t,x)$ and X_t for the vector field at time t. For such a time dependent Lipschitz vector field, the flow is defined for all time. We will use the notation $(t,x) \in \mathbb{R} \times \mathbb{R}^n \mapsto \phi_t(x) \in \mathbb{R}^n$ for the flow generated by X. In the semi-differentiation result, if X is continuous in time, as we can expect, we will get X_0 instead of X. Obviously, we cannot expect such a result to be true if X is not continuous.

Theorem 13. *Let X be a Lipschitz time dependent vector field and ϕ_t be its associated flow. Take $f \in \mathcal{S}_1 = \mathrm{B}$ and $g \in \mathcal{S}_2 = \mathrm{B}$ with $\mathrm{B} = BV_c(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ the set of bounded*

BV functions with compact support. We define the functional

$$J_t(f,g) = \int_{\mathbb{R}^n} f \circ \phi_t^{-1}(x) g(x) dx \,, \tag{2.10}$$

then we have

$$\partial_{t=0^+} J_t = \int g_{X_0}(x) \langle -X_0(x), \partial f(dx) \rangle, \tag{2.11}$$

where ∂f stands for the distributional derivative of f, which is a finite vector measure.

The same result will also be extended to more general situation than simply the product fg but we can first state the theorem for our initial problem which was the differentiation of the L^2 distance $\int_U |f \circ \phi_t^{-1} - g|^2 d\mu$.

Theorem 14. *Let X be a Lipschitz time dependent vector field C^1 in space and ϕ_t be its associated flow, $(f,g) \in \mathrm{B} \times \mathrm{B}$ with $\mathrm{B} = BV_c(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$. We define the functional:*

$$J_t(f,g) = \int_{\mathbb{R}^n} |f \circ \phi_t^{-1}(x) - g(x)|^2 dx \,, \tag{2.12}$$

then we have

$$\partial_{t=0^+} J_t = -2 \int (f(x) - g_{X_0}(x) \langle \partial f, -X_0 \rangle dx \,.$$

In this equation, the notation f stands for the precise representative of f i.e. at a discontinuity point $f(x) = \frac{f^+(x)+f^-(x)}{2}$, which can be also written

$$\partial_{t=0^+} J_t = 2 \int_{\mathbb{R}^n} (f(x) - g(x)) \langle \nabla f, -X_0 \rangle dx + \int_{J_f} (f^+(x) + f^-(x) - 2 g_{X_0}(x)) \langle j(f), -X_0 \rangle d\mathcal{H}^{n-1} \,.$$

The generalization for a more general penalty term is the following

Theorem 15. *Let H be a locally Lipschitz function $H : \mathbb{R}^2 \mapsto \mathbb{R}$ and C^1 in the first variable, $(f,g) \in \mathrm{B} \times \mathrm{B}$ be two functions (with $\mathrm{B} = BV_c(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$), X be a Lipschitz time dependent vector field C^1 in space and ϕ_t its associated flow. We define the functional*

$$J_t(f,g) = \int_{\mathbb{R}^n} H(f \circ \phi_t^{-1}(x), g(x)) dx \,, \tag{2.13}$$

then we have

$$\partial_{t=0^+} J_t = \int \langle \partial_1 H(f(x), g_{X_0}(x)), -X_0 \rangle dx \,, \tag{2.14}$$

where $\partial_1 H(f, g_{X_0})$ is a part of the BV derivative of $H(f,g)$, defined by

$$\partial_1 H(f(x), l) = \partial_x H(f(x), l)(\nabla f(x) + D^c f(x)) + j_H(x) \mathcal{H}^{n-1} \llcorner J_f,$$

with $j_H(x) = \left(H\left(f^+(x), l\right) - H\left(f^-(x), l\right)\right) \nu_f(x)$ (the derivative of H w.r.t. to the first variable is denoted by $\partial_x H$).

Remark 6. *The notation $\partial_1 H(f, g_{X_0})$ is not the usual one and can be understood as follows: let f be a BV function and for any $l \in \mathbb{R}$, the function $x \to H(f(x), l)$ is a BV function as the composition of a Lipschitz function with a BV function. Hence we denote its distributional derivative $\partial_1 H(f(x), l)$.*

Remark first that since f and g are bounded, we can replace H with $\tilde{H}$ such that $\tilde{H}$ is Lipschitz on $\mathbb{R}^n$ and $H = \tilde{H}$ on $\mathrm{Im}(f) \times \mathrm{Im}(g)$. This allows to deal only with the case where H is globally Lipschitz on $\mathbb{R}^n$.

Back to the functional (2.1), the attachment term is the square of the L^2 norm. It is of importance for the applications to be able to differentiate other terms and this is the goal of Theorem 15. Yet, it could be interesting to cover a larger set of penalty terms. For example, our theorem allows to deal with all the L^p norms for $p > 1$, but it does not include $p = 1$ (even if the result seems true in this case as well).

On the contrary, the function $G^\alpha : (x, y) \mapsto |x - y|^\alpha$ is not locally Lipschitz on $\mathbb{R}^2$ for $\alpha \in]0, 1[$. This penalty term cannot be dealt with in a BV framework because the composition of a non Lipschitz function with a BV function may not be a BV function any more; in the case $\alpha < 1$, here there is an example of non differentiability. Take

$$H(x, y) = x^{\frac{1}{2}}, \quad f(t) = \begin{cases} \frac{1}{n^2} & \text{if } t \in [\frac{1}{2n+1}, \frac{1}{2n}], \\ 0 & \text{otherwise,} \end{cases}$$

then we have $J_0(f, f) = 0$ and for any fixed n_0, for t small enough we have

$$J_t(f, f) = \int_0^1 \sqrt{|f(x - t) - f(x)|} dx \geq 2t \sum_{n \leq n_0} \frac{1}{n},$$

which implies $\partial_{t=0^+} J_t = \sum_{n=1}^{\infty} \frac{1}{n} = +\infty$.

We will prove that this class of functions is dense in the strong BV topology in the space $SBV(U)$. This will be useful for extending the result by approximation. Actually, extending derivative results by approximation is always very delicate and there is in general no hope to succeed if no uniform estimate is shown. This is why weak approximation by regular functions will not be sufficient to prove the result for more general f and g.

Then we will present as a first result a suitable uniform estimate of $J_t(f, g) - J_0(f, g)$ (a very similar estimate will be used to prove that we can generalize to the case of continuous time-dependent vector fields). This estimate involves the BV and L^∞ norms of f and g, and justifies the need for strong approximations. All the section will be devoted to extensions of the semi-differentiation result thanks to density and approximation. Since we use strong convergence and SBV is a closed subspace of BV we could not hope, by means of this strategy, for more general results than SBV. Yet, there is sort of a duality between BV and L^∞ in this framework and it turns out that uniform approximations may work as well. This allows to present other extended results with continuous functions. In the end, we get the derivative result for functions which are sum of a continuous one and a SBV one. By the way, this raises an interesting question: can we hope for a decomposition result for arbitrary BV functions into the sum of an SBV and a continuous one? Obviously this is true in the one-dimensional case and it is what we exploit hereafter and we summarize the main results in dimension one as a consequence of what previously proven in any dimension. Thanks to the uniform approximation technique we may also extend the result to one of most natural frameworks in dimension one: the space of functions admitting right and left limit at any point (which is required to define g_X).

The result is now proven for a wide class of BV or BV+continuous functions in any dimension, and up to now the vector field X has been supposed to be Lipschitz continuous. Yet, there is some possibilities of extending it to arbitrary BV functions through a different technique.
The integral curves of the vector field X actually determine a partition of the space $\mathbb{R}^n$ into one-dimensional slices (this only works for autonomous vector fields) and it is worthwhile trying to prove the general result by slicing, through a suitable trivializing diffeomorphism that can transform these curves into straight lines. Yet, we need a slightly stronger assumption on X, which has to be C^1 in space. This is required because a Jacobian factor will appear and we would like to apply the one-dimensional semi-differentiation result to a product which will involve this factor: should it only be L^∞ (as it is the case for Lipschitz vector fields X) we could not, while if it is continuous we can, since we know the result in one dimension for all functions admitting right and left limits.
Thus, if we could recover the general BV result in any dimension through the one-dimensional one, what is the reason of previous sections ?

- First: we must notice that the proof in one dimension, up to some notational simplification, cannot be performed (to our knowledge) through techniques really simpler than the ones we presented in those sections;
- Second: the assumptions in the first part are stronger as far as f and g are concerned (SBV instead of BV) but weaker on X, since we require less regularity;
- Third: the generalization to time-dependent vector fields is more natural if we do not pass through the trivialization of the flow (since this is possible for autonomous systems only), and it could be performed without Lemma 8.
- Last, but not least: we feel that the results in these sections and especially the density results deserve their own attention and are interesting in themselves. Moreover, the whole framework let some questions on BV functions (such as the decomposition SBV + continuous) arise.

Finally, the last section presents an extension to the framework of Theorem 15.

2.3.2 Density lemmas

In this section, we prove the density of the set of Lipschitz piecewise functions in the SBV space. We suppose that U is a Lipschitz domain and all the functions we deal with are of compact support in U.
We want to prove that any SBV function g may be approximated strongly in the BV norm by some functions g_n in $Lip^p(U)$ and that the same sequence of function also gives $\mathcal{H}^{d-1}-$a.e. pointwise convergence of the functions $(g_n)_X$ to g_X (for a fixed vector field

X). Actually this last point will be stated with almost-everywhere convergence with respect to an arbitrary finite measure, absolutely continuous w.r.t. $\mathcal{H}^{d-1}$, due to metrizability conditions. Anyway, this is sufficient for letting the approximation procedure of the next section work.

If G is a Lipschitz graph in U such that $d(G, \partial U) > 0$, we denote by $SBV_G(U) = \{f \in SBV(U) | J(f) \subset G\}$ the set of functions of $SBV(U)$ whose jump set is included in G. We also denote by $SBV_{G,\delta}(U) = \{f \in SBV(U) || D^s f|(U \setminus G) < \delta\}$ the functions in $SBV(U)$ whose jumps occur on G up to small measure jumps.
There are basically two steps in the proof:

1. $Lip^p(U)$ is dense in $SBV_G(U)$.

2. $\mathrm{Vect}(SBV_G(U), G)$ is dense in $SBV(U)$.

Lemma 4. *Let f be a function in $SBV_G(U)$: there exists a sequence $(u_n)_{n\in\mathbb{N}} \subset Lip^p(U)$ such that*

$$u_n \to f \text{ in } BV \text{ and } (u_n)_X \to f_X \text{ pointwisely } \mathcal{H}^{d-1} - a.e.$$

Moreover, if $f \in SBV_{G,\delta}(U)$, then there exists a sequence $(u_n)_{n\in\mathbb{N}} \subset Lip^p(U)$ such that

$$\limsup_n ||f - u_n||_{BV} \leq 2\delta \text{ and } (u_n)_X \to f_X \text{ pointwisely } \mathcal{H}^{d-1} - a.e. \text{ outside } Jf \setminus G.$$

Proof. Let us start from the case $f \in SBV_G(U)$.
First, we can find a partition in Lipschitz domains (V_1, V_2) of U such that $G \subset \partial V_1 \cap \partial V_2 := \Gamma$. This is an application of the Lipschitz extension theorem: indeed, there exists an extension of the Lipschitz graph on $\mathbb{R}^n$, which gives a partition on $\mathbb{R}^n$ and by restriction to U, a candidate partition. Be careful that this partition may not be a Lipschitz partition (but this can only be the case if ∇G and $\nabla \partial U$ are collinear). The hypothesis $d(G, \partial U) > 0$ ensures that a small perturbation in a neighborhood of ∂U of such an extension gives the desired partition.

Consider now $f \in SBV_G(U)$, it gives by restriction $f_i = f_{|V_i} \in W^{1,1}(V_i)$ for $i = 1, 2$.
From Meyers-Serrin theorem we know that any $W^{1,1}$ function may be approximated in $W^{1,1}$ by means of its convolutions (after performing an extension of the function itself beyond the boundary, which requires Lipschitz assumption on the domain, which is exactly the situation we face here). Moreover, by the continuity of the trace map from $W^{1,1}(\Omega)$ to $L^1(\partial\Omega, \mathcal{H}^{d-1})$, we may infer from the strong convergence in $W^{1,1}$ the strong L^1 convergence on the boundary. Let us perform this convolution separately on the two domains V_1 ad V_2, thus obtaining two sequences of regular functions which do not coincide on the common boundary Γ. Then, we glue the two functions and, with [AFP00], we get functions u_n in $Lip^p(U) \subset SBV(U)$. The BV distance of this function to f may be estimated by the sum of the BV distances in the domains V_1 and V_2 (which are actually

$W^{1,1}$ distances, since no singular part of the derivative is involved) and the L^1 distance between the jumps of f and of u_n on Γ, i.e.

$$||(u_n^+ - u_n^-) - (f^+ - f^-)||_1 = || \, |u_n^{(1)} - u_n^{(2)}| - |f^{(1)} - f^{(2)}| \, ||_1 \\ \leq ||u_n^{(1)} - f^{(1)}||_1 + ||u_n^{(2)} - f^{(2)}||_1$$

(the superscripts (1) and (2) at u and f stand for the values on the two different sides of Γ, while the superscripts $+$ and $-$ stand for the upper and the lower of these two values, respectively). Since these last L^1 norms converge to zero as the $W^{1,1}$ norm goes to zero, we get the that u_n converges to f in the BV distance. Obviously, u_n belongs to $Lip^p(U)$ since it is composed of two regular functions glued together on a Lipschitz boundary. Moreover, from the general theory on BV functions (see Evans and Gariepy, [EG92]), we know that the convolution regularizations also converge pointwisely $\mathcal{H}^{d-1}-$a.e. outside the jump set of the limit function g (and on this set they converge to the average between the upper and lower value of g). This gives $\mathcal{H}^{d-1}-$a.e. convergence of $(u_n)_X$ to g_X outside Γ (since outside Γ the functions u_n and g agree with $(u_n)_X$ and g_X, respectively). The convergence on Γ is easily deduced from the fact that the values $(u_n)_X$ and g_X agree with the values of u_n and g on one side of Γ (up to the points where Γ and X are parallel, where these quantities are not well defined): the boundary values of u_n converge to those of g as a consequence of the L^1 convergence on Γ (with respect to the $(d-1)-$dimensional measure).
Hence the thesis is easily obtained in the first case $f \in SBV_G(U)$.
The general case is considered in the same way: divide U into two domains and consider again convolutions. The only difference lies in the fact that we do not have any more $W^{1,1}$ but BV functions. This means that the convolutions do not converge strongly neither pointwisely. Yet, the pointwise convergence stays true outside Jf, and Lemma 5 gives the estimate of the BV distance between a suitable convolution u_n and f. □

Remark that we don't have obtained $Ju_n \subset G$. Yet, we could have slightly modified the proof to get $J(u_n) \subset G$. Anyway, we do not need it in the following.

The following Lemma has been applied in Lemma 4 to the case $\Omega = V_1$ and $\Omega = V_2$.

Lemma 5. *Let $f \in SBV(\Omega)$ be a function with jump set denoted by Jf: then we may obtain by convolution a sequence of smooth function u_n on Ω, such that $\limsup_n |f - u_n|_{BV} \leq 2|D^s f|$ and $u_n \to f$ pointwisely $\mathcal{H}^{d-1}-$a.e. outside Jf.*

Proof. The convolution with a sequence of mollifiers ρ_n gives the result. First, remark that we have $\lim_n ||f * \rho_n - f||_{L^1} = 0$, and then:

$$|\partial(f * \rho_n) - \partial f|(U) = |(D^a f) * \rho_n + (D^s f) * \rho_n - \partial f|(U)$$
$$|\partial(f * \rho_n) - \partial f|(U) \leq ||(D^a f) * \rho_n - D^a f||_{L^1} + ||(D^s f) * \rho_n||_{L^1} + |D^s f|(U)$$

With the theorem of approximation of BV functions by smooth functions ([EG92]), we know that $\lim_n |\partial(f*\rho_n)|(U) = |\partial f|(U)$. Moreover, $\lim ||(D^a f)*\rho_n - D^a f||_{L^1} = 0$ (thanks to the behaviour of convolutions on L^1 functions), and, as a result of the two preceding assertions, we get $\lim ||(D^s f)*\rho_n||_{L^1} = |D^s f|(U)$. Finally we get

$$\limsup_n |\partial(f*\rho_n) - \partial f|(U) \leq \lim_n ||(D^a f)*\rho_n - D^a f||_{L^1} + \lim_n ||(D^s f)*\rho_n||_{L^1} + |D^s f|(U) = 2|D^s f|(U).$$

As in the previous Lemma, we know that the mollified functions $u_n = f*\rho_n$ converge $\mathcal{H}^{d-1}-$a.e. to f outside Jf. **If we notice that outside the jump set f and f_X coincide (and, since the functions u_n are regular, they coincide everywhere with $(u_n)_X$), we get the thesis.** □

Now, we can prove the second lemma:

Lemma 6. *For any function f in $SBV(U)$ there exists a sequence $(u_n)_n \subset Lip^p(U)$ such that $u_n \to f$ in BV. Moreover, for any finite measure $\mu << \mathcal{H}^{d-1}$ this sequence may be chosen so that $(u_n)_X \to f_X$ pointwisely $\mu-$a.e.*

Proof. We know ([EG92]), that the jump set of f is a countable union of compact Lipschitz graph. Hence, we enumerate the Lipschitz graphs involved: $(G_i)_{i\in\mathbb{N}}$. Take $\delta > 0$: there exists an integer N such that $\sum \int_{G_i:\, i>N} |D^s f| \leq \delta$.
We can suppose that $G_i \cap G_j = \emptyset$ for all couples $(i,j) \in [0,N]^2$ with $i \neq j$ (actually, two Lipschitz graphs G and H may be always replaced with a new pair G and $\tilde{H}$, so that $\tilde{H}$ is a finite union of Lipschitz graphs, $G \cap \tilde{H} = \emptyset$ and $\mathcal{H}^{d-1}(\{H \setminus \tilde{H}\})$ is as small as we want).
Take a smooth partition of unity $(\psi_0, \psi_1, \ldots, \psi_N)$ such that, for $i \geq 1$, we have $\psi_i = 1$ on G_i and $\psi_i = 0$ on G_k for $k \neq i$, $1 \leq k \leq N$. Writing $f = \sum_{i=0}^N f\psi_i$, we can restrict our interest to $f_i = f\psi_i$, which is again an SBV function.
Each function f_i has a jump set which is composed by two parts: one part is included in G_i, while the second one has a "small" jump, since $|D^s f_i|(U \setminus G_i) = \int_{U\setminus G_i} \psi_i |D^s f| = \int_{U\setminus \bigcup_{j=1}^N G_j} \psi_i |D^s f|$. Moreover $\sum_{i=0}^N \int_{U\setminus\bigcup_{j=1}^N G_j} \psi_i |D^s f| = \sum_{i=0}^N |D^s f_i|(U\setminus\bigcup_{j=1}^N G_j) \leq \delta$.
Thanks to Lemma 4, we may approximate each function f_i through functions $u_n^{(i)} \in Lip^p(U)$. By summing together the results (and the functions) we get a sequence of functions u_n in $Lip^p(U)$ such that

$$\limsup_n ||u_n - f||_{BV} \leq 2\delta, \quad \text{and } u_n \to f \text{ pointwisely } \mathcal{H}^{d-1} - a.e. \text{ outside } \bigcup_{j>N} G_j.$$

The result can then be obtained if we take a function u_n from this sequence so that $||u_n - f||_{BV} < 3\delta$ and then we repeat the same construction with smaller values of δ, thus getting a sequence converging to f in BV. Yet, this cannot be performed for the

pointwise convergence, because we need a metric to do that. Actually, we know that almost everywhere convergence w.r.t. a measure is a metrizable convergence, provided the measure is finite (or $\sigma-$finite), which is not the case for $\mathcal{H}^{d-1}$. This is why we introduced μ.
If, from the very beginning, we choose N so that $\mathcal{H}^{d-1}\left(\bigcup_{j>N} G_j\right)$ is sufficiently small (and hence $\mu\left(\bigcup_{j>N} G_j\right)$ is sufficiently small as well, say smaller than δ, then we can select a function u_n so that $d(u_n, g) < 2\delta$ (d being for instance the distance in probability $d(f,g) = \inf\{\varepsilon : \mu(\{|f-g| > \varepsilon\}) < \varepsilon\}$) and go on with the same procedure as before. □

2.3.3 SBV and continuous functions in any dimension

This Section presents wider and wider generalizations of the semi-differentiation result thanks to approximations techniques. The first tool we will use is the following lemma.

Lemma 7. *For $t \le t_0$, if $g \in L^\infty(U)$ and $f \in BV(U)$, we have*

$$|J_t(f,g) - J_0(f,g)| \le \int_U |f\circ(\phi_t^{-1}) - f|\,|g|\,dx \le Ct||g||_\infty||f||_{BV}$$

for a constant C which only depends on the vector field X and on t_0.
On a subset $A \subset U$, the same result is true for the functional $J_t(A; f, g)$:

$$\int_A |f\circ(\phi_t^{-1}) - f|\,|g|\,dx \le Ct||X||_{L^\infty(A_{Ct})}||g||_\infty||f||_{BV},$$

where A_ε is $\{x \in U : d(x,A) < \varepsilon\}$ and C is again a constant which only depends on the vector field X and on t_0.
Analogously, if on the contrary $f \in L^\infty(U)$ and $g \in BV(U)$, then we have

$$|J_t(f,g) - J_0(f,g)| \le Ct||f||_\infty||g||_{BV}.$$

Proof. Let us start from the case $g \in L^\infty$ and $f \in C^1$. Just consider

$$\int_U |f\circ(\phi_t^{-1}) - f|\,|g|\,dx \le \int_0^t\int_U |\nabla f\circ\phi_s|\,|X\circ\phi_s|\,|g|\,dx\,ds = \int_0^t\int_U |\nabla f|\,|X|\,|g\circ(\phi_s)^{-1}|\,|\mathrm{Jac}(\phi_s)^{-1}|\,dx\,ds$$

and then use the fact that, since for small s the map $(\phi_s)^{-1}$ is close to the identity, the Jacobian $|J(\phi_s)^{-1}|$ is close to one, and hence bounded. Then estimate the second member by

$$||X||_\infty \sup_{s\in[0,t]} ||J(\phi_s)^{-1}||_\infty||g||_\infty \int |\nabla f|.$$

The proof in the case $A \subset U$ is similar, since we estimate

$$\int_0^t\int_A |\nabla f\circ\phi_s||X\circ\phi_s|\,|g|\,dx\,ds \le \int_0^t\int_{A_{Ct}} |\nabla f||X|\,|g\circ(\phi_s)^{-1}||\mathrm{Jac}(\phi_s)^{-1}|\,dx\,ds$$

where, in the change of variable, we do not exit the set A_{Ct}, provided $||X||_\infty \leq C$. Then we go on with the same estimates.
To pass to the general case $f \in BV(U)$ it is always sufficient to choose a sequence of regular functions $(f_k)_k$ which converges to f in L^1 with the additional property $||f_k||_{BV} \to ||f||_{BV}$ (see [EG92]) and let the previous estimate pass to the limit.
The estimate in the opposite case is only a little bit trickier. Let us perform a change of variables so that

$$J_t(f,g) = \int_U f \circ \phi_t^{-1} g dx = \int_U fg \circ \phi_t \mathrm{Jac}(\phi_t) dx$$
$$= \int_U fg \circ \phi_t dx + \int_U fg \circ \phi_t (\mathrm{Jac}(\phi_t) - 1) dx.$$

The difference between first term in the last sum and $J_0(f,g)$ may estimated as before (replacing the vector field X with $-X$) by $Ct||f||_\infty ||g||_{BV}$. The second may be estimated, after a new change of variable with bounded Jacobian, by $||f||_\infty ||g||_1 ||\mathrm{Jac}(\phi_t) - 1||_\infty$. The thesis is obtained as far as one notices $|\mathrm{Jac}(\phi_t) - 1| \leq Ct$ and $||g||_1 \leq ||g||_{BV}$. □

In a very analogous way we can prove the following reduction lemma, that we referred to in the introductory sections. Its aim is proving that the result is true for time-dependent vector fields (continous in t), if it is true for the autonomous case.

Lemma 8. *If ψ_t denotes the usual flow associated to a time dependent vector field $X = X(t,x)$ (that we suppose continuous in time and C^1 in space) and ϕ_t the flow associated to the (constant in time) vector field $X_0 = X(0,\cdot)$, then we have*

$$\int_U \left(f \circ (\phi_t^{-1}) - f \circ (\psi_t^{-1})\right) \, g \, dx = o(t).$$

Proof. Set $\chi_t := \phi_t^{-1} \circ \psi_t$ and let Y be the vector field hidden behind the flow χ_t, i.e.

$$\dot{\chi_t} = Y(t, \chi_t)$$

(such a field Y exists since χ_t is a diffeomorphism). As usual, we estimate the difference by

$$\int_U \left(f \circ (\phi_t^{-1}) - f \circ (\psi_t^{-1})\right) \, g \, dx = \int_U \left(f \circ (\phi_t^{-1} \circ \psi_t) - f\right) \, g \circ \psi_t \, |\mathrm{Jac}(\psi_t)| \, dx$$
$$\leq \int_U \int_0^t |\nabla f \circ \chi_s| ||Y \circ \chi_s|| g \circ \psi_t| \, \mathrm{Jac}(\psi_t) \, dx$$
$$\leq t||g||_{L^\infty} ||\mathrm{Jac}\psi||_{L^\infty} ||\mathrm{Jac}\chi^{-1}||_{L^\infty} ||f||_{BV} ||Y||_{L^\infty(U \times [0,t])}.$$

The only thing that we need to conclude is to prove that $||Y||_{L^\infty(U\times[0,t])} \to 0$. Let us look for a while at the regularity of ϕ_t and ψ_t: as a consequence of the assumptions on X they are both C^1 functions. Hence $\chi \in C^1$ and $Y \in C^0$. This implies that the condition $Y(0,\cdot) = 0$ is sufficient to imply $\lim_{t\to 0} ||Y||_{L^\infty(U\times[0,t])} = 0$. We can compute

$$\dot{\chi_t} = \frac{\partial(\phi_t^{-1})}{\partial t} + \nabla_x(\phi_t^{-1}) \cdot X(t, \psi_t).$$

If we take $t = 0$, remembering $\phi_t^{-1} = \phi_{-t}$ (since ϕ is the flow of an autonomous vector field) and $\phi_0 = id$, we get

$$Y(0,\cdot) = \dot{\chi}_{t|t=0} = -X_0 + Id \cdot X(0,\cdot) = 0.$$

and this concludes the proof. □

Lemma 9. *If $g \in Lip^p(U)$ and $f \in SBV(U)$, then the semi-differentiation result is true.*

Proof. Take a sequence $f_k \in Lip^p(U)$ converging to f in $BV(U)$. Write $f = f_k + r_k$ with $||r_k||_{BV} \to 0$. By linearity, we have $J_t(f,g) = J_t(f_k,g) + J_t(r_k,g)$. Since we know the derivative of the first term and we can estimate the second we have

$$\limsup_{t\to 0} \frac{J_t(f,g) - J_0(f,g)}{t} \leq \int_U \langle \partial f_k, -X\rangle \, g_X \, dx + C||g||_\infty ||r_k||_{BV}$$

and analogously

$$\liminf_{t\to 0} \frac{J_t(f,g) - J_0(f,g)}{t} \geq \int_U \langle \partial f_k, X\rangle \, g_X \, dx - C||g||_\infty ||r_k||_{BV}.$$

When we let k go to infinity, the last term of both inequalities vanishes, while for the first we have the convergence

$$\int_U \langle \partial f_k, X\rangle \, g_X \, dx \to \int_U \langle \partial f, X\rangle \, g_X \, dx.$$

This convergence is a consequence of the fact that the derivatives of f_k strongly converge as measures to the derivative of f, and this is enough to integrate it against any measurable bounded function (such as Xg_X).
These estimates finally imply the existence of the derivative of $J_t(f,g)$ and its equality with the desired formula. □

Lemma 10. *If $g \in SBV(U) \cap L^\infty(U)$ and $f \in SBV(U) \cap L^\infty(U)$, then the differentiation result is true.*

Proof. The proof of this last step is very similar to the previous one. But now f will be fixed and we will strongly approximate g in BV by a sequence g_k of functions of $Lip^p(U)$. As before, we need to check two facts; first that the remainders $(J_t(f, g - g_k) - J_0(f, g - g_k))/t$ may be made as small as we want; second that the quantities $\int_U \partial f \cdot X(g_k)_X$ actually converge to $\int_U \partial f \cdot Xg_X$. For the first point, we will use the second estimate in Lemma 7. For the second, we only need pointwise convergence $|\partial f|-$ a.e. of $(g_k)_X$ to g_X. The fact that we can satisfy this condition by properly choosing the sequence is ensured by Lemma 6. □

Lemma 11. *If $g \in SBV(U) \cap L^\infty(U)$ and $f \in SBV(U)$, then the differentiation result is true.*

Proof. In this case we fix g and approximate f by a sequence of bounded BV functions. Take $f_k = H_k \circ f$, for some functions H_k satisfying: $H_k(z) = z$ for $|z| \leq k-1,\ 0 \leq H_k' \leq 1; |H_k(z)| \leq k \vee |z|,\ H_k \in C^1$.
We will use the chain rule for BV functions (see Ambrosio Fusco Pallara, Theorem 3.96) so that we can easily get $f_k \to f$ in $BV(U)$. This is enough to use the same arguments as in Lemma 9 and extend the result to such a framework. □

Up to now we have provided results only in the case where both functions f and g belong to $SBV(U)$ and the main reason lies in the fact that we used strong approximation in BV by means of functions in $Lip^p(U)$ and those functions all belong to $SBV(U)$, which is a closed subset of $BV(U)$. Anyway, by means of different and much simpler method it is possible to handle the case where f is a generic BV function and g is continuous.

Lemma 12. *Suppose that $f \in BV(U)$ and $g \in C^\infty(U)$, then the differentiation result is true.*

Proof. We have, by change of variables

$$\begin{aligned} J_t(g) &= \int_U f(x)\ g \circ \phi_t(x)\ \mathrm{Jac}(\phi_t) d\mu(x), \\ \partial_{|t=0} J_t(g) &= \int_U f \langle \nabla g, X \rangle + g\ (\nabla \cdot X)\, dx, \\ \partial_{|t=0} J_t(g) &= \int_U [f \nabla \cdot (gX)]\, dx, \\ \partial_{|t=0} J_t(g) &= -\int \langle \partial f,\, gX \rangle. \quad \square \end{aligned}$$

This proves, as we may have expected, that a sufficiently strong regularity in one of the two functions can compensate weaker assumptions on the other. We go on extending the result to functions g which are only continuous.

Lemma 13. *Suppose that $f \in BV(U)$ and $g \in C^0(U)$, then the differentiation result is true.*

Proof. Approximate uniformly g by a sequence g_k of C^∞ functions. As usual, one only needs to manage the remainder (and this is done thanks to Lemma 7 since we have $||g_k - g||_\infty \to 0$), and to have convergence of the derivative terms. This last convergence is true since ∂f is a fixed finite measure and hence uniform convergence is sufficient (actually pointwise dominated convergence would have been enough). □

A possible interest of the extension to a BV-continuous setting lies in the following question: is it true that all BV functions may be decomposed as the sum of a continuous and an SBV function? this is true in dimension one and it has a priori no hope to be true in higher dimension where even $W^{1,1}$ do not need to be continuous. Yet, possible discontinuities due to this kind of behavior could be inserted in the SBV part. This question is obviously interesting in itself and does not seem being treated in the literature. We thank

Giovanni Alberti for a brief discussion on the subject.
Here in this context a decomposition $f = f_c + f_s$ such as this one, even if with no additional property mimicking what happens in dimension one (i.e. without requiring neither estimates on the dependence of f_c and f_s on f nor any properties on the derivatives of the two addends), would allow to generalize the differentiation result to any pair (f, g) of BV functions, thanks to the following statement.

Theorem 16. *Suppose that $f = f_c + f_s$ and $g = g_c + g_s$ with $f_c,\ g_c \in BV(U) \cap C^0(U)$ and $f_s,\ g_s \in SBV(U) \cap L^\infty(U)$, then the differentiation result is true.*

Proof. We have to manage and differentiate four terms which are precisely $J_t(f_c, g_c)$, $J_t(f_c, g_s)$, $J_t(f_s, g_c)$ and $J_t(f_s, g_s)$. The thesis is proven if we prove, for all of them, that the derivative is given by the $\int_U \langle \partial f, X\rangle\, g_X$, being f and g replaced by their continuous or SBV parts.
The term $J_t(f_s, g_s)$ does not give any problem since its derivative has been the object of the proof of Lemma 10. The terms $J_t(f_s, g_c)$ and $J_t(f_c, g_c)$ can be dealt with thanks to Lemma 13. We need to look at the term $J_t(f_c, g_s)$ which does not fit into the the frameworks we considered so far. The idea is to switch the roles of f_c and g_s.
Notice

$$\begin{aligned} J_t(f_c, g_s) &= \int_U f_c \circ \phi_t^{-1} g_s dx = \int_U f_c\ g_s \circ \phi_t\ \mathrm{Jac}(\phi_t) dx \\ &= \int_U f_c\, g_s \circ \phi_t dx \\ &+ \int_U f_c g_s (\mathrm{Jac}(\phi_t) - 1) dx + \int_U f_c (g_s \circ \phi_t - g_s)(\mathrm{Jac}(\phi_t) - 1) dx. \end{aligned}$$

The first term in the last sum may be differentiated as usual, replacing the vector field X with $-X$, thanks to the result in Lemma 13. Its derivative gives $\int_U \langle \partial g_s, f_c\rangle$. The second may be differentiated pointwisely since the only part depending on t is the Jacobian, and the derivative is $\int_U f_c\, g_s\, (\nabla \cdot X)\ dx$. For the third term we have (thanks to the second estimate in Lemma 7)

$$\left| \int_U f_c (g_s \circ \phi_t - g_s)(\mathrm{Jac}(\phi_t) - 1) dx \right| \leq Ct \int_U |f_c|\, |g_s \circ \phi_t - g_s|\, dx \leq Ct^2 ||f_c||_\infty ||g_s||_{BV},$$

and thus its contribution to the derivative at $t = 0$ is zero.
This means that we have

$$\frac{d}{dt} J_t(f_c, g_s) = \int_U \langle \partial g_s, f_c\rangle + \int_U f_c\, g_s\, (\nabla \cdot X)\,.$$

We want this sum to equal $\int \langle \partial f_c, -X\rangle (g_s)_X = \int \langle \partial f_c, -X\rangle g_s$ (replacing $(g_s)_X$ with g_s itself is allowed since ∂f_c does not give mass to $(d-1)-$dimensional sets, as a consequence of f_c being continuous). To get this equality it is sufficient to integrate by part and use the product rules for the derivatives of $g_s\, X$ (a product of a BV function and a Lipschitz vector field). □

Remark 7. *The same techniques of the last proofs could be used to prove a statement such as the following: if the differentiation result is true for $f \in BV$ and g belonging to a certain functional class $\mathcal{S}$, then the same result stays true if g belongs to the closure of $\mathcal{S}$ for the uniform convergence.*

Yet, we will not develop this remark here in this section, since to be precise we should check that the meaning of g_X stays well-defined for the functions that we obtain as uniform limits of BV functions. Actually, g_X was defined for BV functions g and, by uniform convergence, we may go out of this functional space. We will develop the same concept in Section 4 in the one-dimensional case, which is easier to treat and simpler definitions may be given.

2.4 The one-dimensional case

The results from Section 3 may obviously be applied to the case of dimension one, which is actually much simpler. The main peculiarity is that in dimension one it is true and easy that every BV function is the sum of SBV and continuous functions.
In this section U will be a compact interval of $\mathbb{R}$ and we will always replace a BV function on U with its precise representative. Equivalently, we will use the definition of one-dimensional bounded variation functions through the total variation as a supremum over partitions (and not as a distributional object). This means that one-dimensional BV functions will be defined pointwisely, not only almost everywhere.
For a function $g \in BV(U)$, an equivalent definition for g_X will be the following:

$$g_X(x_0) = \begin{cases} \lim_{x \to x_0^+} g(x) & \text{if } X(x_0) > 0, \\ \lim_{x \to x_0^-} g(x) & \text{if } X(x_0) < 0, \\ g(x_0) & \text{if } X(x_0) = 0. \end{cases} \tag{2.15}$$

Theorem 17. *Suppose that f, $g \in BV(U)$ ($U \subset \mathbb{R}$): then the differentiation result is true.*

Proof. It is sufficient to apply Theorem 16 since any BV function in one dimension is the sum of an SBV function and a continuous one. This can be easily obtained by taking the cumulative distribution function of the Cantor part of the derivative, which will be a continuous BV function whose derivative is exactly the Cantor part of the original derivative. The remainder is an SBV function by construction. □

In this one-dimensional setting, we will be able to extend further the result thanks to Remark 7. The following closure result is quite known.

Theorem 18. *The set of functions on U which are uniform limits of BV functions is the following vector space $RL(U)$:*

$$RL(U) = \{f : U \to \mathbb{R} \,:\, f \text{ admits right and left limits at every point of } U\}.$$

Proof. In dimension one the property of admitting limits on the two sides is satisfied by any BV function (since, if x_h converges monotonely to x then $\sum_h |f(x_{h+1}) - f(x_h)|$ is finite and $f(x_h)$ is a Cauchy sequence).
This property is preserved by uniform convergence and hence any function which is a uniform limit of BV functions belongs to $RL(U)$.
On the other hand, if $f \in RL(U)$, for every $\varepsilon > 0$ and every $x \in U$, there exists a neighborhood $V_{\varepsilon,x} =]a_{\varepsilon,x}, b_{\varepsilon,x}[$ of x such that the oscillations of f on $]a_{\varepsilon,x}, x[$ and on $]x, b_{\varepsilon,x}[$ are smaller than ε. By compactness one can cover U by a finite number of these intervals and hence, by considering the left parts of these intervals, the right ones and the central points, one covers U by a finite number of intervals (or points) where the oscillation of f is smaller than ε.
A function g constant on each one of these intervals may be built so that $||f - g||_\infty < \varepsilon$ and g is BV. This proves the density of BV in $RL(U)$. □

Corollary 1. *Suppose that $f \in BV(U)$ and $g \in RL(U)$: then the differentiation result is true.*

Proof. It is sufficient to proceed by approximations as in Lemmas 9 and 10. the function f belongs to $BV(U)$ and is fixed and g will be approximated by a sequence $(g_k)_k$ of BV functions thanks to Theorem 18. □

It is interesting to notice that the space $RL(U)$ is exactly the natural space for the function g, since it is the largest space where the definition of g_X given in (2.15) makes sense.

2.5 Reduction to the one-dimensional case

We turn now to the proof of the differentiation results for BV functions in dimension d, thanks to a one-dimensional reduction technique. The proof of the reduction is based on two arguments:

- The box flow Theorem **??**,
- the one dimensional restriction of BV functions.

We begin with a remark which states that the property is local. Thus we only have to focus on the open set of non-equilibrium points of the vector field (i.e. where the vector field is non zero).

Remark 8. *If for each $x \in U$ such that $X(x) \neq 0$, there exists a neighborhood V of x such that the result is true for $J_t(V; f, g)$, with*

$$J_t(V; f, g) = \int_V f \circ \phi_t^{-1}(x) g(x) dx,$$

then the result is true.

To prove this remark, consider $Z_\varepsilon = \{y \in U; |X(y)| \geq \varepsilon\}$ which is a compact subset (we denote by Z_ε^c its complement in U). By means of a finite covering the result is true on Z_ε. Then, we need to control what happens on Z_ε^c: this is easy by lemma 7 for $t \leq t_0$:

$$\left|\frac{J_t(Z_\varepsilon^c; f, g) - J_0(Z_\varepsilon^c; f, g)}{t}\right| \leq C(\varepsilon + L||X||_\infty 2t),$$

where L is the Lipschitz constant of X and we used the fact that, if we move from Z_ε^c no more than $t||X||_\infty$, then the value of $|X|$ does not increase more than $tL||X||_\infty$.
This implies that, when we divide by t and let $t \to 0$, we get

$$\limsup_{t\to 0} \left|\frac{J_t(U; f, g) - J_0(U; f, g)}{t} - \int_U \langle \partial f, -X\rangle g_X\right| \leq \left|\int_U \langle \partial f, -X\rangle g_X - \int_{Z_\varepsilon} \langle \partial f, -X\rangle g_X\right| + C\varepsilon.$$

Letting $\varepsilon \to 0$ we get in the end the result we want, since $\int_{Z_\varepsilon} \partial f \cdot X g_X \to \int_{\{X\neq 0\}} \partial f \cdot X g_X = \int_U \partial f \cdot X g_X$.

It is now sufficient to prove the result in the neighborhood of a point $z_0 \in U$ such that $X(z_0) \neq 0$. In that case, we would like to use the standard reduction of the flow of vector field: up to a C^1 diffeomorphism, we just need to treat the case where $\phi_t(x) = x + t\nu$ in a neighborhood of z_0. It has even been proven in the Lipschitz case by Calcaterra and Boldt ([CB03]). (Notice by the way that, with standard arguments, a time dependent vector field can be viewed as a vector field on $\mathbb{R} \times \mathbb{R}^n$ of the same regularity).

Here a slightly different version of the box-flow theorem is presented.

Theorem 19. *If X is a C^1 vector field on $\mathbb{R}^n$ and $x_0 \in \mathbb{R}^n$ is a point such that $X(x_0) \neq 0$, then there exist a neighborhood V of x_0, a neighborhood U of 0, a vector $v \in \mathbb{R}^n$, and a C^1 diffeomorphism $\psi : U \mapsto V$ such that for any $x \in U$ we have*

$$\phi_t \circ \psi(x) = \psi(x + tv) \tag{2.16}$$

for t such that $x + tv \in U$.

Proof. Since this result is well known, we just give here ideas of the proof. Through a diffeomorphism we can assume that $x_0 = 0$, $X(0) = v = e_1$ where $(e_1, \ldots, e_n)$ a basis of $\mathbb{R}^n$. Defining

$$\psi(x_1, \ldots, x_n) = \phi_{x_1}(0, x_2, \ldots, x_n)\,, \tag{2.17}$$

we have the result: the regularity of ψ is obtained due to the regularity of the flow ϕ_t which is C^1. It can be checked that differentiating ψ at 0 gives the identity map Id and by the inverse function theorem, there exists a neighborhood of 0 such that ψ is a local C^1 diffeomorphism.
We get the desired result by writing,

$$\phi_t \circ \psi(x_1, \ldots, x_n) = \phi_t \circ \phi_{x_1}(0, x_2, \ldots, x_n) = \phi_{t+x_1}(0, x_2, \ldots, x_n) = \psi(x + te_1)\,.$$

□

Deriving in time the equality (2.16) we have

$$(X \circ \phi_t \circ \psi) = \nabla\psi(x+tv)\cdot v\,. \tag{2.18}$$

Proposition 9. *Let $(f,g) \in BV(U)$ be two functions, X be a C^1 vector field on $\mathbb{R}^n$ and ϕ_t be the associated flow. Set*

$$J_t = \int_U f\circ\phi_t^{-1}(x)g(x)dx,$$

then the differentiation of J_t gives:

$$\partial_{t|t=0+}J_t = \int < \partial f, -X > g_X\,. \tag{2.19}$$

Proof. Trivializing the flow with the box-flow theorem, we obtain through a change of variables in the functional J_t,

$$J_t = \int_{\mathbb{R}^n} f\circ\psi(x-tv)\, g\circ\psi\,\mathrm{Jac}(\psi)dx.$$

The product $h = g\circ\psi\,\mathrm{Jac}(\psi)$ is not a BV function because $\mathrm{Jac}(\psi)$ is only continuous. However, we have extended our results to continuous functions.

Set $H = \nu^\perp$, the orthogonal hyperplane to ν, and by using coordinates $x = (x', h)$, write

$$\begin{aligned} J_t - J_0 &= \int_{\mathbb{R}^n} (f\circ\psi(x-t\nu) - f\circ\psi(x))h(x)dx\,, \\ &= \int_{H=\mathbb{R}^{n-1}}\int_{\mathbb{R}} (f\circ\psi(x'+(h-t)\nu) - f\circ\psi(x'+h\nu))h(x'+h\nu)dx'dh\,. \end{aligned}$$

We want to differentiate under the integral in dx'. Set

$$\delta j_t(x) = \int_{\mathbb{R}} \frac{f\circ\psi(x+(h-t)\nu) - f\circ\psi(x+h\nu)}{t} h(x+h\nu)dh.$$

We have

$$|\delta j_t(x)| \le \int_{\mathbb{R}} |\partial(f\circ\psi)^\nu| \|h\|_\infty,$$

$$\int_H\int_{\mathbb{R}} |\partial(f\circ\psi)^\nu| \le |\partial(f\circ\psi)|(\mathbb{R}^n) < +\infty$$

This allows to apply dominated convergence on δj_t. Moreover, since we can apply our differentiation result in dimension one, we have

$$\lim_{t\mapsto 0+} \delta j_t(x) = \int_{\mathbb{R}} -\partial(f\circ\psi)^\nu (g\circ\psi)_\nu \mathrm{Jac}(\psi)\,,$$

remark that the definition of $(g\circ\psi)_\nu$ is the one dimensional one. A priori, it may not coincide with the definition on $\mathbb{R}^n$. The Theorem (3.108 in [AFP00]) of the continuity of the precise representative tells us that the one dimensional restriction of $(g\circ\psi)_\nu$ gives $\mathcal{H}^{d-1}$-a.e. the same function. Thus

$$\begin{aligned} \partial_{t=0^+} J_t &= \int_H\int_{\mathbb{R}} -\partial(f\circ\psi)^\nu (g\circ\psi)_\nu \mathrm{Jac}(\psi)\,, \\ &= \int \langle \partial(f\circ\psi), -\nu\rangle (g\circ\psi)_\nu \mathrm{Jac}(\psi)\,, \\ &= \int \langle \partial f, -(\nabla\psi)\circ\psi^{-1}\cdot\nu\rangle (g\circ\psi)_\nu\circ\psi^{-1}\,. \end{aligned}$$

The last equality comes from Lemma 14 below, applied to $\phi = v(g \circ \psi)_\nu \mathrm{Jac}(\psi)$. Using the remark 5 and the equality (2.18) we have the result:

$$\partial_{t=0^+} J_t = \int \langle \partial f, -X_0 \rangle g_{X_0} \, . \qquad \square$$

Lemma 14. *If $f \in BV(U)$ and ψ is a diffeomorphism of U, then, for any bounded measurable function $\phi : U \to \mathbb{R}^d$, we have*

$$\int_U \langle \partial(f \circ \psi), \phi \rangle = \int_U \langle \partial f, \frac{(D\psi) \cdot \phi}{\mathrm{Jac}\psi} \circ \psi^{-1} \rangle.$$

Proof. Suppose f and ϕ regular. From the chain rule from regular function we get

$$\int_U \langle \partial(f \circ \psi), \phi \rangle = \int_U \langle (\nabla f) \circ \psi, (D\psi) \cdot \phi \rangle dx.$$

Thanks to the change of variable $x = \psi^{-1}(y)$ we get the desired formula. This is valid for $f \in C^1$ but we can recover the same result for $f \in BV$ by weak approximating f with a sequence of regular functions f_n (notice that in this case $\partial f_n \rightharpoonup \partial f$ as measures and the other factor is a continuous function, if $\phi \in C^0(U; \mathbb{R}^d)$).
Once the result is obtained for ϕ continuous, one can approximate pointwisely any measurable function ϕ by a sequence of continuous functions ϕ_n, and the results stays true for ϕ as well. $\square$

2.5.1 Extension to other penalty terms

In this last section, we will extend, for the sake of applications, the differentiation result to more general functions $H(f, g)$. Here again, we will only prove that what we have shown so far (i.e. the result for $(x, y) = xy$) is sufficient. We present the easy estimation we need in order to follow with a proof of the reduction based on a density argument.
Let us give a simple definition:

Definition 5. *If H is a Lipschitz function in two variables, we denote by*

$$Lip_1(H) = \inf\{M \in \mathbb{R}^+ | \forall (x, x', y) \in \mathbb{R}^3 \, |H(x, y) - H(x', y)| \leq M|x - x'|_1\}.$$

Exchanging the two variables, we define $Lip_2(H)$ as well.

By definition, we trivially have for each $i = 1, 2$, $\mathrm{Lip}_i(H) \leq \mathrm{Lip}(H)$.
To avoid summability problems and to have the possibility to replace locally Lipschitz functions H with globally Lipschitz ones, we will stick to the case $f, \, g \in \mathrm{B}$, i.e. we suppose them to be bounded.

Lemma 15. *If $g \in L^\infty(\mathbb{R}^n)$ with compact support and $f \in \mathrm{B}$, then if $0 \leq t \leq t_0$*

$$|J_t(f, g) - J_0(f, g)| \leq C t Lip_1(H) \|f\|_{BV},$$

for a constant C which only depends on the vector field X and on t_0.

Proof. We will follow again the same ideas as in Lemma 7.
Let us start from the case $g \in L^\infty$ and $f \in C^1$. Just consider

$$|J_t(f,g) - J_0(f,g)| \leq \text{Lip}_1(H) \int_U |f \circ \phi_t^{-1} - f| dx$$
$$\leq \text{Lip}_1(H) ||X||_\infty \int_0^t \int_U |\nabla (f \circ \phi_s^{-1})| dx\, ds.$$

Then, with a change of variables whose Jacobian is bounded, we get

$$|J_t(f,g) - J_0(f,g)| \leq c\text{Lip}_1(H) ||X||_\infty \sup_{s\in[0,t]} ||J(\phi_s)^{-1}||_\infty \int |\nabla f|.$$

Generalizing to $f \in BV$ works the same as in Lemma 7. □

Remark 9. *Remark that the Lipschitz constant Lip$_1(H)$ can be restricted on Im$(f) \times$ Im(g) and that will be used in the next proof.*

In the next paragraph we will see how the Theorem 13 implies its generalization (Theorem 15).

Proof of Theorem 15:
The idea of the proof is that the case $H(x,y) = xy$ implies the Theorem 15.
The result is true for all the polynomial functions since B is an algebra and we can easily check that the formula coming from Theorem 13 in the case $f^k g^h$ is exactly the one we want. Here with $H(x,y) = x^k y^h$,

$$\partial_{t=0^+} J_t = \int \langle D(f^k)(g^h)_{X_0}(x), -X_0\rangle dx =$$
$$\int \langle D(f^k)(g_{X_0}(x))^h, -X_0\rangle dx = \int \langle \partial_1 H(f(x), g_{X_0}(x)), -X_0\rangle dx\,.$$

The last inequality follows from the definition of $\partial_1 H(f(x), l)$ which is the distributional derivative of the composition of a locally Lipschitz function $x \to H(x,l)$ and with f a bounded BV function (see 29 in appendix).
Remark also that we can suppose that $H(0,y) = 0$ for any $y \in \mathbb{R}$ using the fact that the replacing H by $H(x,y) - H(0,y)$ gives the same result for $\partial_{t=0^+} J_t$. Then we approximate $\partial_x H(x,y)$ with a polynomial function P_ε such that, on Im$(f) \times$ Im(g), we have $|\partial_x H - P_\varepsilon|_{\infty,K} < \varepsilon$ (the sup norm on the compact K). This approximation is allowed since Im$(f) \times$ Im$(g) \subset K$ with K a compact subset of $\mathbb{R}^2$ (here we use the fact that f and g are in L^∞). Since $H(0,y) = 0$ we can integrate P_ε w.r.t. the first variable to obtain a polynomial function Q_ε such that $\partial_x Q_\varepsilon = P_\varepsilon$. This implies $\text{Lip}_{1,K}(H - Q_\varepsilon) < \varepsilon$ where $\text{Lip}_{1,K}$ denotes the Lipschitz constant of the function w.r.t. the first variable on the compact K.
For notation convenience we write (redefining some previous notations):

$$J_t(H) \doteq \int_{\mathbb{R}^n} H(f \circ \phi_t^{-1}(x), g(x)) dx\,.$$

The theorem is valid for Q_ε then we have by the lemma 15 and applying the theorem to Q_ε:

$$|\limsup_{t\to 0^+} \frac{J_t(H) - J_0(H)}{t} - \partial_{s=0^+} J_s(Q_\varepsilon)| \leq C\varepsilon \|f\|_{BV}$$

Now by triangular inequality we have,

$$\begin{aligned}|\limsup_{x\to 0^+} \frac{J_t(H) - J_0(H)}{t} - \partial_{t=0^+} J_t(H)| \leq & |\limsup_{t\to 0^+} \frac{J_t(H) - J_0(H)}{t} - \partial_{s=0^+} J_s(Q_\varepsilon)| \\ & + |\partial_{s=0^+} J_s(Q_\varepsilon) - \partial_{t=0^+} J_t(H)|\,.\end{aligned}$$

The last step is to prove that we can control the second term

$$|\partial_{s=0^+} J_s(Q_\varepsilon) - \partial_{t=0^+} J_t(H)|\,.$$

To this end remark that the formula defining $\partial_{t=0^+} J_t(H)$ is linear in H and we have

$$|\partial_{t=0^+} J_t(Q_\varepsilon - H)| \leq \mathrm{Lip}_{1,K}(Q_\varepsilon - H)|f|_{BV}|X|_\infty \leq \varepsilon |f|_{BV}|X|_\infty\,,$$

so that we get the result. □

2.6 Toward the application to geodesic equations

We aim to apply these previous results to derive the geodesic equations. To this end, we need to ensure that the vector field which appears in the differentiation of the penalty term in the functional (1.27) can be treated in our framework. We study the perturbation with respect to ε for the vector field $v_t + \varepsilon u_t$. Then we need to look at the regularity of the vector field $\partial_\varepsilon \phi_1 \circ (\phi_1^\varepsilon)^{-1}$. The equation (1.10) can be generalized for any ε as follows,

$$\partial_\varepsilon \phi_{0,1}^\varepsilon = d\phi_{0,1}^\varepsilon (\int_0^1 Ad_{\phi_{t,0}^\varepsilon}(u_t)\,dt)\,. \tag{2.20}$$

With the assumption that $p \geq 2$ is the order of regularity of the RKHS of vector fields, we have that for any $t \in [0,1]$ and $u \in V$, $Ad_{\phi_{t,1}^\varepsilon}(u) = [d\phi_{t,1}^\varepsilon]_{\phi_{1,t}^\varepsilon}(u \circ \phi_{1,t}^\varepsilon)$ is a C^1 vector field. Furthermore, thanks to equation (1.8) and the fact that $\phi_{t,1}^\varepsilon$ is C^2, this vector field is Lipschitz in ε on every compact subset in U. By integration w.r.t the time, the same regularity is valid for $\partial_{\varepsilon|}\phi_1 \circ (\phi_1^\varepsilon)^{-1}$. And the hypothesis required to apply Theorem 15 are verified.

We guess that we can still apply the differentiation result to $1-$admissible RKHS of vector fields. A possible way to prove it is by approximation with smooth vector fields since the only important point is the uniform convergence of the vector fields to get estimations such as in lemma 8.

Chapter 3

Application to geodesic equations

Contents

3.1 Introduction 58
3.2 A new model 59
3.3 Applications to metamorphosis 70
3.4 Simple examples 75

3.1 Introduction

In this Chapter, we develop the geodesic equations obtained in two different models. The first one is a new model to deal with the difference of intensity between the two images to match. This model takes into account the change of contrast between the two images through a diffeomorphism on the intensity. We will only deal with the case of a gray level image but this approach can easily be extended to multidimensional valued functions. The diffeomorphism on the intensity will be the equivalent of the diffeomorphism for the support of the image and thus the Lie group action and the invariant metric will be conserved. Hence, we will derive the Euler equations for this model within a large set of functions (SBV functions). We will also prove that geodesics exist in all time and in Chapter 4, we will give the Hamiltonian formulation of these equations. To this purpose, we will give a particular emphasize to the description of the initial momentum.
The results presented in Chapter 2 can also be applied to metamorphosis within the framework developed in [TY05]. It will enlighten the attempt already tackled by the authors to treat discontinuities. Apart from the introduction of the framework of metamorphosis, there are two important results in this article. The first one is the description of the geodesic equations for smooth initial image and the existence in all time of the geodesic for a large set of initial conditions. This first property enables the authors to prove an injectivity result for the exponential mapping between smooth images, namely in H^1: there exists a radius for which the exponential mapping is injective. This property was at the core motivation of this work. Our aim is to detail the geodesic equations in the SBV framework and to state that the Riemannian infinite dimensional manifold is complete in a sense described hereafter. However, we will show that the injectivity radius property

does not hold any more for the BV norm. This counter example gives room to add further conditions to obtain the injectivity property.
The important common feature of the two models is to introduce a photometric deformation. In the last section 3.4 of this Chapter, we will present toy examples in order to compare the models and to better understand both.

3.2 A new model

This model arose from the attempt to develop the multi-modal image matching in the framework of *large deformation diffeomorphisms*. The standard action of the group of diffeomorphism is usually on the template support. If we want to deal with a change of intensity in the image, we can as a first approach consider that this change of intensity can be modeled as a homeomorphism on the gray levels of the image: if I is the image, the action of a change of intensity is modeled as $\eta \circ I$ with η a homeomorphism of the template co-domain. It is then natural to introduce an action of a group of diffeomorphism on the template co-domain. In the same time, we will also take into account an action on the template domain.
Matching in this context, is to find a couple (η, ϕ) which minimizes the energy

$$E(\eta, \phi) = D(Id, (\eta, \phi))^2 + \frac{1}{\sigma^2}\|\eta \circ I_0 \circ \phi^{-1} - I_{targ}\|^2_{L^2}, \tag{3.1}$$

with I_0 the initial function and I_{targ} the target function, Id is the identity map in the product of groups, σ is a calibration parameter. The distance D is the direct product of the metric on each admissible space of vector fields V (on the template domain) and S (on the template co-domain). We will derive in section 3.2.1 the geodesic equations and ensure the existence of a solution for all time from an initial momentum.

3.2.1 Geodesic equations

To completely use the results of Chapter 2, we will deal with a more general penalty term. The support of the image will be denoted by M a Lipschitz domain and the space of images will be denoted by $Im(M) := SBV(M) \cap L^\infty(M)$.

Theorem 20. *Let $I_0, I_{targ} \in Im(M)$ be respectively the initial image and the target image, H be a locally Lipschitz function $H : \mathbb{R}^2 \mapsto \mathbb{R}$ and C^1 in the first variable and V, S two $2-$admissible spaces of vector fields respectively on M and $\mathbb{R}$.*
Let $\mathcal{J} : L^2([0,1], V \times S) \mapsto \mathbb{R}$ be the functional defined by,

$$\mathcal{J}(v_t, s_t) = \frac{\lambda}{2}\int_0^1 \|v_t\|^2_V dt + \frac{\beta}{2}\int_0^1 \|s_t\|^2_S dt + \int_M H(\eta_{0,1} \circ I_0 \circ \phi_{1,0}(x), I_{targ}(x))dx, \tag{3.2}$$

with $\phi_{0,t}$ and $\eta_{0,t}$ the flows of the vector fields $v_t \in L^2([0,1], V)$ and $s_t \in L^2([0,1], S)$. Then there exists $(v, s) \in H$ such that $\mathcal{J}(v, s) = \min_{(v,s)\in H} \mathcal{J}(v, s)$. For such a mini-

mizer, there exists $(p_a, p_b, p_c) \in L^1(M, \mathbb{R}^n) \times L^1(J_{I_0}, \mathbb{R}^n) \times L^1(M, \mathbb{R})$ *such that:*

$$\beta s_t = \int_M p_c(y) d[\eta^*_{t,1}]_{I_0^t(y)} k_S(I_0^t(y), .) d\mu(y), \tag{3.3}$$

$$\begin{aligned}\lambda v_t = &\int_M k_V(\phi_{0,t}(x), .)[d\phi_{0,t}]_x^{-1*}(p_a(x)) d\mu(x) \\ &+ \int_{J_{I_0}} k_V(\phi_{0,t}(x), .)[d\phi_{0,t}]_x^{-1*}(p_b(x)) d\mathcal{H}^{n-1}(x),\end{aligned} \tag{3.4}$$

with:

$$I_t^{t'} = \eta_{0,t'} \circ I_0 \circ \phi_{t,0},$$

and J_{I_0} *the jump set of* I_0*. More precisely for the* (p_a, p_c) *we show, we have the equation:*

$$p_a(x) + \nabla_{|x} I_0^1 \; p_c(x) = 0. \tag{3.5}$$

Before proving the theorem, we give simple remarks to underline the result in the case the contrast part is not taken into account ($\beta = 0$ in the functional (3.2)). What was already known among all is the case of smooth images ([TY05]), the momentum is only a L^2 function on the support of the image: $p_b = 0$ in the theorem.

For the discontinuous case, the case of closed curves in [GTL06] gave the first step to understand discontinuous images. In this paper, the momentum appears to be a normal vector field along the closed curve: $p_a = 0$ in the theorem because the matching of the closed curves is in fact the matching of binary shapes which are piecewise constant. Our result extends it in any dimension for a piecewise constant functions. Again the absolute continuous part of the momentum vanishes and the momentum is reduced to a normal vector field along the jump set of the image.

Another well known case, which is of interest for this article, is the case of the piecewise constant functions in one dimension. In this case, the problem ends up with a Landmark matching problem. Let $(x_0 = 0, x_1, \ldots, x_{n-1}, x_n = 1) \in [0, 1]$ be a group of points, through a diffeomorphism from $]0, 1[$ to $\mathbb{R}$, we reduce the study to the case of $n - 1$ Landmark $(x_1, \ldots, x_{n-1})$ in $\mathbb{R}$. The momentum has only a singular part, and is the vector of landmark's momentum defined on the cotangent space of $\mathbb{R}^{n-1}$. (This case is the simplest implementation for the contrast term (the geodesic term on S) with for each level of intensity a landmark.)

Proof of the theorem: **Existence of a minimizer:** The closed balls in $L^2([0, 1], V \times S)$ are compact for the weak topology. The first term of the functional $\frac{\lambda}{2} \int_0^1 \|v_t\|_V^2 dt + \frac{\beta}{2} \int_0^1 \|s_t\|_S^2 dt$ is classically lower semi-continuous for the weak topology (see [Bré94] as reference). To prove the existence of a minimizer, we then just need to prove that the penalty term is lower semi-continuous.

To this end we will prove in fact that the penalty term is continuous for the weak topology. Let $(v_n, s_n)_{n\in\mathbb{N}}$ be a sequence in $L^2([0, 1], V \times S)$ weakly converging to a limit point (v, s) i.e.

$$(v_n, s_n) \underset{n\mapsto\infty}{\rightharpoonup} (v, s)\,. \tag{3.6}$$

For any $n \geq 0$, we denote

$$m_n \doteq \sup_{x \in M,\ l \in [0,1]} \left\{ d(\phi^{v_n}_{1,0}(x), \phi^{v}_{1,0}(x)) + d(\eta^{s_n}_{0,1}(l), \eta^{s}_{0,1}(l)) \right\}$$

so that from (3.6) and Theorem 2 we get that

$$\lim_{N \to \infty} m_n = 0\,. \tag{3.7}$$

We start from the triangle inequality :

$$\begin{aligned} DH_n &\doteq \int_M |H(\eta^{s_n}_{0,1} \circ I_0 \circ \phi^{v_n}_{1,0}(x), I_{\text{targ}}(x)) - H(\eta^{s}_{0,1} \circ I_0 \circ \phi^{v}_{1,0}(x), I_{\text{targ}}(x))| dx \\ &\leq \int_M \underbrace{|H(\eta^{s_n}_{0,1} \circ I_0 \circ \phi^{v_n}_{1,0}(x), I_{\text{targ}}(x)) - H(\eta^{s}_{0,1} \circ I_0 \circ \phi^{v_n}_{1,0}(x), I_{\text{targ}}(x))|}_{E_1(x)}\, dx \\ &+ \int_M \underbrace{|H(\eta^{s}_{0,1} \circ I_0 \circ \phi^{v_n}_{1,0}(x), I_{\text{targ}}(x)) - H(\eta^{s}_{0,1} \circ I_0 \circ \phi^{v}_{1,0}(x), I_{\text{targ}}(x))|}_{E_2(x)}\, dx\,. \end{aligned}$$

Concerning the first term of the right hand side, we get from the Lipschitz property the following upper bound for $x \in M$:

$$E_1(x) \leq \text{Lip}_1(H) d(\eta^{s_n}_{0,1} \circ I_0 \circ \phi^{v_n}_{1,0}, \eta^{s}_{0,1} \circ I_0 \circ \phi^{v_n}_{1,0}(x)) \leq \text{Lip}_1(H) m_n\,.$$

Thus,

$$\int_M E_1(x) dx \leq \text{Lip}_1(H) \mu(M) m_n \tag{3.8}$$

where μ denotes the Lebesgue measure.

Now, concerning the second term, let us note that by the Lipschitz approximation theorem (see [EG92]), there exists $\tilde{I}_0$ a Lipschitz function such that $\mu(D) \leq \varepsilon$ for $D \doteq \{x \in M : \tilde{I}_0(x) \neq I_0(x)\}$. Then, if we denote

$$A_n \doteq \phi^{v_n}_{0,1}(M \setminus D) \cap \phi^{v}_{0,1}(M \setminus D)\,,$$

we have for any $x \in A_n$

$$\begin{aligned} E_2(x) &= |H(\eta^{s}_{0,1} \circ \tilde{I}_0 \circ \phi^{v_n}_{1,0}(x), I_{\text{targ}}(x)) - H(\eta^{s}_{0,1} \circ \tilde{I}_0 \circ \phi^{v}_{1,0}(x), I_{\text{targ}}(x))| \\ &\leq \text{Lip}_1(H) \text{Lip}(\eta^{s}_{0,1} \circ \tilde{I}_0) d(\phi^{v_n}_{1,0}(x), \phi^{v}_{1,0}(x)) \\ &\leq \text{Lip}_1(H) \text{Lip}(\eta^{s}_{0,1} \circ \tilde{I}_0) m_n \end{aligned}$$

and

$$\int_M E_2(x) dx \leq \text{Lip}_1(H) \text{Lip}(\eta^{s}_{0,1} \circ \tilde{I}_0) \mu(A_n) m_n + 2|H|_\infty \mu(M \setminus A_n) \tag{3.9}$$

with $|H|_\infty = \sup_{(x,y) \in [0,1]^2} |H(x,y)|$. Since $M \setminus A_n \subset \phi^{v_n}_{0,1}(D) \cup \phi^{v}_{0,1}(D) \cup \partial M$, we get

$$\mu(M \setminus A_n) \leq 2K\mu(D) \leq 2K\varepsilon \tag{3.10}$$

where $K \doteq |\text{Jac}(\phi^v_{0,1})|_\infty \vee \sup_{n\geq 0} |\text{Jac}(\phi^{v_n}_{0,1})|_\infty$ (note that since v_n is bounded in $L^2([0,1],V)$, we get from 1.6 that K is finite). Using (3.9) and (3.10), we get

$$\int_M E_2(x)dx \leq \text{Lip}_1(H)\text{Lip}(\eta^s_{0,1}\circ\tilde{I}_0)\mu(M)m_n + 4K|H|_\infty\varepsilon. \tag{3.11}$$

so that with (3.8), we deduce

$$DH_n \leq \text{Lip}_1(H)\mu(M)m_n + \text{Lip}_1(H)\text{Lip}(\eta^s_{0,1}\circ\tilde{I}_0)\mu(M)m_n + 4K|H|_\infty\varepsilon \tag{3.12}$$

and from (3.7) that

$$\limsup DH_n \leq 4K|H|_\infty\varepsilon\,.$$

Since ε is arbitrary, we get $DH_n \to 0$ and the continuity of the penalty term for the weak topology.

Geodesic equation for the contrast term: For such a minimizer, we derive the geodesic equation by differentiating the functional with respect to a variation of the vector field. As we have seen in Chapter 2, the functional $\mathcal{J}$ is only semi-differentiable with respect to the variable v but differentiable w.r.t. the variable s. We first differentiate w.r.t. the vector field $s \in L^2([0,T],S)$, we denote by $\tilde{s}$ a perturbation of s. Using the Theorem 6,

$$\partial_{\tilde{s}}\eta_{0,1}(x) = \int_0^1 [d\eta_{t,1}]_{|\eta_{0,t}(x)}\tilde{s}_t(\eta_{0,t}(x))dt\,.$$

We already introduce the kernel and with $I_1 = \eta^s_{0,1}\circ I_0\circ\phi^v_{1,0}$,

$$\int_0^1 [\beta < s_t,\tilde{s}_t >_S + \int_M \partial_1 H(I_1(y),I_{targ}(y))[d\eta^*_{t,1}]_{|I^t_1(y)} < k_S(I^t_1(y),.),\tilde{s}_t >_S d\mu(y)]dt = 0,$$

it leads to, with the change of variable $x = \phi_{0,1}(y)$

$$\beta s_t + \int_M \text{Jac}(\phi_{0,1})(x)[\partial_1 H](I^1_0(x), I_{targ}\circ\phi_{0,1}(x))[d\eta^*_{t,1}]_{|I^t_0(x)}k_S(I^t_0(x),.)d\mu(x) = 0.$$

With the notation $p_c(x) = -\text{Jac}(\phi_{0,1})(x)[\partial_1 H](I^1_0(x), I_{targ}\circ\phi_{0,1}(x))$, we have the first equation announced and p_c is a bounded function since H is Lipschitz on $[0,1]^2$.

Geodesic equation for the geometric term: For the second equation, we will need the semi-differentiation lemma detailed in Chapter 2. We refine the notation adopted in this Chapter, $j_H(f(x),l) = (H(f^+(x),l) - H(f^-(x),l))\nu_f(x)$. For a minimizer v_t, a perturbation u_t and the notation $\delta u = -\int_0^1 Ad_{\phi_{t,1}}(u)\,dt$ we have,

$$\lambda\int_0^1 \langle v_t,u_t\rangle dt + \int\langle\partial_1 H(I_1,I_{targ}),\delta u\rangle + \int_{J_{I_1}}\langle j_H(I_1(x),[I_{targ}]_{-\delta u}(x)) - j_H(I_1(x),I_{targ}(x)),\delta u\rangle d\mathcal{H}^{n-1}(x) \geq 0. \tag{3.13}$$

In the inequality (3.13), we have decomposed the semi-differentiation formula in a linear part and a non-linear one. The non-linear part is an integration on the jump set of f. As H is Lipschitz on $[0,1]^2$, we have for $y \in [0,1]$,

$$|H(I_1(x)^+, y) - H(I_1^-(x), y)| \leq Lip_1(H)|I_1^+(x) - I_1^-(x)|. \tag{3.14}$$

As a consequence, the second term lies in $L^\infty(J_{I_1}, |D^s I_1|) \subset L^2(J_{I_1}, |D^s I_1|)$.
Let us write the differentiation of the functional as follows,

$$\partial_u \mathcal{J} = \partial_u \mathcal{J}_1 + \int_{J_{I_1}} \langle j_H(I_1(x), [I_{targ}]_{-\delta u}(x)) - j_H(I_1(x), I_{targ}(x)), \delta u \rangle d\mathcal{H}^{n-1}(x),$$

with $\partial_u \mathcal{J}_1$ the linear part of the expression written in (3.13).
We can now apply a version of the Lemma in [GTL06], which is detailed in appendix (see 19) on the Hilbert space $\mathbb{R} \times L^2(J_{I_1}, |D^s I_1|)$, with $H_0 = \{(\partial_u \mathcal{J}_1, \delta u) | u \in L^2([0,1], V)\}$. If $a = (\partial_u \mathcal{J}_1, \delta u)$ and $b = (1, j_H(I_1(.), [I_{targ}]_{-\delta u}(.)) - j_H(I_1(.), I_{targ}(.)))$, we get

$$\partial_u \mathcal{J} = \langle a, b \rangle_H \geq 0\,.$$

Thus, with

$$B = \{(1, j_H(., [I_{targ}]_{-\delta u}(.)) - j_H(., I_{targ}(.))) | u \in L^2([0,1], V)\}\,,$$

we have the needed bounded set since the inequality (3.14) does not depend on $y \in [0,1]$. Hence we get with the Lemma 19 the existence of $(1, \tilde{j}) \in \overline{\mathrm{Conv}(B)}$ such that

$$\partial_u \mathcal{J}_1 + \int_{J_{I_1}} \langle \tilde{j}, \delta u \rangle d\mathcal{H}^{n-1} = 0. \tag{3.15}$$

To go further in the description of $\tilde{j}$, we know that $\tilde{j} \in \mathrm{Conv}(B)$. Let us denote

$$\Delta H(x) = \langle j_H(I_1(x), [I_{targ}]_{-\delta u}(x)) - j_H(I_1(x), I_{targ}(x)), \nu_f(x) \rangle\,.$$

Then, for each $x \in J_{I_1}$, if $m(x) := \min(\Delta H(I_1(x), I^+_{targ}(x)), \Delta H(I_1(x), I^-_{targ}(x)))$ and $M(x) := \max(\Delta H(I_1(x), I^+_{targ}(x)), \Delta H(I_1(x), I^-_{targ}(x)))$, the convexity property implies $\langle \tilde{j}(x), \nu_f(x) \rangle \in [m(x), M(x)]$.
As $H(x, .)$ is continuous in the second variable, by the intermediate value theorem for each $x \in J_{I_1}$, $\{t \in \mathbb{R} | \Delta H(x,t) = \langle \tilde{j}(x), \nu_f(x) \rangle\}$ is a non-empty set and compact by the continuity in the second variable of H. The measurability of $\mathcal{H} : (x,y) \mapsto (x, \Delta H(x,y))$ comes from the stability by composition of measurable maps. By the Lemma 20 in appendix, there exists

$$\tilde{I}_{targ}(x) \in [I^-_{targ}(x), I^+_{targ}(x)]$$

such that $\tilde{j}(x) = j_H(I_1(x), \tilde{I}_{targ}(x)) - j_H(I_1(x), I_{targ}(x))$ and the function $x \in J_{I_1} \mapsto \tilde{I}_{targ}(x)$ is measurable.
With L the adjoint operator, we obtain

$$\lambda L v_t = \int Ad^*_{\phi_{t,1}} \partial_1 F \left(I_1, \tilde{I}_{targ}\right)\,, \tag{3.16}$$

remark that $Ad^*_\phi u = \mathrm{Jac}(\phi) d\phi^*(u \circ \phi)$. Using the kernel, we have (with a change of variable in the second equation)

$$\begin{aligned}
\lambda v_t &= \int k(.,\phi_{1,t}(x)) d\phi^*_{1,t} \partial_1 H(I_1(x), I_{targ}(x)),\\
\lambda v_t &= \int \mathrm{Jac}(\phi_1)(x) k(.,\phi_{0,t}(x)) d\phi^*_{t,0} \partial_1 [H(I^1_0(x), \tilde{I}_{targ} \circ \phi_{0,1}(x))],\\
\lambda v_t &= \int \mathrm{Jac}(\phi_1)(x) k(.,\phi_{0,t}(x)) [d\phi_{0,t}]^{-1*}_x \partial_1 [H(I^v_0(x), \tilde{I}_{targ} \circ \phi_{0,1}(x))].
\end{aligned}$$

The last equation is the desired result if we expand the derivative $\partial_1 H(I^v_0, \tilde{I}_{targ} \circ \phi_{0,1})$.

$$\begin{aligned}
\lambda v_t + \int_M \mathrm{Jac}(\phi_1)(x) k(.,\phi_{0,t}(x)) [d\phi_{0,t}]^{-1*}_x \nabla_1 H(I^1_0(x), \tilde{I}_{targ} \circ \phi_{0,1}(x)) dx \\
+ \int_{J_{I^1_0}} \mathrm{Jac}(\phi_1)(x) k(.,\phi_{0,t}(x)) [d\phi_{0,t}]^{-1*}_x j(x) d\mathcal{H}^{n-1}(x), \quad (3.17)
\end{aligned}$$

with $j(x) := j_H(I^1_0(x), \tilde{I}_{targ} \circ \phi_{0,1}(x)) v_{I^1_0}(x)$. Since $I^1_0 = \eta_{0,1} \circ I_0$ and $\eta_{0,1}$ is a diffeomorphism, we have $J_{I_0} = J_{I^1_0}$ and the integration on the $J_{I^1_0}$ can be replaced by an integration on J_{I_0}. We thus have $p_b(x) := \mathrm{Jac}(\phi_{0,1})(x) j(x)$. We have also,

$$p_a = \mathrm{Jac}(\phi_{0,1}) \nabla_1 H(I^1_0, \tilde{I}_{targ} \circ \phi_{0,1}) = -\nabla I^1_0 p_c,$$

which is the last assertion of the theorem. □

Note that the last relation $p_a + \nabla I^1_0 p_c$ is independent of H and another important point, it shows that I_0, I_{targ}, p_c, p_b is the signature of the geodesic. The study of the geodesics generated from these initial data is the subject of the next Subsection.

3.2.2 Geodesic flow

In this paragraph, we will prove that if p_a, p_b, p_c are given, the geodesic is locally defined, then that the geodesic travels at constant speed and finally that the geodesic is defined for all time. Namely, we answer to existence and uniqueness of solutions to:

$$\begin{aligned}
\eta_{0,t} &= Id + \int_0^t s_u \circ \eta_u du,\\
\beta s_t(.) &= \int_M p_c(y) d[\eta_{t,0}]_{I^t_0(y)} k_S(I^t_0(y)), .) d\mu(y),\\
\phi_{0,t} &= Id + \int_0^t v_u \circ \phi_u du,\\
\lambda v_t(.) &= \int_M k_V(.,\phi_{0,t}(x)) [d\phi_{0,t}]^{-1*}_x (p_a(x)) d\mu(x)\\
&+ \int_{J_{I_0}} k_V(.,\phi_{0,t}(x)) [d\phi_{0,t}]^{-1*}_x (p_b(x)) d\mathcal{H}^{n-1}(x) dt. \quad (3.18)
\end{aligned}$$

On purpose, this system of equations is decoupled in v and s. We will detail the case of the geometric vector field in the following proof.

Proposition 10. *For T sufficiently small, the system of equations* (3.18) *with*

$$(p_a, p_b, p_c) \in L^1(M, \mathbb{R}^n) \times L^1(J_{I_0}, |D^s I_0|) \times L^1(M, \mathbb{R})$$

has a unique solution if both RKHS (geometric and contrast) are 2 *admissible. (see 1.1.1 for the definition)*

Proof:
We aim to apply the fixed point theorem on the Banach space $L^2([0,T], V \times S)$. We will prove that for a given ball in $L^2([0,T], V \times S)$, there exists $T > 0$ such that the following map is a contraction:

$$\begin{aligned} \Xi : L^2([0,T], V \times S) &\mapsto L^2([0,T], V \times S) \\ (v, s) &\mapsto (\xi(v), \xi(s)), \end{aligned} \tag{3.19}$$

with

$$\begin{aligned} \xi(v)_t &= \int_M k_V(., \phi_{0,t}(x))[d\phi_{0,t}]_x^{-1*}(p_a(x))d\mu(x) \\ &\quad + \int_{J_{I_0}} k_V(., \phi_{0,t}(x))[d\phi_{0,t}]_x^{-1*}(p_b(x))d\mathcal{H}^{n-1}(x)dt, \\ \xi(s)_t &= \int_M p_c(x) d[\eta_{t,0}]_{I_0^t(x)} k_S(I_0^t(x)), .)d\mu(x). \end{aligned}$$

For the space $L^2([0,T], V)$, if we have:

$$\|\xi(v)_t - \xi(u)_t\|^2 \leq M\|v - u\|_{L^1[0,T]},$$

the result is then proven with Cauchy-Schwarz inequality:

$$\|\xi(v)_t - \xi(u)_t\|^2 \leq M\sqrt{T}\|v - u\|_{L^2[0,T]},$$

and integrating this inequality:

$$\|\xi(v) - \xi(u)\|^2_{L^2[0,T]} \leq MT^{\frac{3}{2}}\|v - u\|_{L^2[0,T]}.$$

This can be obtained with:

$$\begin{aligned} \|\xi(v)_t - \xi(u)_t\|^2 &= < \xi(v)_t, \xi(v)_t - \xi(u)_t > - < \xi(u)_t, \xi(v)_t - \xi(u)_t >, \\ \|\xi(v)_t - \xi(u)_t\|^2 &\leq 2\max(| < \xi(v)_t, \xi(v)_t - \xi(u)_t > |, | < \xi(u)_t, \xi(v)_t - \xi(u)_t > |). \end{aligned}$$

We adopt the previous notation for the measure ν and for the momentum, we define $p(x) := p_a(x)\mathbf{1}_{x \notin J_{I_0}} + p_b(x)\mathbf{1}_{x \in J_{I_0}}$. For one of the two terms in the equation above:

$$\begin{aligned} < \xi(v)_t, \xi(v)_t - \xi(u)_t > &= \int\int [d\phi^v_{0,t}]_x^{-1*}(p(x))[k(\phi^v_{0,t}(x), \phi^v_{0,t}(y))[d\phi^v_{0,t}]_y^{-1*}(p(y)) \\ &- k(\phi^v_{0,t}(x), \phi^u_{0,t}(y))[d\phi^u_{0,t}]_y^{-1*}(p(y))]d\nu(y)d\nu(x). \end{aligned}$$

On the ball of radius r in $L^2([0,T], V \times S)$ denoted by $B(r)$, and with the inequalities (1.7) and (1.8), we control the diffeomorphisms:

$$\|\phi^u_{0,t} - \phi^v_{0,t}\|_\infty \leq c_V \|v-u\|_{L^1[0,T]} \exp(c_V r\sqrt{T}), \tag{3.20}$$

$$\|d\phi^u_{0,t} - d\phi^v_{0,t}\|_\infty \leq C'' \|v-u\|_{L^1[0,T]} \exp(c_V r\sqrt{T}), \tag{3.21}$$

$$\|[d\phi^u_{0,t}]^{-1} - [d\phi^v_{0,t}]^{-1}\|_\infty \leq \frac{\|d\phi^u_{0,t} - d\phi^v_{0,t}\|_\infty}{1 - \|d\phi^u_{0,t} - d\phi^v_{0,t}\|_\infty}, \tag{3.22}$$

the last inequality is a standard inequality for the norm of the inverse in a Banach algebra and is true as long as $\|d\phi^u_{0,t} - d\phi^v_{0,t}\|_\infty < 1$. With the triangle inequality, we get:

$$\begin{aligned}
&| < \xi(v)_t, \xi(v)_t - \xi(u)_t > | \leq \int\int [|d\phi^v_{0,t}]^{-1*}_x(p(x))| \\
&|[k(\phi^v_{0,t}(x), \phi^v_{0,t}(y))[d\phi^v_{0,t}]^{-1*}_y(p(y)) - k(\phi^v_{0,t}(x), \phi^v_{0,t}(y))[d\phi^u_{0,t}]^{-1*}_y(p(y))]| \\
&+ |[k(\phi^v_{0,t}(x), \phi^v_{0,t}(y))[d\phi^u_{0,t}]^{-1*}_y(p(y)) - k(\phi^v_{0,t}(x), \phi^u_{0,t}(y))[d\phi^u_{0,t}]^{-1*}_y(p(y))]| d\nu(x) d\nu(y).
\end{aligned}$$

On the ball $B(r)$ we have, denoting $K(r,T) := \max(C'', c_V) r\sqrt{T} \exp(c_V r\sqrt{T})$:

$$\begin{aligned}
\|d\phi^u_{0,t} - Id\|_\infty &\leq 2K(r,T), \\
\|[d\phi^u_{0,t}]^{-1}\|_\infty &\leq \frac{1}{1-2K(r,T)}.
\end{aligned}$$

Let $M_k \in \mathbb{R}$ a bound for the kernel and its first derivative on the ball $B(r)$. Such a constant exists thanks to the hypothesis on the kernel and its first derivative.
A bound for the first term can be found with the inequality (3.22) if

$$\max(c_V, C'')\|v-u\|_{L^1[0,T]} \exp(c_V r\sqrt{T}) < \frac{1}{2},$$

$$2\max(C'', c_V)\|v-u\|_{L^1[0,T]} \exp(c_V r\sqrt{T}) M_k \frac{1}{1-2K(r,T)} (\|p_a\| + \|p_b\|)^2,$$

then, if $4K(r,T) < \frac{1}{2}$ the bound is now:

$$4\max(C'', c_V)\|v-u\|_{L^1[0,T]} \exp(c_V r\sqrt{T}) M_k (\|p_a\| + \|p_b\|)^2.$$

The second term is controlled with the inequality (3.20) with the Lipschitz hypothesis on the kernel:

$$\max(C'', c_V)\|v-u\|_{L^1[0,T]} \exp(c_V r\sqrt{T}) M_k [\frac{1}{1-2K(r,T)}]^2 (\|p_a\| + \|p_b\|)^2,$$

with the same assumption $4K(r,T) < \frac{1}{2}$ we have $[\frac{1}{1-2K(r,T)}]^2 \leq 2$ and

$$2\max(C'', c_V)\|v-u\|_{L^1[0,T]} \exp(c_V r\sqrt{T}) M_k (\|p_a\| + \|p_b\|)^2.$$

Finally we get,

$$\|\xi(v)_t - \xi(u)_t\|^2 \leq 10\max(C'', c_V)\|v-u\|_{L^1[0,T]} \exp(c_V r\sqrt{T}) M_k (\|p_a\| + \|p_b\|)^2. \tag{3.23}$$

The inequality (3.23) shows that for any $r > 0$ there exists $T > 0$ such that Ξ is a contraction on the first component. We can get the same conclusion for the second term with a similar proof and we do not give the details. □
As solution of a fixed point problem, there is a nice dependency of geodesics w.r.t. the initial momentums:

Proposition 11. *Let $p^1 := (p_a^1, p_b^1, p_c^1)$ and $p^2 := (p_a^2, p_b^2, p_c^2)$ two initial momentums in $L^1(M, \mathbb{R}^n) \times L^1(J_{I_0}, |D^s I_0|) \times L^1(M, \mathbb{R})$ with $w_1 := (v^1, s^1)$ and $w_2 := (v^2, s^2)$ the associated solutions to the fixed point problem. Then there exists $K > 0$ and $T > 0$ such that*

$$|v_t^1 - v_t^2|_V \leq K(\|p_a^1 - p_a^2\| + \|p_b^1 - p_b^2\|),$$
$$|s_t^1 - s_t^2|_S \leq K\|p_c^1 - p_c^2\|,$$

if $t < T$.

Proof. Let us denote Ξ_{p^i} for $i = 1, 2$ the application introduced previously for one momentum. Remark that w_i for $i = 1, 2$ is the unique solution of $\Xi_{p_i}(w_i) = w_i$.

$$\begin{aligned}\|w_1 - w_2\| &= \|\Xi_{p^1}(w_1) - \Xi_{p^2}(w_2)\| \\ &\leq \|\Xi_{p^1}(w_1) - \Xi_{p^1}(w_2)\| + \|\Xi_{p^1}(w_2) - \Xi_{p^2}(w_2)\| \\ &\leq \text{Lip}(\Xi_{p^1})\|w_1 - w_2\| + \|\Xi_{p^1 - p^2}(w_2)\|.\end{aligned}$$

We have for T sufficiently small, with the inequality (3.23)

$$\|\Xi_{p^1-p^2}(w_2)_t\|^2 \leq 20 c_V \|w_2\|_{L^1[0,T]} \exp(c_V r\sqrt{T}) M_k \|p^1 - p^2\|^2,$$

and then we obtain the result with $M_0^2 = 20 c_V \|w_2\|_{L^1[0,T]} \exp(c_V r\sqrt{T}) M_k$,

$$\|[w_1]_t - [w_2]_t\| \leq \frac{M_0}{1 - \text{Lip}(\Xi_{p^1})}\|p^1 - p^2\|.$$

□

We have proven that there exists $T > 0$ such that we have existence and uniqueness to the system (3.18), we aim to prove that we can choose $T = +\infty$ in the last proposition.
We first demonstrate that if a vector field is a solution to the geodesic equations, then the norm of the vector field is constant in time.

Proposition 12. ***Constant speed curves in vector fields spaces***
If the kernels (contrast and geometric) are admissible, we have

- *If a vector field s_t is a solution of equation* (3.3) *then $\|s_t\|^2$ is constant.*
- *If a vector field v_t is a solution of equation* (3.4) *then $\|v_t\|^2$ is constant.*

Remark that we do not need higher regularity assumptions here.

Proof:

We will prove that the quantity $\|s_t\|^2$ is differentiable (w.r.t. the time variable) everywhere.

$$\|s_t\|_S^2 = \iint_{M^2} p(y')d[\eta_{t,1}]_{I_t^s(y')}k_S(I_t^s(y'), I_t^s(y))p(y)d[\eta_{t,1}]_{I_t^s(y)}d\mu(y')d\mu(y).$$

Remark that

$$\partial_t(d[\eta_{t,1}]_{I_0^t(y)}) = -d[s_t]_{I_0^t(y)}d[\eta_{t,1}]_{I_0^t(y)}\,. \tag{3.24}$$

This equation is obtained by differentiation of the group relation: $d[\eta_{0,1}]_{I_0} = d[\eta_{t,1} \circ \eta_{0,t}]_{I_0}$, and with the differentiation of the equation (3.3):

$$d[s_t]_x = \int_M d[\eta_{t,1}]^*_{I_0^t(y)}p_c(y)\partial_1 k_S(x, I_0^t(y))d\mu(y).$$

These relations are verified almost everywhere a priori. As s_t is a geodesic

$$s_t = \frac{1}{\beta}\int_M p_c(y)d[\eta_{t,0}]_{I_0^t(y)}k_S(I_0^t(y)), .)d\mu(y)$$

and the second term is defined for all time (since the flow and its differential are defined for all time if $s_t \in L^2([0,T], S)$), the equation (3.24) is verified for all time.

To obtain the result, it is sufficient to prove that the derivative vanishes everywhere. The differentiation under the integral sign is allowed since $\partial_t(d[\eta_{t,1}]_{I_0^t(y)})$ is bounded in the space variable.

$$\begin{aligned}\partial_t\|s_t\|^2 = &-\iint_{M^2} p(y')ds_t(I_0^t(y'))d[\eta_{t,1}]_{I_0^t(y')}k_S(I_0^t(y')), I_0^t(y)))p(y)d[\eta_{t,1}]_{I_0^t(y)}d\mu(y')d\mu(y)\\ &+\iint_{M^2} p(y')d[\eta_{t,1}]_{I_0^t(y')}\partial_1 k_S(I_0^t(y')), I_0^t(y)))s_t(I_0^t(y'))p(y)d[\eta_{t,1}]_{I_0^t(y)}d\mu(y')d\mu(y) = 0.\end{aligned}$$

The second point is very similar. We underline that the equation (3.4) is a particular case of the following, with a measure ν which has a Lebesgue part and a singular part on the set J_{I_0} of discontinuities of the function I_0. We also define:

$$p_t(x) = (d[\phi_{0,t}]^*_x)^{-1}(p_a(x)\mathbf{1}_{x\notin J_{I_0}} + p_b(x)\mathbf{1}_{x\in J_{I_0}}),$$

By the definition,

$$\|v_t\|^2 = \iint p_t(x)k_V(\phi_{0,t}(x), \phi_{0,t}(y))p_t(y)d\nu(x)d\nu(y).$$

Remark that $\partial_t p_t(x) = -d_{\phi_{0,t}(x)}v_t^* p_t(x)$, and we differentiate:

$$\begin{aligned}\partial_t\|v_t\|^2 = &-\iiint p_t(x)\partial k_V(\phi_{0,t}(x), \phi_{0,t}(z))p_t(z)k_V(\phi_{0,t}(x), \phi_{0,t}(y))p_t(y)d\nu(x)d\nu(y)d\nu(z)\\ &+\iiint p_t(x)\partial k_V(\phi_{0,t}(x), \phi_{0,t}(y))p_t(y)k_V(\phi_{0,t}(x), \phi_{0,t}(z))p_t(z)d\nu(x)d\nu(y)d\nu(z) = 0.\end{aligned}$$

□

We prove now that the solutions do not blow up. This property shows that the associated Riemannian manifold of infinite dimension is complete, since the exponential map is defined for all time. Without the hypothesis on the kernel, we can find simple counter-examples to this fact (landmarks in one dimension with an exponential kernel).

Proposition 13. *The solution of proposition 10 is defined for all time (with RKHS $2-$admissible).*

Proof. Thanks to proposition 12, the norm of each solution (respectively v_t and s_t) is constant in time. That will enable the extension for all time. Consider a maximal solution with interval of definition $[0, T_{max}]$ with $T_{max} < +\infty$, then with the inequalities from the introduction 1 and especially the inequality (1.6), we define the limit $\lim_{t \mapsto T_{max}} \phi_{0,t} := \phi_{0,T_{max}}$, since for all x, $\phi_{0,t}(x)$ is a Cauchy sequence. This is the same for $\lim_{t \mapsto T_{max}} d\phi_{0,t}(x)$. This limit is also a diffeomorphism, since we can define the limit of the inverse as well. With a change of variable $\phi_{0,T_{max}}$, the geodesic equation can be written with a measure ν:

$$\lambda v_{T_{max}+s}(.) = \int_M k_V(., y)[d\phi_{T_{max},T_{max}+s}]_y^{-1*}[d\phi_{0,T_{max}}]^{-1*}_{\phi_{T_{max},0}(y)}(p_a(\phi_{T_{max},0}(y)))d(\phi_{0,T^*_{max}}\mu)(y).$$

Applying to $p(y) = [d\phi_{0,T_{max}}]^{-1*}_{\phi_{T_{max},0}(y)}(p_a(\phi_{T_{max},0}(y)))$ the proposition 10, we obtain a vector field $v_{T_{max}+s}$ satisfying the geodesic equation for $s \in [0, S]$ with $S > 0$. This vector field gives a solution v_t for $t \in [0, T_{max} + S]$. This is a contradiction.
For the contrast term, we can apply the proposition to the current image I_0^t instead of I_0 and the initial momentum $p_c(.)d[\eta_{t,0}]_{I_0^t(.)}$. As previously, the solution can be extended and it gives a contradiction. □

3.2.3 Comments and conclusion

All these results could be extended for BV functions as it is done only for the geometric term in [VS09]. However, we chose to restrict ourselves in this thesis to SBV functions which covers a large range of cases of interest. It also seems more relevant to deal with SBV functions in order to give the Hamiltonian formulation of the equations. Roughly speaking, the more complex the derivative of the template image I_0 is, the more complex is the structure of the momentum. As discussed in [VS09], the momentum should be now defined as a measurable function of the derivative measure. Since the derivative has a Cantor part, the momentum should be also defined on the Cantor part. Although the evolution of the momentum for the Cantor part follows the same evolution than the classical dense part, it is clearly more demanding to know in addition the value of the momentum on the Cantor set. To make it clear, recall that in the case of just a geometric action, we have (proposition 4)

$$p_t = \mathrm{Jac}(\phi_{t,1})\Delta \circ \phi_{t,1}\,,$$

and $\Delta = I_1 - I_{targ}$ in the case of a template image. In the BV case, this equation is also verified, however we need to know the value of Δ on a Cantor set. Finally our opinion is that the Cantor part should not be taken into account in the template since the template ought to be somehow smooth enough. BV functions are not smooth enough to be a good candidate to the space of template images, whereas SBV functions seem more relevant.
Last, the Hamiltonian formulation of these equations are detailed in 4 for the case of piecewise Lipschitz functions. Again, this setting could have been extended to SBV functions, but it is a convenient framework to develop numerical algorithm and to introduce ideas such as of a multilayer approach.

3.3 Applications to metamorphosis

We chose as a direct application to develop the metamorphosis framework introduced in [TY05]. This framework has been numerically implemented for the case of smooth images in [GY05]. The discontinuous case was partially understood in the previous reference: the case of functions such as the indicator of a smooth domain was addressed. It was a first attempt to deal with discontinuous images and thanks to our previous result, the extension to a much larger set of discontinuous functions was at hand. After a short introduction to the metamorphosis framework, we will fully understand the case of BV functions. The original motivation of this article was to introduce a system of coordinates which explains the deformations and variations around a given template: the obtained result is the local bijectivity of the exponential map on smooth functions (namely H^1). In this Chapter, we will prove that this result can not be extended to BV functions by showing a function around which the injectivity radius of the exponential map is zero.

3.3.1 Introduction to metamorphosis

Roughly speaking, the starting point of this approach is to put a metric on the tangent space of images which takes into account a geometric variation and a photometric variation. At the infinitesimal level, let I_0 be an initial image defined on $\mathbb{R}^n$, σ is a real positive parameter, v a vector field on $\mathbb{R}^n$ and z an image, we consider the following transformation with a real parameter t:

$$I_t(x) = I_0(x - tv(x)) + t\sigma^2 z(x) + o(t)\,.$$

This transformation is divided into two parts, the first one is the composition with the vector field which is called the geometric transformation and the second one is the addition of $\sigma^2 z$ which is called the photometric transformation. Obviously this transformation is not injective i.e. there are different ways of obtaining the same final image. A trade-off can be made by choosing a metric on the tangent space. We now introduce in more details the framework we will consider.
Let $U \subset \mathbb{R}^k$ be a Lipschitz open domain and $L^2(U)$ be the space of images, we reduce

our work to deal with real valued functions. The general case with $\mathbb{R}^d$ valued functions do not introduce any further complications but heavier notations. We chose to develop the case of SBV functions since at this time it seems to be relevant for applications without extending it to BV functions. We want to draw the reader's attention on the singular part of the derivative and not on the Cantor part. We recall the definitions and the central results which can be found in [TY05] with detailed proofs and explanations.
First we define continuously differentiable curves in $L^2(U)$ which are continuous curves for the L^2 topology and weakly continuous differentiable curves:

Definition 6. *(C^1 curves in $L^2(U)$) Let J be an interval in $\mathbb{R}$, σ be a positive real number and W be the Hilbert space $V \times L^2(U)$ with V an admissible space of vector fields. The inner product on W is defined by $\langle (v,z),(v',z')\rangle_W = \langle v,v'\rangle_V + \sigma^2\langle z,z'\rangle_{L^2(U)}$ for $(v,z) \in W$ and $(v',z') \in W$.*
The curve $I : t \in J \mapsto I_t \in L^2(U)$ is said to be C^1 if there exists a continuous map $w : J \ni t \mapsto (v_t, w_t) \in W$ such that

1. *$I \in C(J, L^2(U,\mathbb{R}))$ for the L^2 topology,*
2. *for any $u \in C_c^\infty(U,\mathbb{R})$, $t \mapsto \langle I_t, u\rangle$ is C^1 and $\partial_t\langle I_t,u\rangle = \sigma^2\langle z_t,u\rangle_2 + \langle I_t, \mathrm{div}(uv_t)\rangle_2$.*

Remark that the derivative can also be written, if V is continuously injected in $(C^1(U,\mathbb{R}^n), \|\|_{1,\infty})$ and $I_t \in SBV(U)$,

$$\partial_t\langle I_t,u\rangle = \sigma^2\langle z_t,u\rangle_2 + \langle DI_t, uv_t\rangle.$$

The tangent space is then defined by the equivalence class of the weak derivative.

Definition 7. *Let $I \in Im(U) := SBV(U) \cap L^\infty(U)$ be an image and V an admissible RKHS of vector fields on U, then the map for any $u \in C_c^\infty(U,\mathbb{R})$*

$$\begin{aligned} T_u : W &\mapsto \mathbb{R} \\ (v,z) &\mapsto \langle DI, vu\rangle + \sigma^2\langle z,u\rangle_2 \end{aligned}$$

is linear continuous. Hence $\ker T_u \subset W$ is a closed subspace and introducing $E_I := \cap_{u\in C_c^\infty(U)} \ker T_u$, we define

$$T_I Im(U) := \{I\} \times W/E_I.$$

Remark that E_I is also a closed subspace of W as intersection of closed subspaces. Then the orthogonal projection on E_I is well defined in the Hilbert space W and we denote it by p_I. The tangent space at j is endowed with the induced metric by W, namely $|\bar{w}|_{T_I} = \inf\{|w|_W | p(w) = \bar{w}\}$. Thanks to the projection theorem on closed convex nonempty sets in Hilbert spaces, there exists a unique $w \in W$ such that $p(w) = \bar{w}$ and $|\bar{w}|_{T_I} = |w|_W$.

Definition 8. *The geodesic distance between two images j_0 and j_1 in $L^2(U)$ is defined by,*

$$d(J_0, J_1) = \inf\{ \int_0^1 |p_I(\frac{dj}{dt})|_W dt \,|\, I \in C^1_{pw}([0,1], L^2(U)), I_0 = j_0,\ I_1 = j_1 \}, \quad (3.25)$$

with $C^1_{pw}([0,1], L^2(U))$ is the set of piecewise C^1 curves in $L^2(U)$ (as naturally defined from C^1 curves).

A geodesic is a C^1 curve which minimizes the distance introduced above. It can be proven (Theorem 6 in [TY05]) that there always exists a geodesic between two images if the space of vector fields is admissible.

Theorem 21. *Assume that the space of vector field V is compactly embedded in $C^1_0(U, \mathbb{R}^d)$, and let $j_0, j_1 \in L^2(U)$ be two functions then there exists a C^1 curve I of length $d(j_0, j_1)$ with $I_0 = j_0$ and $I_1 = j_1$.*

3.3.2 Geodesic equations and existence result

The geodesic can be characterized by studying the variation of the energy of the geodesic with respect to the variables (v, z). It is done in [TY05] by first studying a variation with respect to the photometric variable z. The first characterization obtained is that if a C^1 path is a geodesic then it can be written such that:

$$\begin{cases} I_t = I_0 \circ \phi_{t,0} + \sigma^2 \int_0^t z_s \circ \phi^v_{t,s} ds\,, \\ z_t = z_0 \circ \phi_{t,0} \operatorname{Jac}(\phi^v_{t,0})\,, \\ (v_t, z_t) = p_I(\frac{dI}{dt})\,. \end{cases} \quad (3.26)$$

To understand the last condition in system (3.26), we need to introduce two operators that characterize the action of the weak derivative of I.

Definition 9. *Let $I \in L^2$ be a function,*

1. *The operator $\nabla_I : \mathcal{D}_I \mapsto L^2(U, \mathbb{R})$ is defined by*

$$\mathcal{D}_I := \{v \in V | \exists C,\ s.t. \forall u \in C^\infty_c(U, \mathbb{R})\, |\langle I, \operatorname{div}(uv)\rangle_2| \leq C|u|_2\}\,,$$

and for any $v \in \mathcal{D}_I$, $\nabla_I.v$ is the unique element in $L^2(U, \mathbb{R})$ such that,

$$\langle \nabla_I v, u\rangle_2 = \langle I, \operatorname{div}(uv)\rangle_2$$

for any $u \in C^\infty_c(U, \mathbb{R})$.

2. *the adjoint operator $D^*_I : \mathcal{D}^*_I \mapsto V$ is defined by*

$$\mathcal{D}^*_I := \{u \in L^2(U, \mathbb{R}) | \exists C,\ s.t. \forall v \in \mathcal{D}_I\, |\langle \nabla_I v, u\rangle_2| \leq C|v|_V\}\,,$$

and for any $u \in \mathcal{D}^*_I$, $D^*_I.u$ (or $\nabla^*_I u$) is the unique element in $\bar{\mathcal{D}_I}$ such that,

$$\langle D^*_I u, v\rangle_V = \langle u, \nabla_I.v\rangle_2$$

for any $v \in D_I$.

The last condition in (3.26) $(v_t, z_t) = p(\frac{dI}{dt})$ is equivalent to $z_t \in \mathcal{D}^*_{I_t}$ and $v_t - \nabla^*_I z_t \in \mathcal{D}^{\perp}_{I_t}$. To complete the system of equation (3.26) we need to differentiate with respect to the vector field variable v the energy of the geodesic which can be written, using the first two conditions in (3.26):

$$\mathcal{F} = \int_0^1 |v_s|^2_V ds + \frac{1}{\sigma^2} \int_U \frac{|I_1 - I_0 \circ \phi^v_{1,0}|^2}{\int_0^1 \mathrm{Jac}(\phi^v_{1,s})^{-1} ds}\,. \tag{3.27}$$

In the article [TY05], the differentiation was done with the additional assumption that I_0 is H^1. By our previous work in Chapter 2, we can relax this assumption to take into account discontinuities for the initial image as well as discontinuities for the target image. We will assume that I_0 and I_1 belong to $Im(U)$ (seed definition 7). We will see that the last condition in (3.26) is always satisfied for extremals of the geodesic energy $\mathcal{E}$. This last condition will be replaced by the equation obtained when differentiating (3.27) which implies the last condition. We first prove the semi-differentiation Lemma,

Lemma 16. *The geodesic energy $\mathcal{F}(v)$ is semi-differentiable with respect to the vector field $v \in L^2([0,1], V)$ and we have for $\mathcal{F}_\varepsilon = \mathcal{F}(v + \varepsilon h)$,*

$$\frac{d}{d\varepsilon}_{|\varepsilon=0_+} \mathcal{F}_\varepsilon = 2\int_0^1 \langle v_t + KDI_t z_t, h_t\rangle_V dt + \int_0^1 2\langle KD(I_0 \circ \phi^v_{t,0})(z_t^{Ad_{\phi_{1,t}} H} - z_t), h_t\rangle_V dt. \tag{3.28}$$

with $H = -\int_0^1 Ad_{\phi_{t,1}} h_t dt = \int_0^1 (d\phi_{t,1} h_t) \circ \phi_{1,t} dt$.

Remark 10. *In the Lemma 16, the notation $z_t^{Ad_{\phi_{1,t}} H}$ stands for the quantity defined in the previous Chapter 2 $[z_t]_{Ad_{\phi_{1,t}} H}$.*
In the equation 3.28, we distinguish between the linear part and the non-linear part (respectively the first and the second term on the right hand side of the equality).

Proof. The first term is quadratic and its derivative with respect to the time dependent vector field $(h_t) \in L^2([0,1], V)$ is

$$2\int_0^1 \langle v_s, h_s\rangle ds.$$

To study the second term, we introduce the quantity q^v_t in order to simplify notations,

$$\begin{aligned} q^v_t &:= \int_0^t \mathrm{Jac}(\phi^v_{t,s})^{-1} ds\,, \\ I_t &= I_0 \circ \phi^v_{t,0} + \sigma^2 z_t q^v_t\,, \\ H &= -\int_0^1 Ad_{\phi_{t,1}} h_t dt = -\int_0^1 (d\phi_{t,1} h_t) \circ \phi_{1,t} dt. \end{aligned} \tag{3.29}$$

We differentiate the second term denoted by $U_\varepsilon := \frac{1}{\sigma^2}\int_U \frac{|I_1 - I_0 \circ \phi^{v+\varepsilon h}_{1,0}|^2}{q_1^{v+\varepsilon h} ds}$. We use the Theorem 15 to differentiate this expression,

$$\frac{d}{d\varepsilon}_{|\varepsilon=0_+} U_\varepsilon = \int_0^1 \int \langle 2D(I_0 \circ \phi^v_{1,0}), z_1^H Ad_{\phi_{t,1}} h_t\rangle - \int_U \sigma^2 |z_1|^2 \frac{d}{d\varepsilon}[\mathrm{Jac}(\phi^{v+\varepsilon h}_{1,s})^{-1}] dx\, dt. \tag{3.30}$$

The first term in equation (3.30) can be rewritten as with the change of variable $u = \phi_{1,t}(x)$ and the identity $z_t = z_1 \circ \phi^v_{t,1}\mathrm{Jac}(\phi^v_{t,1})$,

$$\int \langle 2D(I_0 \circ \phi^v_{1,0}), z_1^H Ad_{\phi^v_{t,1}} h_t\rangle = \int \langle 2d\phi^*_{t,1} D(I_0 \circ \phi^v_{1,0}) \circ \phi^v_{t,1}, \mathrm{Jac}(\phi^v_{t,1}) z_1^H \circ \phi_{t,1} h_t\rangle, \tag{3.31}$$

$$= \int \langle 2D(I_0 \circ \phi^v_{t,0}), \mathrm{Jac}(\phi^v_{t,1}) z_1^H \circ \phi_{t,1} h_t\rangle, \tag{3.32}$$

$$= \int \langle 2D(I_0 \circ \phi^v_{t,0}), z_t^{Ad_{\phi_{1,t}} H} h_t\rangle, \tag{3.33}$$

$$= \langle 2KD(I_0 \circ \phi^v_{t,0}) z_t^{Ad_{\phi_{1,t}} H}, h_t\rangle_V. \tag{3.34}$$

Note that equalities (3.32) and (3.33) use respectively Lemma 14 and remark 5 in Chapter 2. From now on, we just mimic the proof in [TY05] using the additional assumption $I_0, I_1 \in \mathrm{SBV}(U)$ which implies that $z_t \in \mathrm{SBV}(U)$. The second term involves the derivative of $\mathrm{Jac}(\phi^{v+\varepsilon h}_{t,s})^{-1}$. We recall that the Jacobian of the flow verifies the following ODE in the smooth case,

$$\frac{\partial}{\partial t}\mathrm{Jac}(\phi^v_{0,t}) = \mathrm{div}(v_t) \circ \phi^v_{0,t}\mathrm{Jac}(\phi^v_{0,t}).$$

This equality is also valid a.e. if $v_t \in L^2([0,1], V)$ and consequently we have,

$$\mathrm{Jac}(\phi^{v+\varepsilon h}_{t,s})^{-1} = \exp\left[-\int_t^s \mathrm{div}(v_u + \varepsilon h_u) \circ \phi^{v+\varepsilon h}_{t,u} du\right]. \tag{3.35}$$

Then,

$$\frac{d}{d\varepsilon}\mathrm{Jac}(\phi^{v+\varepsilon h}_{t,s})^{-1}) = \mathrm{Jac}(\phi^v_{t,s})^{-1}\left[\int_s^t \mathrm{div}(h_u) \circ \phi_{t,u} - \mathrm{Jac}(\phi^v_{t,u})^{-1}\langle \nabla_{\phi^v_{t,u}} \mathrm{Jac}{\phi^v_{t,u}}^{-1}, h_u \circ \phi^v_{t,u}\rangle du\right].$$

The second term of equation (3.30) can be rewritten,

$$\int_0^1 \int_U \sigma^2 |z_1|^2 \frac{d}{d\varepsilon}[\mathrm{Jac}(\phi^{v+\varepsilon h}_{1,s})^{-1}] dx\, dt = \int_0^1 \langle q^v_u |z_u|^2, \mathrm{div}(h_u)\rangle_2 - \langle |z_u|^2, \langle \nabla q^v_u, h_u\rangle\rangle_2 du,$$

$$= \int_0^1 -\langle 2\nabla q^v_u |z_u|^2 + D[z_u^2]\nabla q^v_u, h_u\rangle_2 du.$$

Finally, we get injecting the two preceding results in 3.30:

$$\frac{d}{d\varepsilon}_{|\varepsilon=0_+} U_\varepsilon = \int_0^1 2\langle KD(I_0 \circ \phi^v_{t,0}) z_t^{Ad_{\phi_{1,t}} H}, h_t\rangle_V + \langle K[2\nabla q^v_t |z_t|^2 + D[z_t^2]\nabla q^v_t], h_t\rangle_V dt. \tag{3.36}$$

To simplify this equation (3.36) we recall that the pairing $\langle Df, g\rangle$ if $f, g \in SBV(U)$ is defined as $\langle Df, g^0\rangle$ with the 0 the null vector field defining g on the sump set of f. Actually, g^0 is nothing else than the precise representative of g. As used in Theorem 14, remark that $Dz_u^2 = 2z_u Dz_u$ which enables to develop the last right hand side term in equation 3.36.

We use also the following decomposition of the distributional derivative

$$DI_t = D[I_0 \circ \phi^v_{t,0}] + \sigma^2 Dz_t q^v_t + \sigma^2 z_t \nabla q^v_t\,,$$

to decompose the equation in a linear part and a 'semi-linear' one as claimed,

$$\frac{d}{d\varepsilon}_{|\varepsilon=0_+} U_\varepsilon = \int_0^1 2\langle KDI_t z_t, h_t\rangle_V dt + \int_0^1 2\langle KD(I_0\circ\phi_{t,0}^v)(z_t^{Ad_{\phi_{1,t}}H} - z_t), h_t\rangle_V dt.$$

□

In the identity (3.28), the second term boils down to an integration on the jump set of $i_0\circ\phi_{t,0}^v$ since outside this jump set, we only need to take into account the dense part definition of $z_t^{Ad_{\phi_{1,t}}H} - z_t$ which is null. Applying the projection Lemma 19 to the semi-differentiation (3.28) (or on the formulation (3.30) as well), we obtain the existence of Z_0 which is a bounded Borel function defined on J_{I_0} the jump set of I_0 such that,

$$v_t + KDI_t z_t + KD^s[I_0\circ\phi_{t,0}^v]Z_0\circ\phi_{t,0}^v \mathrm{Jac}(\phi_{t,0}) = 0. \tag{3.37}$$

The last term can be also written as,

$$KD^s[I_0\circ\phi_{t,0}^v]Z_0\circ\phi_{t,0}^v\mathrm{Jac}(\phi_{t,0}) = KAd_{\phi_{t,0}^*}(Z_0 D^s I_0),$$

which is a transport equation of the singular part. The characterization $v_t - \nabla_I^* z_t \in \mathcal{D}_{I_t}^\perp$ is fulfilled and $z_t \in \mathcal{D}_{I_t}^*$ is clear since $L^\infty(U)\subset\mathcal{D}_{I_t}^*$.
As a consequence,

Theorem 22. *The geodesic equations on $SBV(U)$ are given by,*

$$\begin{cases} I_t = I_0\circ\phi_{t,0} + \sigma^2\int_0^t z_s\circ\phi_{t,s}^v ds\,, \\ z_t = z_0\circ\phi_{t,0}\,\mathrm{Jac}(\phi_{t,0}^v)\,, \\ v_t + KDI_t z_t + KAd_{\phi_{t,0}^*}(Z_0 D^s I_0) = 0. \end{cases} \tag{3.38}$$

Now, Theorem 7 in [TY05] says that the initial data (I_0, z_0, Z_0) uniquely specify a solution for the system (3.38) under smoothness assumptions on the RKHS of vector fields (namely V needs to be $3-$admissible). Moreover the solution is defined for all time.
To sum up at this point, endowed with this geodesic flow the space $SBV(U)$ is complete in this sense: between two images, there always exists a geodesic between them and geodesics are defined for all time. Now, the last result in [TY05] which motivated the paper was the injectivity of the exponential map restricted on $H^1(U)$. We will see in the next section that it is no longer the case in $SBV(U)$.

3.4 Simple examples

We aim to prove with a simple example that in the case of metamorphoses, the exponential map may not be injective on any neighborhood of a well-chosen image. This example shows that the introduction of discontinuities can lead to a non trivial complication of the geometry. Yet there is room to add conditions on the template to reach the injectivity of the exponential map. This example is based on the simple idea of exploiting the two different ways of transformation: the geometric deformation and the photometric one; the

discontinuities allow fine situations that can not take place in the continuous (H^1) case. Then it can be reproduced at any scale to obtain the counter example to the non injectivity of the exponential map exactly as for the case of an infinite product of spheres with radius decreasing to 0.

One of the simplest case to deal with is the matching between two characteristic functions of an interval. Let us denote by $\mathbb{1}_{]x,y[}$ the characteristic function of the interval $[x, y] \in \mathbb{R}$, we will also write $B(z, \delta) =]x, y[$ the Euclidean ball in $\mathbb{R}$. Let $z_0 < z_1 \in \mathbb{R}$ be two distinct points, we denote by $d(z_0, z_1)$ the Euclidean distance. Let us consider a translation invariant kernel (such as the Gaussian kernel), there exists $r_0 > 0$ such that if $x, y \in \mathbb{R}$ and $d(x, y) < r_0$ there exists a unique geodesic between x and y for the metric induced by the kernel.

If $z_1 \in B(z_0, r_0)$ then there exists a unique vector field v_0 such that $\psi_{0,1}^{v_0}(z_0) = z_1$ and such that $|v_0|_{L^2}^2 \doteq \int_0^1 |v_0(t)|_V^2 dt$ minimizes the deformation cost. This vector field is also the unique minimizer of the matching problem (minimizing $|v|_{L^2}^2$ for $v \in L^2([0, 1], V)$)

$$\begin{cases} \phi_{0,1}^v(z_0) = z_1 \\ \phi_{0,1}^v(z_0 + \delta) = \psi_{0,1}^{v_0}(z_0 + \delta)\,, \end{cases}$$

and this is valid for every $\delta > 0$. Moreover the distance $d(z_1, \phi_{0,1}^v(z_0 + \delta))$ can be controlled with the inequalities (1.4),

$$e^{-c_V |v|_{L^2}} \delta \leq d(z_1, \phi_{0,1}^v(z_0 + \delta)) \leq e^{c_V |v|_{L^2}} \delta\,. \tag{3.39}$$

We will study the metamorphosis problem between $M\mathbb{1}_{]z_0,z_0+\delta[}$ (also denoted by (M, z_0, δ)) and $M\mathbb{1}_{]z_1,\psi_{0,1}^{v_0}(z_0+\delta)[}$. To this end, recall that the geodesics are extremums of the functional $\mathcal{F}$ defined in (3.27). We will prove that for $d(z_0, z_1)$ small enough, there exists $M > 0$ such that the minimization of (3.27) leads at least to two different geodesics. One with a geometric deformation mode and the other with a photometric deformation mode. The idea comes from the fact that when M is small, the minimizing geodesic will be closed to a pure photometric transformation. For the case M large enough the minimizing geodesic will be closed to a pure geometric transformation. To be more precise we define in this particular case of matching of step functions the photometric and geometric type for a minimizing geodesic:

Definition 10. *A minimizing geodesic between $I_0 = (M_0, z_0, \delta_0)$ and $I_1 = (M_1, z_1, \delta_1)$ is said to be of photometric type if Supp$(I_0 \circ \phi_{1,0}) \cap$Supp$(I_1) = \emptyset$. Otherwise the geodesic is said to be of geometric type.*

For any couple of images I_0, I_1, we associate the set of types of minimizing geodesics (geometric, photometric or both).

We begin with simple inequalities on the energy function $\mathcal{F}$. First remark that the minimizer u_0 of the metamorphosis problem for the two step functions verifies (for the parameter $\sigma = 1$)

$$\mathcal{F}(u_0) \leq \min(|v_0|_{L^2}^2, 2\delta M^2)\,. \tag{3.40}$$

This inequality comes from the pure deformation solution associated to the cost $|v_0|^2_{L^2}$ and from the pure photometric transformation associated to the cost $2\delta M^2$.
Let us study a minimizing geodesic of photometric type between $I_0^M = M\mathbb{1}_{]z_0,z_0+\delta[}$ and $I_1^M = M\mathbb{1}_{]z_1,\psi_{0,1}^{v_0}(z_0+\delta)[}$. We get the inequality:

$$|v|^2_{L^2} + 2M^2\delta(1-\gamma|v_0|_{L^2}) \leq \mathcal{F}(v) = |v|^2_{L^2} + \int_U \frac{|I_1|^2 + |I_0 \circ \phi^v_{1,0}|^2}{\int_0^1 \mathrm{Jac}(\phi^v_{1,s})^{-1} ds}\,, \tag{3.41}$$

where γ is a constant which comes from the control of the Jacobian thanks to inequalities on the derivatives of the vector field (we use here the $2-$admissibility of the kernel) in introduction 1.1.1:

$$1 - c|v|_{L^2} \leq |\mathrm{Jac}(\phi_{s,t})| \leq 1 + c|v|_{L^2}\,,$$

for a constant c if $|v|_{L^2} \leq \varepsilon$. Remark that this is the case thanks to the inequality (3.40) if $|v_0|_{L^2} \leq \varepsilon$ since $|v|^2_{L^2} \leq \mathcal{F}(v)$. We also used (3.39) to control the integral of $|I_0 \circ \phi^v_{1,0}|^2$. We now change the parameterization to make things clearer. Let us introduce $\alpha \geq 0$ such that $|v|^2_{L^2} = \alpha|v_0|^2_{L^2}$ and $\beta > 0$ such that $2M^2\delta = \beta|v_0|^2_{L^2}$. The inequality (3.41) can be rewritten

$$\alpha + \beta(1 - \gamma|v_0|_{L^2}) \leq \min(1, \beta)\,. \tag{3.42}$$

It implies $\alpha \leq \beta\gamma|v_0|_{L^2}$ and if we assume $\gamma|v_0|_{L^2} \leq \frac{1}{N}$ for $N \in \mathbb{N}^*$ we get

$$\begin{aligned}
&\alpha \leq \frac{\beta}{N}\,,\\
&\alpha + N\alpha(1 - \frac{1}{N}) \leq 1\,,\\
&\alpha \leq \frac{1}{N}\,.
\end{aligned}$$

It means that the deformation has small energy if the supports of the final step and the target one do not intersect. Choosing N big enough, in the **photometric case**

$$d(\phi^v_{0,1}(z_0+\delta), z_1) \geq \frac{d(z_0,z_1)}{2}\,. \tag{3.43}$$

We now give some inequalities in the case of a minimizing geodesic of **geometric type**. Thanks to inequality 3.39, we have for δ small enough

$$d((\phi^v_{0,1}(z_0), z_1)) < \frac{d(z_0,z_1)}{2}\,.$$

Observe now that when M is relatively big enough, a minimizing geodesic is of geometric type otherwise it would contradict (3.42). And if M is small enough, a minimizing geodesic is of photometric type thanks to the same inequality.

We denote by P the set of real numbers M such that every minimizing geodesic between I_0 and I_1 is of photometric type. We have proved that P contains a neighborhood of 0. Consider $M_0 \doteq \sup\{M \,|\,]0, M[\subset P\}$. By the remarks above $0 < M_0 < \infty$. We aim to prove that for M_0 there exists at least one minimizing geodesic of each type.
The distance induced by a Riemannian metric is continuous, then it implies that

$$\lim_{M \to M_0} \mathcal{F}(M) = \mathcal{F}(M_0)$$

(where $\mathcal{F}(M)$ denotes the minimum of $\mathcal{F}$ for the metamorphosis problem between I_0^M and I_1^M). As there exists at least a minimizing geodesic for M_0, it is sufficient to prove that there exists a minimizing geodesic of the other type. We can face two cases according to the known type (photometric for instance) of a minimizing geodesic for M_0 but the argument will be the same: there exists a sequence $M_n \to M_0$ such that the geodesic is of the other type (geometric for instance). By lower semi-continuity of the functional and weak compactness of balls in $L^2([0,1], V)$ there exists v_0 such that

$$\mathcal{F}(v_0) \leq \liminf_{n \to \infty} \mathcal{F}(M_n) = \mathcal{F}(M_0)\,,$$

up to the extraction of a subsequence. Hence we get $\mathcal{F}(v_0) = \mathcal{F}(M_0)$. In our case (geometric type) as the supports are closed sets, to be of geometric type for a minimizing geodesic is a closed condition. Hence the vector field v_0 gives rise to a minimizing geodesic of geometric type.
In the other case we would have needed to take a limit of photometric geodesics. This condition is also closed thanks to the inequality 3.43. Last, thanks to the inequality 3.42 we have necessarily

$$2M_0^2\delta \leq |v_0|_{L^2}^2\,. \tag{3.44}$$

A lower bound could also have been given.
Everything is done to reproduce this case at smaller scales since this inequality remains true if $|v_0|_{L^2} \to 0$. As the cost of the initial deformation $|v_0|_{L^2}^2$ goes to 0, M_0 goes also to 0. It enables the control of the BV norm of our counter-example.
An example of function in which the exponential map is non injective is the following,

$$f = \sum_{i=1}^{\infty} m_i \mathbb{1}_{]2^i, 2^i + \frac{1}{i^4}[} \tag{3.45}$$

with m_i corresponds to the M_0 defined above if it exists and $m_i = 1$ otherwise. We have in particular using the inequality (3.44)

$$|f|_{BV} = 2\sum_{i=1}^{\infty} m_i \leq 2\sum_{i=1}^{\infty} \frac{1}{i^2} < \infty\,.$$

We do not give details but only the ideas to prove that f is actually a counter-example to the local injectivity property of the exponential map. The Gaussian kernel being fast decreasing, for i large enough f behaves like a point on a product of Riemannian manifolds

and the cut locus of each projection (i.e. each step function $m_i \mathbb{1}_{]2^i, 2^i + \frac{1}{i^4}[}$) gets smaller and smaller.
This counter-example uses the fact that $\mathbb{R}$ is not bounded. Yet the reader can convince himself that same examples can be found on a bounded domain and it can be obviously generalized in every dimension.

Despite this counter-example, we can hope for a more positive result adding some assumptions on the template (the initial image I_0). Adding some smoothness assumptions (H^1) was a way to obtain the injectivity. However if the discontinuities of the template are large enough (pointwisely) it does not seem difficult to discriminate between the photometric transformation and the geometric one around a point of discontinuity.
We have the following proposition

Proposition 14. *Let $\epsilon > 0$ be a positive real number. If f is a piecewise Lipschitz function on $[a, b] \subset \mathbb{R}$ then there exists a neighborhood of 0 such that the exponential map $\exp_f$ defined*

$$\exp_f : L^\infty(J_f) \times SBV(U) \cap L^\infty \mapsto SBV(U)$$
$$(Z, g) \mapsto f \circ \phi_{1,0} + \sigma^2 \int_0^1 g_s \circ \phi_{t,s}^v ds$$

(with $\phi_{t,s}$ and g_t characterized by system (3.38)*) is injective.*

Remark that the condition on the size of the jumps of the function f is implied by the Lipschitz hypothesis.

Proof. Let $w_a \doteq (Z_a, g_a)$ and $w_b \doteq (Z_b, g_b)$ be two initial conditions and ϕ^a and ϕ^b be their associated flows with underlying vector fields v_t^a and v_t^b.
Let us define $\varepsilon = \min\{|j_f(x)| \,;\, x \in \mathbb{R}\}$. Since the number of discontinuities is finite we have $\varepsilon > 0$. We obtain $|\exp_f(tw_a) - \exp_f(tw_b) \circ \phi_{0,t}^a|_{BV} \geq 2\varepsilon \mathcal{H}^0(S)$ for $S \doteq \{x \in j_f \,|\, \langle v^a(x) - v^b(x), j_f(x) \rangle \neq 0\}$ and t small enough (here we use the fact that the jump set of f is a finite number of points). We get

$$v_0^a - v_0^b \in D_f^\perp \,,$$

hence we have $Z_a = Z_b$ if the there exists w_a, w_b such that $\exp_f(tw_a) = \exp_f(tw_b)$.
Thanks to our semi-differentiation Lemma and the fact that $v_a - v_b$ vanishes on the jump set of f, we have

$$|\exp_f(tw_a) - \exp_f(tw_b) \circ \phi_{0,t}^a|_{L^2}^2 \approx Ct |g_a - g_b|_{L^2}^2 \,,$$

with C a positive constant (see Lemma 5 in [TY05]).
It gives the result when $t \to 0$. □

We guess that this proposition can be extended in any dimension. An additional assumption such as $\inf\{|j_f(x)| \,;\, x \in J_f\} \geq \varepsilon > 0$ should turn the proof easier.

Chapter 4

Hamiltonian formulation of image matching

Contents

4.1 Heuristic derivation of the Hamiltonian formulation 80
4.2 The case of discontinuous images 83
4.3 Conclusion 89

In numerous papers on *large deformation diffeomorphisms*, the Hamiltonian framework arises. The simplest example is probably the Landmark matching problem for which the geodesic equations and the Hamiltonian version of the evolution are well known ([VMTY04],[ATY05]). Our goal is to provide a Hamiltonian interpretation of the initial variational problem in the case of discontinuous images. The first step was done in [GTL06] where Hamiltonian equations were formulated on closed curves. Our work generalizes this approach to the space of images. We use the point of view of the optimal control theory (as it is developed in [ATY05]) to formally introduce the Hamiltonian. The first section presents the heuristic method to find the Hamiltonian via the Pontryagin maximum principle. Then, we state a weak Hamiltonian formulation of the equations obtained by the variational approach for the model developed in 3.2. We prove uniqueness for the solutions to these equations. At the end of this section, we discuss the existence of solutions with the help of the existence result in proposition 13.
The weak Hamiltonian system is written on piecewise Lipschitz functions, yet such a function is considered as the superposition of several functions having in mind the introduction of a multilayer approach. This general formulation could be extended to other matching problems of interest.

4.1 Heuristic derivation of the Hamiltonian formulation

We present here a heuristic approach which was a motivation to prepare the derivation of the Hamiltonian formulation in the case of discontinuous images. As recalled in the introduction (chapter 1), the Hamiltonian formulation is well-known for the case of landmarks (proved in 1.3) and was developed recently for the case of curves ([GTL06]). In

the case of landmarks, the derivation of the equation can be done from an optimal control point of view as said in [ATY05]. In [GTL06], the optimal control point of view is not followed. One obstacle is the infinite dimension of the space, which makes the generalization of the Pontryagin maximum principle not straightforward. Despite certain generalizations in infinite dimension, we prefer to stay at a heuristic level and to properly derive the Hamiltonian formulation in our context. Recall that we gave an overview of the Hamiltonian formulation in the case of smooth images in the introduction 1.3.
In all the three cases (landmarks, curves, images), the objects on which acts the group of diffeomorphisms lie in a linear space. Let Q be a Banach space (the space of objects) and P its dual. Let us introduce the action of G_V on Q,

$$G_V \times Q \mapsto Q$$
$$(\phi, q) \mapsto \phi.q$$

For the landmark case, if $\mathbf{x} = (x_1, \dots, x_n)$ represents the group of landmarks, the action is defined by $\phi.\mathbf{x} = (\phi(x_1), \dots, \phi(x_n))$. We assume that the action is differentiable in the following sense, there exists a map

$$V \times Q \mapsto Q$$
$$(v, q) \mapsto v.q$$

such that for any $v \in L^2([0,1], V)$ we have

$$\frac{d}{dt}[\phi_{0,t}.q] = v_t.[\phi_{0,t}.q]\,,$$
$$\phi_{0,1}.q = q + \int_0^1 v_t.[\phi_{0,t}.q]\,dt\,.$$

The problem of matching two objects is to minimize the usual functional

$$\frac{1}{2}\int_0^1 |v_t|_V^2\,dt + G(\phi_{0,1}.q_0, q_{targ})\,, \tag{4.1}$$

with G a matching term. Once we know the existence of a solution v_0, it minimizes the energy $\frac{1}{2}\int_0^1 |v_t|_V^2\,dt$ with fixed endpoint $\phi_{0,1}^v.q_0 = \phi_{0,1}^{v_0}.q_0$. Before formally applying the optimal control theory, we get a little more general by assuming that we are interested in the solutions of the minimization of:

$$\begin{cases} \inf \frac{1}{2}\int_0^1 |v_t|^2 dt \\ q(0) \in M_0 \\ q(1) \in M_1 \end{cases} \tag{4.2}$$

with M_0, M_1 two subsets of Q with tangent spaces at a point $q_i \in M_i$ denoted by $T_{q_i}M_i$ for $i = 0, 1$. The case of M_0 and M_1 can be found in the case of curves considered up to reparameterization. The functional (4.1) may be invariant for the action of the group of diffeomorphisms of S_1 and we aim to retrieve orthogonality relations for the momentum from the Pontryagin maximum principle ([AS04], [Tré05]).

With P the dual of Q, the control on $q \in Q$ is $v \in V$ with an instantaneous cost function $\frac{1}{2}|v|_V^2$ and we have $\dot{q} = v.q$. Then, the Hamiltonian system associated to this minimization is

$$H(p,q,v) = (p, v.q)_{P\times Q} - \frac{1}{2}\langle v, v\rangle_V \,. \tag{4.3}$$

Before minimizing in v, we need to assume that

$$V \mapsto \mathbb{R}$$
$$v \mapsto (p, v.q)_{P\times Q}$$

is a continuous linear form on V for any $p, q \in P \times Q$ (hypothesis (H1)).
For example in the case of landmarks the hypothesis (H1) just says that the pointwise evaluation on V is continuous, i.e. V is a RKHS of vector fields.
Then, by Riesz theorem it defines an element in V^* that we denote $p \diamond q$ defined by the equation:

$$(p, v.q)_{P\times Q} = -(p \diamond q, v)_{V^*\times V} \,.$$

As far as we know, this diamond operator was introduced in this domain by D.Holm and L.Younes. This notation appears naturally in [HTY08] when deriving the Euler equations once the metamorphosis problem is re-casted into an Euler-Poincaré variational framework. It is also well exlained and used in [CH] for the application of the Clebsch method to EPDiff.
The second term of equation (4.3) can be rewritten as $(Lv, v)_{V^*\times V}$. Then, at a minimum we can differentiate in v to obtain

$$p \diamond q + Lv = 0\,,$$
$$v = -K(p \diamond q)\,.$$

Hence, the 'minimized' Hamiltonian is,

$$H(p,q) = (p \diamond q, Kp \diamond q)_{V^*\times V} - \frac{1}{2}\langle v, v\rangle_V = \frac{1}{2}(p \diamond q, Kp \diamond q)\,. \tag{4.4}$$

The Pontryagin maximum principle says that a minimizer of the problem (4.2) verifies the following Hamiltonian system

$$\begin{cases} \dot{p} = -\partial_q H(p,q) \\ \dot{q} = \partial_p H(p,q)\,, \end{cases}$$

with orthogonality conditions

$$p(0) \perp T_{q(0)}\,,$$
$$p(1) \perp T_{q(1)}\,.$$

At this point, we need to give a sense to $\partial_q H(p,q)$ in (4.5). Assuming that $q \mapsto (p, v.q)$ is differentiable for every $q \in Q$, we write $\delta q \mapsto \partial_q(p, v.q)(\delta q)$. In addition, we assume that $\partial_q(p, v.q)(\delta q)$ is a linear form on V (hypothesis (H2)) that we denote

$$\partial_q(p, v.q)(\delta q) = -(\partial_q(p \diamond q)(\delta q), v)_{V^*\times V}\,.$$

The differentiation of $H(p,q)$ is now clear,

$$\partial_q H(p,q) = (\partial_q(p \diamond q), K(p \diamond q)) .$$

The second hypothesis (H2) says that the pointwise evaluation for the first derivative of the vector fields is continuous on V. In particular, if V is admissible this condition is fulfilled.

Let us detail the case of curves following the point of view adopted in [GTL06]. We consider generalized closed curves i.e. we will work on $Q = L^2(S_1, \mathbb{R}^2)$. Of course, this framework can be extended to $Q = L^2(M, \mathbb{R}^d, \mu)$ with M a compact Riemannian manifold and μ its associated measure, for instance $M = S_n$ the n dimensional sphere or $M = T_n$ the n dimensional torus. The action by G_V is simply the left composition on Q. The map on $V \times Q$ induced by the action of G_V is also the left composition with $v \in V$: $(v, q) \mapsto v \circ q$ and the hypothesis (H1) is verified if V is an admissible space of vector fields since:

$$\int_M \langle v(q(s)), p(s) \rangle d\mu(s) \leq |p|_{L^2} |v \circ q|_{L^2} \leq |p|_{L^2} |v|_\infty \sqrt{\mu(M)} .$$

The second hypothesis (H2) is verified in this case too replacing v by dv. Remark that $[\partial_q v \circ q].\delta q = [dv \circ q](\delta q) \in L^2(M, \mathbb{R}^d, \mu)$.

The transversality conditions are interesting in the case of curves considered up to reparameterization. If the initial curve c_0 and the final one c_1 are smooth enough the action of the diffeomorphism group of S_1 gives tangent vectors at c_i for $i = 0, 1$. Let w be a smooth vector field on S_1, if ψ_t is the flow generated by w, we have $\frac{d}{dt}_{t=0} \phi_t(x) = w(x)$ then,

$$\{ s \mapsto w(s) c_i'(s) \,|\, w \in \mathfrak{X}^\infty(S_1) \} \subset T_{c_i} .$$

The orthogonality condition says that $p_i \perp T_{c_i}$ and considering all the choices for w (any smooth vector field on S_1) we obtain that

$$\langle p_i(s), c_i'(s) \rangle = 0 \; a.e. \; s \in S_1 .$$

The case of images does not belong to this framework since the differentiation of $t \mapsto I_t = I \circ \phi_{0,t}^{-1}$ leads to $\langle \nabla I, v \rangle \in L^2(U)$ if I is H^1 (see the geodesic equation for smooth images derived in 1.2.2). Hence this path is not differentiable in H^1 but only in L^2. However we can hope for a weak formulation of the Hamiltonian system as showed in 1.2.2. Assuming I to be a SBV function, we would get by differentiation of the path $t \mapsto I_t$ a Borel function defined on the jump set of I, J_I. It means that the deformation of jump set J_I can be identified and consequently the jump set should be treated as a variable position.

4.2 The case of discontinuous images

Since this Chapter is devoted to the Hamiltonian formulation of the model presented in section 3.2, we will keep the same notations. In the whole section, we maintain our pre-

vious assumptions on the kernel for the existence of solutions for all time of proposition 13.

4.2.1 Weak formulation

In this paragraph, we present the approach of decomposing an image in "more simple parts". Let I be a piecewise Lipschitz function. Let us introduce the position variables. We will consider in the following that the discontinuity boundary is one of the position variables. Instead of considering the function I_t as the second position variable (that would be redundant), we introduce a product space which can be projected on the space $\mathrm{Im}(M)$. Let $(U_1, \dots, U_n)$ be a partition in Lipschitz domain of M (to be less demanding, we could take the partition introduced in remark 3). We denote by $\Sigma_0 = \cup_{i=1}^n \partial U_i$ the union of the boundaries of the Lipschitz domains. We consider the projection:

$$\begin{aligned} p : \prod_{i=1}^n W^{1,\infty}(M) &\mapsto \mathrm{Im}(M) \\ (I^i)_{i=1,\dots,n} &\mapsto I = \sum_{i=1}^n I^i \mathbf{1}_{U_i} . \end{aligned}$$

Of course, we could have treated this case adopting only I_t as position variable, but we find the idea of decomposing an image into more simple parts rich enough to be studied. Observe that Σ_0 is endowed with an important role in the definition of the projection: to write down a Hamiltonian system on the large space, we need to introduce the deformation of Σ_0 and the deformation of each function in the product space. We will derive the Hamiltonian equations from this optimal control problem: (the position variable is Q and the control variable is U, $c(U)$ is the instantaneous cost function)

$$\begin{cases} Q = (Q_i)_{0\le i\le r} = (\Sigma, (I^i)_{1\le i\le r}) \in L^1(\Sigma_0, M) \times W^{1,\infty}(M)^r , \\ U = (v, s) \in V \times S , \\ \dot{Q} = f(Q, U) = (v \circ Q^0, (-\langle \nabla Q^i, v\rangle + s(Q^i))_{1\le i\le r}) , \\ c(U) = \frac{\lambda}{2}|v|_V^2 + \frac{\beta}{2}|s|_S^2 . \end{cases}$$

The cotangent space of the position variable contains $F = L^\infty(\Sigma, \mathbb{R}^n) \times L^1(M, \mathbb{R})^r$. We write the **formal** minimized Hamiltonian of the control system on the subspace F, with $P \in F$:

$$H(P, Q) = \min_U \int_{\Sigma_0} \langle P_0(x), \dot{Q}_0(x)\rangle d\mu_{|\Sigma_0}(x) + \sum_{i=1}^r \int_M P^i(x)\dot{Q}_i(x)d\mu(x) - c(U) . \tag{4.5}$$

Minimizing in U, we obtain optimality conditions in (u, v) a minimizer such that for any perturbation $(\delta v, \delta s)$:

$$\begin{aligned}\lambda\langle v,\delta v\rangle &= \int_{\Sigma_0}\langle P^0,\delta v\circ Q^0\rangle d\mu_{|\Sigma_0} - \sum_{i=1}^{r}\int_M P^i\langle\nabla Q^i,\delta v\rangle d\mu\,,\\ \beta\langle s,\delta s\rangle &= \sum_{i=1}^{r}\int_M P^i\delta s(I^i)d\mu\,.\end{aligned}$$

Using the kernel, it can be rewritten,

$$\begin{aligned}\lambda v(.) &= \int_{\Sigma_0} k(Q^0(x),.)P^0(x)d\mu_{|\Sigma_0}(x) - \sum_{i=1}^{r}\int_M k_V(x,.)P^i(x)\nabla Q^i(x)d\mu(x)\,,\\ \beta s(.) &= \sum_{i=1}^{r}\int_M k_S(Q^i(x),.)P^i(x)d\mu\,.\end{aligned} \tag{4.6}$$

We deduce the expression of the Hamiltonian,

$$\begin{aligned}H(P,Q) &= \frac{1}{2\lambda}[\int_{\Sigma_0}\int_{\Sigma_0} P^0(x)k_V(Q^0(x),Q^0(y))P^0(y)d\mu_{|\Sigma_0}(x)d\mu_{|\Sigma_0}(y)\\ &+ \int_M\int_M P^j(y)\nabla Q^j(y)k_S(y,x)\nabla Q^i(x)P^i(x)d\mu(x)d\mu(y)\\ &- 2\sum_{1\le i\le r}\int_M\int_{\Sigma_0} P^0(y)k_V(Q^0(y),x)P^i(x)\nabla Q^i(x)d\mu(x)d\mu_{|\Sigma_0}(y)]\\ &+ \frac{1}{2\beta}\sum_{1\le i,j\le r}\int_M\int_M P^j(y)k_S(Q^j(y),Q^i(x))P^i(x)d\mu(x)d\mu(y)\,.\end{aligned}$$

Now, we want to give a sense to the Hamiltonian equations, $\forall\ i\in[1,r]$:

$$\begin{cases}\dot{Q}^0_t = \partial_{P^0}H(P_t,Q_t)(.)\,,\\ \dot{Q}^i_t = \partial_{P^i}H(P_t,Q_t)\ \forall\ i\in[1,r]\,,\\ \dot{P}^0_t = -\partial_{Q^0}H(P_t,Q_t),\\ \dot{P}^i_t = -\partial_{Q^i}H(P_t,Q_t)\ \forall\ i\in[1,r]\,.\end{cases} \tag{4.7}$$

These derivatives should be understood as distributions, for $\Psi\in C^\infty_0(M,\mathbb{R})$ and $u\in C^\infty_0(M,\mathbb{R}^n)$ and with the notation introduced in (4.6), $\forall\ i\in[1,r]$:

$$\begin{aligned}\partial_{P^0}H(P,Q)(u) &= \int_{\Sigma_0}\langle v\circ Q^0(y),u(y)\rangle d\mu_{|\Sigma_0}(y)\,,\\ \partial_{P^i}H(P,Q)(\Psi) &= \int_M \Psi(y)\left(s(Q^i(y)) - \langle v(y),\nabla Q^i(y)\rangle\right)d\mu(y)\,,\\ \partial_{Q^0}H(P,Q)(u) &= \int_{\Sigma_0}\langle[dv]_{|Q^0(y)}(u(y)),P^0(y)\rangle d\mu_{|\Sigma_0}(y)\,,\\ \partial_{Q^i}H(P,Q)(\Psi) &= \int_M \Psi(y)[ds]_{|Q^i(y)}P^i(y) - \langle v(y),P^i(y)\nabla\Psi(y)\rangle d\mu(y)\,.\end{aligned}$$

Remark that only the last equation really needs to be defined as a distribution and not as a function. Now we can give a sense to the Hamiltonian equations in a weak sense:

Definition 11. *A map* $\chi \in C^0([0,T], L^1(\Sigma_0, M)\times W^{1,\infty}(M)^r \times L^1(\Sigma_0,\mathbb{R}^n)\times L^1(M,\mathbb{R})^r)$ *is said to be a weak solution if it verifies for* $\Psi \in C_0^\infty([0,T]\times M,\mathbb{R})$ *and* $u \in C_0^\infty([0,T]\times M,\mathbb{R}^n)$*: (we denote* $\chi(t) = (Q_t, P_t)$*.)*

$$\int_0^T \int_M -\partial_t\Psi\ Q_t^i d\mu dt = \int_0^T \partial_{P^i} H(P_t,Q_t)(\Psi)\ dt \ \forall\ i \in [1,r]\,,$$
$$\int_0^T \int_M -\partial_t\Psi\ P_t^i d\mu dt = -\int_0^T \partial_{Q^i} H(P_t,Q_t)(\Psi)\ dt \ \forall\ i \in [1,r]\,,$$
$$\int_0^T \int_{\Sigma_0} -\partial_t u\ Q_t^0 d\mu_{|\Sigma_0} dt = \int_0^T \partial_{P^0} H(P_t,Q_t)(u)\ dt\,,$$
$$\int_0^T \int_{\Sigma_0} -\partial_t u\ P_t^0 d\mu_{|\Sigma_0} dt = -\int_0^T \partial_{Q^0} H(P_t,Q_t)(u)\ dt\,.$$

4.2.2 Uniqueness of the weak solutions

In this paragraph, the uniqueness to the weak Hamiltonian equations is proven and the proof gives also the general form of the solutions.

Theorem 23. *There exists at most only one weak solution* χ *verifying* $\chi(0) = (Q_0, P_0)$*. Moreover, if this solution exists, there exists an element of* $L^2([0,T], V\times S)$ *which generates the flow* $(\phi_{0,t}, \eta_{0,t})$ *such that:*

$$\begin{aligned} Q_t^0(x) &= \phi_{0,t}(x), x \in \Sigma_0\,,\\ Q_t^i(u) &= \eta_{0,t}\circ Q_0^i \circ \phi_{t,0}(u), u \notin \phi_{0,t}(\Sigma_0), i \in [0,n]\,, \end{aligned}$$

and for the momentum variables:

$$\begin{aligned} P_t^0(x) &= d[\phi_{0,t}]_x^{-1*}(P_0^0(x)), x \in \Sigma_0\,,\\ P_t^i(u) &= P_0^i \circ \phi_{t,0}(u)\mathrm{Jac}(\phi_{t,0}(u)) d[\eta_{t,0}]_{Q_t^i(u)}, u \notin \phi_{0,t}(\Sigma_0)\,. \end{aligned}$$

Proof: Let χ be a weak solution on $[0,T]$, we introduce

$$t \mapsto v_t(.) = \frac{1}{\lambda}\int_{\Sigma_0} k(Q_t^0(x),.)P_t^0(x) d\mu_{|\Sigma_0}(x) - \sum_{i=1}^r \int_M k_V(x,.)P_t^i(x)\nabla Q_t^i(x) d\mu(x),$$

which lies in $L^2([0,T],V)$. This vector field is uniquely determined by the weak solution. Thanks to the result recalled in introduction (chapter 2, section 1.1), we deduce that $\phi_{0,t}(x) = \int_0^t v_s(\phi_{s,0}(x))ds$ is well defined. We introduce also,

$$t \mapsto s_t(.) = \sum_{i=1}^r \int_M k_S(Q_t^i(x),.)P_t^i(x) d\mu(x)\,.$$

The flow of η is defined as well, $\eta_{0,t}(x) = \int_0^t s_r(\eta_{r,0}(x))dr$. Introducing $\tilde{Q}_t^i(x) = \eta_{t,0}\circ Q_t^i \circ \phi_{0,t}(x)$ for $i \in [1,r]$, we obtain with $S_t \circ \eta_{t,0}(x) = \partial_t \eta_{t,0}(x)$ and $V_t \circ \phi_{0,t}(x) = \partial_t \phi_{0,t}(x)$:

$$\begin{aligned}
\int_0^T \int_M -\partial_t\Psi\ \tilde{Q}_t^i d\mu &= \int_0^1 \int_M -\partial_t\Psi\ \eta_{t,0}\circ Q_t^i\circ\phi_{0,t} d\mu dt \\
&= \int_0^T \int_M -\eta_{t,0}\circ Q_t^i\ [\partial_t\Psi]\circ\phi_{t,0}\mathrm{Jac}(\phi_{t,0}) d\mu dt \\
&= \int_0^T \int_M -\eta_{t,0}\circ Q_t^i\ (\partial_t[\Psi\circ\phi_{t,0}] - <\nabla\Psi\circ\phi_{t,0}, v_t\circ\phi_{t,0}>)\mathrm{Jac}(\phi_{t,0}) d\mu dt \\
&= \int_0^T \int_M (S_t(\tilde{Q}_t^i) - <\nabla\tilde{Q}_t^i, V_t> + d\eta_{t,0}(\dot{Q_t^i}\circ\phi_{0,t}))\Psi\ d\mu dt \\
&= \int_0^T \int_M (S_t(\tilde{Q}_t^i) - <\nabla\tilde{Q}_t^i, V_t> + \\
&\quad d\eta_{t,0}(- <\nabla Q^i\circ\phi_{0,t}, v\circ\phi_{0,t}> + s(Q^i\circ\phi_{0,t})))\Psi\ d\mu dt\,.
\end{aligned}$$

Thanks to the group relation of flows of vector fields:

$$S_t + d\eta_{t,0}(s_t\circ\eta_{0,t}) = 0\,,$$

the first and last terms vanish together. The remaining terms vanish too because of the relations:

$$\begin{aligned}
\nabla\tilde{Q}_t^i &= d\phi_{0,t}^*\left(d\eta_{t,0}(\nabla Q_t^i\circ\phi_{t,0})\right), \\
v_t &+ d\phi_{0,t}(V_t\circ\phi_{t,0}) = 0.
\end{aligned}$$

Then, we conclude:

$$\int_0^T \int_M -\partial_t\Psi\ \tilde{Q}_t^i\ d\mu dt = 0\,.$$

Introducing $\Psi(t,x) = \lambda(t)\gamma(x)$, with $\lambda \in C^\infty([0,T],\mathbb{R})$ and $\gamma \in C^\infty(M)$, we have: $\int_0^T -\lambda'(t)(\int_M \gamma(x)\ \tilde{Q}_t^i d\mu)dt = 0$, hence: $\int_M \gamma(x)\ \tilde{Q}_t^i d\mu = \int_M \gamma(x)\ \tilde{Q}_t^i d\mu$, i.e. $\tilde{Q}_t^i = \tilde{Q}_0^i = Q_0^i$, and

$$Q_t^i = \eta_{0,t}\circ Q_0^i\circ\phi_{t,0}\,.$$

Now, we introduce for $i \in [1,r]$, $\tilde{P}_t^i(.) = \frac{P_t^i\circ\phi_{0,t}\mathrm{Jac}(\phi_{0,t}(.))}{d[\eta_{t,0}]_{\eta_{0,t}\circ Q_0^i(.)}}$, this quantity is well defined because of the invertibility of the flow of s_t and v_t. Remark that $\frac{\mathrm{Jac}(\phi_{0,t}(.))}{d[\eta_{t,0}]_{\eta_{0,t}\circ Q_0^i(.)}}$ is differentiable almost everywhere because Q_0^i is Lipschitz on M. We want to prove that $\int_0^1\int_M -\partial_t\Psi\ \tilde{P}_t^i d\mu dt = 0$ with $\Psi \in C_0^\infty(M)$, which leads to: $\tilde{P}_t^i = \tilde{P}_0^i = P_0^i$, and then we are done.

To prove the result, we first use the change of variable $y = \phi_{0,t}(x)$, this is a straightfor-

ward calculation. We will also use the equality: $d[\eta_{t,0}]_{\eta_{0,t}(.)} d\eta_{0,t}(.) = 1$.

$$\begin{aligned}
\int_0^T \int_M \Psi \partial_t \tilde{P}_t^i \, d\mu dt &= -\int_0^T \int_M \tilde{P}_t^i \partial_t \Psi \circ \phi_{t,0} \mathrm{Jac}(\phi_{t,0}) \, d\mu dt \,, \\
\int_0^T \int_M \Psi \partial_t \tilde{P}_t^i \, d\mu dt &= -\int_0^T \int_M \frac{P_t^i}{d[\eta_{t,0}]_{Q_t^i(.)}} \partial_t \Psi \circ \phi_{t,0} \, d\mu dt \,, \\
\int_0^T \int_M \Psi \partial_t \tilde{P}_t^i \, d\mu dt &= \int_0^T \int_M -P_t^i \partial_t \left(\frac{\Psi \circ \phi_{t,0}}{d[\eta_{t,0}]_{Q_t^i(.)}}\right) + \frac{P_t^i}{d[\eta_{t,0}]_{Q_t^i(.)}} < \nabla \Psi_{|\phi_{t,0}}, -d\phi_{t,0}(v_t) > \\
&+ P_t^i \Psi \circ \phi_{t,0} \, \partial_t (d[\eta_{0,t}]_{Q_0^i \circ \phi_{0,t}}) d\mu dt \,, \qquad (4.8)
\end{aligned}$$

The third term of the last equation (4.8) can be rewritten:

$$\begin{aligned}
&\int_0^T \int_M P_t^i \Psi \circ \phi_{t,0} \, \partial_t (d[\eta_{0,t}]_{Q_0^i \circ \phi_{0,t}}) \, d\mu dt = \\
&\int_0^T \int_M P_t^i \Psi \circ \phi_{t,0} \, (< \nabla (d[\eta_{0,t}]_{Q_0^i}), -d\phi_{t,0}(v_t) > + d[s_t]_{Q_t^i} d[\eta_{0,t}]_{Q_0^i \circ \phi_{0,t}}) \, d\mu dt \,. \qquad (4.9)
\end{aligned}$$

Now, we can apply the hypothesis on ∂P_t^i, the first term of the expression is equal to:

$$\begin{aligned}
\int_0^T \int_M -P_t^i \partial_t \left(\frac{\Psi \circ \phi_{t,0}}{d[\eta_{t,0}]_{Q_t^i(.)}}\right) d\mu dt = \int_0^T \int_M P_t^i < \nabla (\Psi \circ \phi_{t,0} d[\eta_{0,t}]_{Q_0^i \circ \phi_{0,t}}), v_t > \\
- \Psi \circ \phi_{t,0} \, d[\eta_{0,t}]_{Q_0^i \circ \phi_{0,t}} d[s_t]_{Q_t^i} P_t^i \, d\mu dt,
\end{aligned}$$

We have also,

$$\begin{aligned}
&\int_0^1 \int_M P_t^i < \nabla (\Psi \circ \phi_{t,0} d[\eta_{0,t}]_{Q_0^i \circ \phi_{0,t}}), v_t > \, d\mu dt = \\
&\int_0^1 \int_M P_t^i < d\phi_{t,0}^* (\nabla \Psi_{|\phi_{t,0}}), v_t > d[\eta_{0,t}]_{Q_0^i \circ \phi_{0,t}} + P_t^i < d\phi_{0,t}^* \nabla (d[\eta_{0,t}]_{Q_0^i}), v_t > \, d\mu dt.
\end{aligned}$$

All the terms of the equation vanish together, so we obtain the result.

With $\tilde{P}_t^0(x) = d[\phi_{0,t}]_x^* P_t^0(x)$ and $\tilde{Q}_t^0 = \phi_{t,0} \circ Q_0^0$, we get the result for the last two terms of the system in the same way than the preceding equations, but it is even easier. □

Remark that we only need to assume the weak solution to be L^2 to obtain the result. Yet we would need to give a sense to the initial condition with L^2 paths, that's why we prefer to deal with the continuity for weak solutions.

4.2.3 On the existence of weak solutions

For any initial data (Q_0, P_0) does there exist a weak solution (definition 11) to the Hamiltonian equations (4.7)? We can give a partial answer thanks to our previous work on the existence of geodesics in section 3.2.2. The answer is only partial since the Hamiltonian system we are studying is a little more general than the situation which arises in the variational minimization of the functional (3.1). We reformulate this result in,

Theorem 24. *Let $(U_1, \dots, U_n)$ be a partition in Lipschitz domain of M, $\Sigma_0 = \cup_{i=1}^n \partial U_i$. For any initial data, $I_0 \in W^{1,\infty}(M)^r$, $Q_0 = (\Sigma_0, I_0)$ and $P_0 \in L^\infty(\Sigma_0, \mathbb{R}^n) \times L^1(M, \mathbb{R})^r$ such that Supp$(P_0^i) \subset U_i$, then there exists a solution to the Hamiltonian equations which is defined for all time.*

The assumption on the support of each part of the momentum is verified whenever the cost function is invariant w.r.t. any variation of the image Q^i on $M \setminus U_i$. Remark that this solution has the same structure than a minimizer of (3.1). In this case, the momentum variable can be decomposed in two variables: one function defined on the whole domain by $P_t = \sum_{i=1}^r P_t^i$ and the momentum on the Σ.
Now, what can we say about the existence in all time of the existence of weak solutions in the general case (when we relax the condition on the support of P_0^i)? We guess that the existence is somehow a by-product of the last section, but we will not give any proof of such result.
In this context, a multilayer approach will take advantage of the decomposition of the position variable in several functions. It seems interesting to explore the case where several objects are identified in the image (even in the medical imaging case) with their own mode of deformation. For example, a single object can evolve on a background and we can assume that the background and the object have two different deformations. In our optimal control point of view, the control is driven by two vector fields $(v_1, v_2) \in V_1 \times V_2$ (the same for the contrast term) and we would have the open set U_1 of the object on the background which determines the Lipschitz partition. Then, the control would be,

$$\dot{Q} = f(Q, U) = (v_1 \circ Q^0, (-\langle \nabla Q^1, v_i \rangle + s_i(Q^i))_{i=1,2}) \,.$$

The induced Hamiltonian system would be roughly the same but changing the kernel. A variational problem that can be associated to this case would be

$$\begin{aligned} \mathcal{E} = D(Id, (\eta_{0,1}^1, \phi_{0,1}^1))^2 + D(Id, (\eta_{0,1}^2, \phi_{0,1}^2))^2 \\ + \|\eta_{0,1}^1 \circ [Q_1 \mathbf{1}_{U_1}] \circ \phi_{1,0}^1 + \eta_{0,1}^2 \circ [Q_2 \mathbf{1}_{U_2}] \circ \phi_{1,0}^2 - I_{targ}\|_{L^2}^2 \,, \end{aligned} \quad (4.10)$$

with $U_2 = M \setminus U_1$.

4.3 Conclusion

Summing up our work at this point, from a precise variational problem (minimizing the functional (3.1)) we obtained generalized Hamiltonian equations, for which we proved uniqueness of the solutions, and some results about their existence. This formulation may have interesting extensions (multilayer approach) which do not increase the complexity of this framework. Apart from a multilayer approach, the introduction of an other vector field for the contrast is also a generalization to deal with multi modal matching, which is the next step of the model. Before such developments, we need to implement the model with adapted numerical algorithms.

Chapter 5

A stochastic model for growth of shapes

Contents

5.1 Introduction 90
5.2 The landmark case 93
5.3 On the limit of the stochastic model 95
5.4 The spaces P and Q 98
5.5 Cylindrical Brownian motion and stochastic integral 109
5.6 Solutions to the SDE on $P \times Q$ 113
5.7 Approximations Lemmas 118
5.8 Applications and numerical simulations 121
5.9 Possible extensions and conclusion 127

5.1 Introduction

Our aim is to better understand the statistical analysis of shape spaces. Studying the geometric properties of an infinite dimensional manifold of shape spaces such as geodesics or curvature is meaningful and was done in many papers in [Mic08], [MM07] and [MM06]. These works will provide a better understanding of how to perform statistics on the tangent spaces (see [Fle04]). At least it should give a restriction on the domain where Karcher mean or other statistical quantities can be estimated in a sound way. This statistical issue comes once the deterministic problem is solved: the minimization of a functional which is the sum of a distance on the group of diffeomorphisms and a penalty term. Yet this matching procedure can be motivated by a Bayesian approach as presented in [DUM98]: the rigorous asymptotic theory that connects the two procedures was developed very recently in [BDV07]. The point of view taken in that article is to study the large deviations properties of stochastic flows of diffeomorphisms. However this approach does not lead to a generative model and study asymptotic deviations in small time. The probabilistic foundations for this work are given by the properties of stochastic flows of diffeomorphisms deeply developed in [Kun97]. The starting point is to interpret the

minimization of the following functional

$$\mathcal{J}(u) = \int_0^1 \|u_t\|_V^2 dt + d(\phi_1.A, B)\,, \tag{5.1}$$

as a maximum a posteriori. In this equation, $u_t \in L^2([0,1], V)$ with V a reproducing kernel Hilbert space of vector fields and d a distance on a class of objects provided with an action of the flow at time 1 of the vector field. The first term in 5.1 should reflect the likelihood of the transformation ϕ_1. If on the path of the flow $(\phi_t)_{t\in[0,1]}$ we assume the prior to be modeled as follows: $P((\phi_t)_{t\in[0,1]}) \approx \exp(-\int_0^1 \|u_t\|_V^2 dt)$, then as for the Brownian motion, the random object associated to this prior is the stochastic flow defined by,

$$\begin{aligned} \phi_t &= \int_0^t W_s(\circ ds, \phi_s(x))\,, \\ W_t &= \sum_{i=0}^{\infty} B_i(t) e_i\,. \end{aligned} \tag{5.2}$$

In these equations $(B_i)_{i\in\mathbb{N}}$ are i.i.d Brownian motions, $(e_i)_{i\in\mathbb{N}}$ is an orthonormal basis of V and the symbol $\circ$ stands for the Stratonovich integral. If the space of vector fields is smooth enough (regularity assumptions on the kernel), the proof of the existence of the stochastic flow can be found in [Kun97] (Theorem 4.6.5). This approach gives a probabilistic framework with non smooth trajectories with respect to the time variable: the evolution of the shape for almost trajectory through the action of the diffeomorphism will be non smooth almost everywhere in time due to the Brownian motion. We would prefer a probabilistic framework for smooth evolution of shapes, which is closer to a biological growth evolution. One important property of the model would be the smoothness (i.e. not as rough as a standard path of the Brownian motion) of the trajectories. Such a growth model could be somehow related to the approach of U.Grenander in [GSS07] but the most important difference is that our approach is non-parametric.

Back to the deterministic equations of geodesics for the landmark case, it is well known the equations can be written in a Hamiltonian form. To perturb this Hamiltonian system has at least two interesting aspects: one is the numerical tractability of these equations (for example [CH06] or [ATY05]). By introducing a white noise in this system, we expect to keep its numerical tractability. Yet the main interesting feature concerning the modeling aspect of our work is the physical interpretation of these equations. If we consider the evolution of landmarks as a physical system of particles, it seems natural to introduce a random force on their evolution: an additive white noise is added to the evolution equation of the momentum. Integrating these equations, we will obtain an evolution with the desired smoothness property which are perturbations of geodesics (evolutions without perturbation term). This idea of perturbing the evolution equations with a random force has been introduced for a long time in the stochastic fluid dynamics community ([BT73]). Our stochastic system is a stochastic perturbation of Euler-Poincaré equation (EPDiff, [CH06]). And from this point of view, this work may have some relations with the study

of stochastic perturbation of the vortex model ([AK02], for a brief survey [SFS08]). Yet we do not develop these links in this work and we will focus on this model of growth of shape as a good candidate to deal with a collection of shape evolutions at different times and to perform statistical studies on these data.

We briefly recall the Hamiltonian equations, with the classical notations: for a system of n landmarks in $\mathbb{R}^d$, p is the momentum variable and q is the position of the landmarks. With k the reproducing kernel, the energy of the system depends on q and p: $H(p,q) = \frac{1}{2}\sum_{i,j} p_i k(q_i,q_j)p_j$. We also denote the vector field $v(.) = pk(q,.)$, $dv(.) = pdk(q,.)$. We need as for in the deterministic case a regularity assumption for the underlying vector field Hilbert space.

Assumption 1. *There exists a continuous injection of the Hilbert space V of vector fields into C^1, i.e. $|dv|_\infty \leq K|v|_V$.*

We also say according to Subsection 1.1.1 that V is $1-$admissible. The stochastic differential system is:

$$dp_t = -\partial_q H(p_t,q_t)\,dt + \varepsilon dB_t\,, \tag{5.3a}$$

$$dq_t = \partial_p H(p_t,q_t)\,dt. \tag{5.3b}$$

Here, ε is a parameter and B_t is a Brownian motion on $\mathbb{R}^{dn}$ and we can think to the kernel as a diagonal kernel, for instance Gaussian or Cauchy kernel. The notation $\partial_p H$ (respectively $\partial_q H$) stands for the partial derivative of H with respect to the first variable p (respectively the second variable q).

We will first show in section 5.2 that the system 5.3 does not blow up in finite time a.s. This is the first required property to develop our probabilistic approach. The second essential property for this model is that there exists a limit model for continuously parameterized shapes such as for example closed curves in the plane. To this end we will discuss in section 5.3 the extension of this simple model on landmarks to a stochastic model of shape evolution: first in 5.3.1 we will detail a nice property of the deterministic equations which enables us to get the convergence results of the landmark case to the continuous case. Then in 5.3.2 we will derive the properties we need on the underlying functional spaces to fit the convergence condition of our stochastic model and to get the desired convergence results.
In 5.4, we will introduce the suitable spaces on which to study the stochastic system: the shape space is closed to a Sobolev space on the Haar basis. We verify the required properties and the results are extended in any dimension.
The section 5.5 is devoted to the presentation of the cylindrical Brownian motion and we give a brief presentation of the stochastic integral in infinite dimension. This section is a self-contained presentation of the material we need to prove the convergence properties in

section 5.6. In this section, we prove that solutions of the stochastic system on the shape space are defined for all time and that the finite dimensional approximations converge almost surely to the solutions of the infinite dimensional system using approximations Lemmas developed in section 5.7.
In Section 5.8, we apply the previous results to the simplest model and we present some simulations. The last Section describes possible extensions of the model and future work.

5.2 The landmark case

When ε is constant, there is no difference studying these stochastic differential equations with the Ito integral or Stratonovich one. However, when we will deal with a general variance term, we will use the Ito stochastic integral. From the theorem of existence and uniqueness of solution of stochastic differential equation under the linear growth conditions, we can work on the solutions of such equations for a large range of kernels.
Yet in our case the Hamiltonian is quadratic, and the classical results for existence and uniqueness of stochastic differential equations only prove that the solution is locally defined. Even in the deterministic case, this quadratic property can imply existence of solutions that blow up in finite time. To prove that the solutions do not blow up in finite time in the deterministic case ($\varepsilon = 0$), we use the fact that the Hamiltonian of the system is constant in time. By enhancing the deterministic case proof and controlling the Hamiltonian, we will prove that the solutions are defined for all time.
Thus, we introduce the stopping times defined as follows: let $M > 0$ be a constant and

$$\tau_M = \{t \geq 0 \,|\, \max(|q_t|, |p_t|) \geq M\}\,, \tag{5.4}$$

let also $\tau_\infty = \lim_{M\to\infty} \uparrow \tau_M$ be the explosion time.
Differentiating $H(p_{t\wedge\tau_M}, q_{t\wedge\tau_M})$ with respect to t, we get on $(t < \tau_M)$:

$$dH(t) = \partial_q H(p_t, q_t)dq_t + \partial_p H(p_t, q_t)dp_t + \frac{\varepsilon^2 nd}{2}dt\,.$$

In the deterministic case the Hamiltonian is constant, whereas here the stochastic perturbation gives

$$\partial_q H(p_t, q_t)dq_t + \partial_p H(p_t, q_t)dp_t = \varepsilon\partial_p H(p_t, q_t)dB_t\,.$$

$$\int_0^{T\wedge\tau_M} dH(t) = \int_0^{T\wedge\tau_M} \varepsilon\langle\partial_p H(p_t, q_t), dB_t\rangle + \int_0^{T\wedge\tau_M} \frac{\varepsilon^2 nd}{2}dt\,,$$

$$E[H(p_{T\wedge\tau_M}, q_{T\wedge\tau_M})] \leq H(0) + E(\frac{\varepsilon^2}{2}dn\,T\wedge\tau_M) \leq H(0) + \varepsilon^2 dnT\,.$$

Now, we aim to control $q_{t\wedge\tau_M}$ using the control on dq_t given by $|\partial_p H(p_t, q_t)|_\infty \leq K\sqrt{H(p_t, q_t)}$:

$$\begin{aligned}|q_{\tau_M\wedge t}| \leq |q_0| + \int_0^{\tau_M\wedge t} KH(p_s, q_s)^{1/2}ds &\leq |q_0| + \int_0^{\tau_M\wedge t} KH(p_{s\wedge\tau_M}, q_{s\wedge\tau_M})^{1/2}ds \\ &\leq A_t \doteq |q_0| + \int_0^{\tau_\infty\wedge t} KH(p_{s\wedge\tau_\infty}, q_{s\wedge\tau_\infty})^{1/2}ds\,. \end{aligned} \tag{5.5}$$

However, $0 \leq A_t$ P a.s. and by monotone convergence theorem (recall that H is non negative),

$$E(A_t) = \lim_{M\to\infty}(|q_0| + E\left(\int_0^{t\wedge\tau_M} KH(p_{s\wedge\tau_M}, q_{s\wedge\tau_M})^{1/2}ds\right).$$

Also,

$$\begin{aligned} E\left(\int_0^{t\wedge\tau_M} H(p_{s\wedge\tau_M}, q_{s\wedge\tau_M})^{1/2}ds\right) &\leq E\left(\int_0^{t} H(p_{s\wedge\tau_M}, q_{s\wedge\tau_M})^{1/2}ds\right) \\ &\overset{Fub.}{=} \int_0^t E\left(H(p_{s\wedge\tau_M}, q_{s\wedge\tau_M})^{1/2}\right)ds \overset{Jen.}{\leq} \int_0^t E\left(H(p_{s\wedge\tau_M}, q_{s\wedge\tau_M})\right)^{1/2}ds \\ &\overset{CS+(5.5)}{\leq} \sqrt{t}\left(\int_0^t (H(0) + \epsilon^2 dns)\,ds\right)^{1/2}. \quad (5.6) \end{aligned}$$

We deduce

$$E(A_t) \leq |q_0| + K\sqrt{t}\left(\int_0^t (H(0) + \epsilon^2 dns)\,ds\right)^{1/2} < \infty \text{ and } A_t < \infty\ P\ a.s.$$

and as a consequence

$$\limsup_{M\to\infty} |q_{t\wedge\tau_M}| < +\infty\, P\, a.s.$$

We also control the evolution equation of the momentum as follows,

$$|p_{t\wedge\tau_M}| \leq \int_0^{t\wedge\tau_M} |\partial_q H(p_s, q_s)|\,ds + |p_0 + \int_0^{t\wedge\tau_M} \varepsilon dB_s\,|. \quad (5.7)$$

Now we use the assumption 1 to control $\partial_q H(p,q)$:

$$|\partial_q H(p,q)| \leq |p||dv(q)| \leq K|p|H^{1/2}.$$

We rewrite inequality 5.7 and we use Gronwall Lemma 18 to get:

$$\begin{aligned} |p_{t\wedge\tau_M}| &\leq \int_0^{t\wedge\tau_M} K|p_s|\,H(p_s,q_s)^{1/2}\,ds + |p_0 + \int_0^{t\wedge\tau_M} \varepsilon dB_s\,|, \\ |p_{t\wedge\tau_M}| &\leq \left(|p_0| + \sup_{u\leq t}|\int_0^{u\wedge\tau_M} \varepsilon dB_s\,|\right) e^{\int_0^{t\wedge\tau_M} KH(p_s,q_s)^{1/2}ds}, \\ |p_{t\wedge\tau_M}| &\leq \left(|p_0| + \sup_{u\leq t\wedge\tau_\infty}|\int_0^{u} \varepsilon dB_s\,|\right) e^{\int_0^{t\wedge\tau_\infty} KH(p_s,q_s)^{1/2}ds}. \end{aligned}$$

The first term on the right hand side $|p_0| + \sup_{u\leq t\wedge\tau_\infty}|\int_0^u \varepsilon dB_s\,|$ is bounded by $|p_0| + \sup_{u\leq t}|\int_0^u \varepsilon dB_s\,| < \infty\, P\, a.s.$ and with inequation 5.6 we have that

$$e^{\int_0^{t\wedge\tau_\infty} KH(p_s,q_s)^{1/2}ds} < \infty\, P\, a.s.$$

Since on $(\tau_\infty \leq t)$ one has

$$\lim_{M\to\infty} \max(|q_{t\wedge\tau_M}|, |p_{t\wedge\tau_M}|) = \lim_{M\to\infty} |p_t| = \infty,$$

we deduce $P(\tau_\infty \leq t) = 0$ and $\tau_\infty = +\infty$ almost surely.

Remark 11. *Note that the proof is the same if ε would have been replaced by a bounded Lipschitz function of (p, q). The local existence is ensured by the Lipschitz condition and the property of non blow up is conserved thanks to the bound on ε.*

We have proven for the case $\varepsilon(p, q) = \varepsilon Id$,

Proposition 15. *Under assumption 1, the solutions of the stochastic differential equation defined by*

$$\begin{aligned} dp_t &= -\partial_q H(p_t, q_t)dt + \varepsilon(p_t, q_t)dB_t \\ dq_t &= \partial_p H(p_t, q_t)dt. \end{aligned}$$

do not blow up, with $\varepsilon : \mathbb{R}^{nd} \times \mathbb{R}^{nd} \mapsto L(\mathbb{R}^{nd})$ a Lipschitz and bounded function.

Proof. To extend the proof to the case ε a Lipschitz and bounded function of p and q, we can prove that the preceding inequalities are still valid.
First, thanks to the Lipschitz property of ε the solutions are still defined locally. The Ito formula is now written as, on $(t < \tau_M)$

$$dH(t) = \partial_q H(p_t, q_t)dq_t + \partial_p H(p_t, q_t)dp_t + \frac{1}{2}\text{tr}(\varepsilon^t(p_t, q_t)\varepsilon(p_t, q_t))dt\,,$$

We still have the inequality (5.5) with $\text{tr}(\varepsilon^t(p_t, q_t)\varepsilon(p_t, q_t)) \leq C|\varepsilon|^2_\infty$ (C is a constant which only depends on the chosen norm on $L(\mathbb{R}^{nd})$) and $|\varepsilon(p, q)w|^2 \leq C|\varepsilon|^2_\infty |w|_\infty$,

$$\begin{aligned} \int_0^{T\wedge\tau_M} dH(t) &\leq \int_0^{T\wedge\tau_M} \langle \partial_p H(p_t, q_t), \varepsilon(p_t, q_t)dB_t\rangle + \int_0^{T\wedge\tau_M} \frac{|\varepsilon|^2_\infty nd}{2}dt\,, \\ E[H(T \wedge \tau_M)] &\leq H(0) + E(\frac{C}{2}|\varepsilon|^2_\infty dn\, T \wedge \tau_M) \leq H(0) + C|\varepsilon|^2_\infty dn\, T\,. \end{aligned}$$

and all the remaining inequalities follow easily thanks to the control on H and the bound on ε. □

The first comment is that this model perturbs the trajectories of the deterministic system. When $\varepsilon \to 0$, the solutions of the system 5.8 converge to the corresponding geodesic $\varepsilon = 0$.

5.3 On the limit of the stochastic model

Our aim is to design a model for growth of shapes. We need to have a consistency property to extend the model to continuous shapes. Has the model got a limit when increasing the number of landmarks? The kind of properties we are looking for is inspired by the deterministic case developed in the previous Chapter 4. We will detail a very simple but important convergence property of the deterministic case in 5.3.1. Then, we will discuss in 5.3.2 informally the kind of properties we need to pass to the limit in the stochastic system.

5.3.1 On the deterministic case of curves in the plane

We develop a consequence of the Hamiltonian formulation of the equations originally written in [GTL06]. This point was not explicit in the article. This paper presents a proof of the existence in all time of the geodesic equations based on the Hamiltonian formulation and using the Banach fixed point theorem. We briefly recall the Hamiltonian formulation. The space of closed curves is the Hilbert space $H = L^2(S_1, \mathbb{R}^2)$. The moment variable should lie in the dual space of H, identified to H. This Hamiltonian system

$$\partial_q H(p_t, q_t) = -p_t(.) \int_{S_1} \partial_1 k(q_t(.), q_t(s)) p_t(s) ds, \tag{5.8a}$$

$$\partial_p H(p_t, q_t) = \int_{S_1} k(q_t(.), q_t(s)) p_t(s) ds, \tag{5.8b}$$

has solutions for all time for any initial conditions $(p_0, q_0) \in H^2$. The proof is close to the one presented for the case of landmarks in Section 5.2.
An important and simple remark is

Remark 12. *The ODE* (5.8) *conserves the common structure of p and q: if p and q are both constant (in space) on an interval (resp. a measurable set on S_1) then the solution (p, q) will be constant on this interval (resp. on this measurable set).*

The consequence of this remark is that the landmark case is a special case of the ODE (5.8). Consider the n dimensional subspace of H for $n \geq 1$, $G_n = \mathrm{Vect}_{i \in [0, n-1]}(1_{[\frac{i}{n}, \frac{i+1}{n}[})$, then G_n is one candidate to describe the trajectories of n landmarks, taking initial conditions in $(p_0, q_0) \in G_n \times G_n$. It gives also a convergence property by the continuity of solutions of a Lipschitz ODE system. With stronger assumption on the convergence of q^n but still the same assumption on the convergence for p^n, we obtain strong convergence of q^n.

Proposition 16. *Let $(p_0^n, q_0^n) \in G_n \times G_n$ be initial conditions for the system 5.8 with $\lim_{n \mapsto \infty}(p_0^n, q_0^n) = (p, q)$ then, the solutions (p_t^n, q_t^n) converge in $H \times H$ to (p_t, q_t) uniformly for t in a compact set. If $\lim_{n \mapsto \infty} \|q^n - q\|_\infty = 0$ then $\lim_{n \mapsto \infty} \|q_t^n - q_t\|_\infty = 0$ uniformly for t in a compact set.*

Proof. The first fact is the direct application of the continuity theorem for the Banach fixed point theorem with parameter. For the second point, remark that the L^2 convergence implies the L^∞ convergence of the vector field $v_t(.) = \int_{S_1} k(., q_t(s)) p_t(s) ds$ on every compact. □

5.3.2 Toward the stochastic extension

Back to the stochastic Hamiltonian system, it shows a natural limit of the system 5.3 when increasing the number of landmarks: the Brownian motion in the system (5.3) could be interpreted as the projection of a cylindrical Brownian motion on $L^2(S_1, \mathbb{R}^2)$. A white

noise on the particles can be extended to a white noise on the parameterization of the shape. Though the white noise on the parameterization space may not be a white noise on the shape (geometric). This is the reason why we will introduce a variance term to account for white noises with respect to a different measure than Lebesgue measure (which are white noises after reparameterization, see 5.8).

Let us first discuss this extension from a heuristic point of view.
We give a short definition of the white noise on $H = L^2(S_1, \mu, \mathbb{R}^2)$ with μ the Lebesgue measure. We will come back on this definition later on.

Definition 12. *Let $(B^i)_{i=0}^{+\infty}$ be a family of independent real valued Brownian motions and $(e_i)_{i=0}^{+\infty}$ an orthonormal and complete basis of H a separable Hilbert space. The process $B_t = \sum_{i=0}^{+\infty} B_t^i e_i$ is a cylindrical Brownian motion on H.*

Our approach leads to the following equations,

$$dp_t = -[p_t \int_{S_1} \partial_1 k(q_t, q_t(s)) p_t(s) ds]dt + \varepsilon dB_t, \tag{5.9a}$$

$$dq_t = [\int_{S_1} k(q_t, q_t(s)) p_t(s) ds]dt. \tag{5.9b}$$

There are important issues in the structure of this system to be discussed. We study these Hamiltonian equations (5.9) with $q \in Q$ and $p \in P$ for (P, Q) some undefined Hilbert spaces. This study will give informations on candidate spaces P, Q to develop our approach.
The first property is Q to be as 'big' as possible. Especially, we want Q to contain piecewise constant functions to account for the landmark case as described above.

Property 1. *Piecewise constant (on some finite partition of S_1 in intervals for instance) functions are contained in Q.*

The quantity $\int_{S_1} k(., q_t(s)) p_t(s) ds$ (in equation (5.9b)) does not mean anything on the space P, Q. However, it is still correctly defined for piecewise constant functions or L^2 functions. Hence a natural hypothesis on P, Q is

Property 2. *P and Q are dual and we have the injections*

$$Q \hookrightarrow L^2 \hookrightarrow P\,.$$

Still, we would need to have $k(., q_t(s)) \in Q$. Then an additional condition would be

Property 3.

If K is a smooth function, $f \in Q \mapsto K \circ f \in Q$ is locally Lipschitz.

This property is verified for example for the Sobolev space H^1.
Now the quantity $\int_{S_1} k(., q_t(s)) p_t(s) ds$ is well defined by $k_q p(.) \doteq (p, k(., q))_{P,Q}$. If $k_q p(.)$ is sufficiently smooth, then the last property gives a sense to (5.9b):

$$dq_t = k_q p \circ q dt\,.$$

Let us study equation (5.9a) which can be rewritten as

$$dp_t = -p_t dv_q p \circ q \, dt + \varepsilon dB_t \,,$$

with $dv_q p(.) \doteq (p, \partial_1 k(., q))_{P,Q}$. If this map is smooth enough, $dv_q p \circ q \in Q$ thanks to property 3. However, the equation still does not have a sense since there is the multiplication with p. The question is what kind of property do we need to give a sense to: $p(dv_q p \circ q) \in P$. An answer is

Property 4. *Q is an algebra, the multiplication is continuous (for the norm on Q).*

Indeed if $p \in Q$ and $q_0 \in Q$, then we can define $p.q_0 \in P$ defined by:

$$(p.q_0, q)_{P,Q} = (p, q_0 q)_{P,Q} \,,$$

the right hand term being continuous w.r.t. q since the product is continuous. Again, this property is verified for example for $H^1(S_1, \mathbb{R})$.
Finally the noise term should belong to P as follows

Property 5. *The paths of the cylindrical Brownian motion $t \to B_t$ lie almost surely in $C([0,T], P)$ for all $T > 0$.*

It is well known that the white noise on $L^2(S_1, \mathbb{R})$ belongs for example to H^{-1} the dual of H^1. This last property ends to give a sense to equation (5.9a).
This set of conditions is a guide to get a proper space to prove the results. We will present in the next Section a candidate for (P, Q) that fulfills all the previous properties in any dimension (curves, surfaces, ...).

Once the system is well-posed, the other issue to be discussed is the existence of solutions to this stochastic system on P, Q and the convergence of the projections (landmark case) to the infinite dimensional case. If ε is constant, the use of the fixed point theorem could work. We want however to be slightly more general and tackle the case of

$$\varepsilon : P \times Q \mapsto L(P) \,,$$

is Lipschitz and bounded which would be a natural extension of the system in proposition 15.

5.4 The spaces P and Q

In this Section, we present the spaces P, Q that verify all the properties of 5.3.2. We have seen that a Sobolev space $H^1(S_1)$ satisfies most of the required properties but it does not contain piecewise constant functions. As a consequence, we could think to a Sobolev (or Besov) space on the Haar basis. Hopefully the properties we need would be verified. However, it is not convenient to work with Sobolev spaces on Haar functions to prove the Lipschitz property of the composition with a smooth function. We slightly modify this

space to define the space $F_s = Q$ for $s > 0$ and $F_{-s} = P$ which are well suited to easily obtain the required properties of Subsection 5.3.2. This modification was inspired by the very definition of the Besov spaces and the book [LT77].

Recall that $H = L^2(S_1, \mathbb{R})$, we consider the Haar orthonormal basis with $\psi_0(x) = \chi_{[0,\frac{1}{2}[} - \chi_{[\frac{1}{2},1[}$ and $\psi_{n,k}(x) = 2^{\frac{n}{2}}\psi_0\left(2^n(x - \frac{k}{2^n})\right)$ for $n \geq 1$ and $0 \leq k \leq 2^n - 1$ and the constant function $\psi_{-1,0} := 1$.
We define the Haar coefficients of a function $f \in H$ by

$$f_{n,k} = \langle f, \psi_{n,k} \rangle_H ,$$

for $n \in [-1, +\infty[$ and $k \in A_n \doteq [0, 2^n - 1]$ if $n \geq 0$ and $A_{-1} = \{0\}$.
Let us start with a simple remark.

Remark 13. *Let $g \in L^\infty(S_1, \mathbb{R})$ be a function, we have*

$$|g_{n,k}| \leq |g|_\infty |\psi_{n,k}|_{L^1} = 2^{-\frac{n}{2}}|g|_\infty .$$

Definition 13. *We define $H_s = \{f \in H; |f|^2_{H_s} = \sum_{n=-1}^{\infty} \sum_{k\in A_n} 2^{ns}|f_{n,k}|^2 < +\infty\}$, with $s \geq 0$ a nonnegative real number. For $s < 0$, we define H_s as the dual of H_{-s}.*

We study some properties of an element in this Hilbert space.

Proposition 17. *An element $f \in H_s$ for $s > 1$ is continuous in every $x \in [0, 1]$ which is not a dyadic number, more precisely, if $a = \frac{2k+1}{2^{n+1}}$ with an integer k such that $2k + 1 \in [0, 2^{n+1} - 1]$ and $(x, y) \in B(a, \frac{1}{2^{n+1}})^2$, then with $C_s^2 = \sum_{i=0}^{+\infty} 2^{-i(s-1)} = \frac{2^{s-1}}{2^{s-1}-1}$*

$$|f(x) - f(y)| \leq C_s \frac{2}{2^{n\frac{(s-1)}{2}}} |f|_{H_s} . \tag{5.10}$$

Proof. We define for $x \in [0, 1]$ and $n \geq -1$,

$$A_{x,n} = \{k \in A_n \,|\, x \in \text{Supp}(\psi_{n,k})\}.$$

Remark that for each n there exists one and only one k in $A_{x,n}$. We will use this remark in the next inequalities. Also, the difference $|f(x) - f(y)|$ does not involve terms in the sequence that are constant on the ball $B(a, \frac{1}{2^{n+1}})$, thus we have with Cauchy-Schwarz inequality and using the remark 13,

$$\begin{aligned}
|f(x) - f(y)| &\leq \sum_{l\geq n} \left(\sum_{k\in A_{x,l}} |f_{l,k}| 2^{\frac{l}{2}} + \sum_{k\in A_{y,l}} |f_{l,k}| 2^{\frac{l}{2}} \right) , \\
|f(x) - f(y)| &\leq 2|f|_{H_s} \sqrt{\sum_{l\geq n} 2^{-l(s-1)}} , \\
|f(x) - f(y)| &\leq C_s \frac{2}{2^{n\frac{(s-1)}{2}}} |f|_{H_s},
\end{aligned}$$

which is the result. □

Remark 14. *This proof gives also that the sup norm is bounded by the H_s norm for $s > 1$:*

$$|f|_\infty \leq C_s |f|_{H_s}.$$

Now, we introduce the space F_s very close to the space H_s.

Definition 14. *We define the Hilbert space for $s \geq 0$,*

$$F_s = \{f \in H \,||f|^2 = \int_0^1 f^2\,dx + \sum_{n,k} 2^{ns-1} \int_{I_{n,k}} |f(x + \frac{1}{2^{n+1}}) - f(x)|^2 dx < \infty\},$$

with $I_{n,k} = [\frac{k}{2^n}, \frac{k}{2^n} + \frac{1}{2^{n+1}}]$. The dual for the L^2 product defines H_{-s}.

The first comment is the following proposition (a further comparison between these two spaces is developed in appendix with proposition 42).

Proposition 18. *We have the inclusion $F_s \subset H_s$ and if $s > 1$ and*

$$|f|_{H_s} \leq |f|_{F_s}.$$

Proof. To see this fact, we apply the Cauchy-Schwarz inequality

$$f_{n,k}^2 = 2^n (\int_{I_{n,k}} f(x + \frac{1}{2^{n+1}}) - f(x) dx)^2,$$
$$f_{n,k}^2 \leq \frac{1}{2} \int_{I_{n,k}} |f(x + \frac{1}{2^{n+1}}) - f(x)|^2 dx.$$

Moreover,

$$(\int_{S_1} f(x)dx)^2 \leq \int_{S_1} f(x)^2 dx,$$

so that we have $|f|_{H_s} \leq |f|_{F_s}$. □

We want our space to be big enough to contain usual functions. In the following, we prove that F_s contains Lipschitz functions for $s < 2$ and also, if $s < 2$:

$$\text{Lip}_{dyad}(S_1) \doteq \{f \in L^2(S_1, \mathbb{R}) | \exists n,\, f_{|I_{n,k}} \in \text{Lip}(I_{n,k}, \mathbb{R})\, \forall k \in [0, 2^n - 1]\} \subset F_s.$$

This fact (further developed later) is important since it means that we can deal with a wide range of shapes in this space.

Proposition 19. *If $s < 2$, F_s contains piecewise Lipschitz functions,*

$$Lip_{dyad}(S_1) \subset F_s.$$

Proof. Let $f \in \text{Lip}(S_1, \mathbb{R})$be a function and denote by M its Lipschitz constant. Then we have

$$2^{ns-1} \int_{I_{n,k}} |f(x + \frac{1}{2^{n+1}}) - f(x)|^2 dx \leq \frac{M 2^{ns-1}}{2^{3n+3}}$$

so that if $s < 2$, $f \in F_s$. □

The following proposition, though straightforward by applying the definitions, is of importance to ensure the stability of our system. For example, if G is Lipschitz and bounded, the growth of $G \circ f$ is linear:

Proposition 20. *If $s > 1$, G a real Lipschitz function and $f \in F_s$, then $G \circ f \in F_s$ and also,*

$$\begin{aligned} |G \circ f|_{F_s} &\leq \mathit{Lip}(G)|f|_{F_s} + |G \circ f|_{L^2}, \\ |G \circ f|_{F_s} &\leq \mathit{Lip}(G)|f|_{F_s} + |G(0)|. \end{aligned} \tag{5.11}$$

with $\mathit{Lip}(G)$ the Lipschitz constant for G.

Proof. Applying the Lipschitz property we have,

$$\int_{I_{n,k}} |G \circ f(x + \frac{1}{2^{n+1}}) - G \circ f(x)|^2 \, dx \leq \int_{I_{n,k}} \mathrm{Lip}(G)^2 |f(x + \frac{1}{2^{n+1}}) - f(x)|^2 \, dx \, ,$$

then,

$$|G \circ f|^2_{F_s} - |G \circ f|^2_{L^2} \leq \mathrm{Lip}(G)^2 (|f|^2_{F_s} - |f|^2_{L^2}) \, . \tag{5.12}$$

Since for any couple (a, b) of nonnegative real numbers $a^2 + b^2 \leq (a + b)^2$, we have the first inequality.
The second inequality is the application of inequality 5.12 to the function $g = G \circ f - G(0)$ using the inequality $|g|_{L^2} \leq \mathrm{Lip}(G)|f|_{L^2}$. □

We need to go further by proving that the composition is locally Lipschitz if G' is Lipschitz.

Proposition 21. *If $s > 1$, G a real C^1 function for which G' is Lipschitz (or locally Lipschitz only) then, for every $r > 0$ there exists $M > 0$ such that*

$$|G \circ f_1 - G \circ f_2| \leq M |f_1 - f_2| \, ,$$

if $(f_1, f_2) \in B(0, r)^2$. If G' and G'' are bounded then the Lipschitz constant M has linear growth,

$$M \leq \sqrt{2}(|G'|_\infty + 3C_s r |G''|_\infty).$$

Proof. For notation convenience, we will denote by for $i = 1, 2$ and $\delta > 0$,

$$\Delta f_i(x) = f_i(x + \delta) - f_i(x).$$

We will control the quantity

$$\int_{I_{n,k}} |(G \circ f_1 - G \circ f_2)(x + \frac{1}{2^{n+1}}) - (G \circ f_1 - G \circ f_2)(x)|^2 dx.$$

We will divide by Δf_1, it is permitted in this situation, since we can extend the definition with the equation (5.14). Let $a \leq b$ be two real valued numbers. We have, with μ the Lebesgue measure

$$\begin{aligned} \int_a^b |\Delta(G \circ f_1 - G \circ f_2)|^2 d\mu \leq 2 \int_a^b |\frac{\Delta(G \circ f_1)}{\Delta f_1} \Delta f_1 - \frac{\Delta(G \circ f_2)}{\Delta f_2} \Delta f_1|^2 d\mu \\ + 2 \int_a^b |\frac{\Delta(G \circ f_2)}{\Delta f_2} \Delta f_1 - \frac{\Delta(G \circ f_2)}{\Delta f_2} \Delta f_2|^2 d\mu. \end{aligned} \tag{5.13}$$

Observe that the second term can be bounded easily:

$$\int_a^b |\frac{\Delta(G\circ f_2)}{\Delta f_2}\Delta f_1 - \frac{\Delta(G\circ f_2)}{\Delta f_2}\Delta f_2|^2 d\mu \leq \int_a^b (\sup_{|y|\leq C_s r} |G'(y)|)^2 \, |\Delta f_1 - \Delta f_2|^2 d\mu,$$

and we have the Lipschitz property on this term. Now we bound the first term. Remark that

$$\frac{\Delta(G\circ f_i)}{\Delta f_i}(x) = \int_0^1 G'(t\Delta f_i(x) + f_i(x))dt\,, \tag{5.14}$$

we get,

$$|\frac{\Delta(G\circ f_1)}{\Delta f_1} - \frac{\Delta(G\circ f_2)}{\Delta f_2}|_\infty \leq \int_0^1 \sup_{|y|\leq 3C_s r} |G''(y)||t\Delta(f_1-f_2) + f_1 - f_2|_\infty \, dt.$$

Since for $|t|\leq 1$, $|t\Delta(f_1-f_2)+f_1-f_2|_\infty \leq 3|f_1-f_2|_\infty \leq 3C_s|f_1-f_2|_{F_s}$, we obtain,

$$|\frac{\Delta(G\circ f_1)}{\Delta f_1} - \frac{\Delta(G\circ f_2)}{\Delta f_2}|_\infty \leq \sup_{|y|\leq 3C_s r} |G''(y)|3C_s|f_1-f_2|_{F_s}\,.$$

Back to the inequality (5.13), we obtain

$$\int_a^b |\Delta(G\circ f_1 - G\circ f_2)|^2 dx \leq 2\int_a^b (\sup_{|y|\leq C_s r} |G'(y)|)^2 \, |\Delta(f_1-f_2)|^2 d\mu + 2\int_a^b \left(\sup_{|y|\leq 3C_s r} |G''(y)|3C_s|f_1-f_2|_{F_s}\right)^2 |\Delta f_1|^2 \, d\mu.$$

Remark also that on the L^2 norm, applying the Lipschitz property,

$$|G\circ f_1 - G\circ f_2|_{L^2}^2 \leq \sup_{|y|\leq C_s r} |G'(y)|^2 |f_1-f_2|_{L^2}^2.$$

We get the result,

$$|G\circ f_1 - G\circ f_2|_{F_s}^2 \leq 2\left((\sup_{|y|\leq C_s r} |G'(y)|)^2 + (\sup_{|y|\leq 3C_s r} |G''(y)|3C_s)^2 |f_1|_{F_s}^2\right)|f_1-f_2|_{F_s}^2.$$

To be more precise in the proposition, we obtain the following inequality on the Lipschitz constant of the composition on every F_s ball of radius $r>0$:

$$M \leq \sqrt{2(\sup_{|y|\leq C_s r} |G'(y)|)^2 + 2(\sup_{|y|\leq 3C_s r} |G''(y)|3C_s r)^2}. \tag{5.15}$$

The linear growth of the Lipschitz constant is the direct application of this inequality. □

Proposition 22. *Let $K : \mathbb{R}^j \mapsto \mathbb{R}$ be a $C^1(\mathbb{R}^n,\mathbb{R})$ function with K' locally Lipschitz, $s>1$, then $(f_1,\ldots,f_j)\in (F_s)^j \mapsto K(f_1,\ldots,f_j)\in F_s$ is Lipschitz on every bounded ball.*

We do not detail the proof of this proposition since it is a particular case of a generalization of this proposition which will be stated in proposition 27. A direct consequence of this proposition is that F_s is an algebra. In the next proposition, we give an explicit bound for the continuity of the multiplication. Even if it is a direct application of the last proposition, we can give a better bound.

Proposition 23. *The product in F_s is continuous for $s > 1$ and,*

$$|fg|_{F_s} \leq 2\, C_s\, |f|_{F_s} |g|_{F_s}.$$

Proof. First we bound the L^2 norm,

$$|fg|_{L^2}^2 \leq |f|_\infty^2 |g|_{L^2}^2, |fg|_{L^2}^2 \leq C_s^2 |f|_{F_s}^2 |g|_{L^2}^2.$$

Now, with the inequality,

$$|\Delta(fg)|^2 \leq 2(|g|_\infty^2 |\Delta f|^2 + |f|_\infty^2 |\Delta g|^2),$$

we obtain the result $|fg|_{F_s}^2 \leq 4C_s^2 |f|_{F_s}^2 |g|_{F_s}^2$. □

Once we have dealt with the one dimensional case, a natural question arises on the generalization to higher dimensions. The natural generalization in two dimensions of H_s is the tensor product $H_s \otimes H_{s'}$. We could do the same for F_s but we prefer to define a space $F_{s,s'}$ which is in fact the tensor product $F_s \otimes F_{s'}$ from a slightly different point of view. We will take advantage of this definition to extend the composition property.

Definition 15. *Let (s, s') be two positive real numbers, the space $F_{s,s'} \subset L^2(T_2)$ is defined by*

$$F_{s,s'} = F_s(S_1, F_{s'}),$$

in the following sense with $I_{n,k} = [\frac{k}{2^n}, \frac{k}{2^n} + \frac{1}{2^{n+1}}]$,

$$|f|_{F_{s,s'}}^2 = \int_{S_1} |f|_{F_{s'}}^2\, dx + \sum_{n,k} 2^{ns-1} \int_{I_{n,k}} |f(x + \frac{1}{2^{n+1}})(.) - f(x)(.)|_{F_{s'}}^2 dx < \infty.$$

Remark 15. *This definition can be rewritten in a more explicit form,*

$$\begin{aligned} F_{s,s'}(T) \;=\; & \{f \in L^2(T) |\, |f|_{F_{s,s'}}^2 = \int_{S_1} |f(x,.)|_{F_{s'}}^2\, dx + \int_{S_1} |f(.,y)|_{F_s}^2\, dy + \\ & \sum_{n,k,n',k'} 2^{ns+n's'-2} \int_{I_{n,k}} \int_{I_{n',k'}} |\Delta_{2,n'}(\Delta_{1,n} f)|^2 dx\, dy \;< \infty\}, \quad (5.16) \end{aligned}$$

with $\Delta_{1,n} f = f(x + \frac{1}{2^{n+1}}, y) - f(x,y)$ and $\Delta_{2,n} f = f(x, y + \frac{1}{2^{n+1}}) - f(x,y)$.

In what follows, we will denote $c_{(n,k),(n',k')}^2 := \int_{I_{n,k}} \int_{I_{n',k'}} |\Delta_{2,n'}(\Delta_{1,n} f)|^2 dx\, dy$.

As in the one dimensional case, the injection $F_{s,s'} \hookrightarrow H_{s,s'}$ would be again a straightforward application of Cauchy-Schwarz inequality. Yet we provide a different proof using the following proposition:

Proposition 24. *As defined in 15, we have*

$$F_{s,s'} = F_s \otimes F_{s'}\,,$$

and as a consequence, $F_{s,s'} \hookrightarrow H_{s,s'}$.

Remark that the identification is here allowed since the two spaces are each one L^2 subspace.

Proof. Let $(n,k),(n',k')$ be two couples of integers, then $(\phi_{n,k}\otimes\psi_{n',k'})_{(n,k),(n',k')}$ is an orthonormal Hilbert basis in $F_s\otimes F_{s'}$ by the definition. From the remark 15, this is also true in $F_{s,s'}$.
If $u\otimes v\in F_s\otimes F_{s'}$, then

$$|u\otimes v|_{H_{s,s'}}=|u|_{H_s}|v|_{H_{s'}}\leq|u|_{F_s}|v|_{F_{s'}}\,,$$

and the second assertion is verified. □

Now we prove that if f is sufficiently smooth then f belongs to $F_{s,s'}$. It requires to control one more derivative that in the one dimensional case.

Proposition 25. *Let f be a C^1 function such that $y\to\partial_1 f(x,y)$ is Lipschitz uniformly in the first variable x, then $f\in F_{s,s'}$ for $s<2$ and $s'<2$.*

Proof. We will denote by $\mathrm{Lip}_{1,2}$ the Lipschitz constant in the second variable of $\partial_1 f$. We need to estimate the integrals detailed in (5.16), we will denote by $c_{(n,k),(n',k')}$ each integral. Using the Lipschitz property, we have

$$|\Delta_{2,n'}(\Delta_{1,n}f)|\leq\int_x^{x+\frac{1}{2^{n+1}}}|\partial_1 f(u,y+\frac{1}{2^{n'+1}})-\partial_1 f(u,y)|\,du\leq\mathrm{Lip}_{1,2}2^{-(n'+n+2)}\,.$$

Then, we have

$$c_{(n,k),(n',k')}=\int_{I_{n,k}}\int_{I_{n',k'}}|\Delta_{2,n'}(\Delta_{1,n}f)|^2dx\,dy\leq\mathrm{Lip}_{1,2}^2\,2^{-3(n'+n+2)}.$$

Summing up on (k,k'), we get that

$$\begin{aligned}\sum_{(n,k),(n',k')}2^{ns+n's'-2}c^2_{(n,k),(n',k')}&\leq\sum_{(n,k),(n',k')}\mathrm{Lip}_{1,2}^2\,2^{n(s-3)+n'(s'-3)-8},\\&\leq\sum_{n,n'}\mathrm{Lip}_{1,2}^2\,2^{n(s-2)+n'(s'-2)-8}.\end{aligned}$$

In the equation (5.17), the right member converge if $s<2$ and $s'<2$. To conclude the proof, The first two terms in the $F_{s,s'}$ norm according to the expression (5.16) are well defined thanks to the proposition 19 which requires $s<2$ and $s'<2$. □

Remark 16. *Again, any function for which there exists a finite dyadic partition of T such that the restriction of f on each domain of the partition satisfies the condition in the proposition 25 belongs to $F_{s,s'}$.*

Next, we will state the norm inequalities that are relevant to prove the needed properties. At this point, it is worthwhile to generalize a little the spaces we introduced in order to extend our results to any dimension. Observe that F_s is a Hilbert algebra of functions. A way to generalize our result is to study $F_s(S_1,F)$ where F is a Hilbert space of functions

which is also an algebra. Though we could be more general, we will assume further that $F \subset L^2(T_n)$ where T_n is the n dimensional torus: $T_n := \mathbb{R}^n/\mathbb{Z}^n$. As in the previous definition,

Definition 16. *Let $s > 0$ be a positive real number, the space $F_s(F)$ is by*

$$F_s(F) = F_s(S_1, F)\,,$$

in the following sense:
with $I_{n,k} = [\frac{k}{2^n}, \frac{k}{2^n} + \frac{1}{2^{n+1}}]$, a function $f \in L^2(T_{n+1})$ belongs to $F_s(F)$ if

$$|f|^2_{F_s(F)} := \int_{S_1} |f(x)|^2_F\, dx + \sum_{n,k} 2^{ns-1} \int_{I_{n,k}} |f(x + \frac{1}{2^{n+1}}) - f(x)|^2_F dx < \infty.$$

We have

$$F_s(F) = F_s \otimes F\,. \tag{5.17}$$

The last equality (5.17) uses the same argument as in proposition 24.

Proposition 26. *Let us assume that F is a RKHS on a space X and that there exists a constant C_F such that $\sup_{x\in X} |f(x)| \le C_F |f|_F$. If $s > 1$ then the following inequalities hold for $f \in F_s(F) = F_s \otimes F$,*

$$\begin{aligned}
&\sup_{s\in S_1} |f(s)|_F \le C_s |f|_{F_s(F)} \\
&\sup_{s\in S_1} |\Delta_1 f|_F \le C_s |\Delta_1 f|_{F_s(F)} \\
&\sup_{(s,x)\in S_1\times X} |f(s,x)|_\infty \le C_s\, C_F |f|_{F_s(F)}\,,
\end{aligned}$$

where Δ_1 is a difference operator defined for $\delta > 0$ as $(\Delta_1 g)(x) \doteq g(x+\delta) - g(x)$. Moreover proposition 17 is also verified.

Proof. If $u \in F_s \otimes F$, then $u = \sum_{n,m\in\mathbb{N}^2} \alpha_{n,m} e_n \otimes f_n$, with $(e_n)_{n\in\mathbb{N}}$ and $(f_n)_{n\in\mathbb{N}}$ Hilbert basis respectively for F_s and F. By definition, $\sum_{n,m\in\mathbb{N}^2} \alpha^2_{n,m} < \infty$. Hence, we have,

$$u = \sum_{m\in\mathbb{N}} (\sum_{n\in\mathbb{N}} \alpha_{n,m} e_n) \otimes f_m = \sum_{m\in\mathbb{N}} E_m \otimes f_m\,.$$

Then, we can apply the evaluation at point $z \in S_1$ since F_s is also a RKHS then we denote by C_z the norm of the evaluation at point $z \in S_1$: the sequence $\sum_{m=0}^N E_m(s) f_m$ is a Cauchy sequence in F since

$$|\sum_{n=p}^q E_n(z) f_n|^2_F = \sum_{n=p}^q E_n(z)^2 \overset{\text{rkhs prop.}}{\le} C_z^2 \sum_{n=p}^q |E_n|^2_{F_s}\,.$$

As a consequence, the evaluation at point $z \in S_1$ is well defined and it makes the space $F_s \otimes F$ a RKHS on S_1 with values in F. Furthermore, as $C_z \le C_s$, we have:

$$\sup_{z\in S_1} |f(z)|_H \le C_s |f|_{F_s(F)}\,.$$

The second inequality is the application of the first one to $\Delta_1 f$.
The last one just uses the assumption on the RKHS F. □

Now, we can easily generalize the work done in the one dimensional case.

Proposition 27. *Let $F \subset L^2(T_n, \mathbb{R}^k)$ for $k \geq 1$ be a RKHS algebra with a continuous injection in $L^\infty(T_n, \mathbb{R}^k)$. Assume that the left composition with an element $H \in C^l(\mathbb{R}^k, \mathbb{R}^k)$*

$$F \ni g \mapsto H \circ g \in F$$

is Lipschitz on every ball $B(0,r)$ of constant $C_H(r)$. Then,

- *if $G \in C^{l+1}(\mathbb{R}^k, \mathbb{R}^k)$, the composition*

$$F_s(F) \ni f \mapsto G \circ f \in F_s(F),$$

  is Lipschitz on every ball,

- *with the additional assumption that the left composition with G' and G'' in F are locally Lipschitz such that there exists a polynomial real function P verifying $\max(C_{G'}(r), C_G(r)) \leq P(r)$ for $r > 0$, then there exists a constant depending on F and s, $a \in \mathbb{R}^+$ such that the Lipschitz constant $C_{G,F_s(F)}$ for the left composition with G on $F_s(F)$*

$$C_{G,F_s(F)}(r) \leq arP(r).$$

Proof. We first need to prove that if $f \in F_s(F)$ then $G \circ f \in F_s(F)$. With the proposition 26, we have that

$$\sup_{x \in S_1} |f(x)|_H \leq C_s |f|_{F_s(F)}. \tag{5.18}$$

Then, we obtain for the first term in the norm,

$$\int_{S_1} |G \circ f - G(0)|_F^2 d\mu \leq C_s^2 C_G^2 (C_s |f|_{F_s(F)}) |f|_{F_s(F)}^2.$$

For the term involving the difference, we need to introduce again the formula,

$$\Delta_x (G \circ f) = \left(\int_0^1 G'(t\Delta_x f + f)(\Delta_x f) dt \right), \tag{5.19}$$

which is now allowed since G' is C^l. The formula (5.19) uses the fact that $F \subset L^2(T_n, \mathbb{R}^k)$ to give a sense to the composition. As F is an algebra, we have $G'(t\Delta_x f + f)\Delta_x f \in F$. Obviously we have also $t\Delta_x f + f \in B_F(0, 3r_0)$ for $|t| \leq 1$ and $r_0 = C_s |f|_{F_s(F)}$. With the inequality (5.18), we get

$$|G'(t\Delta_x f + f)\Delta_x f|_F \leq 3Mr_0 |\Delta_x f|_F C_{G'}(3r_0), \tag{5.20}$$

with M the continuity constant of multiplication in F:

$$\forall (f,g) \in F^2 \,,\, |fg|_F \leq M |f|_F |g|_F \,.$$

Remark that $G'(t\Delta_x f + f)$ can be seen as a matrix valued function. We use the matrix norm implied by the underlying norm on $\mathbb{R}^k$. The inequality (5.20) directly proves that $G \circ f \in F_s(F)$ with in addition:

$$|G \circ f - G(0)|_{F_s(F)}^2 \leq \left[\max(3Mr_0 C_{G'}(3r_0), C_s C_G(r_0)) \right]^2 |f|_{F_s(F)}^2.$$

We now prove the Lipschitz property.
Let f_1 and f_2 be two elements in $F_s(F)^2$ with $\max(|f_1|_{F_s(F)}, |f_2|_{F_s(F)}) \leq r_1$. With the Lipschitz property of the composition on F we have with $r_2 = C_s r_1$,

$$\int_{S_1} |G \circ f_1(x) - G \circ f_2(x)|_F^2 dx \leq \int_{S_1} C_G^2(r_2)|f_1(x) - f_2(x)|_F^2 dx. \tag{5.21}$$

Now, for the other terms, we use again the formula (5.19):

$$\begin{aligned}|\Delta_x(G \circ f_1 - G \circ f_2)|_F^2 \leq 2M^2[|\Delta_x(f_1 - f_2)|_F^2(\int_0^1 |G'(t\Delta_x f_1 + f_1)|_F \, dt)^2 \\ + |\Delta_x f_2|_F^2(\int_0^1 |G'(t\Delta_x f_1 + f_1) - G'(t\Delta_x f_2 + f_2)|_F \, dt)^2].\end{aligned} \tag{5.22}$$

We pay attention to the last term of inequality (5.24):

$$\int_0^1 |G'(t\Delta_x f_1 + f_1) - G'(t\Delta_x f_2 + f_2)|_F \, dt \leq 3C_s|f_1 - f_2|_{F_s(F)} C_{G'}(3r_2). \tag{5.23}$$

We then obtain,

$$\begin{aligned}|\Delta_x(G \circ f_1 - G \circ f_2)|_F^2 \leq 2M^2[|\Delta_x(f_1 - f_2)|_F^2(3r_2 C_{G'}(3r_2))^2 \\ + |\Delta_x f_2|_F^2(3C_s|f_1 - f_2|_{F_s(F)} C_{G'}(3r_2))^2].\end{aligned} \tag{5.24}$$

For notation convenience, we define $K(r_1) := 3r_2 C_{G'}(3r_2)$ and we finally get, combining equations (5.21) and (5.24)

$$|G \circ f_1 - G \circ f_2|_{F_s(F)}^2 \leq \max(4M^2K^2(r_1), C_G^2(r_2))|f_1 - f_2|_{F_s(F)}^2 . \tag{5.25}$$

This inequality implies directly the last item in the proposition. □

We have presented all the necessary material to generalize easily our results. We generalize F_s in dimension $n \geq 3$ as it is already done in one and two dimensions.

Definition 17. *We define by recurrence F_s where $s = (s_1, \ldots, s_n) \in \mathbb{R}_+^n$ by for $n \geq 3$,*

$$F_{s_1}(S_1, F_{(s_2,\ldots,s_n)}) = F_{s_1} \otimes \ldots \otimes F_{s_n} .$$

We denote its dual F_{-s}, $s_* = \min_{i\in[1,n]} s_i$ and $s^* = \max_{i\in[1,n]} s_i$.

To sum up our work to this point, we have defined a RKHS algebra F_s for a multi-index s which is stable under the composition with smooth functions. The continuity of the product is detailed in appendix 43 (it is also a byproduct of the previous result on the composition with smooth functions in proposition 27). We will now prove that F_s is 'big enough'. Provided that linear functions are in F_s, the proposition of the composition 27 answers this question. We can have a better result:

Proposition 28. *If $s^* < 2$ and $f \in C^n(T_n, \mathbb{R}^k)$ then $f \in H_s$ and $|f|_{H_s} \leq c_n|f|_{n,\infty}$.*

Proof. By recurrence, this true for $n = 1$. With the inequality,

$$|f|^2_{F_s} \leq \int_{S_1} c^2_{n-1}|f(x)|^2_{n-1,\infty}dx + \sum_{n,k} 2^{ns-1} \int_{I_{n,k}} (\int_x^{x+\frac{1}{2^{n+1}}} c_{n-1}|\partial_1 f|_{n-1,\infty}dx)^2,$$

$$|f|^2_{F_s} \leq c^2_{n-1}|f|^2_{n,\infty}(1 + \sum_{n=0}^{\infty} 2^{n(s-2)-4}). \tag{5.26}$$

We have the result with $c^2_n = c^2_{n-1}(1 + \sum_{n=0}^{\infty} 2^{n(s-2)-4})$. □

As usual, we have the direct application of this proposition:

Definition 18. *If $n \geq 2$, a dyadic partition of T_n is a product of a dyadic partitions in one dimension.*

Proposition 29. *If $\max s < 2$ and $f \in L^2(T_n, \mathbb{R}^k)$ such that there exists a dyadic partition on which the restriction of f is C^n then $f \in H_s$.*

The space of functions such that the restriction is C^p on a dyadic will be denoted $C^p_{\text{dyad}}(T_n, \mathbb{R}^k)$. In the next Section, we will detail the cylindrical Brownian motion, we will prove that almost surely its trajectories are continuous paths in H_{-s}, hence in F_{-s}. With this result, we can summarize the properties of F_s:

Theorem 25. *The Hilbert space $F_s \subset L^2(T_n, \mathbb{R}^k)$ satisfies the following properties*

- *the left composition with a function $G \in C^{n+1}(\mathbb{R}^k, \mathbb{R}^k)$ is locally Lipschitz,*
- *for $s_* > 1$, F_s is an algebra with continuous product,*
- *the cylindrical Brownian motion defines a continuous random process in F_{-s},*
- *$C^p_{dyad}(T_n, \mathbb{R}^k) \subset F_s$ for $s^* < 2$.*

We can now deduce an important property to prove the existence for all time of the SDE solutions.

Proposition 30. *If the kernel k has continuous derivatives,*

$$\frac{\partial^{l+n} k(x,y)}{\partial_l x \partial_n y}\, l, n \in [1, m+2]\,,$$

a couple $(p, q) \in F_{-s} \times F_s$ defines an element of V by $k_q p(x) = \langle k(x, q), p\rangle_{F_s \times F_{-s}}$, with $s = (s_1, \ldots, s_m)$. Moreover, the following mappings are Lipschitz

$$F_{-s} \times F_s \ni (p, q) \mapsto k_q p \circ q \in F_s\,, \tag{5.27}$$

$$F_{-s} \times F_s \ni (p, q) \mapsto \langle p, (\partial_x k_q p) \circ q\rangle_{F_{-s} \times F_s} \in F_{-s}\,. \tag{5.28}$$

Proof. First, for any x, $k_q p(x)$ is well defined: since $y \mapsto k(x, y)$ is C^{m+2}, we apply the Theorem 25 to get $k(x, q) \in F_s$. Hence, $k_q p(x)$ is well defined. Our aim is to prove

that $k_q p(x)$ is C^{m+2} and $\partial_x k_q p$ is C^{m+1}. Hence, we would obtain that $k_q p \in V$. The composition with a C^{m+1} function being locally Lipschitz on F_s, the results will follow.

To prove the continuity of $x \mapsto k_q p(x)$, we just need the weak convergence:

$$k(x_n, q(.)) \rightharpoonup_{n\mapsto\infty} k(x, q(.))$$

when $\lim_{n\mapsto\infty} x_n = x$. We first prove that $k(x_n, q(.)) \to_{L^\infty} k(x, q(.))$: thanks to the injection $F_s \hookrightarrow L^\infty(T_m, \mathbb{R}^d)$, $|q|_\infty \leq r_0$. Since k is continuous, it is uniformly continuous on $W \times \overline{B(0, r_0)}$ with W a compact neighborhood of x.
Then, for any $\varepsilon > 0$, there exists $\delta > 0$ such that if $|x_n - x| \leq \delta$, we have $|k(x_n, y) - k(x, y)| \leq \varepsilon$ and as a consequence $|k(x_n, q) - k(x, q)|_\infty \leq \varepsilon$.

We now prove that $k(x_n, q)$ is bounded in F_s: as $\partial_y^n k(x, y)$ for $n \in [1, m+1]$ is bounded on any compact set, we get that $k(x_n, q(.)) \in F_s$ is bounded in F_s.
Since $L^1 \subset L^\infty(T_m)' \subset F_{-s}$ (thanks to $F_s \hookrightarrow L^\infty(T_m)$) every weak sub sequence of $k(x_n, q(.))$ converges to $k(x, q(.))$. Then, $\lim_{n\mapsto\infty} k_q p(x_n) = k_q p(x)$. Hence, $k_q p$ is continuous. By the same proof, $k_q p$ is a C^1 function:
as k is C^{m+1}, one can apply the same argument to $\frac{k_q p(x+t_n v) - k_q p(x)}{t_n} - \partial_1 k_q p(x)(v)$ for $v \in \mathbb{R}^d$ and $t \neq 0$. We have,

$$\frac{k_q p(x + t_n v) - k_q p(x)}{t_n} - \partial_1 k_q p(x)(v) = \int_0^1 \partial_1 k_q p(x + s t_n v, q)(v) p - \partial_1 k_q p(x)(v) ds\,.$$

The sequence $\int_0^1 \partial_1 k_q p(x + s t_n v, q(.))(v) - \partial_1 k(x, q(.))(v) ds$ converges in L^∞ to 0 by uniform continuity of $\partial_1 k$ on every compact set. It is also bounded in F_s since $\partial_1 k$ is C^{m+1} in the second variable. We get the same conclusion as above.
Since the pointwise derivative $\partial_1 k_q p(v)$ is continuous (by the same argument than for $k_q p$ we have that $\partial_1 k_q p(v)$ is continuous) $d[k_q p] = (\partial_1 k)_q p$.
By recurrence, the result is extended to $\partial_x^n k q_p$ for $n \in [1, n+2]$: we obtain that $H = \frac{1}{2}\langle p, k_q p\rangle < \infty$ and $k_q p \in V$.

To prove that the mapping $(p, q) \mapsto k_q p$ is Lipschitz on every compact, the composition is Lipschitz on every bounded ball if $\partial_1 k$ is C^{m+1} in the second variable. Hence we deduce that the maps for each x_0, $q \in F_s \mapsto k(x_0, q) \in F_s$ and $q \in F_s \mapsto \partial_1 k(x_0, q) \in F_s$ are Lipschitz. The Lipschitz constant can be bounded for $x_0 \in K$ by continuity of the kernel derivatives.
As the dual pairing is Lipschitz, we obtain the result. Then, by triangular inequality, we also obtain that $k_q p \circ q$ is Lipschitz in both variables and so is $\langle p, (\partial_x k_q p) \circ q\rangle$. □

5.5 Cylindrical Brownian motion and stochastic integral

The goal of this Section is to define the stochastic integral $\int_0^T u(x) dB_x$ for u a suitable random variable with values in F_s. We will give a self-contained presentation inspired by [DPZ08] provided basic knowledge of the Ito stochastic integral. This construction is of

importance since we do not want to be limited to the white noise on $L^2(S_1, \mathbb{R}^2)$.
First we provide an elementary and self-contained introduction to the cylindrical Brownian motion in H. The construction here puts the emphasis on the finite dimensional approximations obtained by projection on finite dimensional subspaces which are the counterpart in the noise model of the finite dimensional approach with landmarks. We will then detail in 5.38 the stochastic integral.

5.5.1 Cylindrical Brownian motion in $L^2(S^1, \mathbb{R})$

We start with the simplest situation where the underlying space in the one dimensional torus S^1 ie $m = 1$ and $H = L^2(S^1, \mathbb{R})$.
Let $(B^{n,k})_{n\geq 0, k\in A_n}$ be a collection of continuous independent standard one dimensional Brownian motions (BM) on $(\Omega, \mathcal{F}, P)$ a probability space. For any $n \geq 0$, and consider the H valued random process

$$W_t^n \doteq \sum_{l=-1}^{n-1} \sum_{k\in A_l} B_t^{l,k} \psi_{l,k} \,.$$

At a given time, the coefficients of W_t^n in the orthonormal basis of H are i.i.d. Gaussian variable with variance t (truncated at rank n). Moreover, since

$$H^n \doteq \text{Span}\{\, \psi_{l,k}, -1 \leq l < n,\ k \in A_n \,\} \tag{5.29}$$

is the 2^n dimensional space of piecewise constant on the dyadic partition of S^1 at scale 2^{-n}, $W_t^n \in H^n$ is obviously a random piecewise constant function. Moreover, for any $f \in H^n$ $f \doteq \sum_{l=-1}^{n-1} \sum_{k\in A_l} f_{l,k}\psi_{l,k}$, we get

$$\langle f, W_t^n\rangle_H = \sum_{l=-1}^{n-1} \sum_{k\in A_l} f_{l,k} B_t^{l,k} \sim \mathcal{N}(0, |f|_H^2) \,. \tag{5.30}$$

More generally, for any $f_1, f_2 \in H^n$, $(\langle f_1, W_t^n\rangle_H, \langle f_2, W_t^n\rangle_H)$ are jointly Gaussian, centred with covariance

$$\Gamma_{i,j} \doteq E(\langle f_i, W_t^n\rangle_H \langle f_j, W_t^n\rangle_H) = \langle f_i, f_j\rangle_H \,. \tag{5.31}$$

In particular, if we introduce $\phi_{n,k} \doteq 2^{n/2}\mathbf{1}_{[k2^{-n},(k+1)2^{-n}[}$ for any $0 \leq k \leq 2^n - 1$, the $\phi_{n,k}$'s define an orthonormal basis of H^n. Denoting $\gamma_t^{n,k} \doteq \langle \phi_{n,k}, W_t^n\rangle_H$, we get from (5.30) and (5.31) that

$$W_t^n = 2^{n/2} \sum_{k=0}^{2^n-1} \gamma_t^{n,k} \mathbf{1}_{[k2^{-n},(k+1)2^{-n}[} \tag{5.32}$$

where $(\gamma_t^{n,k})_{t\geq 0}$ is a i.i.d. family of 2^n standard Brownian motions indexed by k.
A cylindrical Brownian motion on H is the limit of W_t^n when $n \to \infty$. A well known but important fact is that this limit is not defined in H since $E(|W_t^{n+j'}|_n^2) = t2^n \to +\infty$ but in any H_{-s} for $s > 1$.

Indeed, $|W_t^{n+j} - W_t^{n+j'}|^2_{H_{-s}} \leq R^2_{n,t} \doteq \sum_{m=n+1}^{\infty} 2^{-ms} \sum_{k\in A_m} |B_t^{m,k}|^2$ so that

$$E(\sup_{j,j'\geq 0} |W_t^{n+j} - W_t^{n+j'}|^2_{H_{-s}}) \leq E(R^2_{n,t}) = C_{n,s}t$$

with $C_{n,s} \doteq 2^{(n+1)(1-s)}/(1-2^{1-s})$. Hence, $a.s.$, W_t^n is a Cauchy sequence in H_{-s} and one can define W_t as the limit in H_{-s} of W_t^n. In fact, it will be helpful to do a little more. Since the process $t \to W_t^n$ has continuous trajectories in H_s, one can look for a limit in $C(\mathbb{R}_+, H_s)$ for the uniform topology.
For any $T > 0$, we have

$$\sup_{j,j'\geq 0} \sup_{0\leq t\leq T} |W_t^{n+j} - W_t^{n+j'}|^2_{H_{-s}} \leq R_n^2 \doteq \sum_{l=n+1}^{\infty} 2^{-ls} \sum_{k\in A_l} \sup_{0\leq t\leq T} |B_t^{l,k}|^2$$

so that using the Doob inequality $E(\sup_{0\leq t\leq T} B_t^2) \leq 4E(B_t^2)$ for the standard Brownian motion, we get

$$E\left(\sup_{j\geq 0} \sup_{0\leq t\leq T} |W_t^{n+j} - W_t^n|^2_{H_{-s}}\right) \leq 4C_{n,s}T\,.$$

Hence a.s. $t \to W_t^n$ is a Cauchy sequence in $C([0,T], H_{-s})$. Since $T > 0$ is arbitrary, we can define a limit process W living in $C(\mathbb{R}_+, H_{-s})$ such that for any $T \geq 0$

$$E(\sup_{0\leq t\leq T} |W_t - W_t^n|^2_{H_{-s}}) \leq 4C_{n,s}T\,.$$

5.5.2 Cylindrical Brownian motion in $L^2(T_m, \mathbb{R})$

For a general $m \geq 1$, since $L^2(T_m, \mathbb{R}) = L^2(T_1, \mathbb{R}) \otimes \cdots \otimes L^2(T_1, \mathbb{R})$, the construction of the W^n is built from the Hilbert basis obtained by usual tensorisation. To be more explicit, we denote by

$$\psi^m_{l,k} \doteq \otimes_{i=1}^m \psi_{l^{(i)},k^{(i)}}$$

for any $l = (l^{(i)})_{1\leq i\leq m}$ and $k = (k^{(i)})_{1\leq i\leq m}$ such that $l^{(i)} \geq -1$ and $k^{(i)} \in A_{l^{(i)}}$ for $l^{(i)} \geq -1$. Now, if $I_n \doteq \{\ (l,k)\ |\ l^{(i)} \leq n,\ 1 \leq i \leq m\}$

$$H^n = \text{span}\{\psi^m_{l,k}\ |\ (l,k) \in I_n\}$$

and $I_\infty = \cup_{n\geq 0} I_n$, we define from a family $(B^{l,k})_{(l,k)\in I_\infty}$ of i.i.d. standard BM

$$W_t^n \doteq \sum_{(l,k)\in I_n} B_t^{l,k} \psi^m_{l,k}\,. \tag{5.33}$$

As previously, if $f \in H^n$, we have $t \to |f|_H^{-1/2} \langle f, W_t^n\rangle$ is a standard BM and for any $f, g \in H^n$,

$$(\langle f, W_t^n\rangle \langle g, W_t^n\rangle) = t\langle f, g\rangle_H\,. \tag{5.34}$$

In particular, if

$$\phi^m_{n,k} \doteq 2^{mn/2} \mathbf{1}_{\prod_{i=1}^m [k^{(i)}2^{-n}, (k^{(i)}+1)2^{-n}[},\ k^{(i)} \in [\![0, 2^n[\![}$$

the family $(\phi^m_{n,k})_{k\in[\![0,2^n[\![^m}$ is an orthonormal basis of H^n based on a dyadic partition of T_m in cells of size $2^{-n}\times\cdots\times 2^{-n}$. As previously, we have

$$W^n_t = \sum_{k\in\{0,\cdots 2^n-1\}^m} \langle W^n_t, \phi^m_{n,k}\rangle \phi^m_{n,k} = 2^{nm/2} \sum_{k\in\{0,\cdots 2^n-1\}^m} \gamma^{n,k}_t \mathbf{1}_{\prod_{i=1}^m [k^{(i)}2^{-n},(k^{(i)}+1)2^{-n}[} \tag{5.35}$$

where $(\gamma^{n,k})$ is a family of i.i.d. BM.

For $H_s = \otimes_{i=1}^m H_{-s_i}$, with $s=(s_1,\cdots,s_m)$, we get immediately that W^n_t is a Cauchy sequence in H_s as soon as $s_* \doteq \min_i s_i > 1$ converging uniformly on any time interval $[0,T]$ to a process $W\in C(\mathbb{R}_+, H_{-s})$. More precisely,

$$E(\sup_{0\le t\le T} |W_t - W^n_t|_{H_{-s}}) \le 4C_{n,s}T \tag{5.36}$$

with $C_{n,s_*} = (\sum_{l=n+1}^\infty 2^{-l(s_*-1)})^m$.

5.5.3 Cylindrical Brownian motion in $H = L^2(T_m,\mathbb{R}^d)$

For a general $d\ge 1$, the previous definition on cylindrical Brownian motion can be extended easily in the more general situation where $H=L^2(T_m,\mathbb{R}^d)$. Indeed, we define $W \doteq (W^{(1)},\cdots,W^{(d)})$ where $(W^{(i)})_{1\le i\le d}$ is a family of i.i.d. cylindrical Brownian motions in $L^2(T_m,\mathbb{R})$ as defined previously. The finite dimension approximations are defined accordingly on

$$H^n \doteq \text{span}\{(\psi^m_{l_1,k_1},\cdots,\psi^m_{l_d,k_d}) \mid (l_j,k_j)\in I_n\ 1\le j\le d\}. \tag{5.37}$$

In this case, there is an analog of inequality (5.37), with the constant

$$C_{n,s_*} = d(\sum_{l=n+1}^\infty 2^{-l(s_*-1)})^m.$$

5.5.4 Stochastic integral

We assume basic knowledge of the Ito integral (CITE) and we directly deal with the general case on $H=L^2(T_m,\mathbb{R}^d)$. We recall that $F_s \hookrightarrow H \hookrightarrow F_{-s}$. Having in mind applications that we will develop later, we need to introduce the space of integrands.

Let us denote by $L(F_{-s})$ the continuous endomorphisms of F_{-s}. If $u\in L(F_{-s})$, then there exists a constant denoted by $|u|$ such that

$$|u(e)|_{F_{-s}} \le |u||e|_{F_{-s}}.$$

Definition 19. *The set L_T contains all random variables $u:[0,T]\times\Omega\mapsto L(F_{-s})$ verifying,*

- *$(t,\omega)\to u(t,\omega)$ is $\mathcal{B}[0,T]\otimes\mathcal{A}$ measurable,*
- *$\omega\to u(t,\omega)$ is $\mathcal{F}_t-$measurable for $t\in[0,T]$,*
- *$\int_0^T E[|u(t)|^2]\,dt<\infty$.*

Now, we want to give a sense to

$$\int_0^T u(t)dW_t\,, \tag{5.38}$$

for $u \in L_T$. To this end, we first define

$$\int_0^T u(t)dW_t^n = \sum_{l'\in\mathbb{N}}(\int_0^T \sum_{l,k}^{l<n} u_{l,k}^{l'}dB_t^{l,k})e_{l'}\,, \tag{5.39}$$

with $(e_{l'})_{l'\in\mathbb{N}}$ an orthonormal basis of F_{-s}. Each term $\int_0^T \sum_{l,k}^{l<n} u_{l,k}^{l'}dB_t^{l,k}$ are well-defined since it is a finite sum of Ito integrals and we have thanks to Doob inequality

$$E[\sup_{t\in[0,T]} |\int_0^t \sum_{l,k}^{l<n} u_{l,k}^{l',k'}dB_t^{l,k}|^2] \leq 4\sum_{l,k}^{l<n}\int_0^T E[(u_{l,k}^{l'})^2]ds\,. \tag{5.40}$$

Hence we get,

$$\begin{aligned} E[\sup_{i'\geq i\geq 0}\sup_{t\in[0,T]} |\int_0^t u(t)dW_t^{n+i'} - \int_0^T u(t)dW_t^{n+i}|_{F_{-s}}^2] &\leq 4\sum_{l'\in\mathbb{N}}\sum_{l=n+i}^{n+i'-1}\sum_{k\in A_l}(\int_0^T E[(u_{l,k}^{l'})^2]ds)|e_{l'}|_{F_{-s}}^2 \\ &\leq (\sum_{l=n+i}^{n+i'-1}\sum_{k\in A_l}|\psi_{n,k}|_{F_{-s}}^2)\int_0^T E[|u(s)|^2]ds\,, \end{aligned} \tag{5.41}$$

since

$$\sum_{l'\in\mathbb{N}}(\int_0^T E[(u_{l,k}^{l'})^2]ds)|e_{l'}|_{F_{-s}}^2 = \int_0^T E[|u(s)(\psi_{l,k})|^2]ds\,.$$

We have also $|\psi_{n,k}|_{F_{-s}} \leq |\psi_{n,k}|_{H_{-s}}$, and then

$$\sum_{l=n+i}^{n+i'-1}\sum_{k\in A_l}|\psi_{n,k}|_{F_{-s}}^2 \leq \sum_{l=n+i}^{n+i'-1}\sum_{k\in A_l}|\psi_{n,k}|_{H_{-s}}^2 \leq (C_{n+i,s_*} - C_{n+i',s_*})\,.$$

Hence, $\int_0^T u(t)dW_t^n$ is a Cauchy sequence in $C([0,T],F_{-s},|.|_\infty)$.
The next property is simply the application of the previous Doob inequality (5.41) in the case of σ bounded,

Proposition 31. *Assume that $\sigma \in L_T$ is bounded by $|\sigma|_\infty$ then we have,*

$$E[\sup_{t\in[0,T]}\left(\int_0^t \sigma(s)(dW_s^{n+l} - dW_s^n)\right)^2] \leq 4|\sigma|_\infty^2 T(C_{n+i,s_*} - C_{n+i',s_*})\,. \tag{5.42}$$

5.6 Solutions to the SDE on $P \times Q$

Recall that $P = F_{-s}$ and $Q = F_s$. Let $E = P \times Q$ be the phase space equipped with the product Hilbert structure. Considering the injection $i : F_{-s} \to E$ defined by $w \mapsto (w, 0)$ and identifying W with $i \circ W$ and W^n with $i \circ W^n$, we can assume that W and the projections W^n are $C(\mathbb{R}, E)$-valued. Now, for any $n \geq 0$, we introduce the finite dimensional subspace $E_n \doteq H_n \times H_n \subset E$ where H_n is given by (5.29). We denote also

$E_\infty \doteq \cup_{n\geq 0} E_n$ which defines a dense subspace of E. The space E_n is finite dimensional and the restriction of the Hamiltonian H on E_n is well defined. Moreover, if the kernel $K(a,b)$ is C^2 on each variable, then $H(x) = H(p,q)$ is C^2 in the variable $x \in E_n$. We can define the C^1 function f on E_n as

$$x \mapsto f(x) \doteq (-\partial_q H(x), \partial_p H(x))^T \in E_n, \quad x \in E_n\,. \tag{5.43}$$

Let $\sigma : E_\infty \to l_E$ be Lipschitz on any ball of E_∞ and let $(X^n)_{0\leq t<\tau}$ be the pathwise continuous solution of the SDE

$$dY_t = f(Y_t)dt + \sigma(Y_t)dW_t^n, \quad Y_0 = x_0^n \tag{5.44}$$

defined until explosion time τ^n.
We need to consider the following hypothesis.

H0 The space V can be continuously embedded in $C_b^{m+1}(\mathbb{R}^d, \mathbb{R}^d)$ ie there exists $C > 0$ such that $|v|_{m+1,\infty} \leq C|v|_V$ for any $v \in V$.

H0' The trace of the operator induced by on H_n by $\partial_p^2 H(X_s^n)$ can be controlled as

$$\mathrm{tr}(\sigma^T k(Q^n, Q^n)\sigma) \leq c\,. \tag{5.45}$$

Note that if **H0** holds, then for any $b, b' \in \mathbb{R}^d$, $K(.,b)b' \in V$ and K is C^{m+1} in each variable. Moreover, the second hypothesis **H0'** will be verified (17) for σ a Lipschitz mapping in $L(F_s)$ with the additional assumption that for every $X \in E$, $\sigma_X(L^1) \subset L^1$ and the norm of this restriction (with the L^1 norm) is bounded uniformly in X. However it can be interesting to keep this hypothesis for slightly different models.

Proposition 32. *Under assumption **H0**, the explosion time of the SDE (5.44) is almost surely infinite ie X^n is defined for $t \geq 0$ a.s.*

Proof. Let $R > 0$ be a positive real number and $\tau_R^n = \inf\{\, t \geq 0 \mid |X_t^n| \geq R \,\}$ (which is well defined since X^n exists and is continuous until explosion time). We denote by $\tau^n = \lim_{R\to\infty} \tau_R^n$, so that on the event $(\tau^n < \infty)$ the solution X_t^n blows up for $t \to \tau^n$. Using Itô formula for the process $H(X^n_{t\wedge\tau_R^n})$ we get for $X_t^n = (P_t^n, Q_t^n)$

$$H(X^n_{t\wedge\tau_R^n}) = H(x_0^n) + \int_0^{t\wedge\tau_R^n} (\partial_q H(X_s^n)dQ_s^n + \partial_p H(X_s^n)dP_s^n) + \frac{1}{2}\langle (\sigma(X_s^n)dW^n)^T(\partial_p^2 H(X_s^n)\sigma(X_s^n)dW^n\rangle_s\,.$$

Since we have

- $\partial_q H(X_t^n)dQ_t^n + \partial_p H(X_t^n)dP_t^n = \partial H_p(X_t^n)\sigma(X_t^n)dW_t^n$
- from **H0'**, we have that almost surely, for all t

$$\int_0^t \langle (\sigma(X_s^n)dW^n)^T \partial_p^2 H(X_s^n)\sigma(X_s^n)dW^n\rangle_s \leq ct\,.$$

we get

$$H(X^n_{t\wedge\tau^n_R}) \le H(x^n_0) + M_{t\wedge\tau^n_R} + \int_0^{t\wedge\tau^n_R} \frac{1}{2} c\, ds$$

where $M_{t\wedge\tau^n_R}$ is a bounded continuous martingale. So that with the hypothesis **H0'** we have,

$$E(H(X^n_{t\wedge\tau^n_R})) \le H(x^n_0) + \frac{1}{2} c\,(t \wedge \tau^n_R)\,.$$

In particular, $E(\int_0^{t\wedge\tau^n_R} H(X^n_s)ds) \le tH(x^n_0) + \frac{c}{4}\, t^2 < \infty$ and using Fatou Lemma

$$E(\int_0^{t\wedge\tau^n} H(X^n_s)ds) \le tH(x^n_0) + \frac{c}{4}\, t^2 < \infty \tag{5.46}$$

so that almost surely

$$\int_0^{t\wedge\tau^n_R} H(X^n_s)ds < \infty\,. \tag{5.47}$$

Now, for $x = (p, q) \in E_n$, we consider

$$v_x(z) \doteq \int_{T_m} K(z, q(z'))p(z')dz' \in V$$

for which

$$\frac{1}{2}|v_x|^2_V = \frac{1}{2}\int_{T_m\times T_m} p(z)^T K(q(z), q(z'))p(z')dzdz' = H(x)\,. \tag{5.48}$$

From (5.47) and (5.48), we can define by pathwise integration a continuous random process $\Phi^n \doteq (\Phi^n_t)_{0\le t<\tau^n_R}$ solution of the flow equation

$$\frac{\partial}{\partial t}\Phi^n_t = v_{X^n_t}(\Phi^n_t)\,. \tag{5.49}$$

Assuming a continuous embedding $V \hookrightarrow C^{m+1}(\mathbb{R}^d, \mathbb{R}^d)$, $\Phi^n_{t\wedge\tau^n_R}$ is almost surely a C^{m+1} diffeomorphism and there exists a constant D such that almost surely

$$|\Phi^n_{t\wedge\tau^n_R}|_{m+1,\infty} \le C\exp(C\sqrt{t}(\int_0^{t\wedge\tau^n_R} H(X^n_s)ds)^{1/2})\,. \tag{5.50}$$

(*In the sequel, we denote by C a generic constant non depending on n, t and R possibly changing from line to line.*).

Thus, since $Q^n_{t\wedge\tau^n_R} = \Phi^n_{t\wedge\tau^n_R}(q^n_0)$ and $\sup_{n\ge0}|q^n_0|_{F_s} < \infty$, we get from Theorem 25 and proposition 30 that, uniformly in n, we have almost surely (for maybe a different but still universal constant C, see above)

$$K^n_t \doteq \limsup_{R\to+\infty} |Q^n_{t\wedge\tau^n_R}|_{F_s} \le C\exp(C\sqrt{t}(\int_0^{t\wedge\tau^n} H(X^n_s)ds)^{1/2}) < +\infty\,. \tag{5.51}$$

In particular, Q^n_t does not blow up for $t \to \tau^n$ on $\tau^n < \infty$. Therefore it is sufficient to show that P^n_t does not blow up as well to get by contradiction that $\tau = \infty$ almost surely. As from the continuous embedding on V in C^{m+1}, $|dv_x|_{m,\infty} \le C|v_x|_V$, we get from proposition 30 $|dv_{X_{t\wedge\tau^n_R}}(Q^n_{t\wedge\tau^n})|_{F_s} \le C|H(X_{t\wedge\tau^n_R})|^{1/2}K^n_t$ and using the continuity of

the product in F_s (i.e. there exists $M > 0$ such that $|ff'|_{F_s} \leq M|f|_{F_s}|f'|_{F_s}$ for any $f, f' \in F_s$) we obtain for any $\delta q \in F_s$

$$|dv_{X_{t\wedge\tau_R^n}}(Q^n_{t\wedge\tau^n})\delta q|_{F_s} \leq C|H(X_{t\wedge\tau_R^n})|^{1/2}K^n_t|\delta q|_{F_s}\,.$$

Therefore,

$$\begin{aligned}\left|\int_{T_m}\langle dv_{X_{t\wedge\tau_R^n}}(Q^n_{t\wedge\tau^n}(z))^* P^n_{s\wedge\tau_R^n}(z), \delta q(z)\rangle_{\mathbb{R}^d}dz\right| \\ = \left|\int_{T_m}\langle P^n_{s\wedge\tau_R^n}(z), dv_{X_{t\wedge\tau_R^n}}(Q^n_{t\wedge\tau^n}(z))\delta q(z)\rangle_{\mathbb{R}^d}dz\right| \\ \leq C|H(X_{t\wedge\tau_R^n})|^{1/2}K^n_t|P^n_{t\wedge\tau_R^n}|_{F_{-s}}|\delta q|_{F_s}\end{aligned} \tag{5.52}$$

and

$$|\partial_q H(X^n_{t\wedge\tau_R^n})|_{F_{-s}} \leq C|H(X_{t\wedge\tau_R^n})|^{1/2}K^n_t|P^n_{t\wedge\tau_R^n}|_{F_{-s}}$$

since $\partial_q H(X^n_{t\wedge\tau_R^n}) = dv_{X_{t\wedge\tau_R^n}}(Q^n_{t\wedge\tau^n})^* P^n_{s\wedge\tau_R^n}$. We deduce that

$$|P^n_{t\wedge\tau_R^n}|_{F_{-s}} \leq |p^n_0|_{F_{-s}} + CK^n_t\int_0^{t\wedge\tau_R^n} H(X_{s\wedge\tau_R^n})^{1/2}|P^n_{s\wedge\tau_R^n}|_{F_{-s}} + |\int_0^{t\wedge\tau_R^n}\sigma(X^n_s)dW^n_s|_{F_{-s}}$$

and from Gronwall's Lemma

$$|P^n_{t\wedge\tau_R^n}|_{F_{-s}} \leq \left(|p^n_0|_{F_{-s}} + \sup_{u\leq t\wedge\tau_R^n}|\int_0^{u\wedge\tau_R^n}\sigma(X^n_s)dW^n_s|_{F_{-s}}\right)e^{MCK^n_t\sqrt{t}(\int_0^{t\wedge\tau_R^n}H(X_{s\wedge\tau_R^n})ds)^{1/2}}\,. \tag{5.53}$$

Since from Doob inequality we have for $s_* = \inf_{1\leq i\leq m} s_i$

$$E(\sup_{u\leq t\wedge\tau_R^n}|\int_0^{u\wedge\tau_R^n}\sigma(X^n_s)dW^n_s|^2_{H_{-s}}) \leq 4|\sigma|^2_\infty tC_{-1,s_*} \tag{5.54}$$

with a right hand side independent of R, we deduce that almost surely

$$\sup_{u\leq t\wedge\tau}|\int_0^{u\wedge\tau_R^n}\sigma(X^n_u)dW^n_u|_{F_{-s}} \leq \sup_{u\leq t\wedge\tau}|\int_0^{u\wedge\tau_R^n}\sigma(X^n_u)dW^n_u|_{H_{-s}} < +\infty\,. \tag{5.55}$$

and from (5.47), (5.51) and (5.55), we get almost surely

$$\sup_{R\to\infty}|P^n_{t\wedge\tau_R^n}|_{F_{-s}} < +\infty \text{ and } \tau^n > t\,.$$

□

Proposition 33. *Let f be defined by (5.43) and assume that **H0-H'1-H2** hold. Then for any $n \geq 0$ there exists a unique strong solution $X^n = (P^n, Q^n) : \Omega \to C(\mathbb{R}_+, E)$ to*

$$X^n_t = X^n_0 + \int_0^t f(X^n_s)ds + \int_0^t \sigma(X^n_s)dW^n_s,\ X^n_0 \equiv x^n_0 \in E_n \tag{5.56}$$

and a random solution $X = (P, Q) : \Omega \to C(\mathbb{R}_+, E)$ to

$$X_t = X_0 + \int_0^t f(X_s)ds + \int_0^t \sigma(X_s)dW_s,\ X_0 \equiv x_0 \in E \tag{5.57}$$

such that almost surely:

$$\sup_{0\leq t\leq T}|X_t - X^n_t| \to 0\,.$$

Proof. From proposition 33, we know the existence of the finite dimensional approximation solution $(X_t^n)_{t\geq 0}$ for $t \geq 0$. Moreover, we know from proposition 35 the existence of maximal solution $(X_t)_{0\leq t<\tau}$ of the SDE (5.57) $(X_t)_{0\leq t<\tau}$ up to a possibly finite explosion stopping time τ. Moreover, for any $T > 0$ and any $r > 0$ we have almost surely

$$\sup_{t\leq T} |X^n_{t\wedge\tau_r} - X_{t\wedge\tau_r}| \to 0 \quad \text{and} \sup_{0\leq t\leq T} |\int_0^{t\wedge\tau_r} \sigma(X_s^n)dW_s^n - \int_0^{t\wedge\tau_r} \sigma(X_s)dW_s| \to 0 \tag{5.58}$$

where $\tau_r = \inf\{t \geq 0 \mid |X_t| \geq r\}$. What we need to prove is that there is no blow up ie $P(\tau < +\infty) = 0$ or equivalently, $\tau_r \to +\infty$ almost surely.
We start from inequality (5.46) in the proof of proposition 32. Using the uniform convergence (5.58) and Fatou Lemma, we get

$$E(\int_0^{t\wedge\tau} H(X_s)ds) \leq tH(x_0) + \frac{c}{4}t^2 < \infty\,. \tag{5.59}$$

Similarly, starting from (5.51), we get that

$$K_t \doteq \limsup_{R\to+\infty} |Q_{t\wedge\tau_R}|_{F_s} \leq C\exp(C\sqrt{t}(\int_0^{t\wedge\tau} H(X_s)ds)^{1/2}) < +\infty\,. \tag{5.60}$$

Moreover, from (5.53), we get for $n \to \infty$,

$$|P_{t\wedge\tau_r}|_{F_{-s}} \leq \left(|p_0|_{F_{-s}} + \sup_{u\leq t\wedge\tau_r} |\int_0^{u\wedge\tau_r} \sigma(X_s)dW_s|_{F_{-s}}\right) e^{MCK_t\sqrt{t}(\int_0^{t\wedge\tau_r} H(X_{s\wedge\tau_r})ds)^{1/2}}\,. \tag{5.61}$$

Since as in (5.54) Doob inequality gives for $s_* = \inf_{1\leq i\leq m} s_i$

$$E(\sup_{u\leq t\wedge\tau_r} |\int_0^{u\wedge\tau_r} \sigma(X_s)dW_s|^2_{F_{-s}}) \leq 4|\sigma|^2_\infty tC_{-1,s_*}$$

there exist a random constant $A_t > 0$ independent of r such that almost surely

$$|P_{t\wedge\tau_r}|_{F_{-s}} \leq A_t\,.$$

In particular $(\tau \leq t) \subset (\tau_r \leq t) \subset (\max\{K_t, A_t\} \geq r)$ and considering the limit $r \to \infty$, we get $P(\tau \leq t) = 0$. □

5.6.1 A trace Lemma

We now present a sufficient condition to fulfill the hypothesis **H0'**. With additional assumptions on the kernel and on the variance term, we give a bound for the bracket of the stochastic term of the SDE on finite dimensional subspaces H_n.

Lemma 17. *Let k be a kernel bounded on the diagonal i.e. there exists $c > 0$ such that $b^TK(a,a)b \leq c|b|^2$ for any $a, b \in \mathbb{R}^d$ or equivalently $K(a,a) \leq cId_d$ as a symmetric non-negative bilinear form on $\mathbb{R}^d$. We assume also that $\sigma(L^1) \subset L^1$ and this restriction is continuous i.e. there exists $M > 0$ such that $|\sigma(f)|_{L_1} \leq M|f|_{L^1}$. Then we have with H the usual Hamiltonian*

$$\int_0^t \langle(\sigma dW_s^n)^T\partial_p^2 H\sigma dW_s^n\rangle_s \leq c^2M^2\,t\,.$$

Proof. We consider the orthonormal basis $(\phi_{n,k})_{k\in[0,2^{n-1}]^r}$ with $A_n = [0, 2^{n-1}]^r$ to write the Hamiltonian as:

$$H = \frac{1}{2^{2nr+1}} \sum_{(i,j)\in A_n^2} p_i^T k(q_i, q_j) p_j \,,$$

with $q = \sum_{i\in A_n} q_i 2^{-nr/2}\phi_{n,i}$ and $p = \sum_{i\in A_n} p_i 2^{-nr/2}\phi_{n,i}$. In this basis, the L^2 scalar product can be written as $\langle p, q\rangle = 2^{-nr}\sum_{i\in A_n} p_i q_i$. We can write $\sigma dW^n = \sum_{i\in A_n}(\sum_{j\in A_n}\alpha_{i,j}dW^j)2^{nr/2}\phi_{n,i}$ with $\alpha_{i,j} \in L(\mathbb{R}^d)$ and $(W^j)_{j\in A_n}$ i.i.d. standard BM with values in $\mathbb{R}^d$. Then we have,

$$\langle(\sigma dW_s^n)^T\partial_p^2 H\sigma dW_s^n\rangle_s = \frac{1}{2^{2nr}}[\sum_{i,j\in A_n^2}\sum_{h\in A_n}(\alpha_{i,h}^T k(q_i, q_j)\alpha_{i,h})]ds\,.$$

Thanks to the hypothesis on the kernel, we have for any $x, y \in \mathbb{R}^d$ that

$$|a^T k(x, y)b| \le \sqrt{a^T k(x, x)a}\sqrt{b^T k(y, y)b} \le c^2|a||b|$$

and then,

$$|\alpha_{i,h}^T k(q_i, q_j)\alpha_{i,h}| \le c^2|\alpha_{i,h}||\alpha_{j,h}\,|.$$

Now, we can write with Cauchy Schwarz inequality

$$\begin{aligned} d\langle(\sigma dW_s^n)^T\partial_p^2 H\sigma dW_s^n\rangle_s \le c^2\frac{1}{2^{2nr}}\sum_{h\in A_n}(\sum_{i\in A_n}|\alpha_{i,h}|)^2 ds &\le \sum_{h\in A_n} c^2|\sigma(\phi_{n,h})|_{L^1}^2\, ds\,, \\ &\le c^2M^2(\sum_{h\in A_n}|\phi_{n,h}|_{L^1}^2)\, ds \le c^2M^2\, ds\,, \quad (5.62) \end{aligned}$$

since for any $h \in A_n$, $|\phi_{n,h}|_{L^1}^2 = 2^{-nr}$. Note that the inequality 5.62 is a little abusive but it is to be understood as an inequality on measures with density w.r.t. the Lebesgue measure. □

remark that we do not need to assume that $\sigma : P \times Q \mapsto L_T$ is bounded by $|\sigma|_\infty$, $\sigma(H_n) \subset H_n$. This hypothesis is only required for the existence and uniqueness in all time but not to bound the trace of the operator.

The assumption on the kernel is not restrictive in our range of applications with kernels such as Gaussian kernel or Cauchy kernel. However, the assumption on σ is much more demanding. Yet a wide range of linear maps can be reached. For instance, the convolution with a smooth function is a continuous operator on F_s then by duality it gives a continuous operator on F_{-s}. This operator has a continuous restriction to L^1.

An important point is that this Lemma covers the case where σ is the multiplication by an element of F_s. However, we will give a simpler proof in this case (in 26).

5.7 Approximations Lemmas

Let $f : E_\infty \to E$ be a function on E_∞ such that $f(E_n) \subset E_n$ for any $n \ge 0$. Let also $\sigma : P \times Q \mapsto L(P)$ be a Lipschitz function. Assume that for any $n \ge 0$, we have a

random variable $X^n : \Omega \to C(\mathbb{R}_+, E)$ solution of the stochastic integral equation

$$X_t^n = X_0^n + \int_0^t f(X_s^n)ds + \int_0^t \sigma(X_s^n)dW_s^n,\ X_0^n \equiv x_0^n \in E_n\,. \tag{5.63}$$

H1 The functions f and σ are Lipschitz on E_∞ and can be uniquely extended as Lipschitz functions on E. Moreover σ is bounded.

H2 For some $\alpha > 1$, we have $\sum_{n\geq 0} n^{2\alpha}|x_0^{n+1} - x_0^n|^2 < \infty$.

Proposition 34. *Let $s > 1$ be a real number. Under the hypothesis $(\mathbf{H1} - \mathbf{H2})$, there exists a random solution $X : \Omega \to C(\mathbb{R}_+, E)$ to*

$$X_t = X_0 + \int_0^t f(X_s)ds + \int_0^t \sigma(X_s)dW_s,\ X_0 \equiv x_0 \in E$$

such that for any $T > 0$, we have almost surely:

$$\begin{cases} \sup_{0\leq t\leq T} |X_t^n - X_t| \to 0\,, \\ \\ \sup_{0\leq t\leq T} |\int_0^t \sigma(X_s^n)dW_s^n - \int_0^t \sigma(X_s)dW_s| \to 0\,. \end{cases} \tag{5.64}$$

Proof. Let $n \geq 0$ be a positive integer and K an upper bound of the Lipschitz constant for f and σ. We have

$$|X_t^{n+1} - X_t^n| \leq |x_0^{n+1} - x_0^n| + K\int_0^t |X_s^{n+1} - X_s^n|ds + |M_t^{n+1} - M_t^n| \tag{5.65}$$

with $M_t^n = \int_0^t \sigma(X_s^n)dW_s^n$. Let us consider the last right-hand martingale term. We have

$$M_t^{n+1} - M_t^n = \int_0^t (\sigma(X_s^{n+1}) - \sigma(X_s^n))dW_s^{n+1} + \int_0^t \sigma(X_s^n)d(W^{n+1} - W^n)_s\,.$$

Using the Doob inequality, we get

$$\begin{aligned} E(\sup_{u\leq t} |\int_0^u (\sigma(X_s^{n+1}) - \sigma(X_s^n))dW_s^{n+1}|_{F_{-s}}^2) &\leq E(\sup_{u\leq t} |\int_0^u (\sigma(X_s^{n+1}) - \sigma(X_s^n))dW_s^{n+1}|_{H_{-s}}^2) \\ &\leq 4C_{-1,s_*}K^2E(\int_0^u |X_s^{n+1} - X_s^n|^2 ds) \end{aligned} \tag{5.66}$$

and with $\delta_{n,s_*} = C_{n,s_*} - C_{n+1,s_*}$

$$\begin{aligned} E(\sup_{u\leq t} |\int_0^u \sigma(X_s^n)d(W^{n+1} - W^n)_s|_{F_{-s}}^2) &\leq E(\sup_{u\leq t} |\int_0^u \sigma(X_s^n)d(W^{n+1} - W^n)_s|_{H_{-s}}^2) \\ &\leq 4|\sigma|_\infty^2\delta_{n,s_*}\,. \end{aligned} \tag{5.67}$$

Thus, for $Z_t^n \doteq \sup_{u\leq t} |X_u^{n+1} - X_u^n|^2$, we have

$$E(\sup_{u\leq t} |M_u^{n+1} - M_u^n|^2) \leq 8(K^2C_s\int_0^u E(Z_s^n)ds + |\sigma|_\infty^2\delta_{n,s_*}) \tag{5.68}$$

and using the inequality $(a+b+c)^2 \leq 3(a^2+b^2+c^2)$ and (5.65), we get

$$E(Z_t^n) \leq 3(|x_0^{n+1} - x_0^n|^2 + 8|\sigma|_\infty^2\delta_{n,s_*} + K^2(8C_{-1,s_*} + 1)\int_0^t E(Z_s^n)ds)\,. \tag{5.69}$$

Applying Gronwall's Lemma, we get for a sufficiently large constant A

$$E(Z_t^n) \leq A(|x_0^{n+1} - x_0^n|^2 + |\sigma|_\infty^2 \delta_{n,s_*}) \exp(At) \tag{5.70}$$

and from **H2**

$$\sum_{n\geq 0} P(\sup_{s\leq t} |X_s^{n+1} - X_s^n| \geq n^{-\alpha}) \leq \sum_{n\geq 0} n^{2\alpha} E(Z_t^n) < \infty\,.$$

Borel-Cantelli Lemma gives a.s. $\sup_{s\leq t} |X_s^{n+1} - X_s^n| < n^{-\alpha}$ for n large enough so that X^n converge uniformly on any compact interval $[0,t]$ to a $C(\mathbb{R}_+, E)$-valued process X. Similarly, from (5.68), (5.70) and **H2** we get

$$\sum_{n\geq 0} P(\sup_{s\leq t} |M_s^{n+1} - M_s^n| \geq n^{-\alpha}) \leq \sum_{n\geq 0} n^{2\alpha} E(\sup_{s\leq t} |M_s^{n+1} - M_s^n|^2) < \infty$$

and M^n converges uniformly on any compact interval $[0,t]$ to a limit $C(\mathbb{R}_+, E)$-valued process M for which

$$X_t = x_0 + \int_0^t f(X_s)ds + M_t\,.$$

Let us check that

$$M_t = \int_0^t \sigma(X_s)dW_s\,.$$

Indeed, since

$$E(\sup_{u\leq t} |\int_0^u \sigma(X_s)dW^s - M_u^n|^2) \leq 8(E(\int_0^t K^2 |X_s - X_s^n|^2 ds) + |\sigma|_\infty^2 C_{n,s_*}$$

we get for $n \to 0$, $E(\sup_{u\leq t} |\int_0^u \sigma(X_s)dW^s - M_u^n|^2) = 0$. □

We extend now the previous result to locally Lipschitz drift f and diffusion σ.

H1' The functions f and σ are Lipschitz on any ball of E_∞ and can be uniquely extended as locally Lipschitz functions on E.

Proposition 35. *Let $s > 1$ be a positive real number. Under the hypothesis $(\mathbf{H1'} - \mathbf{H2})$, there exists a stopping time τ and a continuous adapted process $(X_t)_{0\leq t<\tau}$ with value in E such that*

1. $\limsup_{t\to\tau^-} |X_t| = +\infty$ *on* $(\tau < \infty)$ *(τ is the explosion time) ;*
2. *for any $r > 0$ and any $T > 0$, we have almost surely*

$$\begin{cases} \sup_{0\leq t\leq T} |X^n_{t\wedge\tau_r} - X_{t\wedge\tau_r}| \to 0 \\ \sup_{0\leq t\leq T} |\int_0^{t\wedge\tau_r} \sigma(X_s^n)dW_s^n - \int_0^{t\wedge\tau_r} \sigma(X_s)dW_s| \to 0 \end{cases} \tag{5.71}$$

where $\tau_r \doteq \inf\{t \geq 0 \mid |X_t| \geq r\}$.

Moreover, for any $t \geq 0$

$$X_{t\wedge\tau_r} = x_0 + \int_0^{t\wedge\tau_r} f(X_s)ds + \int_0^{t\wedge\tau_r} \sigma(X_s)dW_s \quad a.s. \tag{5.72}$$

Proof. The proof is quite straightforward. Let $R > r > 0$ be two positive real numbers and

$$g(x) \doteq \max(\min(1, R - |x|), 0)$$

be a Lipschitz function such that $g^R(x) = 1$ if $|x| \leq R$ and $g^R(x) = 0$ if $|x| \geq R + 1$. We introduce also $f^R = g^R f$ and $\sigma^R = g^R \sigma$. From **H1'** f^R and σ^R are Lipschitz and we get from standard results on existence of finite dimensional SDE a solution $X^{R,n} \in C(\mathbb{R}_+, E_n)$ of $dY_t = f^R(Y_s)ds + \sigma^R(Y_s)dW_s^n$ and $Y_0 = x_0^n$. From Proposition 34 applied to f^R and σ^R, there exists a $C(\mathbb{R}_+, E)$-valued process X^R solution of $dY_t = f^R(Y_s)ds + \sigma^R(Y_s)dW_s$ and $Y_0 = x_0$ such that $P(\sup_{s\leq t} |X_s^R - X_s^{R,n}| \to 0) = 1$ and $P(\sup_{s\leq t} |\int_0^s \sigma(X_u^R)dW_u - \int_0^s \sigma(X_u^{R,n})dW_u^n| \to 0) = 1$. Since f^R and f (resp. σ^R and σ) coincide on $|x| \leq R$, we get for any $n \geq 0$ that

$$X^{R,n}_{t\wedge\tau_R^{R,n}} = X^n_{t\wedge\tau_R^{R,n}} \; a.s.$$

where $\tau_R^{R,n} = \inf\{t \geq 0 \mid |X_t^{R,n}| \geq R\}$. In particular $\tau_R^{R,n} = \tau_R^n \doteq \inf\{t \geq 0 \mid |X_t^n| \geq R\}$ almost surely and for any $T \geq 0$, $P(\sup_{0\leq t\leq T} |X^n_{t\wedge\tau_r^R} - X^R_{t\wedge\tau_r^R}| \to 0) = 1$ with $\tau_r^R \doteq \inf\{t \geq 0 \mid |X_t^R| \geq R\}$ since the uniform convergence of $X^{R,n}$ to X^R on compact set implies that a.s. $\tau_R^n > \tau_r^R$ for n large enough. As a consequence, for two solutions X^R and $X^{R'}$ for $R' > r$, we have $\tau_r^R = \tau_r^{R'}$ a.s. and the trajectories before the common value τ_r^R are equal. Let R_k be an increasing sequence converging to $+\infty$ and $\tau = \lim_{k\to\infty} \tau_{R_k}^{R_k}$. If $X_t = \sum_{k=0}^{\infty} X_t^{R_{k+1}} \mathbf{1}_{\tau_{R_k}^{R_k} \leq t < \tau_{R_{k+1}}^{R_{k+1}}}$ for $t \leq \tau$, the process $(X_t)_{0\leq t<\tau}$ verifies 1), 2) and (5.72). □

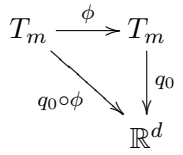
5.8 Applications and numerical simulations

This Section will present a direct application of the SDE we have studied. In this simplest model, we suppose a shape to be given with an initial momentum and we model the perturbation term with a white noise on the initial shape. Hence the variance of the noise term is constant in time.

Let (q_0, p_0) be respectively the initial shape and the initial momentum of the system. As in the deterministic matching with a sufficiently smoothing attachment term, the momentum is a normal L^2 vector field on the shape, it is relevant enough for applications to consider $q_0 \in F_s$ and $p_0 \in F_{-s}$. To assume $q_0 \in F_s$ means that we chose a parameterization of the shape by T_m. We would like our stochastic system to be independent of this initial parameterization. Then we need to understand what is the reparameterization transformation on the deterministic system and on the white noise.

Assume that ϕ is a diffeomorphism of T_m, then we detail the correspondence between the solution (p_t, q_t) from initial conditions (p_0, q_0) and $(\tilde{p}_t, \tilde{q}_t)$ with the initial position

variable $q_0 \circ \phi$.



As the trajectory is entirely determined by the vector field v_{p_t,q_t}, the change of variable by ϕ gives the correspondence:

$$\tilde{q}_t = q_t \circ \phi\,,$$
$$\tilde{p}_t = \mathrm{Jac}(\phi)\, p_t \circ \phi\,.$$

We will denote by $\phi^*(p)$ the pull-back of p under ϕ.
The stochastic system verifies the same transformation and the pull-back of the cylindrical Brownian motion is given by

$$\tilde{B} \doteq \mathrm{Jac}(\phi)\, B \circ \phi\,.$$

We need the following proposition,

Proposition 36. *If B is a cylindrical Brownian motion on $L^2(T_n, \nu)$ with $\nu \ll \mu$ (μ is the Lebesgue measure) then $\phi^*(B_t) = \mathrm{Jac}(\phi) B_t(\phi(s))$ is a cylindrical Brownian motion on $L^2(T_n, \nu')$ with $\frac{d\nu'}{d\nu} = \frac{1}{\mathrm{Jac}(\phi)}$.*
Moreover if $f \in L^2$ with $f \neq 0$ a.e., $f\,B$ is a cylindrical Brownian motion for the measure $\frac{1}{f^2} d\nu$.

Proof. If $(e_i)_{i\in\mathbb{N}}$ is an orthonormal basis of $L^2(T_n, \nu)$, then $\sqrt{\mathrm{Jac}(\phi)} e_i \circ \phi$ is also an orthonormal basis for $L^2(T_n, \nu)$ (change of variable with ϕ). As a result, $\mathrm{Jac}(\phi) e_i \circ \phi$ is an orthonormal basis for $L^2(T_n, \nu')$ with $\frac{d\nu'}{d\nu} = \frac{1}{\mathrm{Jac}(\phi)}$.
The second part of the proposition is also straightforward: if $(e_i)_{i\in\mathbb{N}}$ is an orthonormal basis for $L^2(T_m, \mathbb{R}^d)$, then $(f e_i)_{i\in\mathbb{N}}$ is an orthonormal basis for $L^2(T_m, \frac{1}{f^2} d\nu)$. □

Corollary 1. *The random process $\phi^*(B_t)$ is equal to $\sqrt{\mathrm{Jac}\phi} W_t$ with W_t a cylindrical Brownian motion on $L^2(T_m, \mathbb{R}^d)$.*

Back to our framework with $F_s \times F_{-s}$, we remark that the space F_s is not invariant under a change of coordinates: let ϕ be a diffeomorphism of T_n then a priori, if $q \in F_s$ then $q \circ \phi$ may not belong to F_s. However, if q is sufficiently smooth then $T(q) := q \circ \phi$ belongs to F_s if ϕ is smooth enough. Hence there exist subspaces in F_s invariant under this transformation. To go further we could prove that if $s < s'$ then $T(F_{s'}) \subset F_s$. We would have the same result for the dual spaces: $T(F_{-s}) \subset F_{-s'}$. Furthermore we would like to deal with piecewise diffeomorphisms, this is why the formulation of the following proposition is a little more general,

Proposition 37. *Let $q \in F_s(T_n, \mathbb{R}^k)$ and $\phi : T_n \mapsto T_n$ be a measurable invertible mapping (i.e. there exists $\phi^{-1} : T_n \mapsto T_n$ measurable such that $\phi \circ \phi^{-1} = Id$). We*

assume also that $q \circ \phi \in F_s$, $p \circ \phi \in F_{-s}$ *and* $J := \frac{d\phi_*^{-1}\mu}{d\mu} \in F_s(T_n, \mathbb{R})$ *such that* $J > \varepsilon > 0$ *a.e. Finally, let* B *be a cylindrical Brownian motion. If* (p_t, q_t) *is the solution of the system*

$$\begin{cases} dp_t = -\partial_q H(p_t, q_t) + \varepsilon dB_t \\ dq_t = \partial_p H(p_t, q_t)\,, \end{cases} \tag{5.73}$$

(with ε *a constant parameter) for initial data* (p_0, q_0) *and for the path of the white noise* B *then* $(\phi^*(p_t), q_t \circ \phi)$ *is the solution of the system*

$$\begin{cases} dp_t = -\partial_q H(p_t, q_t) + \varepsilon\sqrt{\mathrm{Jac}(\phi)}d\tilde{B}_t \\ dq_t = \partial_p H(p_t, q_t)\,, \end{cases} \tag{5.74}$$

for initial data $(J\phi^*(p_0), q_0 \circ \phi)$ *and for the random process* $\sqrt{J}\tilde{B}$*, with* $\tilde{B}$ *a cylindrical Brownian motion.*

The random process $\sqrt{J}\tilde{B}$ can be treated in our framework with the map $\sigma : E_n \mapsto L(F_s)$ given by the multiplication with $p_n(\sqrt{J})$. Remark that as $J > \varepsilon$ we have that $\sqrt{J} \in F_s$ by smoothness of the square root outside 0. Thus σ is a Lipschitz map such that $\sigma(H_n) \subset H_n$ since the projection is Lipschitz and the multiplication in F_{-s} by an element of F_s is Lipschitz too. It leads to

Theorem 26. *Under assumption* ***H0****, let* $f \in F_s$ *and* $(p_0, q_0) \in F_{-s} \times F_s$ *be initial conditions verifying that there exists* $s'' < s$ *and* $s' > s$ *such that* $q_0 \in F_{s'}$ *and* $p_0 \in F_{s''}$*. Then random solution of the following system with initial conditions* (p_0, q_0)

$$\begin{cases} dp_t = -\partial_q H(p_t, q_t) + f dB_t \\ dq_t = \partial_p H(p_t, q_t)\,, \end{cases} \tag{5.75}$$

are defined for all time and there is an almost surely convergence of the approximations (also defined for all time)

$$\begin{cases} dp_t^n = -\partial_q H(p_t^n, q_t^n) + f^n dB_t \\ dq_t^n = \partial_p H(p_t^n, q_t^n)\,, \end{cases} \tag{5.76}$$

to the previous random solution with initial conditions (p_0^n, q_0^n) *the projection on* E_n *of* (p_0, q_0)*.*

Proof. This is the application of proposition 33. We verify hypothesis **H0'** and control the trace of the operator. Remark that we need to control the Hilbert-Schmidt norm of the operator $\sigma^T k(q,q)\sigma$ with σ the multiplication by an element of F_s. This is a consequence of Lemma 17, but we give here a simpler proof, since the multiplication is a diagonal operator. We have

$$|f^n|_\infty \leq C_s |f^n|_{F_s} \leq C_s |f|_{F_s}\,,$$

then we get for any $q \in F_s$,

$$0 \leq \int_{T_m} f^n(s)^T k(q(s), q(s)) f^n(s) ds \leq C_s^2 |k|_\infty |f|_{F_s}^2 \mathrm{Vol}(T_m)\,.$$

Hence hypothesis **H0'** is verified.
From the assumption on the initial conditions, if $q_0 \in F_{s'}$ with $s' > s$ we have

$$|q_0|^2_{F_{s'}} = \sum_{n=0}^{\infty} 2^{n(s'-s)} |q_0^{n+1} - q_0^n|^2_{F_s} < +\infty \,.$$

Moreover, if $p_0 \in F_{s''}$ then

$$|p_0|^2_{F_{s'}} = \sum_{n=0}^{\infty} 2^{n(s-s'')} |p_0^{n+1} - p_0^n|^2_{F_s} < +\infty \,.$$

Hence **H2** is verified. □

Remark that **H2** is not very demanding as proved in this theorem.

Now we can detail some basic situations to simulate the stochastic system (5.73). The preceding result will enable to deal with a wide family of shapes and white noises. We first develop the case of curves.

Proposition 38. *Let $1 < s < 2$ a real number and $f : S_1 \mapsto \mathbb{R}^2$ be a piecewise C^2 mapping such that $|f'| > \varepsilon > 0$ with ε a real positive number. If we denote by α the arc-length parameterization of f then there exists a piecewise affine homeomorphism $\phi : S_1 \mapsto S_1$ such that $\phi \in F_s$ and $\alpha \circ \phi$ is in F_s.*

Proof. Let us assume that the arc-length parameterization α has p singularities at points $x_1 < \ldots < x_n \in S_1$. Then, we define ϕ as the linear interpolation on n dyadic points in S_1, $d_1 < \ldots < d_n$ with for images respectively $x_1 < \ldots < x_n$. As ϕ is piecewise affine it belongs to F_s thanks to proposition 20. With the same argument we conclude that $\alpha \circ \phi \in F_s$. □

The previous proposition tells us that we can consider our stochastic system on an initial shape with a white noise which is 'white' with respect to the arc length. Hereunder are some simulations to illustrate the convergence of the landmark simulation and the kind of trajectories generated by this model. The figure 5.1 presents the convergence of the image of the unit circle under the flow generated by the system with an increasing number of landmarks. We chose to illustrate this convergence since in some sense it just shows the convergence of the vector fields generated by the stochastic system and does not include the convergence of the projection of the landmark discretization.
To illustrate the fact that this model generate perturbation of geodesic equations, we computed the initial momentum for the matching between the unit circle and an ellipse (figure 5.2). Then we simulated one path of the stochastic system for different values of the variance of the noise. The unit circle is discretized with 100 landmarks equally distributed. The initial momentum is obtained by a Newton method following [ATY05].
The simulations[1] obtained shows that the evolution is smooth in time as desired. As the

[1] We used an Euler scheme to simulate the SDE