

MASTER MVA

Géométrie et espaces de formes

Alain Trouvé
2018-2019

Table des matières

Chapitre I. Une petite introduction sur les espaces de formes	5
1. Formes versus espace de formes	5
2. Premiers exemples d'espaces de formes	6
3. Action de groupe	7
4. D'Arcy Thompson shape spaces	8
5. Espaces homogènes	9
6. Comparaisons et distances	10
7. Structure riemannienne	10
8. Distance quotient	11
9. Métrique invariante à droite sur un groupe	13
10. Le cas des difféomorphismes	13
11. Appendice : Quelques rappels sur les variétés abstraites	14
Chapitre II. Métriques locales à noyaux et appariements linéaires de points	17
1. Actions infinitésimales	17
2. Métrique induite	17
3. Le problème de l'appariement de points indexés <i>à la Bookstein</i>	18
4. Espaces à noyau reproduisant	19
5. Comment construire des noyaux ?	22
6. Résolution du problème de projection	25
Chapitre III. Groupes G_V et appariements non-linéaires	29
1. Reconstruction de G via l'action infinitésimale	29
2. EDO dans les Banachs	30
3. Définition de G_V	33
4. Distance invariante sur G_V	35
5. Appariements non-linéaires	37
Chapitre IV. Un détour par le contrôle optimal	39
1. Perturbation tangente et linéarisé	40

2. Equations hamiltoniennes	44
3. Cas du problème (I')	45
4. Cas du problème (E')	47
Chapitre V. Flots géodésiques et equations hamiltoniennes	49
1. Hamiltonien réduit	49
2. Existence globale en tout temps du flot hamiltonien	50
3. Cas des landmarks	52
4. Cas des courbes	54
5. Cas des difféomorphismes	56
6. Méthode de gradient trajectorien et de tir pour les landmarks	58
Chapitre VI. Continuité faible du flot et théorèmes d'existence	61
1. Complétude de G_V	61
2. Continuité faible du flot	62
3. Existence de géodésiques minimisantes et solution de problèmes variationnels	63
Chapitre VII. Normes duales et appariements de sous-variétés	65
1. p -vecteurs et produit extérieur	65
2. Formes différentielles	66
3. Courants	68
4. Transport de courants	70
5. ENR sur les courants	71
6. Calcul des appariements	72
Présentation orale	75
Bibliographie	79

Chapitre I

Une petite introduction sur les espaces de formes

1. Formes versus espace de formes

Plutôt que d'analyser les formes elles-mêmes individuellement, le principe de ce cours est de dire qu'il vaut mieux considérer des *ensembles* ou des populations de formes et d'essayer de les appréhender comme des *espaces* au sens mathématique du terme avec une structure particulière. En particulier, on prendra un point de vue *géométrique* en essayant de considérer les espaces de formes M comme des variétés mathématiques de très grande dimension puisque on comprend bien que les formes elles-mêmes ont un grand nombre de degrés de libertés.

Rappelons la définition d'une sous-variété de dimension finie :

DÉFINITION I.1 (D'après [8]). *Un ensemble $M \subset \mathbb{R}^{n+k}$ est une sous-variété de dimension n de classe \mathcal{C}^p de \mathbb{R}^{n+k} si pour tout $m \in M$ il existe un ouvert $U \ni m$ dans \mathbb{R}^{n+k} et une submersion \mathcal{C}^p $f : U \rightarrow \mathbb{R}^k$ telle que $U \cap M = f^{-1}(0)$ (on rappelle que f est une submersion si sa différentielle est surjective en tout point).*

EXERCICE I.1. Vérifier que la sphère $S^d = \{ (x_0, \dots, x_d) \in \mathbb{R}^{d+1} \mid \sum_{i=0}^d x_i^2 = 1 \}$ est une sous variété C^∞ de dimension d de \mathbb{R}^{d+1} .

EXERCICE I.2. Montrer que $SO(n) = \{ A \in M_n(\mathbb{R}) \mid \det A = 1 \text{ et } A^T A = \text{Id} \}$ est une sous-variété de dimension $n(n-1)/2$ de $M_n(\mathbb{R}) \simeq \mathbb{R}^{n^2}$.

Souvent le cadre naturel sera de considérer des variétés de dimension infinie où des approximations de dimension finie d'espaces intrinsèquement de dimension infinie.

Les plus :

- On se donne un cadre cohérent pour parler de *variations* d'une forme m au travers de l'espace tangent $T_m M$ (défini par $\text{Ker } df(m)$)
- On a un *modèle numérique local* par une paramétrisation dans une carte.

En effet

PROPOSITION I.1. *Les deux assertions sont équivalentes :*

- (1) *M est une sous-variété de dimension n de \mathbb{R}^{n+k}*
- (2) *Pour tout $m \in M$, il existe un ouvert U de m dans M et un ouvert Ω de 0 dans \mathbb{R}^n et une application \mathcal{C}^p $g : \Omega \rightarrow \mathbb{R}^{n+k}$ telle que (Ω, g) est une paramétrisation de $M \cap U$ au voisinage de m (c'est à dire que g est un homéomorphisme de Ω dans $M \cap U$ et $dg(0)$ est injective).*

EXERCICE I.3. Construire une paramétrisation de la sphère S^d au voisinage d'un pôle.

2. Premiers exemples d'espaces de formes

Voici quelques exemples d'espaces de formes intéressants pour fixer les idées :

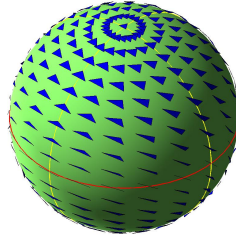


FIGURE 1. Exemple d'espace quotient : les espaces de Kendall. Ici $k = 3$, $d = 2$ (triangles à rotation et échelle près). A ce stade, il n'est pas trivial que l'espace Σ_3^2 soit homotope à une sphère. En fait, on mettra tout à l'heure une structure métrique pour laquelle l'espace sera isométrique à la sphère S^2

- (1) les k -uplet de points $m = (x_1, \dots, x_k)$ dans \mathbb{R}^d tel que $x_i \neq x_j$ pour au moins une paire i, j (on enlève la diagonale correspondant à des points tous confondus).
- (2) les classes d'équivalences $[m]$ à rotation directe, translation, échelle près noté Σ_d^k . On arrive alors sur les *espaces de Kendall* et au domaine de la morphométrie classique
- (3) les courbes dans \mathbb{R}^p et plus généralement les immersions $f : P \rightarrow \mathbb{R}^d$ de classe C^k où P est un ouvert de \mathbb{R}^p ou encore les quotients $[f]$ à reparamétrisation près $\psi : P \rightarrow P$ où ψ est un C^k difféomorphisme sur P (préservant éventuellement l'orientation) ce qui permet par exemple de considérer des formes planes à partir de contours fermés ou ouverts ou des surfaces.
- (4) les images : variations autour d'une template ou encore l'espace L^2 tout entier.
- (5) les distributions d'objets élémentaires comme les distributions de points (mesures), distributions d'éléments de lignes ou de surfaces (courants)

On remarque rapidement que si (1) est trivialement une sous-variété (de dimension nd comme ouvert de \mathbb{R}^{nd}) les autres exemples ne sont pas facilement vu comme des sous-variétés : (2) parce que l'on travaille sur un espace quotient, (3) et (4) parce qu'il s'agit d'un exemple de dimension infinie (ou encore d'un quotient d'une variété de dimension infinie) et (5) parce que nous n'avons pas vraiment défini les objets en questions.

3. Action de groupe

On a déjà vu dans l'introduction de la notion de quotient. Nous allons rencontrer dans la suite plusieurs situations de quotients que l'on peut construire à partir d'une action de groupe sur un ensemble :

DÉFINITION I.2. Soit G un groupe et X un ensemble. On dit que G agit à gauche sur X si il existe une application $A : G \times X \rightarrow X$ telle que

- (1) $A(e, x) = x$ pour tout $x \in X$ où e est l'élément neutre de G ,
- (2) pour tout $g, g' \in G$ et $x \in X$

$$A(g, A(g', x)) = A(gg', x)$$

On notera alors $g.x \triangleq A(g, x)$ si bien que la propriété d'action s'écrit $g.(g'.x) = (gg').x$.

REMARQUE I.1. On définit de même une action à droite par la propriété $A(g, A(g', x)) = A(g'g, x)$ pour laquelle on utilisera la notation $x.g \triangleq A(g, x)$.

EXERCICE I.4. si $A : G \times X \rightarrow X$ est une action de G sur X , alors pour tout $g \in G$, $A_g : x \rightarrow A(g, x)$ définit une bijection sur X . La propriété d'action définit en particulier un morphisme $G \rightarrow \mathfrak{S}(X)$ puisque $A_g \circ A_{g'} = A_{gg'}$.

Un exemple particulièrement simple est celui des espaces de Kendall où $G = \text{Sim}^+(\mathbb{R}^d)$ est le groupe des similitudes directes (engendré par les rotations directes, les translations et les homothéties) sur \mathbb{R}^d , X est l'ensemble des n -uplets de \mathbb{R}^d et l'action est donnée par pour $x = (x_1, \dots, x_n) \in \mathbb{R}^{dk}$

$$g.x = (g(x_1), \dots, g(x_n))$$

DÉFINITION I.3 (Orbites, Quotient). (1) Pour tout point $x \in X$, on définit l'orbite de x sous l'action de G , notée, Gx par

$$Gx \triangleq \{ g.x \mid g \in G \}.$$

- (2) Pour tout $x, y \in X$ on définit la relation d'équivalence $x \sim y$ par $x \sim y$ ssi $Gx = Gy$. On note alors X/G le quotient X/\sim .

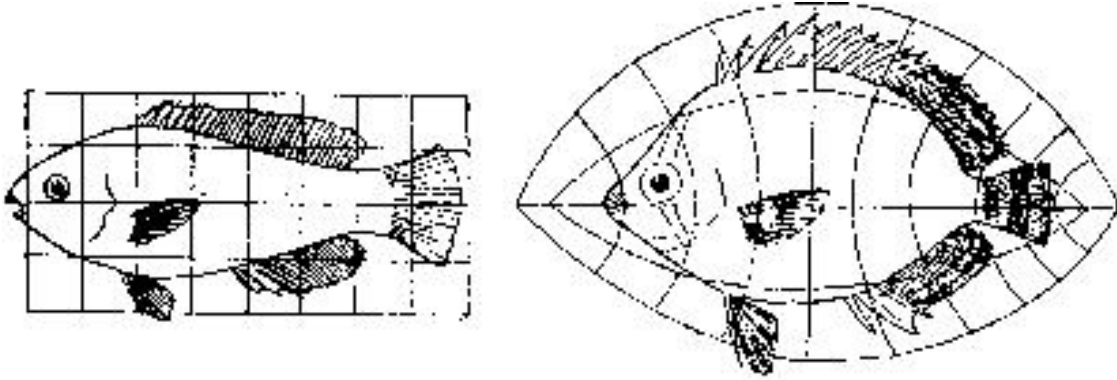
EXERCICE I.5. La propriété de relation d'équivalence de la relation \sim vient directement de la notion la propriété d'action de groupe. Le vérifier.

L'espace de Kendall est donc $X/\text{Sim}^+(\mathbb{R}^d)$ où $X = (\mathbb{R}^d)^k \setminus \Delta$ avec $\Delta = \{x \in (\mathbb{R}^d)^k \mid x_1 = \dots = x_k\}$.

4. D'Arcy Thompson shape spaces

Une autre situation très importante dans laquelle le concept d'action de groupe joue un rôle central est celui de la théorie des transformations de D'Arcy Thompson (1860-1948) dans son livre *Growth and Forms*[13]. En un mot une idée centrale de cette théorie est que si les objets sont complexes, leurs variations observées peuvent souvent se représenter par un changement de coordonnées ou l'action d'une transformation de l'espace.

"In a very large part of morphology, our essential task lies in the comparison of related forms rather in the precise definition of each; and the deformation of a complicated figure may be a phenomenon easy of comprehension, though the figure itself may have to be left unanalyzed and undefined. This process of comparison, of recognizing in one form a definite permutation or deformation of another, apart altogether from a precise and adequate understanding of the original 'type' or standard of comparison, lies within the immediate province of mathematics and finds its solution in the elementary use of a certain method of the mathematician. This method is the Method of Coordinates, on which is based the Theory of Transformations."



Considérons ici le groupe $\text{Diff}_0^1(\mathbb{R}^3)$ des C^1 difféomorphismes homotopes à l'identité. On remarque que l'application

$$(\varphi, I) \rightarrow \varphi.I \triangleq I \circ \varphi^{-1}$$

est une action à gauche de $\text{Diff}_0^1(\mathbb{R}^3)$ sur $\mathcal{F}(\mathbb{R}^d, \mathbb{R})$ l'ensemble des fonctions de \mathbb{R}^d dans \mathbb{R} .

En prenant une image de référence I_0 qui sert de modèle, l'orbite $\text{Diff}_0^1(\mathbb{R}^3)I_0$ contient potentiellement toutes les déformations géométriques du modèle I_0 et constitue potentiellement un espace de forme particulièrement intéressant.

REMARQUE I.2. *L'action restreinte à des déformations géométriques peut paraître un cadre trop restrictif. On peut imaginer de faire agir des groupes plus gros encore qui agissent sur les niveaux de gris de I comme dans le cas de la théorie des métamorphoses. Nous nous focaliserons dans ce cours dans la maîtrise de la situation déjà assez riche des déformations purement géométriques.*

5. Espaces homogènes

La situation de D'Arcy Thompson correspond à la situation mathématique des espaces homogènes.

DÉFINITION I.4. *Soit G opérant sur X . On dit que X est un espace homogène si G agit transitivement sur X i.e. pour tout $x, y \in X$, il existe $g \in G$ tel que $g.x = y$.*

On remarque qu'alors le quotient X/G est réduit à un point et que pour tout $x \in X$ on a $Gx = X$: l'ensemble X est constitué d'une seule orbite.

Pour que X soit un espace homogène il faut que G soit plus gros que X i.e. qu'il y ait suffisamment d'élément dans G . Ce n'est pas la situation des espaces de Kendall puisque $G = \text{Sim}^+(\mathbb{R}^d)$ est un petit groupe mais c'est le cas des espaces de D'Arcy Thompson.

Lorsque G est gros, l'action de G n'est pas *libre* i.e. elle n'agit pas sans point fixe.

DÉFINITION I.5 (Stabilisateur d'un point). *Soit G agissant sur X et $x \in X$. On appelle le stabilisateur de $x \in X$ l'ensemble G_x défini par*

$$G_x \triangleq \{g \in G \mid gx = x\}$$

REMARQUE I.3. *Lorsque l'action est libre, on a que tous les stabilisateurs sont réduits au singleton $\{e\}$.*

THÉORÈME I.1. *Soit G agissant sur X et $x \in X$. Alors*

- (1) $G_x \times G \rightarrow G$ défini par $(h, g) \mapsto gh$ est une action à droite de G_x sur G .
- (2) Le quotient G/G_x est en bijection avec l'orbite Gx par l'application $[g] \mapsto g.x$.

Ainsi si x_0 un élément de X qui joue le rôle du modèle dans la situation de Thompson et $G_0 = G_{x_0}$ est le stabilisateur de x_0 alors on peut voir l'orbite Gx_0 comme un quotient de G . Que gagne-t-on ? Si G est muni d'une structure différentielle, alors on peut construire beaucoup de nouvelles structures par le passage au quotient.

En particulier, si on travaille sur des objets irréguliers (comme des faisceaux de fibres), on peut l'orbite Gx_0 hériter d'une structure naturelle venant de G . Pas besoin de recommencer à chaque fois le travail de construire une structure différentielle.

6. Comparaisons et distances

Évidemment à ce stade, nous sommes encore loin du compte car les espaces de formes doivent permettre de faire de l'analyse de données et se prêter à la construction de modèle. L'utilisation de carte locale ne suffit pas car elle ne permet pas par exemple de répondre à une question aussi simple que celui d'un calcul de moyenne ou d'un centroïde.

Cette notion de moyenne d'une famille $(x_i)_{1 \leq i \leq k}$ apparaît très naturelle dans les espaces euclidiens avec la moyenne arithmétique $\frac{1}{k} \sum_{i=1}^k x_i$ mais n'a pas de sens pour des objets vivant dans des variétés pour laquelle on a pas de structure additive.

Cependant si l'espace est muni d'une distance, on peut définir une notion extension de la moyenne par celle de *moyenne de Fréchet*¹ ou *moyenne de Karcher* on considérant

$$x_* \in \operatorname{argmin} \sum_{i=1}^k d(x, x_i)^2$$

EXERCICE I.6. Vérifier que dans le cas de la métrique euclidienne, on retrouve la notion de moyenne usuelle.

Plus généralement, on veut pouvoir comparer des formes entre elles et avoir une notion de *distances entre formes*. Sinon comment interpoler entre deux formes ?

Une autre façon de poser le problème est de se poser la question d'une paramétrisation locale *intrinsèque* d'une sous-variété. Il est clair qu'il y a une infinité de façon de paramétrer localement une sous-variété mais il nous semble par exemple que ne serait-ce que pour le cas des courbes, que celle induite par l'abscisse curviligne respecte la métrique induite par l'espace euclidien ambiant (par rapport à une reparamétrisation arbitraire du temps). En dimension plus grande, on veut pouvoir préserver dans la carte les notions de distance au moins de façon locale.

7. Structure riemannienne

Nous allons donc dans ce cours considérer les espaces de formes comme des variétés riemanniennes pour lesquelles on dispose en tout point $m \in M$ d'un produit scalaire $\langle \cdot, \cdot \rangle_m$ sur l'espace tangent $T_m M$ (avec des conditions de régularité appropriées). Le cas le plus

1. Il semble que la notion de moyenne de Fréchet soit parfois réservée à la situation où il y a existence et unicité du minimum global dans la fonctionnelle de Fréchet $x \rightarrow \sum d(x, x_i)^2$. Lorsqu'il n'y a pas unicité, on parle de moyenne de Karcher. Nous ne ferons pas de différence dans le cours.

simple est celui où le produit scalaire est induit par l'espace ambiant mais ce n'est pas le plus naturel (par exemple dans la situation d'espaces quotients).

Si $\gamma : [0, 1] \rightarrow M$ est une courbe C^1 dans M , ie un chemin tracé sur M , alors la longueur du chemin est donnée par $l(\gamma) = \int_0^1 \left| \frac{d\gamma}{dt} \right|_\gamma dt$

On peut définir ensuite une distance géodésique $d_M(m, m')$ entre deux points m et m' en minimisant la longueur sur les chemins de m à m' puis la notion de moyenne au travers de la moyenne de Fréchet minimisant la fonctionnelle $m \rightarrow \sum_{i=1}^n d_M(m, m_i)^2$. Enfin, en chaque point m , on peut définir la carte exponentielle qui associe à chaque vecteur $u \in T_m M$ l'unique géodésique de longueur 1 partant de m avec la direction u (en fait cette géodésique peut ne pas être minimisante pour cause de présence de cut-locus). Toutes ces notions dépendent de la structure riemannienne qui définit une généralisation appréciable de l'espace euclidien classique.

Les plus

- On peut parler de distance entre formes (inégalité triangulaire)
- On a une notion de point milieu et d'interpolation naturelle via la notion de géodésique entre deux formes
- On dispose d'une paramétrisation locale naturelle via la carte exponentielle.

REMARQUE I.4. *Cependant la construction d'une métrique locale en chaque point n'est pas une chose évidente en soit et qui comporte son lot d'arbitraire. Deux questions centrales se posent :*

- (1) *Comment choisir les métriques locales ?*
- (2) *Comment calculer avec ?*

8. Distance quotient

On a vu que le cadre des espaces homogènes permettait de construire des espaces des formes intéressants comme l'orbite Gx_0 d'une forme de base par l'action d'un gros groupe et qu'alors l'orbite s'identifie au quotient G/G_{x_0} (Thm I.1) et peut hériter de la structure quotient.

Peut-on pousser l'idée un peu plus loin en partant d'une structure riemannienne unique et donc d'une distance géodésique d_G commune sur G et en définissant une *distance quotient* sur G/G_{x_0} qui donnerait immédiatement une distance riemannienne sur l'orbite Gx_0 ?

THÉORÈME I.2 (distance quotient). *Soit $H \times Y \rightarrow Y$ une action à gauche (resp. à droite) de H sur Y . On suppose que Y est muni d'une distance d_Y qui est équivariante*

à gauche (resp à droite) sous l'action de H au sens où $d_Y(hy, hy') = d_Y(y, y')$ (resp. $d_Y(yh, y'h) = d_Y(y, y')$) pour tous $y, y' \in Y$ et $h \in H$. On définit alors pour tout $[y], [y'] \in Y/H$

$$d_{Y/H}([y], [y']) \triangleq \inf_{z \in [y], z' \in [y']} d_Y(z, z').$$

Alors

- (1) $d_{Y/H}$ est symétrique positive et satisfait l'inégalité triangulaire (i.e. c'est une pseudo-distance sur Y/H).
- (2) Si les orbites $[y]$ sont toutes fermées dans Y (pour la topologie induite par d_Y) alors $d_{Y/H}$ est une distance.

DÉMONSTRATION. On considère le cas ici d'une action à gauche, le cas d'une action à droite se montrant de même. La positivité et la symétrie est quasi immédiate. Pour l'inégalité triangulaire, on commence par remarquer que la propriété d'équivariance implique que $d_{Y/H}([y], [y']) = \inf_{h' \in H} d_Y(y, h'y')$. Puis pour tous $h, h'' \in H$ on a

$$d_{Y/H}([y], [y'']) \leq d_Y(y, h'h''y'') \leq d_Y(y, h'y') + d_Y(h'y', h'h''y'') \leq d_Y(y, h'y') + d_Y(y', h''y'')$$

et donc en prenant l'infimum en $h', h'' \in H$, on obtient le résultat.

Pour le point 2), il suffit que montrer que si $d_{Y/H}([y], [y']) = 0$ alors il existe une suite y_n dans $[y']$ telle que $d_Y(y, y_n) \rightarrow 0$ i.e. $y_n \rightarrow y$. Comme $[y']$ est fermée, $y \in [y']$ d'où $[y] = [y']$. \square

Dans le cas qui nous intéresse, en considérant l'action à droite $G_{x_0} \times G \rightarrow G$ de G_{x_0} sur G on déduit (pour $H = G_{x_0}$ et $Y = G$ que si d_G est G_{x_0} équivariant (à droite) alors $d_{G/G_{x_0}}$ est une distance. De plus comme par le Théoreme I.1, $[g] \mapsto g.x_0$ est une bijection entre G/G_{x_0} et $M = Gx_0$ on a construit ainsi une distance sur Gx_0 définie par

$$(1) \quad d_M(m, m') \triangleq \inf \{ d_G(g, g') \mid gx_0 = m, g'x_0 = m' \}.$$

Si on veut construire facilement des métriques sur des orbites variées au travers d'action de G , une solution particulièrement simple est de partir d'une métrique sur G qui est équivariante à droite sous l'action du plus grand de ses sous-groupes c'est à dire G lui-même !

Considérons donc une distance d_G équivariante à droite sous l'action de G (on dira plus simplement que la distance est *invariante à droite*) alors (1) définit toujours une distance et de plus on a

$$(2) \quad d_M(m, m') = \inf \{ d_G(e, g') \mid g'.m = m' \}.$$

EXERCICE I.7. Vérifier (2).

9. Métrique invariante à droite sur un groupe

Il s'avère que le point de vue riemannien est particulièrement adapté pour construire des distances invariantes à droite sur G . En effet supposons que G soit un groupe de Lie de dimension finie :

DÉFINITION I.6. On dit que G est un groupe de Lie (C^∞) de dimension finie si l'application $G \times G \rightarrow G$ définie par $(g, h) \mapsto gh^{-1}$ est C^∞

Dans le cas d'un groupe de Lie, on dispose d'un isomorphisme naturel entre les espaces tangents donné par les translations à droite $R_g : h \rightarrow hg$ puisque R_g est C^∞ inversible et donc en tout point $g \in G$ on a $dR_g(h) : T_h G \rightarrow T_{gh} G$.

EXERCICE I.8. Vérifier que $d(R_{g_1 g_2})(h)\xi = dR_{g_2}(hg_1)dR_{g_1}(h)\xi$ et donc en notant pour tout $\xi \in T_h G$ $\xi.g \triangleq dR_g(h)\xi$ on a

$$\xi.(g_1 g_2) = (\xi.g_1).g_2.$$

DÉFINITION I.7 (Métrique invariante à droite). Si G est un groupe de Lie (de dimension finie), et $\langle \cdot, \cdot \rangle_e$ un produit scalaire sur $T_e G$ l'espace tangent en l'identité, la métrique invariante à droite associée à $\langle \cdot, \cdot \rangle_e$ est définie pour tout $g \in G$ et $\xi, \xi' \in T_g G$ par

$$\langle \xi, \xi' \rangle_g \triangleq \langle \xi.g^{-1}, \xi'.g^{-1} \rangle_e.$$

On a donc pour tout $u, v \in T_e G$, $\langle u.g, v.g \rangle_g = \langle u, v \rangle_e$.

10. Le cas des difféomorphismes

Dans notre situation, le cas central est celui où G est un groupe de difféomorphismes. C'est un cas de dimension infinie et la construction d'une structure différentiable C^∞ est délicate.

Dans le cas des difféomorphismes C^∞ , l'espace tangent naturel en l'identité est l'espace des champs de vecteurs C^∞ qui n'est pas un espace de Banach (mais un espace de Fréchet). Une théorie infinie dimensionnelle des groupes de Lie est possible ainsi que des outils permettant de faire du calcul des variations ("The convenient setting of global analysis", A. Kriegl et P. Michor). Cependant nous avons surtout besoin pour notre cadre riemannien d'étendre la notion de métrique en l'identité. L'extension la plus naturelle est de considérer que l'espace tangent en l'identité est un espace de Hilbert V de champs de vecteurs muni d'une métrique qui est un produit scalaire.

Quelques questions se posent immédiatement :

- (1) Comment définir des métriques intéressantes sur les champs de vecteurs ?
- (2) Peut-on associer à V un vrai groupe de difféomorphismes G_V muni d'une métrique invariante et une distance géodésique associée ?
- (3) Peut-on induire des métriques sur espaces de D'arcy Thompson dans le cas d'objets d'intérêt lorsque les objets sont des points, des courbes, des images, etc... ?
- (4) Peut-on comprendre les équations géodésiques associées et travailler avec les cartes exponentielles ?
- (5) Comment relier cela avec des problématiques de modélisation statistiques sur les espaces de formes ?

En gros, on souhaite aborder pour des espaces beaucoup plus riches les problématiques initiées par David Kendall [11]. On trouvera dans [15] une référence utile dans l'esprit de ce cours.

11. Appendice : Quelques rappels sur les variétés abstraites

On rappelle la définition d'une variété abstraite de dimension finie :

DÉFINITION I.8 (Variété topologique). *Une variété topologique M est de dimension n est espace topologique séparé ayant à base d'ouverts dénombrable (tout ouvert est une réunion d'ouverts de la base) et telle que pour tout $m \in M$ il existe un ouvert $U \ni x$, Ω un ouvert de \mathbb{R}^n et un homéomorphisme $\varphi : U \rightarrow \Omega$. (U, φ) est appelé une carte locale.*

DÉFINITION I.9 (Atlas). *Un atlas \mathcal{C}^p sur une variété topologique M est la donnée*

- (1) *d'un recouvrement $(U_i)_{i \in I}$ ouvert de M*
- (2) *d'une famille d'homéomorphismes $\varphi_i : U_i \rightarrow \Omega_i$ (où Ω_i est un ouvert de \mathbb{R}^n) telle que pour tout $i, j \in I$, l'application de changement de cartes $\varphi_j \circ \varphi_i^{-1} : \varphi_i(U_i \cap U_j) \rightarrow \varphi_j(U_i \cap U_j)$ est un \mathcal{C}^p difféomorphisme.*

DÉFINITION I.10 (Structure différentiable). *Deux atlas (U_i, φ_i) et (V_j, ψ_j) de classe \mathcal{C}^p sont dit compatibles si leur réunion est un atlas de classe \mathcal{C}^p . Une structure différentiable de classe \mathcal{C}^p est une classe d'équivalence d'atlas compatibles de classe \mathcal{C}^p .*

DÉFINITION I.11 (Variété différentiable de classe \mathcal{C}^p). *Une variété différentiable de classe \mathcal{C}^p de dimension n est une variété topologique de dimension n munie d'une structure différentiable de classe \mathcal{C}^p .*

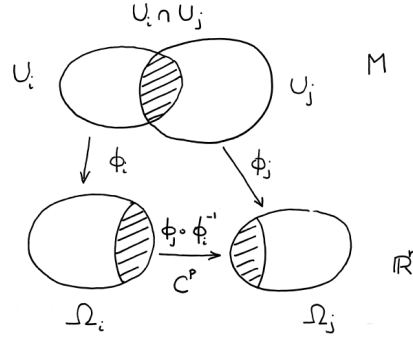


FIGURE 2. Changement de carte pour une variété abstraite

DÉFINITION I.12 (Sous-variété d'une variété). Soit M une variété de classe \mathcal{C}^p . Une sous-ensemble $N \subset M$ est une sous-variété de M si pour tout $n \in N$, il existe un ouvert $U \ni n$ de M est une carte locale (U, φ) telle que $\varphi(N \cap U)$ est une sous-variété \mathcal{C}^p de l'ouvert $\varphi(U)$ (cf Def I.1).

DÉFINITION I.13. Soient X et Y deux variétés différentiables de classe \mathcal{C}^p . Une application $f : X \rightarrow Y$ est de classe \mathcal{C}^k pour $k \leq p$ si pour tout $x \in X$, il existe une carte (U, φ) sur X et une carte (V, ψ) sur Y telles que $U \ni x$, $f(U) \subset V \ni y = f(x)$ et $\psi \circ f \circ \varphi^{-1} : \varphi(U) \rightarrow \varphi(V)$ est \mathcal{C}^k .

DÉFINITION I.14 (Produit de variétés). Soient M et N deux variétés de classes \mathcal{C}^p de dimension respective m et n . Alors $M \times N$ est une variété \mathcal{C}^p de dimension $m + n$ muni de l'atlas $(U_i \times V_j, \varphi_i \times \psi_j)_{(i,j) \in I \times J}$, en notant $\varphi_i \times \psi_j$ l'application $(m, n) \mapsto (\varphi_i(m), \psi_j(n))$ où $(U_i, \varphi_i)_{i \in I}$ (resp. $(V_j, \psi_j)_{j \in J}$) est un atlas compatible avec la structure différentielle de M (resp. N).

EXERCICE I.9. Vérifier que $(U_i \times V_j, \varphi_i \times \psi_j)_{(i,j) \in I \times J}$ engendre bien des applications de changement de cartes \mathcal{C}^p .

Chapitre II

Métriques locales à noyaux et appariements linéaires de points

1. Actions infinitésimales

Reprenons ici le contexte du premier chapitre dans lequel on considère un groupe G muni d'une métrique invariante à droite définie par sa donnée sur l'espace tangent $T_e G$ et une action à gauche $A : G \times M \rightarrow M$ sur un espace M homogène. Supposons que l'action soit différentiable. Alors en différenciant A en $g = e$ par rapport à la première variable, on obtient une application linéaire $d_1 A(e, m) : T_e G \rightarrow T_m M$ qui, dans le cas d'un espace homogène, est une surjection. On va noter ici pour $u \in T_e G$

$$(3) \quad u.m \triangleq d_1 A(e, m).u$$

la valeur en u de $d_1 A(e, m)$. Par suite $u \mapsto u.m$ est linéaire en u . On l'appelle l'*action infinitésimale* de $T_e G$ sur M et on a donc une application linéaire $T_e G \rightarrow TM$ définie par $u \mapsto \rho_u$ où ρ_u est le champ de vecteur $m \mapsto u.m$

Notons que nous avons vu un cas particulier en considérant l'action de G sur lui-même qui s'écrit $A(h, g) = R_g(h)$ où R_g est la multiplication (ou encore la translation à droite) par g . On a donc dans ce cas, $d_1 A(e, g).u = dR_g(e).u = u.g$ comme introduit à la fin du chapitre 1 et la notation est donc cohérente. On a donc maintenant deux actions infinitésimales, l'une sur G et l'autre sur M .

Remarquons maintenant que si g_t est une courbe disons \mathcal{C}^1 sur G alors $m_t = g_t.m$ est une courbe tracée sur M et $\dot{m}_t = d_1 A(g_t, m_t).\dot{g}_t$. En écrivant $A(g_{t+s}g_t^{-1}, A(g_t, m)) = A(g_{t+s}, m)$, on a donc $\dot{m}_t = (\dot{g}_t.g_t^{-1}).m_t$ et si $u_t = \dot{g}_t.g_t^{-1}$, on a $\dot{g}_t = u_t.g_t$ et $\dot{m}_t = u_t.m_t$. Par suite la longueur de la courbe sur G est donnée par

$$\int |\dot{g}_t|_{g_t} dt = \int |u_t|_e dt$$

et tandis que celle sur M est...

2. Métrique induite

Pour cela rappelons que dans notre construction la distance entre deux configurations m_0 et m_1 s'obtient dans la situation d'une métrique invariante à droite sur G par $d_M(m_0, m_1) = \inf\{d_G(e, g) \mid g.m_0 = m_1\}$ puis en considérant les chemins g_t partant de

e à minimiser $\int |\dot{g}_t|_{g_t} dt$ et donc $\int |u_t| dt$ sur les chemins solution de $\dot{g}_t = u_t \cdot g_t$ tel que $g_1 \cdot m_0 = m_1$. Evidemment sur un chemin minimisant, on aura $|u_t|_e = \min\{ |u|_e \mid u \cdot m_t = \dot{m}_t \}$ car sinon on pourrait relever dans G le chemin m_t par un chemin *plus court*¹ Nous avons donc comme définition potentielle de la *métrique induite* sur M

$$(4) \quad |\delta m|_m \triangleq \inf\{ |u|_e \mid u \cdot m = \delta m \}$$

Revenons au cas de nos difféos. Nous avons dit que nous prendrions pour $T_e G$ un espace de Hilbert V de champs de vecteurs muni d'un produit scalaire $\langle \cdot, \cdot \rangle_V$ qui servira de métrique en $e = \text{Id}$. Regardons la situation de l'action sur les configurations de points $m = (x_1, \dots, x_k)$ dans \mathbb{R}^d . On a alors $\varphi \cdot m = (\varphi(x_1), \dots, \varphi(x_k))$ et l'action infinitésimale s'écrit en considérant *formellement* la variation $\text{Id} \rightarrow \text{Id} + u$:

$$(5) \quad u \cdot m \triangleq (u(x_1), \dots, u(x_k)).$$

La construction de la métrique sur M nécessite de résoudre le problème (4) dans ce cas particulier ou de façon équivalente pour $\delta m \in T_m M \simeq (\mathbb{R}^d)^k$, de calculer $|\delta m|_m^2$ solution du problème d'optimisation sous contraintes :

$$(6) \quad (P_1) \begin{cases} \min_V |u|_V^2 \\ \text{avec} \\ u(x_i) = \delta m_i, \quad 1 \leq i \leq k \end{cases}$$

3. Le problème de l'appariement de points indexés à la *Bookstein*

Nous allons voir maintenant que le problème (P_1) est lié à un problème classique que l'on retrouve dans les méthodes introduites par Bookstein [4] dans la morphométrie à base de landmarks par des approches linéaires.

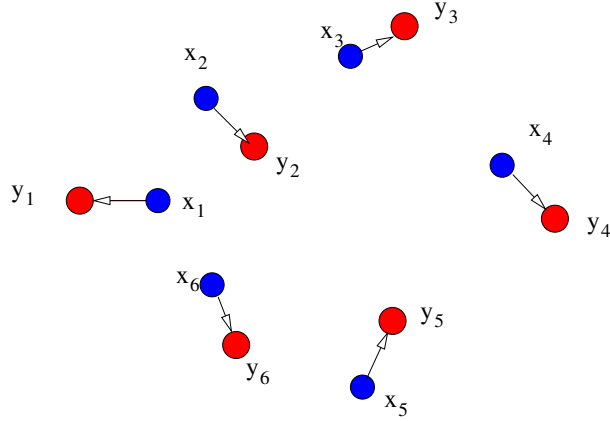
On se place toujours sur \mathbb{R}^d et on considère pour $I = \{1, \dots, k\}$, deux configurations finies $m \triangleq (x_i)_{i \in I}$ et $n \triangleq (y_i)_{i \in I}$ de points sur \mathbb{R}^d . Le problème est de trouver $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ telle que $\varphi(x_i) = y_i$ pour tout $i \in I$. Dans cette approche, on essaie de se ramener dans un contexte linéaire en associant à φ le *champ de déplacements* $u : \mathbb{R}^d \rightarrow \mathbb{R}^d$ tel que

$$\varphi(x) = x + u(x)$$

i.e. $u = \varphi - \text{Id}$.

On voit que la différence est ici que l'on ne considère plus une variation infinitésimale en l'identité mais une déformation macroscopique que l'on plonge directement dans la structure vectorielle des champs de vecteurs. Bien sûr, rien ne garanti a priori l'inversibilité de $\varphi(x) = x + u(x)$ sauf si ∇u reste suffisamment petit. Notons ici que u peut être considéré

1. En fait ceci n'est pas un argument rigoureux à ce stade, en particulier dans le cas où G est de dimension infinie. Pour cela il nous faudra les outils du chapitre suivant. Cependant il peut être justifié dans un cadre approprié et ce raisonnement pour le moment heuristique abouti à des conclusions correctes

FIGURE 1. Exemple de problème d'appariement de points ($k = 5$)

comme un champ de vecteurs du fait de la structure vectoriel de l'espace ambiant. Notre problème est donc :

$$(7) \quad \begin{cases} \text{Trouver } u : \mathbb{R}^d \rightarrow \mathbb{R}^d \\ \text{avec} \\ u(x_i) = \Delta m_i \triangleq y_i - x_i, \quad i \in I \end{cases}$$

REMARQUE II.1. *Les valeurs en dehors de la configuration m sont indéfinies : on a donc un problème mal posé qui admet une infinité de solutions.*

Pour le régulariser, reprenons notre espace de Hilbert V de champs vecteurs ici considéré comme des champs de déplacements² de \mathbb{R}^d dans \mathbb{R}^d . La norme $|u|_V^2$ de u est une mesure de la régularité de u . On transforme le problème en un nouveau problème régularisé du même type que précédemment :

$$(8) \quad (P'_1) \quad \begin{cases} \min_V |u|_V^2 \\ \text{avec} \\ u(x_i) = \Delta m_i, \quad i \in I \end{cases}$$

4. Espaces à noyau reproduisant

Les problèmes (P_1) et (P'_1) se ramènent au problème générique (P_V)

$$(9) \quad (P_V) \quad \begin{cases} \min_V |u|_V^2 \\ \text{avec} \\ u(x_i) = a_i, \quad i \in I \end{cases}$$

2. Notons que la différence entre champ de vecteur et champ de déplacement devient évidente lorsqu'on remplace l'espace \mathbb{R}^d par une variété. Les champs de vecteurs sont alors toujours définis tandis que les champs de déplacements n'ont pas de sens pour une variété abstraite et correspondent à des « cordes » dans le cas des sous-variétés

où $a_I = (a_i)_{i \in I}$ est une famille de vecteurs de \mathbb{R}^d . Il s'agit à un problème de projection sur un sous-espace affine $V_{m,a_I} = \{u \in V \mid u(x_i) = a_i, i \in I\}$. Lorsque le sous-espace est *fermé* et *non vide*, on a existence et unicité de la solution u_* solution de (P'_V) . Pour qu'il soit non vide, il faut que l'espace V soit assez riche. Pour qu'il soit fermé, il suffit que pour tout $x \in \mathbb{R}^d$, l'application $x \rightarrow u(x)$ soit continue c'est à dire qu'il existe $C_x \geq 0$ tel que

$$|u(x)| \leq C_x |u|_V.$$

Ce sont les conséquences de cette propriété que nous allons étudier maintenant dans les deux sections à suivre en explorant le cadre des espaces à noyau reproduisant (ENR)³. Nous reviendrons ensuite avec l'outil des noyaux à la résolution des problèmes (P_V) et (P'_V) aboutissant à l'expression de la métrique locale induite par V sur M .

Les ENR ne sont pas des objets nouveaux [1], mais ils intéressent maintenant une large partie des mathématiciens et informaticiens appliqués en théorie de l'apprentissage et en théorie de l'interpolation pour leur capacité à produire de nouveaux algorithmes. Dans notre cadre, nous avons besoin de définir des ENR de fonctions à valeurs dans \mathbb{R}^m .

Rappelons ici brièvement les principaux points.⁴

4.1. Qu'est-ce qu'un ENR ?. Un ENR H sera pour nous un espace de Hilbert de fonctions $h : \mathcal{X} \rightarrow \mathbb{R}^m$ tel que pour tout $x \in \mathcal{X}$ et tout $\alpha \in \mathbb{R}^m$, la fonction d'évaluation $\delta_x^\alpha : h \rightarrow \langle h(x), \alpha \rangle_{\mathbb{R}^m}$ soit une forme linéaire continue pour la norme de H .

REMARQUE II.2. Lorsque H est un espace de fonctions régulières, H^* n'est généralement pas un espace de fonctions mais de distributions. Une situation bien connue est celle des espaces de Sobolev $H^s(\mathbb{R}^d) \triangleq \{ f \in L^2(\mathbb{R}^d) \mid \int_{\mathbb{R}^d} |\hat{f}(\xi)|^2 (1 + |\xi|^2)^s d\xi < +\infty \}$ pour $s \geq 0$ pour lesquels $(H^s)^* = H^{-s}$ est un ensemble de distributions dont la transformée de Fourier $\hat{f}(\xi) = \int f(x) e^{-i\langle \xi, x \rangle} dx$ peut croître sur les hautes fréquences (pour $s < 0$ remplacer dans la définition $f \in L^2(\mathbb{R})$ par $f \in \mathcal{S}'(\mathbb{R}^d)$ l'espace de Schwartz des distributions tempérées par lequel la transformée de Fourier est bien définie). On voit clairement ici que plus s est grand, plus H^s contient des fonctions régulières (via les injections de Sobolev) et plus H^{-s} contient des objets très irréguliers. Notons que dans le cas $s = 0$, $H^0 = L^2$ et on a une identification triviale entre L^2 et son dual.

EXERCICE II.1. Dans le cas, $p = 1$ (interpolation de fonctions), on a $\hat{\delta}_x(\xi) = e^{-i\langle \xi, x \rangle}$. Vérifier alors à la main que H^s est un ENR ssi $s > d/2$. On remarquera au passage que

3. La terminologie anglo-saxonne est RKHS pour Reproducible Kernel Hilbert Space

4. On pourra consulter la thèse de J. Glaunes [9] pour un traitement très complet des ENR dans le cadre qui nous intéresse ici, ou encore le papier assez complet [10] consultable ici

plus la dimension d est grande, plus il faut contrôler la décroissance de la Transformée de Fourier (et donc la régularité des fonctions).

Un résultat central est le théorème de représentation de Riesz (cf Thm V.5 de [6]) :

THÉORÈME II.1 (Théorème de représentation de Riesz-Fréchet). *Soit H un espace de Hilbert muni d'un produit scalaire $\langle \cdot, \cdot \rangle_H$ et $l \in H^*$ une forme linéaire continue. Alors il existe unique $g \in H$ tel que pour tout $h \in H$*

$$l(h) = \langle g, h \rangle.$$

Dans notre cas, le théorème nous dit que pour tout $l \in H^*$ il existe unique $K(l) \in H$ tel que

$$(10) \quad \langle K(l), h \rangle_H = l(h) \quad \forall h \in H$$

i.e. la forme linéaire l est représentée par le vecteur $K(l) \in H$.

PROPOSITION II.1. *On notant $|l|_H^* = \sup_{h \in H, |h|_H=1} l(h)$ la norme duale, on a que $K : H^* \rightarrow H$ est une isométrie : $|l|_{H^*} = |K(l)|_H$.*

DÉMONSTRATION. On a déjà que K est bijective. De plus, on vérifie facilement que K est linéaire : $\langle K(\lambda l_1 + l_2), h \rangle_H = (\lambda l_1 + l_2)(h) = \lambda l_1(h) + l_2(h) = \lambda \langle K(l_1), h \rangle_H + \langle K(l_2), h \rangle_H = \langle \lambda K(l_1) + K(l_2), h \rangle_H$. Par unicité, on obtient la linéarité de K .

Enfin, $|l|_H^* = \sup_{h \in H, |h|_H=1} l(h) = \sup_{h \in H, |h|_H=1} \langle Kl, h \rangle_H = |Kl|_H$. \square

En considérant l'isométrie $K : H^* \rightarrow H$ donnée par le théorème de représentation de Riesz, on définit pour tous $x, y \in \mathcal{X}$, une matrice $K(x, y) \in \mathcal{M}_m(\mathbb{R})$ unique telle que

$$(11) \quad \langle K\delta_x^\alpha, K\delta_y^\beta \rangle_H = \alpha^* K(x, y) \beta$$

pour tous $\alpha, \beta \in \mathbb{R}^m$.⁵ On appelle alors noyau associé à H , la fonction $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{M}_m(\mathbb{R})$ définie par (11).⁶

4.2. Positivité du noyau. On a les deux propriétés :

$$(12) \quad K(x, y) = K(y, x)^T$$

et pour toute famille finie $(x_i)_{1 \leq i \leq n}$ de \mathcal{X} et toute famille $(\alpha_i)_{1 \leq i \leq n}$ de vecteurs de \mathbb{R}^m :

$$(13) \quad \sum_{1 \leq i, j \leq n} \alpha_i^* K(x_i, x_j) \alpha_j \geq 0.$$

5. Il suffit de vérifier que $(\alpha, \beta) \rightarrow \langle K\delta_x^\alpha, K\delta_y^\beta \rangle_H$ est une application bilinéaire

6. Notons ici un l'abus de notation consistant à identifier opérateur et noyau

Lorsqu'un noyau $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{M}_m(\mathbb{R})$ vérifie (12) et (13), on dit que c'est un *noyau positif*⁷.

Un point important est si a tout ENR on peut associer un noyau positif, la réciproque est aussi vraie :

*A tout noyau positif $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{M}_m(\mathbb{R})$, on peut associer un espace de Hilbert H de fonctions $h : \mathcal{X} \rightarrow \mathbb{R}^m$ dont K est le noyau.*⁸

Ce résultat permet de construire des espaces à partir des noyaux, d'autant plus que les éléments de H qui se décomposent en sommes finies $\sum_i K(x, x_i)\alpha_i$ forment un sous-espace dense H_f de H sur lequel on a une forme explicite des produits scalaires $\langle g, h \rangle_H$ (cf (15)).

5. Comment construire des noyaux ?

Il y a de nombreuses façons et cela constitue encore un sujet de recherche actif en particulier lorsque l'espace \mathcal{X} est un espace de données non conventionnelles (séquences de longueurs variables, arbres, etc...).

5.1. Building blocks. La première façon de construire des noyaux est de partir de noyaux connus et d'utiliser les « recettes » suivantes :

- (1) $K(x, y) = f(x)^T f(y)$ est un noyau pour tout $f : \mathcal{X} \rightarrow \mathbb{R}^d$
- (2) Si K_1 et K_2 sont deux noyaux, $K = K_1 + K_2$ l'est aussi.
- (3) Si K_1 et K_2 sont deux noyaux, $K = K_1 \circ K_2$ l'est aussi où $(K_1 \circ K_2)(x, y) \triangleq K_1(x, y) \circ K_2(x, y)$ désigne le produit d'Hadamard (i.e. le produit coordonnée par

7. ou plutôt un noyau vectoriel positif

8. La preuve consiste à considérer l'ev H_f des $h : \mathcal{X} \rightarrow \mathbb{R}^m$ qui se représentent sous la forme d'une somme finie du type :

$$(14) \quad h(x) = \sum_{1 \leq i \leq n} K(x, x_i)\alpha_i$$

Pour tous $g, h \in H_f$ s'écrivant $g(x) = \sum_i K(x, x_i)\alpha_i$ et $h(x) = \sum_j K(x, y_j)\beta_j$, on déduit de la positivité du noyau que

$$(15) \quad \langle g, h \rangle_H \triangleq \sum_{i,j} \alpha_i^* K(x_i, y_j)\beta_j = \sum_j g(y_j)^* \beta_j = \sum_i \alpha_i^* h(x_i)$$

définit un produit scalaire sur H_f (notons que les deux dernières égalités de (15) montre que la définition de $\langle \cdot, \cdot \rangle_H$ est consistante et ne dépend pas du choix de la décomposition de g et h en somme finie). On construit H par un procédé de complétion : Si (h_n) est une suite de Cauchy dans H_f , alors par continuité des évaluations (h_n) converge simplement vers une fonction h_∞ . Enfin, si (h_n) est de Cauchy et converge simplement vers la fonction nulle, alors $\langle h_n, h_n \rangle \rightarrow 0$ car $\langle h_n, h_n \rangle_H \leq \langle h_n, h_n - h_m \rangle_H + \langle h_n, h_m \rangle_H$, où $|\langle h_n, h_n - h_m \rangle_H| \leq |h_n|_H |h_n - h_m|_H$ qui tend vers zéros lorsque $n, m \rightarrow \infty$ et $\langle h_n, h_m \rangle = \sum_j \langle h_n(y_j), \beta_j \rangle_{\mathbb{R}^m} \rightarrow 0$ si $h_m = \sum_j K(\cdot, y_j)\beta_j$. Ceci nous permet d'identifier la complétion de H_f comme un Hilbert de fonctions de \mathcal{X} dans \mathbb{R}^m .

coordonnée des matrices⁹⁾ de $K_1(x, y)$ avec $K_2(x, y)$ i.e.

$$K(x, y)_{ij} = K_1(x, y)_{ij} K_2(x, y)_{ij}.$$

(4) Si $P(t) = \sum_{k=0}^n p_k t^k$ est un polynôme à coefficients positifs, alors $P(K) \triangleq \sum_{k=0}^n p_k \underbrace{K \circ \cdots \circ K}_{k \text{ fois}}$ est un noyau pour tout noyau K .

(5) Si $K : \mathcal{Y} \times \mathcal{Y} \rightarrow \mathcal{M}_m(\mathbb{R})$ est un noyau et $\varphi : \mathcal{X} \rightarrow \mathcal{Y}$, alors $K_\varphi : \mathcal{X} \times \mathcal{X}$ défini par $K_\varphi(x, y) = K(\varphi(x), \varphi(y))$ est un noyau.

DÉMONSTRATION. La preuve de (1),(2) et (3) est laissée en exercice. Pour le point 4, il suffit de remarquer que la multiplication d'un noyau par un scalaire positif est toujours un noyau puis d'utiliser (3) de façon répétée pour déduire que $\circ_{j=1}^k K$ est un noyau et de terminer en sommant par la règle (1). \square

EXERCICE II.2. Montrer le point (1). Que vaut l'ENR associé à K ?

EXERCICE II.3. Montrer que A et B sont deux matrices symétriques positives, alors $A \circ B$ est symétrique et positive (Indication : commencer à le vérifier dans le cas où $A = uu^*$ et $B = vv^*$ puis déduire le cas général en décomposant A et B dans des bases de vecteurs propres). Etendre ce résultat à la preuve de (3).

Le point (1) permet de construire facilement des noyaux élémentaires que l'on peut combiner facilement à l'aide d'applications répétées de (2) et (3).

Notons que si nous voulons seulement des noyaux diagonaux, $K(x, y) = k(x, y)\text{Id}_d$, il suffit de partir d'un noyau fonctionnel $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$.

5.2. Exemple du noyau gaussien. Vérifions par exemple que

$$(16) \quad K(x, y) = \exp(-|x - y|^2)\text{Id}_d$$

est un noyau, appelé noyau gaussien : il suffit de montrer que $k(x, y) = \exp(-|x - y|^2)$ est un noyau fonctionnel.

Or $(x, y) \rightarrow \langle x, y \rangle$ est évidemment un noyau fonctionnel, d'où en utilisant un développement de $t \rightarrow \exp(2t)$, la règle (4) et un passage à la limite, on déduit que $(x, y) \rightarrow \exp(2\langle x, y \rangle)$ est un noyau. Comme

$$k(x, y) = \underbrace{\exp(-|x|^2) \exp(-|y|^2)}_{\text{noyau (cf (1))}} \underbrace{\exp(2\langle x, y \rangle)}_{\text{noyau}}$$

on obtient en utilisant la règle 3 que k est un noyau.

9. avec des notations "Scilab" ou "Matlab", on a $A \circ B = A .* B$

5.2.1. *Invariance par translation.* Dans le cas où $\mathcal{X} = \mathbb{R}^m$, les noyaux invariants par translation de la forme

$$(17) \quad K(x, y) = \rho(y - x)$$

engendre des ENR pour lesquels les opérateurs de translations $\tau : H \rightarrow H$ définis par $\tau h(x) = h(x + \tau)$ sont des isométries¹⁰ :

$$\langle \tau g, \tau h \rangle_H = \langle g, h \rangle_H$$

EXERCICE II.4. Vérifier que réciproquement, si les translations sont des isométries, alors K est invariant par translation.

La question est alors du choix de ρ dans (17). Le théorème de Bochner, nous donne une caractérisation de la plupart des noyaux invariants par translation au travers de la transformée de Fourier :

Une application $\rho : \mathbb{R}^d \rightarrow \mathcal{M}_m(\mathbb{R})$ ¹¹ définit un noyau positif invariant par translation ssi sa transformée de Fourier $\hat{\rho}$ est telle que $\hat{\rho}(\omega)$ est hermitienne positive pour tout $\omega \in \mathbb{R}^m$. En prenant $\rho(u) = \exp(-|u|^2/2)\text{Id}_m$, on obtient $\hat{\rho}(\xi) = \sqrt{2\pi}^m \exp(-|\xi|^2/2)\text{Id}_m$ et le théorème de Bochner nous permet de retrouver la positivité du noyau gaussien (voir aussi exercice II.7).

5.2.2. *Invariance par isométrie.* Si l'on veut ajouter à l'invariance par translation une invariance par isométrie i.e. pour toute isométrie $R \in O(\mathbb{R}^d)$, l'application $R : H \rightarrow H$ définie par

$$Rh(x) = h(Rx)$$

est une isométrie, alors on montre que $K(x, y) = \rho(|y - x|)$ i.e. ne dépend que de la norme de $x - y$. Les conditions sur ρ pour générer un noyau positif invariant par isométries affines sont données par le théorème de Schoenberg. Nous ne garderons que la condition suffisante suivante qui nous dit :

Si $\rho(r) = \int \exp(-r^2 u^2) d\mu(u)$ où μ est une mesure borélienne positive finie sur \mathbb{R}_+ , alors $K(x, y) = \rho(|x - y|)\text{Id}_{\mathbb{R}^m}$ est un noyau positif invariant par isométries affines

Notons qu'en prenant $\mu = \delta_\lambda$, on retrouve le noyau gaussien et que la forme générale s'obtient en utilisant la règle (2) et en passant à limite.

EXERCICE II.5. Montrer que $\rho(r) = 1/(1 + r^2)$ définit un noyau positif (bonne alternative au noyau gaussien en pratique)

10. cette propriété est intéressante pour construire des méthodes d'appariement qui sont invariantes par translation

11. on suppose que ρ est intégrable, ainsi que sa transformée de Fourier et que ρ s'obtient par transformée de Fourier inverse.

5.3. Que dire sur l'ENR associé ? Si l'on part d'un noyau et qu'on construit l'ENR associé, que peut-on dire de l'espace construit ?

5.3.1. *Régularité en fonction du noyau.* Une réponse est de questionner la régularité. On se doute qu'elle est relié au noyau. Par exemple, considérons un ouvert Ω de \mathbb{R}^d et $\mathcal{C}_0(\Omega, \mathbb{R}^m)$ l'espace des applications h continues de Ω dans \mathbb{R}^m qui tendent vers 0 lorsque $x \rightarrow \partial\Omega$ ¹². Muni de la norme uniforme $\| \cdot \|_\infty$ sur Ω , c'est un Banach. On a le résultat suivant¹³ :

Si le noyau K est continu borné sur $\Omega \times \Omega$ et que pour tout x , $K(x, \cdot) \in \mathcal{C}_0(\Omega, \mathcal{M}_m(\mathbb{R}))$ ¹⁴, alors H s'injecte continûment dans $\mathcal{C}_0(\Omega, \mathbb{R}^m)$.

Ce résultat peut être étendu pour des régularités supérieures. Une extension assez naturelle dans le cas des noyaux $K(x, y) = k(y - x)$ invariants par translation dit que :

Si $k \in \mathcal{C}_0^{2p}(\Omega, \mathcal{M}_m(\mathbb{R}))$ ¹⁵ alors H s'injecte continûment dans $\mathcal{C}_0^p(\mathbb{R}^d, \mathbb{R}^m)$.

En particulier, dans le cas du noyau gaussien, les éléments de H sont \mathcal{C}^∞ et tendent vers 0 en l'infini ainsi que toutes leurs dérivées.

5.3.2. *Séparabilité.* Une dernière chose mérite d'être contrôlée ? Est-ce que l'ENR construit avec un noyau K est séparable (i.e. existence d'une famille dénombrable dense) ? Sans séparabilité, impossible de construire une base Hilbertienne (e_n) . Le résultat suivant nous suffira

Si K est un noyau dans $\mathcal{C}(\mathcal{X} \times \mathcal{X}, \mathcal{M}_m(\mathbb{R}))$, alors H est séparable dès que \mathcal{X} l'est.

EXERCICE II.6. Montrer si K est continue, alors $x \rightarrow K\delta_x^\alpha$ est continue sur H . En déduire le résultat.

6. Résolution du problème de projection

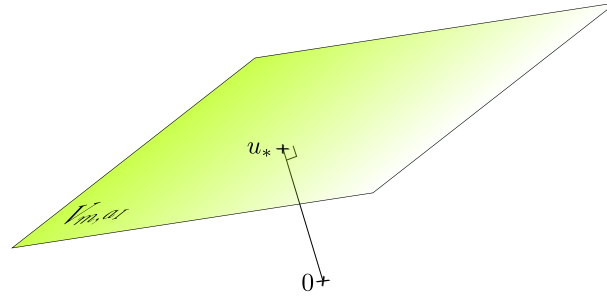
On suppose donc maintenant que V est un ENR de champs de vecteurs sur \mathbb{R}^d . On dispose donc d'une isométrie : $V^* \rightarrow V$ et d'un noyau $K(x, y) \in \mathcal{M}_d(\mathbb{R})$.

12. on considère que si Ω est non borné, $\infty \in \partial\Omega$ et $h(x)$ tends vers 0 lorsque $|x| \rightarrow \infty$. En particulier $\mathcal{C}_0(\mathbb{R}^d, \mathbb{R}^m)$ est l'espace des fonctions continues qui tendent vers 0 à l'infini

13. la preuve en est assez simple : si H_f est l'espace des combinaisons finies $h(x) = \sum_i K(x, x_i)\alpha_i$, alors $H_f \subset \mathcal{C}_0(\Omega, \mathbb{R}^m)$. De plus $\langle h(x), \alpha \rangle_{\mathbb{R}^m} = \langle h, K\delta_x \alpha \rangle_H \leq |h|_H |K\delta_x \alpha|_H \leq |h|_H \sqrt{\alpha^* K(x, x) \alpha} \leq |h|_H |K(x, x)| |\alpha|^2$, d'où si M majore uniformément $|K(x, x)|$ on a $|h|_\infty \leq M|h|_H$. Comme les éléments de H sont les limites simples des suites de Cauchy dans H d'éléments de H_f (et donc de $\mathcal{C}_0(\Omega, \mathbb{R}^m)$), on obtient le résultat

14. $K(x, \cdot)$ désigne ici l'application $y \rightarrow K(x, y)$

15. Cette fois ci, on suppose que les dérivées partielles jusqu'à l'ordre $2p$ sont dans $\mathcal{C}_0(\Omega, \mathcal{M}_m(\mathbb{R}))$

FIGURE 2. Projection sur V_{m,a_I}

6.1. Système linéaire associé.

THÉORÈME II.2. *On suppose que V est un ENR. Si (P_V) admet une solution, alors elle est unique et s'écrit $u_* = K(\sum_{i \in I} \delta_{x_i}^{\alpha_i})$ pour une famille $\alpha = (\alpha_i)_{i \in I}$ de coefficients dans \mathbb{R}^d solution du système linéaire $|I|d \times |I|d$:*

$$(18) \quad \mathbb{K} \alpha_I = a_I$$

où \mathbb{K} est la matrice bloc $\mathbb{K} = (K(x_i, x_j))_{i,j \in I}$, $\alpha_I = (\alpha_i)_{i \in I}$ et $a_I = (a_i)_{i \in I}$.

DÉMONSTRATION. Si (P_V) admet une solution, alors pour $m = (x_i)_{i \in I}$, V_{m,a_I} est un fermé non vide et par le théorème de projection, 0 se projette en un point unique u_* . De plus $V_{m,a_I} = u_* + V_{m,0}$ où $V_{m,0} = \{u \in V \mid u(x_i) = 0, i \in I\}$. Il est clair que $V_{m,0} = \{u \in V \mid \langle u(x_i), \alpha_i \rangle_{\mathbb{R}^d} = 0 \forall \alpha_i \in \mathbb{R}^d \text{ et } i \in I\}$. D'après (7), on déduit que $V_{m,0}$ est l'orthogonal de $\text{Vect}\{K(\delta_{x_i}^{\alpha_i}), i \in I, \alpha_i \in \mathbb{R}^d\}$ et donc $V_{m,0}^\perp = \text{Vect}\{K(\delta_{x_i}^{\alpha_i}), i \in I, \alpha_i \in \mathbb{R}^d\}$.

Si $u_*(x) = \sum_{j \in I} K(x, x_j) \alpha_j$ on a donc, pour tout i ,

$$(19) \quad \sum_{j \in I} K(x_i, x_j) \alpha_j = a_i$$

ce qui conduit au système (18). □

REMARQUE II.3. *L'inversibilité de la matrice \mathbb{K} est vraie dès que V contient suffisamment de champs i.e. $\Pi : V \rightarrow (\mathbb{R}^d)^I$ telle que $\Pi(u) \triangleq (u(x_i))_{i \in I}$ est surjective. En effet, si $\mathbb{K} \alpha_I = 0$ alors $\alpha_I^T \mathbb{K} \alpha_I = \sum_{i,j} \alpha_i^T K(x_i, x_j) \alpha_j = |K(\sum_{i \in I} \delta_{x_i}^{\alpha_i})|_V^2 = |\sum_{i \in I} \delta_{x_i}^{\alpha_i}|_{V^*}^2 = 0$. Par suite $\sum_{i \in I} \delta_{x_i}^{\alpha_i}(u) = \sum_{i \in I} \langle \alpha_i, u(x_i) \rangle_{\mathbb{R}^d} = 0$ et donc par la surjectivité de Π , on a le résultat.*

EXERCICE II.7. Lorsque le noyau $K(x, y) = \rho(x - y) \text{Id}_d$ avec ρ de transformée de Fourier $\widehat{\rho}$ ¹⁶

16. On suppose à nouveau que ρ est intégrable et que ρ s'obtient par transformée de Fourier inverse.

- (1) Montrer que $\sum_{i,j \in I} \rho(x_i - x_j) c_i c_j = \frac{1}{(2\pi)^d} \int |\sum_{j \in I} c_j e^{-i\langle \xi, x_j \rangle}|^2 \widehat{\rho}(\xi) d\xi$ pour toute famille de points x_I dans \mathbb{R}^d et toute famille c_I de scalaires.
- (2) En déduire que \mathbb{K} est inversible dès que $\widehat{\rho}(\xi)$ est réelle et strictement positive puis appliquer au cas gaussien.

6.2. Expression de la métrique locale induite. On déduit des calculs précédents que si u_* est solution de (P_V) alors

$$(20) \quad |u_*|_V^2 = \alpha_I^T \mathbb{K} \alpha_I = a_I^T \mathbb{K}^{-1} a_I$$

la deuxième égalité étant vraie dans la situation classique où \mathbb{K} est inversible. On voit aisément que \mathbb{K}^{-1} définit un produit scalaire sur $(\mathbb{R}^d)^I$ et donc une métrique locale qui dépend de m . Pour rappeler cette dépendance, nous noterons $\mathbb{K} = \mathbb{K}_m$ et nous avons donc l'expression de la métrique locale

$$(21) \quad |\delta m|_m^2 \triangleq \delta m^T \mathbb{K}_m^{-1} \delta m.$$

Notons donc que cette métrique locale est complètement déterminée, *induite*, par le choix du noyau K et donc du ENR V comme prévu dans la construction d'où l'expression de *métrique locale induite*

6.3. Application à l'appariement linéaire. Si on se retourne maintenant vers le problème de l'appariement de points via la décomposition $\varphi = \text{id} + u$ par un vecteur u de déplacements à la *Bookstein*, on obtient la solution en calculant α_I pour $a_i = y_i - x_i$ puis le champ dense u_* par

$$u_*(x) \triangleq \sum_{i \in I} K(x, x_i) \alpha_i.$$

EXERCICE II.8. (Appariement inexact) On cherche un appariement qui tienne compte du bruit sur les données en minimisant

$$(22) \quad J(u) \triangleq \frac{1}{2} |u|_V^2 + \frac{\gamma^2}{2} \sum_{i \in I} |u(x_i) - a_i|^2.$$

Montrer que si la solution existe, elle s'écrit sous la forme $u_*(x) = \sum_{i \in I} K(x, x_i) \alpha_i$ pour un bon choix de α . Vérifier qu'alors $J(u_*) = \frac{1}{2} \alpha_I^T (\mathbb{K} + \text{Id}/\gamma)^2 \alpha_I$ avec $\alpha_i = (\mathbb{K} + \text{Id}/\gamma)^{-1} a_I$. Retrouver l'appariement exact en passant à la limite sur γ .

6.4. Métrique locale et modes propres. Si on fixe une forme de référence x_I , toutes les appariements optimaux vivent dans l'espace $V_{x_I,0}^\perp = \{u(x) = \sum_{i \in I} K(x, x_i) \alpha_i \mid x_i \in \mathbb{R}^d, \forall i \in I\}$ de dimension finie $|I|$.

Pour tout $u \in V_{x_I,0}$, si $u_I = (u(x_i))_{i \in I}$ et $\alpha_I = \mathbb{K}^{-1}u_I$, alors on a

$$(23) \quad |u|_V^2 = \alpha_I^T \mathbb{K} \alpha_I = u_I^T \mathbb{K}^{-1} u_I.$$

La matrice \mathbb{K}^{-1} définit donc une métrique locale sur les u_I qu'il est intéressant de comparer à la métrique euclidienne standard (qui est *grosso modo* la métrique locale dans les espaces Σ_d^k de Kendall). Pour cela, on peut diagonaliser la matrice \mathbb{K} (resp. \mathbb{K}^{-1}) ce qui donne une suite de couples (vecteur propre, valeur propre) (u_I^k, λ^k) (resp. (a_I^k, μ^k)) orthonormés pour la métrique euclidienne. On a alors $\lambda^k = 1/\mu^k$ et $u_I^k = a_I^K = \lambda^k \mathbb{K} a_I^k$.

En définissant $u^k(x) = \sum_{i \in I} K(x, x_i) a_i^k$, les u^k définissent les modes propres de déformations pour la métrique locale (ou les “Principal Warps” selon la terminologie de F. Bookstein [3]).

Chapitre III

Groupes G_V et appariements non-linéaires

1. Reconstruction de G via l'action infinitésimale

Si nous revenons à la discussion en introduction du chapitre précédent, nous savons que lorsqu'on travaille dans le cadre d'une action $G \times M \rightarrow G$ sur un espace homogène, pour un G muni d'une distance invariante à droite, les courbes régulières $t \mapsto g_t$ sur G ont une longueur qu'il est facile de calculer (au moins formellement) : on introduit, via l'action infinitésimale de G sur lui-même, la vitesse "eulérienne" $\xi_t \triangleq \dot{g}_t \cdot g_t^{-1}$ grâce à laquelle la longueur $\ell(g)$ de la courbe s'écrit $\ell(g) = \int |\dot{g}_t|_{g_t} dt = \int |\xi_t|_e dt$ par invariance à droite (cf définition I.7). On peut prendre alors le problème un peu différemment *en partant* de la courbe $t \mapsto \xi_t$ dans $T_e G$ telle que $\|\xi\|_1 \triangleq \int_0^1 |\xi_t| dt < +\infty$ et en construisant la courbe $t \mapsto g_t$ solution de

$$(24) \quad \dot{g}_t = \xi_t \cdot g_t.$$

qui est par construction de longueur $\|\xi\|_1$.

Par suite, si $f_G(g, \xi) \triangleq \xi \cdot g$, $f_G(\cdot, \xi)$ est un champ de vecteur sur G et g_t est solution d'une EDO du type

$$(25) \quad \dot{g}_t = f_G(g_t, \xi_t).$$

On peut donc reconstruire les éléments de G connectés à e en intégrant l'équation (24) pour différents choix de la courbe $t \mapsto \xi_t \in L^1([0, 1], T_e G)$ ¹ Mieux encore, dans le cas où G peut être vu comme plongé dans une variété linéaire ou affine, l'intégration de (24) ne nécessite pas de déployer un outillage mathématique trop délicat mais seulement de savoir intégrer des EDO dans \mathbb{R}^n ou éventuellement dans des espaces de *Hilbert* ou de *Banach*.

Ainsi, si on connaît $T_e G$ et f_G via l'action infinitésimale, on peut alors en supposant G connexe par arc, définir G à partir des solutions de (24) partant de e pour $t \rightarrow \xi_t \in L^1([0, 1], T_e G)$. C'est pour nous une voie royale pour définir concrètement des groupes de

1. Dans le cas des groupe de Lie de dimension finie, en fait il suffit de considérer que ξ_t est constant en temps ce qui permet directement d'identifier la composante connexe de l'identité de G avec $T_e G$ son algèbre de Lie.

difféomorphismes comme nous allons le voir. Dans ce cas, on pourra prendre $T_e G = V$ un ENR de champs de vecteurs sur \mathbb{R}^d et $f(\varphi, v) = v \circ \varphi$.

Par ailleurs, on a pour $m_t = g_t \cdot m_0$ que $\dot{m}_t = \xi_t \cdot m$ et donc $\dot{m}_t = f_M(m_t, \xi_t)$ où cette fois $f_M(\cdot, \xi)$ est un champ de vecteur sur M . On se retrouve ici dans une situation similaire où la courbe tracée sur M par l'action sur m_0 d'une courbe dans G est solution d'une EDO sur M .

2. EDO dans les Banachs

Nous allons ici donner un résultat assez simple mais important d'existence d'unicité de la solution d'une équation différentielle ordinaire (E.D.O) du type

$$\dot{q} = f(q, u)$$

dans laquelle u joue le rôle d'une variable de contrôle qui dépend du temps sans condition de régularité en temps. Par contre, nous supposerons bien sûr une régularité *spatiale* Lipschitz de f en q et u pour avoir unicité des trajectoires. Pour nos besoins, nous devons considérer des EDO à valeurs dans des Banachs et non seulement dans un espace de dimension finie.

Soient donc B_e et B_c deux espaces de Banach munis de leurs normes respectives $|\cdot|_{B_e}$ et $|\cdot|_{B_c}$, $\mathcal{B} = b + B_e$ un espace *affine* de direction B_e et $f : \mathcal{B} \times B_c \rightarrow B_e$ telles qu'il existe $K > 0$ vérifiant

$$(26) \quad |f(q, u) - f(q', u')|_{B_e} \leq K ((|q - b|_{B_e} + 1)|u - u'|_{B_c} + (|u'|_{B_c} + 1)|q - q'|_{B_e})$$

REMARQUE III.1. *Dans la suite, pour alléger les notations, nous noterons $|u|$ pour $|u|_{B_c}$ lorsque $u \in B_c$ et $|q - q'|$ pour $|q - q'|_{B_e}$ lorsque $q, q' \in \mathcal{B}$. De façon plus générale, lorsqu'il n'y a pas d'ambiguïté sur l'appartenance d'un élément x à un e.v.n, $|x|$ désigne la norme de x dans cet espace.*

2.1. Fonctions Bochner mesurables. L'intégration d'EDO dans des Banach ne présente pas vraiment de difficultés comme nous allons le voir. Le seul point technique est de disposer d'une théorie de l'intégration pour des fonctions $f : I \rightarrow B$ d'un intervalle I borné de \mathbb{R} dans un Banach. Celle-ci est connue et due à Bochner [2] que nous retrouvons à nouveau. Nous ouvrons donc ici une petite parenthèse que l'on peut sauter en première lecture en s'appuyant sur la notion classique d'intégrale de Lebesgue. Nous ne donnons ici que quelques éléments et nous renvoyons le lecteur intéressé au chapitre 1 de [12] pour une exposition détaillée.

On commence par définir les fonctions *Bochner mesurables* (ou encore fortement mesurables)².

DÉFINITION III.1. On dit que $f : I \rightarrow B$ est *Bochner mesurable* si f est une limite p.p. de fonctions simples (celles qui s'écrivent $f = \sum_{i=1}^n \mathbf{1}_{E_i} b_i$ où (E_i) est une famille finie de mesurables de I et $b_i \in B$).

On vérifie facilement que si f est Bochner mesurable, alors $t \mapsto |f(t)|_B$ l'est aussi au sens classique. On a de plus cette caractérisation équivalente :

PROPOSITION III.1 (cf [12] Corollary 1.1.8). Une fonction $f : I \rightarrow B$ est Bochner mesurable ssi il existe une suite de fonction mesurables $f_n : I \rightarrow B$ ne prenant qu'un nombre dénombrable de valeurs telle que f_n converge vers f uniformément sur $I \setminus N$ où N est négligeable (i.e. $\lambda(N) = 0$).

EXERCICE III.1. La notion de Bochner intégrabilité est très stable comme on peut le voir dans l'exercice suivant qui regroupe quelques propriétés importantes.

- (1) Vérifier que si $f : I \rightarrow B$ est une limite p.p. de fonctions Bochner mesurables, alors f est Bochner mesurable.
- (2) Montrer que si $f \in \mathcal{C}(I, B_1)$ où B_1 est un Banach, alors f est Bochner mesurable.
- (3) Vérifier que si $g : B_1 \rightarrow B_2$ est une application continue entre B_1 et B_2 et $f : I \rightarrow B_1$ est Bochner mesurable alors $g \circ f$ est aussi Bochner mesurable.
- (4) Vérifier que si $g \in \mathcal{C}(I \times B_1, B_2)$ et $f : I \rightarrow B_1$ est Bochner mesurable alors $t \mapsto g(t, f(t))$ est Bochner mesurable.

2.2. Intégrale de Bochner. On dit alors que f est *Bochner intégrable* si il existe une suite f_n de fonctions simples sur I telle que f_n converge vers f p.p. (f est donc Bochner mesurable) et $\lim \int_I |f - f_n|_B dt \rightarrow 0$. Dans ce cas, on définit

$$(27) \quad \int_I f(t) dt \triangleq \lim_{n \rightarrow +\infty} \int_I f_n(t) dt$$

où l'intégrale des fonctions simples s'obtient par la formule de sommation habituelle.

EXERCICE III.2. Vérifier que la définition précédente de l'intégrale de Bochner à bien un sens (ie que la limite existe et ne dépend pas du choix de la suite (f_n)).

2. La notion de mesurabilité correspond dans le cas où $B = \mathbb{R}^d$ à la notion classique de mesurabilité des cours d'intégration. Par le théorème de Pettis, on montre que si B est séparable, alors la forte mesurabilité est équivalente à la notion de faible mesurabilité définie par la propriété que $t \mapsto l(f(t))$ est mesurable pour toute forme linéaire continue $l \in B^*$ et donc à la notion de mesurabilité classique pour les fonctions entre deux espaces mesurables.

On utilise généralement la caractérisation importante suivante :

THÉORÈME III.1 (cf Theorem 1.4.3 dans [12]). *Les fonctions Bochner intégrables ne sont rien d'autres que les fonctions Bochner mesurables telles que $\int |f|_B dt < \infty$.*

L'utilisation des résultats de l'exercice III.1 et du théorème permet en gros de travailler avec les intégrales de Bochner comme avec les intégrales classiques sur \mathbb{R}^d .

2.3. Espaces $L^p(I, B)$. On peut donc définir l'espace $L^1(I, B)$ comme (à équivalence p.p. près) l'ensemble des fonctions Bochner mesurables telles que $\int_T |f(t)|_B dt < +\infty$. L'espace $L^1(I, B)$ est alors un espace de Banach pour la norme

$$\|f\|_1 = \int_I |f|_B dt$$

qui vérifie

$$\left| \int_I f(t) dt \right|_B \leq \int_I |f|_B(t) dt.$$

On peut définir de la même façon pour $p \geq 1$ l'espace de Banach $L^p(I, B)$ des fonctions Bochner mesurables L^p muni de la norme $\|\cdot\|_p$.

2.4. Existence et unicité de solution. Nous refermons la parenthèse sur l'intégrale de Bochner et nous reprenons maintenant le cours de notre discussion commencée en introduction de cette section.

PROPOSITION III.2. *Pour tout $u \in L^1([0, 1], B_c)$ et tout $q_0 \in \mathcal{B}$, il existe une unique solution $q \in \mathcal{C}([0, 1], \mathcal{B})$ de*

$$(28) \quad q(t) = q_0 + \int_0^t f(q(s), u(s)) ds$$

DÉMONSTRATION. La preuve suit le schéma par point fixe classique. On définit $q^0(t) \equiv q_0$ et pour tout $n \geq 0$ $q^{n+1}(t) \triangleq q_0 + \int_0^t f(q^n(r), u(r)) dr$ ce qui donne pour $n \geq 1$ que

$$|q^{n+1}(t) - q^n(t)| \leq K \int_0^t (|u(r)| + 1) |q^n(r) - q^{n-1}(r)| dr \leq \frac{(K \int_0^1 |u(r)| + 1) dr)^n}{n!} \|q^1 - q^0\|_\infty$$

où $\|q\|_\infty = \sup_{[0,1]} |q(t)|$ est la norme infinie³. Par suite $(q^n)_{n \geq 0}$ est de Cauchy dans $\mathcal{C}([0, 1], \mathcal{B})$ et si q^∞ est la limite des q^n on a $|f(q^n(r), u(r)) - f(q^\infty(r), u(r))| \leq K(|u(r)| + 1)|q^n(r) - q^\infty(r)| \rightarrow 0$ et $|f(q^n(r), u(r))| \leq K(|u(r)| + 1) \sup \|q^n - b\|_\infty \in L^1([0, 1], B_e)$. Ainsi, par application du théorème de convergence dominée on déduit que q^∞ est une solution.

3. Les intégrales que l'on considère ici sont des intégrables de Bochner. Il faudrait donc vérifier que toutes les fonctions sont Bochner intégrables ce qui peut être fait sans difficultés en utilisant les outils introduits dans la parenthèse précédente en particulier l'exercice III.1.

Il nous faut vérifier l'unicité. Pour cela, on remarque que si q et q' sont deux solutions, alors $|q'(t) - q(t)| \leq K \int_0^t (|u(r)| + 1)|q(r) - q'(r)|dr$ et donc d'après le lemme de Grönwall rappelé ci-dessous, on a $q = q'$. \square

LEMME III.1 (Grönwall). *Soient $a, b, c \geq 0$ trois fonctions de $[t_0, t_1]$ dans \mathbb{R} , telles que a est bornée et pour tout $t \in [t_0, t_1]$ on a*

$$a(t) \leq b(t) + \int_{t_0}^t c(s)a(s)ds.$$

Alors pour tout $t \in [t_0, t_1]$,

$$a(t) \leq \left(\sup_{s \in [t_0, t]} b(s) \right) \exp\left(\int_{t_0}^t c(s)ds \right).$$

DÉMONSTRATION. On a $a(t) \leq b(t) + \int_{t_0}^t c(s_1)(b(s_1) + \int_{t_0}^{s_1} c(s_2)a(s_2)ds_2)ds_1$ et par récurrence

$$a(t) \leq \left(\sup_{[t_0, t]} b(s) \right) \sum_{k=0}^n \frac{(\int_{t_0}^t c(s)ds)^k}{k!} + \left(\sup_{[t_0, t]} a(s) \right) \frac{(\int_{t_0}^t c(s)ds)^{n+1}}{(n+1)!}$$

\square

On remarquera que l'équation (28) est une version intégrale de l'équation classique $\dot{q} = f(q, u)$ puisque la solution construite, même dans le cas de la dimension finie, peut être une fonction seulement absolument continue⁴ admettant une dérivée seulement définie dans L^1 .

3. Définition de G_V

Revenons à notre problème de la construction des courbes de difféomorphismes et à celle de la définition d'un groupe de difféomorphismes à partir du choix d'un espace tangent en l'identité V et de l'action infinitésimale pour l'action à gauche $(\varphi', \varphi) \mapsto \varphi' \circ \varphi$ donnée par $v \cdot \varphi = v \circ \varphi$.

Commençons par plonger notre groupe dans une variété affine \mathcal{B} en considérant le Banach $B_e = \mathcal{C}_b(\mathbb{R}^d, \mathbb{R}^d)$ des fonctions continues bornées muni de la norme uniforme et $b = \text{Id}$ si bien que $\mathcal{B} \triangleq \mathcal{C}_{\text{Id}}(\mathbb{R}^d, \mathbb{R}^d) = \{ \varphi : \mathbb{R}^d \rightarrow \mathbb{R}^d \mid \|\varphi - \text{Id}\|_\infty < \infty \}$. Si maintenant on pose $f(\varphi, v) \triangleq v \circ \varphi$, l'équation (24) devient une équation de flot pour lequel on voudrait

4. Nous dirons que $q : [a, b] \rightarrow B$ est absolument continue si il existe $g \in L^1([a, b], B)$ telle que pour tout $t \in [a, b]$ on a $q(t) = q(0) + \int_a^t g(s)ds$. On vérifie alors que q est non seulement continue sur tout $[a, b]$ mais aussi dérivable p.p. de dérivée $\dot{q}(t) = g(t)$. On note généralement $AC([a, b], B)$ l'ensemble des fonctions continues de $[a, b]$ dans B .

pouvoir utiliser la proposition III.2. Pour cela, il faut des conditions de régularité sur f qui dépendent en fait de conditions de régularités sur les champs de vecteurs dans V .

DÉFINITION III.2. *On dit que V est admissible si on a l'injection continue $V \hookrightarrow \mathcal{C}_b^1(\mathbb{R}^d, \mathbb{R}^d)$ c'est-à-dire, si il existe $K > 0$ tel que*

$$(29) \quad \sup_{x \in \mathbb{R}^d} |v(x)| + \sup_{x \in \mathbb{R}^d} |dv(x)| \leq K|v|_V$$

En supposant que V est admissible, on a immédiatement $\|v \circ \varphi - v' \circ \varphi'\|_\infty \leq [\|v - v'\|_\infty + \|dv'\|_\infty \|\varphi - \varphi'\|_\infty] \leq K[\|v - v'\|_V + (\|v'\|_V + 1)\|\varphi - \varphi'\|_\infty]$ ce qui permet de vérifier les hypothèses de la proposition III.2.

NOTATION III.1. *Pour $\varphi \in \mathcal{C}([0, 1], \mathcal{C}_{Id}(\mathbb{R}^d, \mathbb{R}^d))$ et $v \in L^1([0, 1], V)$ on écrira*

$$\varphi_t(x) \triangleq \varphi(t)(x) \text{ et } v_t(x) \triangleq v(t)(x),$$

(la dernière égalité étant définie pour presque tout t) si bien qu'alors $\varphi_t \in \mathcal{C}_{Id}(\mathbb{R}^d, \mathbb{R}^d)$ et $v_t \in V$ p.p.

Considérons à partir de la proposition III.2 l'application $\Phi : L^1([0, 1], V) \rightarrow \mathcal{C}([0, 1], \mathcal{C}_{Id}(\mathbb{R}^d, \mathbb{R}^d))$ définie par $v \mapsto \Phi^v$ où

$$\Phi_t^v(x) = x + \int_0^t v_s \circ \Phi_s^v(x) ds.$$

REMARQUE III.2. *On peut remarquer ici que l'utilisation de la proposition III.2 donne l'application $v \mapsto \Phi^v$ par intégration d'une EDO dans l'espace affine $\mathcal{C}_{Id}(\mathbb{R}^d, \mathbb{R}^d)$. A tout temps $t \in [0, 1]$, on a donc $\Phi_t^v \in \mathcal{C}_{Id}(\mathbb{R}^d, \mathbb{R}^d)$ et on peut donc parler de l'application $x \mapsto \Phi_t^v(x)$. Un petit point technique apparaît ici, celui de vérifier l'on a bien $\Phi_t^v(x) = x + \int_0^t v_s \circ \Phi_s^v(x) ds$ (comme conséquence de la construction de $v \mapsto \Phi^v$ et non comme définition de $\Phi^v(x)$) car sinon on pourrait sans y prendre garde parler de deux objets différents.*

Or pour x fixé, $f \mapsto f(x)$ est linéaire continue de $B_e = \mathcal{C}_b(\mathbb{R}^d, \mathbb{R}^d)$ dans \mathbb{R}^d et pour toute fonction simple $t \mapsto f_t$ de $[0, 1]$ dans B_e , on vérifie que $t \mapsto f_t(x)$ est une fonction simple à valeurs dans \mathbb{R}^d et que $(\int_0^1 f_s ds)(x) = \int_0^1 f_s(x) ds$ et $|\int_0^1 f_s(x) ds| \leq \int_0^1 \|f_s\|_\infty ds = \int_0^1 |f_s|_{B_e} ds$. En passant à la limite, on étend le résultat à $f \in L^1([0, 1], B_e)$. En remarquant que dans notre situation pour $v \in L^1([0, 1], V)$, $t \mapsto f(\Phi_t^v, v_t) \in L^1([0, 1], B_e)$ et que $f(\Phi_t^v, v_t)(x) = v_t \circ \Phi_t^v(x)$, on déduit $(\int_0^1 v_s \circ \Phi_s^v ds)(x) = \int_0^1 v_s \circ \Phi_s^v(x) ds$.

On remarque enfin que si on définit $\tilde{f}(q, v) \triangleq v(q)$ où $\tilde{f} : \mathbb{R}^d \times V \rightarrow \mathbb{R}^d$, alors \tilde{f} satisfait les conditions de la proposition III.2, pour $\mathcal{B} = B_e = \mathbb{R}^d$ et $B_c = V$. On a donc l'existence et l'unicité de la solution de $q(t) = x + \int_0^t v_s(q(s)) ds$ qui vérifie donc $q(t) = \Phi_t^v(x)$.

PROPOSITION III.3. *On suppose que V est admissible. Alors pour tout $v \in L^1([0, 1], V)$ et tout $t \in [0, 1]$, $\varphi \triangleq \Phi_t^v$ est un homéomorphisme de \mathbb{R}^d .*

DÉMONSTRATION. En considérant $\tilde{v}(t, x) = -v(1 - t, x)$ on obtient en faisant le changement de variable $s' = 1 - s$ dans $\Phi_1^v(x) = \Phi_{1-t}^v(x) + \int_{1-t}^1 v_s \circ \Phi_s^v(x) ds$ que $\Phi_1^v(x) = \Phi_{1-t}^v(x) + \int_0^t v_{1-s'} \circ \Phi_{1-s'}^v(x) ds'$ et donc que

$$\Phi_{1-t}^v(x) = \Phi_1^v(x) + \int_0^t \tilde{v}_s \circ \Phi_{1-s}^v(x) ds$$

d'où par unicité des solutions

$$\Phi_{1-t}^v(x) = \Phi_t^{\tilde{v}}(\Phi_1^v(x))$$

i.e. $x = \Phi_1^{\tilde{v}} \circ \Phi_1^v(x)$ et donc Φ_1^v est inversible. On a le même résultat pour tout t et on déduit que $t \mapsto \Phi_t^v \in \mathcal{C}([0, 1], \mathcal{C}_{\text{Id}}([0, 1], \mathbb{R}^d))$ avec $\Phi_t^v \in \text{Hom}(\mathbb{R}^d)$ pour tout $t \in [0, 1]$. \square

On peut donc définir maintenant G_V :

DÉFINITION III.3 (G_V). *On suppose que V est admissible. On définit alors*

$$G_V \triangleq \{ \Phi_1^v \mid v \in L^1([0, 1], V) \}$$

Il nous reste plus qu'à vérifier que G_V est un groupe ce qui n'est pas difficile avec les outils dont nous disposons et l'occasion de faire un important exercice :

EXERCICE III.3. On considère sur $L^1([0, 1], V)$ l'opération $(v, w) \mapsto v \star w$ par $v \star w(t) = 2w_{2t}\mathbf{1}_{t \leq 1/2} + 2v_{2t-1}\mathbf{1}_{t > 1/2}$.

- (1) Vérifier que \star est une loi interne sur $L^1([0, 1]; V)$ que $\|v \star w\|_1 = \|v\|_1 + \|w\|_1$ et que $\Phi_1^{v \star w} = \Phi_1^v \circ \Phi_1^w$. On a donc G_V stable par composition et $v \mapsto \Phi_1^v$ est un morphisme de $(L^1([0, 1], V), \star)$ dans (G_V, \circ) .
- (2) En déduire en utilisant la proposition III.3 que G_V est un groupe d'homéomorphismes sur \mathbb{R}^d .

On voit ici que le groupe G_V est défini *une fois* donné l'espace V et non *a priori*. Ceci nous permet à moindre frais de définir aussi des espaces de courbes sur G_V en considérant les courbes $t \mapsto \Phi_t^v$ pour $v \in L^1([0, 1], V)$.

4. Distance invariante sur G_V

On peut maintenant définir sur V la distance géodésique associée en minimisant la longueur des chemins entre deux éléments du groupe. Pour tout $\varphi \in G_V$, on définit

$$d_G(\text{Id}, \varphi) \triangleq \inf \{ \|v\|_1 \mid v \in L^1([0, 1], V), \Phi_1^v = \varphi \}$$

et pour tout $\varphi, \varphi' \in G_V$

$$d_G(\varphi, \varphi') \triangleq d_G(\text{Id}, \varphi' \circ \varphi^{-1}).$$

par construction, G_V la propriété d'invariance à droite mais c'est bien sûr aussi une distance :

EXERCICE III.4. En utilisant l'opérateur \star introduit dans l'exercice III.3, montrer que d_G vérifie bien l'inégalité triangulaire. Vérifier de plus que si $d_G(\text{Id}, \varphi) = 0$ alors $\varphi = \text{Id}$.

Nous avons donc construit explicitement pour tout V admissible un groupe d'homéomorphismes muni d'une "vraie" distance invariante à droite.

On peut remarquer de plus que si $v \in L^1([0, 1], V)$, alors en prenant $\epsilon > 0$ et en notant $s_\epsilon(t) \triangleq (\int_0^t |v_t| dt + \epsilon t) / (\|v\|_1 + \epsilon)$, alors s_ϵ est une bijection absolument continue strictement croissante de $[0, 1]$ dans $[0, 1]$ d'inverse $s \mapsto t_\epsilon(s)$ absolument continue de dérivée p.p. $t'_\epsilon(s) = (\|v\|_1 + \epsilon) / (|v_{t(s)}|_V + \epsilon)$. On vérifie que si $w_s \triangleq t'_\epsilon(s) v_{t(s)}$, alors $|w_s| \leq (\|v\|_1 + \epsilon)$ pour tout $s \in [0, 1]$ et donc $w \in L^\infty([0, 1], V) \subset L^2([0, 1], V)$. On a de plus que $\Phi_s^w = \Phi_{t_\epsilon(s)}^v$ et $\|w\|_1 \leq \|w\|_2 \leq (\|v\|_1 + \epsilon)$. On a donc montré la proposition :

PROPOSITION III.4. On a $G_V = \{ \Phi_1^v \mid v \in L^2([0, 1], V) \}$ et

$$d_G(\text{Id}, \varphi) = \inf \{ \|v\|_2 \mid v \in L^2([0, 1], V), \Phi_1^v = \varphi \}$$

$$\text{où } \|v\|_2 = \left(\int_0^1 |v_t|_V^2 dt \right)^{1/2}.$$

L'irruption de l'espace $L^2([0, 1], V)$ est une bonne nouvelle car c'est un espace de Hilbert sur lequel on dispose d'un produit scalaire et de toutes les bonnes propriétés de tels espaces. On remarque au passage dans la preuve de la proposition III.4 que l'inégalité $\|w\|_1 \leq \|w\|_2$ n'est atteinte que si $|w_t| = \text{Cte}$ i.e. on se déplace à vitesse constante sur la courbe Φ_t^w . Par ailleurs la fonctionnelle

$$(30) \quad \mathcal{E}(w) \triangleq \frac{1}{2} \int_0^1 |w_t|_V^2 dt = \frac{1}{2} \|v\|_2^2$$

peut être identifiée formellement, en utilisant l'invariance à droite, avec la quantité $\int_0^1 |\dot{\Phi}_t^w|_{\Phi_t^w}^2 dt$ appelée la *fonctionnelle d'énergie* (energy functional) qui est souvent introduite en géométrie Riemannienne [8]. La minimisation de la fonctionnelle d'énergie aux conditions aux limites fixées construit les géodésiques sur lesquels on se déplace à vitesse constante.

5. Appariements non-linéaires

Enfin, si on considère maintenant l'action de G_V sur les k -uplets de landmarks, on peut remarquer qu'en prenant $B_c = V$ et $\mathcal{B} = B_e = (\mathbb{R}^d)^k$ et

$$f(m, v) \triangleq v \cdot m = (v(x_1), \dots, v(x_k))$$

alors $|f(m, v) - f(m', v')| \leq \sum_{i=1}^k |v(x_i) - v'(x'_i)| \leq k\|v - v'\|_\infty + \sum_{i=1}^k \|dv'\|_\infty |x_i - x'_i| \leq K'(|v - v'| + |v'| |m - m'|)$ pour un $K' > 0$ et donc par application de la proposition III.2 on a existence et unicité de la solution de $\dot{m} = v \cdot m$ pour tout $v \in L^1([0, 1], V)$ et le problème est bien posé. Il est par ailleurs clair encore par un argument d'unicité des trajectoires que si $m_t \triangleq \Phi_t^v \cdot m_0$ alors

$$(31) \quad m_t = m_0 + \int_0^t v_s \cdot m_s ds.$$

Si on veut construire une méthode d'appariement non-linéaire à partir de l'approche des espaces de formes que nous avons introduite, on peut chercher étant donné deux configurations de points m_0 et m_1 la courbe de longueur minimale m_t pour la métrique induite joignant les deux configurations. Nous avons vu qu'on peut construire facilement des chemins sur M à partir de chemins sur G_V en considérant l'action de G_V sur M par $m_t = \Phi_t^v \cdot m_0$ et que la longueur de la courbe m_t pour la métrique induite est majorée par celle de Φ^v sur G . Cependant on se rappelle que

$$d_M(m_0, m_1) \triangleq \inf\{d_G(\text{Id}, \varphi) \mid \varphi \in G_V, \varphi \cdot m_0 = m_1\}$$

si bien que par construction de G_V et de d_G on a

$$d_M(m_0, m_1) = \inf\{\|v\|_1 \mid v \in L^1([0, 1], V), \Phi_1^v \cdot m_0 = m_1\}$$

ou encore en utilisant la proposition III.4

$$(32) \quad d_M(m_0, m_1) = \inf\{\|v\|_2 \mid v \in L^2([0, 1], V), \Phi_1^v \cdot m_0 = m_1\}.$$

On aboutit donc en introduisant la fonctionnelle d'énergie $\mathcal{E}(w) = \frac{1}{2}\|w\|_2^2$ au problème variationnel suivant

$$(E) \begin{cases} \text{argmin } \mathcal{E}(v) = \frac{1}{2} \int_0^1 |v_t|_V^2 dt \\ \text{avec } v \in L^2([0, 1], V), \dot{m}_t = v_t \cdot m_t \\ m_0, m_1 \text{ fixés} \end{cases}$$

où l'équation $\dot{m}_t = v_t \cdot m_t$ doit être comprise au sens intégré de (31) équivalent à $m_t = \Phi_t^v \cdot m_0$ pour tout t . On comprend bien ici que la solution v_* de (E) si elle existe (et on verra que c'est le cas) construit une géodésique $t \mapsto \Phi_t^{v_*}$ sur G_V qui est aussi une géodésique $t \mapsto m_t = \Phi_t^{v_*} \cdot m_0$ sur M pour la métrique induite.

On peut construire également une version inexacte lorsque la configuration cible y est bruitée. On introduit une attache aux données $g(m_1)$ du type $g(m) = \sum_{i=1}^k |y_i - x_i|^2$ par exemple et la solution est exprimée comme la solution du problème variationnel

$$(I) \begin{cases} \operatorname{argmin} \mathcal{E}(v) + g(m_1) \\ \text{avec } v \in L^2([0, 1], V), \dot{m}_t = v_t \cdot m_t \\ m_0 \text{ fixé} \end{cases} .$$

Pour conclure ce chapitre, on voit facilement que les problèmes (E) et (I) peuvent être étendus à des espaces homogènes M plus généraux dès que l'on dispose d'une action infinitésimale $(v, m) \mapsto v \cdot m$.

EXERCICE III.5. Dans le cas des courbes $\gamma \in \mathcal{C}([0, 1], \mathbb{R}^d)$, identifier l'action infinitésimale $(v, \gamma) \rightarrow v \cdot \gamma$ et en posant $f(v, \gamma) \triangleq v \cdot \gamma$ vérifier en choisissant les espaces adéquats à l'aide de la proposition III.2 l'existence et l'unicité des solutions de l'EDO $\dot{\gamma}_t = v_t \cdot \gamma_t$ pour $v \in L^2([0, 1], V)$. Montrer enfin que $\Phi_t^v \cdot \gamma_0 = \gamma_t$ pour tout $t \in [0, 1]$.

Chapitre IV

Un détour par le contrôle optimal

Nous avons vu que nos problèmes d'appariement de points pouvaient être vus comme la résolution de l'un des deux problèmes d'optimisation suivant

$$(E) \begin{cases} \operatorname{argmin} \frac{1}{2} \int_0^1 |v_t|_V^2 dt \\ \text{avec } v \in L^2([0, 1], V), \dot{m}_t = v_t \cdot m_t \\ m_0, m_1 \text{ fixés} \end{cases}$$

ou

$$(I) \begin{cases} \operatorname{argmin} \frac{1}{2} \int_0^1 |v_t|_V^2 dt + g(m_1) \\ \text{avec } v \in L^2([0, 1], V), \dot{m}_t = v_t \cdot m_t \\ m_0 \text{ fixé} \end{cases}.$$

Dans ce chapitre, nous proposons de faire un petit détour pour construire des outils qui nous seront utiles pour caractériser les solutions des problèmes (E) et (I) dans le contexte des appariements de points mais aussi dans des situations beaucoup plus diverses d'appariement de courbes, d'images, de surfaces etc.

Pour cela nous devons faire le lien avec le calcul des variations et le contrôle optimal. Nous avons déjà dans le chapitre précédent considéré des problèmes d'évolutions du type $\dot{q} = f(q, u)$ où $q(t) \in \mathcal{B}$ avec \mathcal{B} espace affine de direction $\vec{\mathcal{B}} = B_e$ et $u(t) \in B_c$. Ici $u(t)$ joue le rôle du paramètre de contrôle dans l'évolution de $q(t)$ à l'instant t . Dans les cas qui nous intéressent, $u(t)$ peut être un paramètre de dimension infinie (champs denses par exemple). On a supposé que B_e et B_c sont deux espaces de Banach et que $f : \mathcal{B} \times B_c \rightarrow B_e$ vérifie pour un $b \in \mathcal{B}$ et un $K > 0$ et tous $(q, u), (q', u') \in \mathcal{B} \times B_c$:

$$(33) \quad |f(q', u') - f(q, u)| \leq K((|q - b| + 1)|u' - u| + (|u'| + 1)|q' - q|).$$

Nous avons alors montré l'existence et l'unicité pour tout $u \in L^1([0, 1], B)$ et tout $q_0 \in \mathcal{B}$ d'une solution dans $\mathcal{C}([0, 1], \mathcal{B})$ à l'EDO.

Cela montre, comme $L^2([0, 1], B_c) \subset L^1([0, 1], B_c)$, que sous la condition (33) on peut définir l'application

$$\Phi : L^2([0, 1], B_c) \rightarrow \mathcal{C}([0, 1], \mathcal{B})$$

telle que $t \rightarrow \Phi(u)(t)$ (équivalente à la notion Φ_t^u précédemment introduite) est la solution de $\dot{q} = f(q, u)$ (sous sa forme intégrée $q(t) = q(0) + \int_0^t f(q(s), u(s))ds$) avec $q(0) = q_0$ (on considère ici q_0 fixé).

Dans ce nouveau contexte, les problèmes (E) et (I) peuvent être généralisés sous la forme

$$(E') \begin{cases} \operatorname{argmin} \int_0^1 C(q, u) dt \\ \text{avec } u \in L^2([0, 1], B_c), \quad q = \Phi(u) \\ q(0), \quad q(1) \text{ fixés} \end{cases}$$

ou

$$(I') \begin{cases} \operatorname{argmin} \int_0^1 C(q, u) dt + g(q(1)) \\ \text{avec } u \in L^2([0, 1], B_c), \quad q = \Phi(u) \\ q(0) \text{ fixé} \end{cases}$$

Nous préciserons bientôt les hypothèses sur C et f pour pouvoir traiter ces problèmes.

1. Perturbation tangente et linéarisé

DÉFINITION IV.1. Soient X un Banach, \mathcal{X} un espace affine de direction $\vec{\mathcal{X}} = X$, $x \in \mathcal{X}$ et $\delta x \in X$. On dit qu'une courbe continue $\tilde{x} : U \rightarrow \mathcal{X}$ définie sur un ouvert $U =]a, b[$ contenant 0 est une perturbation de x tangente à δx si $\tilde{x}(0) = x$ et \tilde{x} dérivable en 0 de dérivée δx :

$$\lim_{s \rightarrow 0, s \neq 0} |(\tilde{x}(s) - x)/s - \delta x|_X \rightarrow 0$$

On s'intéresse ici au cas $\mathcal{X} = X = L^1([0, 1], B_c)$, $u(\cdot), \delta u(\cdot) \in X$.

On considère une perturbation \tilde{u} de u tangente à δu et $\tilde{q}(0, s)$ une perturbation de $q(0)$ tangente à $\delta q(0) \in B_e$. On peut définir $\tilde{q}(\cdot, s)$ solution de

$$\tilde{q}(t, s) = \tilde{q}(0, s) + \int_0^t f(\tilde{q}(r, s), \tilde{u}(r, s)) dr.$$

Formellement, en dérivant l'équation précédente par rapport à s et en prenant $s = 0$ on trouve $\delta q(t) = \delta q(0) + \int_0^t \frac{\partial f}{\partial q}(q(r), u(r)) \delta q(r) + \frac{\partial f}{\partial u}(q(r), u(r)) \delta u(r) dr$. Ce calcul est formel car on ne sait rien de la dérivabilité de $\tilde{q}(t, s)$ en s et le théorème de dérivation sous le signe somme ne s'applique pas. Le résultat reste cependant essentiellement correct comme l'indique le théorème suivant.

THÉORÈME IV.1. On suppose que f est \mathcal{C}^1 et qu'il existe $K > 0$ tel que

$$(H_f) \begin{cases} |\frac{\partial f}{\partial q}| \leq K(|u| + 1) \\ |\frac{\partial f}{\partial u}| \leq K(|q - b| + 1) \end{cases}$$

alors il existe une unique solution $\delta q \in \mathcal{C}([0, 1], B_e)$ de l'équation linéaire

$$\delta q(t) = \delta q(0) + \int_0^t \frac{\partial f}{\partial q}(q(r), u(r)) \delta q(r) + \frac{\partial f}{\partial u}(q(r), u(r)) \delta u(r) dr,$$

qu'on appelle parfois le linéarisé de l'équation $\dot{q} = f(q, u)$ ou encore l'équation tangente. De plus $\tilde{q}(\cdot, s)$ est une perturbation de $q(\cdot)$ dans $\mathcal{C}([0, 1], \mathcal{B})$ tangente à $\delta q(\cdot)$.

EXERCICE IV.1. Vérifier que les conditions (H_f) entraînent la condition (33).

EXERCICE IV.2. On considère le cas des landmarks $\mathcal{B} = (\mathbb{R}^d)^k$ et V un ENR de champs de vecteurs sur \mathbb{R}^d admissible¹. On note pour tout $m = (x_1, \dots, x_k) \in \mathcal{B}$ et tout $u \in V$,

$$f(m, u) \triangleq (u(x_1), \dots, u(x_m)).$$

(1) Montrer que $\frac{\partial f}{\partial m}(m, u)\delta m = (du(x_1)\delta x_1, \dots, du(x_k)\delta x_k)$ et $\frac{\partial f}{\partial u}(m, u)\delta u = (\delta u(x_1), \dots, \delta u(x_k))$.

(2) Vérifier que f est \mathcal{C}^1 et satisfait les conditions (H_f) .

EXERCICE IV.3. On reprend l'exercice précédent dans des surfaces paramétrées $\mathcal{B} = L^\infty(U, \mathbb{R}^d)$ où U est un ouvert de \mathbb{R}^p avec $p \leq d$ et V inchangé. On définit pour tout $q \in \mathcal{B}$ et tout $u \in V$, $f(q, u) \triangleq u \circ q$.

(1) Montrer que f est \mathcal{C}^1 et calculer ses dérivées partielles.

(2) Vérifier les conditions (H_f) .

REMARQUE IV.1. Pour tout $u \in L^p([0, 1], B_c)$, on rappelle la notation $\|u\|_p \triangleq (\int_0^1 |u(t)|^p dt)^{1/p}$.

DÉMONSTRATION. On construit facilement la solution comme limite des δq^n construit par récurrence selon la méthode classique pour lesquels on a (avec $\delta q^0(t) \equiv \delta q^0$)

$$|\delta q^{n+1}(t) - \delta q^n(t)| \leq \int_0^t \left| \frac{\partial f}{\partial q}(q, u) \right| |\delta q^n - \delta q^{n-1}|(r) dr \leq \frac{|K \int_0^1 (|u| + 1) dr|^n}{n!} \|\delta q^1 - \delta q^0\|_\infty$$

et

$$\|\delta q^1 - \delta q^0\|_\infty \leq K \left(\int_0^1 |u|(r) dr + 1 \right) \|\delta q(0)\| + K(\|q - b\|_\infty + 1) \int_0^1 |\delta u|(r) dr < \infty.$$

L'unicité s'obtient par une application de Grönwall. La partie la moins évidente est de montrer que \tilde{q} est une perturbation tangente à δq .

On commence par montrer un résultat plus faible : $\|\tilde{q}(\cdot, s) - \tilde{q}(\cdot, 0)\|_\infty = O(|s|(\|\delta u\|_1 + |\delta q(0)|)) + o(|s|)$. Pour alléger un peu les notations, nous noterons $q^s = \tilde{q}(\cdot, s)$ et $u^s = \tilde{u}(\cdot, s)$. Pour $t \in [0, 1]$ on a

$$\begin{aligned} |q^s(t) - q(t)| &\leq \int_0^t |f(q^s, u^s) - f(q, u)|(r) dr + |q^s(0) - q(0)| \\ &\leq K \int_0^t ((|u^s| + 1)|q^s - q| + (|q - b| + 1)|u^s - u|)(r) dr + |q^s(0) - q(0)| \\ &\leq \exp(K(\|u^s\|_1 + 1))(K(\|q - b\|_\infty + 1)\|u^s - u\|_1 + |q^s(0) - q(0)|) \end{aligned}$$

1. V s'injecte continûment dans $\mathcal{C}_b^1(\mathbb{R}^d, \mathbb{R}^d)$ des champs \mathcal{C}^1 bornés à dérivées bornées muni de la norme $|u|_{1,\infty} \triangleq \sup_{\mathbb{R}^d} |u| + \sup_{\mathbb{R}^d} |du|$ et donc il existe $C > 0$ telle que $|u|_{1,\infty} \leq C|u|_V$.

Or $\|u^s - u\|_1 = |s|(\|\delta u\|_1 + o(1))$ et donc comme $\sup_s \|u^s\|_1$ est borné (continuité de $s \rightarrow u^s$) on a $\|q^s - q\|_\infty = O(|s|(\|\delta u\|_1 + |\delta q(0)|) + o(|s|))$.

On considère maintenant $M_{t,s} \triangleq |\frac{q^s - q}{s}(t) - \delta q(t)|$. On a alors

$$M_{t,s} \leq |\frac{q^s - q}{s}(0) - \delta q(0)| + \int_0^t |\frac{f(q^s, u^s) - f(q, u)}{s} - \frac{\partial f}{\partial q}(q, u)\delta q - \frac{\partial f}{\partial u}(q, u)\delta u|(r)dr.$$

En écrivant

$$|f(q^s, u^s) - f(q, u) - \frac{\partial f}{\partial q}(q, u)(q^s - q) - \frac{\partial f}{\partial u}(q, u)(u^s - u)|(r) \leq \epsilon_{r,s}(|q^s - q| + |u^s - u|)$$

où

$$\epsilon_{r,s} = \int_0^1 |\frac{\partial f}{\partial q}(q^{s,h}, u^{s,h}) - \frac{\partial f}{\partial q}(q, u)|(r) + |\frac{\partial f}{\partial u}(q^{s,h}, u^{s,h}) - \frac{\partial f}{\partial u}(q, u)|(r)dh$$

avec $q^{s,h} = q + h(q^s - q)$ et $u^{s,h} = u + h(u^s - u)$ on déduit

$$\begin{aligned} M_{t,s} &\leq \int_0^t (|\frac{\partial f}{\partial q}(q, u)|M_{r,s} + |\frac{\partial f}{\partial u}(q, u)||\frac{u^s - u}{s} - \delta u|(r)dr + |\frac{q^s - q}{s}(0) - \delta q(0)| \\ &\quad + \int_0^t \epsilon_{r,s}(|\frac{q^s - q}{s}| + |\frac{u^s - u}{s}|)(r)dr \\ &\leq \exp(\int_0^1 |\frac{\partial f}{\partial q}(q, r)|dr)B_s \end{aligned}$$

où

$$B_s \triangleq \int_0^1 \epsilon_{r,s}(|\frac{q^s - q}{s}| + |\frac{u^s - u}{s}|)(r) + |\frac{\partial f}{\partial u}(q, u)||\frac{u^s - u}{s} - \delta u|(r)dr + |\frac{q^s - q}{s}(0) - \delta q(0)|.$$

Or si $\|u^{s_n} - u\|_1 \rightarrow 0$, quitte à extraire une sous-suite, on peut supposer que (réciproque de Lebesgue) $|u^{s_n}(r) - u(r)| \rightarrow 0$ presque partout (en r) et par continuité de $\frac{\partial f}{\partial q}$ et $\frac{\partial f}{\partial u}$ et convergence dominée que $\epsilon_{r,s_n} \rightarrow 0$ presque partout.

Comme par hypothèse sur \tilde{u} , $\|\frac{u^s - u}{s} - \delta u\|_1 \rightarrow 0$ et comme $\sup_r \|\frac{\partial f}{\partial u}\|_\infty \leq \sup_r K(|q(r) - b| + 1) < +\infty$, on déduit finalement que $B_{s_n} \rightarrow 0$. Par suite $B_s \rightarrow 0$ et $\|\frac{q^s - q}{s} - \delta q\|_\infty \rightarrow 0$ ce qui prouve que \tilde{q} est une perturbation tangente à δq . \square

THÉORÈME IV.2. *On suppose en outre que C est \mathcal{C}^1 et qu'il existe $K > 0$ tel que*

$$(H_C) \begin{cases} |\frac{\partial C}{\partial u}| \leq K(|q - b| + 1)(|u| + 1) \\ |\frac{\partial C}{\partial q}| \leq K(|u| + 1)^2 \end{cases}$$

Alors, si \tilde{u} est une perturbation dans $L^2([0, 1], B_c)$ de u tangent à δu on a que

$$\tilde{c}(s) = \int_0^1 C(\tilde{q}(t, s), \tilde{u}(t, s))dt$$

est une perturbation de $\tilde{c}(0)$ dans \mathbb{R} tangente à

$$\delta c \triangleq \int_0^1 \frac{\partial C}{\partial q}(q(r), u(r)) \delta q(r) + \frac{\partial C}{\partial u}(q(r), u(r)) \delta u(r) dr.$$

EXERCICE IV.4. Vérifier que si $C(q, u) = \frac{1}{2}|u|_V^2$, alors (H_c) est satisfait.

DÉMONSTRATION. On procède de façon similaire. On a

$$(34) \quad \left| \frac{\tilde{c}(s) - \tilde{c}(0)}{s} - \delta c \right| = \left| \int_0^1 \left(\frac{C(q^s, u^s) - C(q, u)}{s} - \frac{\partial C}{\partial q}(q, u) \delta q - \frac{\partial C}{\partial u}(q, u) \delta u \right) dt \right|$$

$$(35) \quad \leq \int_0^1 (A_t + B_t + C_t + D_t) dt$$

avec

$$A_t = \left| \frac{\partial C}{\partial q}(q, u) \right| \left| \frac{q^s - q}{s} - \delta q \right|(t)$$

$$B_t = \left| \frac{\partial C}{\partial u}(q, u) \right| \left| \frac{u^s - u}{s} - \delta u \right|(t)$$

$$C_t = \epsilon_{t,s}^q \left| \frac{q^s - q}{s} \right|(t)$$

$$D_t = \epsilon_{t,s}^u \left| \frac{u^s - u}{s} \right|(t)$$

où $\epsilon_{t,s}^q = \int_0^1 \left| \frac{\partial C}{\partial q}(z^{s,h}(t)) - \frac{\partial C}{\partial q}(z(t)) \right| dh$ et $\epsilon_{t,s}^u = \int_0^1 \left| \frac{\partial C}{\partial u}(z^{s,h}(t)) - \frac{\partial C}{\partial u}(z(t)) \right| dh$ avec $z^{s,h}(t) = (q, u)(t) + h((q^s, u^s) - (q, u))(t)$. Or on a

$$\int_0^1 A_t dt \leq K \underbrace{\int_0^1 (|u| + 1)^2 dt}_{< +\infty} \left\| \frac{q^s - q}{s} - \delta q \right\|_\infty \rightarrow 0$$

et

$$\int_0^1 B_t dt \leq K(\|q - b\|_\infty + 1) \left(\int_0^1 (|u| + 1)^2(t) dt \right)^{1/2} \left\| \frac{u^s - u}{s} - \delta u \right\|_2 \rightarrow 0.$$

De plus,

$$\int_0^1 C_t dt \leq \left\| \frac{q^s - q}{s} \right\|_\infty \int_0^1 \epsilon_{t,s}^q dt$$

avec

$$\epsilon_{t,s}^q \leq K((|u| + 1)^2 + (|u^s| + 1)^2)(t) \leq K'(|u|^2 + |\delta u|^2 + |R^s|^2 + 1)(t)$$

et $R^s(t) = (u^s - u - s\delta u)(t)$. Comme $\int |R^s|^2 dt \rightarrow 0$, on déduit que la famille $(\epsilon_{t,s}^q)_{s \in [0,1]}$ est uniformément intégrable (u.i.). Par suite comme $\lim_{s \rightarrow 0} \epsilon_{t,s}^q = 0$ presque partout en t on a par uniforme intégrabilité que $\int \epsilon_{t,s}^q dt \rightarrow 0$. Ainsi $C_t \rightarrow 0$.

Enfin,

$$\int_0^1 D_t dt \leq \left\| \frac{u^s - u}{s} \right\|_2 \left(\int_0^1 (\epsilon_{t,s}^u)^2 dt \right)^{1/2}.$$

En utilisant les même arguments que précédemment, on montre que $(\epsilon_{t,s}^u)_s^2$ est u.i. puis que $\int (\epsilon_{t,s}^u)^2 dt \rightarrow 0$. \square

2. Equations hamiltoniennes

Les problèmes variationnels (E') et (I') peuvent être vu comme des problèmes de contrôle optimal (u jouant le rôle d'un contrôle). La caractérisation des u optimaux est un problème de calcul des variations dans lequel le lagrangien $L(q, \dot{q})$ est remplacé par une forme plus générale $L(q, u)$ (ici noté $C(q, u)$ avec un lien entre \dot{q} et u donné par la dynamique $\dot{q} = f(q, u)$). On sait que le calcul classique des variations autour de l'optimisation d'un lagrangien $\int_0^1 L(q, \dot{q}) dt$ avec condition aux bords contient deux facettes, l'une lagrangienne avec les équations d'Euler-Lagrange $\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = 0$ caractérisant les extrémales et l'autre hamiltonienne faisant intervenir le hamiltonien $H(q, p) = (p|\dot{q}) - L(q, \dot{q})$ où $p = \frac{\partial L}{\partial \dot{q}}$ est le vecteur *adjoint* associé (attention ici \dot{q} est une fonction implicite de q et p à travers l'égalité $p = \frac{\partial L}{\partial \dot{q}}$). Les équations d'Euler-Lagrange se traduisent alors sous la forme $\dot{q} = \frac{\partial H}{\partial p}$, $\dot{p} = -\frac{\partial H}{\partial q}$. L'arrivée du hamiltonien permet souvent une étude plus simple des solutions et amènent de nouveaux outils. Cette approche s'étend de la situation classique $L(q, \dot{q})$ à celle plus générale qui nous intéresse et que nous allons développer ici.

On notera dans la suite B_e^* le dual de B_e et pour tout $p \in B_e^*$ et $\delta q \in B_e$, on notera $(p|\delta q) \triangleq p(\delta q)$. On utilisera également la notation $(\frac{\partial C}{\partial q}(q, u)|\delta q)$ pour $\frac{\partial C}{\partial q}(q, u)\delta q$.

Nous pouvons maintenant introduire l'application

$$c : L^2([0, 1], B_e) \rightarrow \mathbb{R}$$

telle que $c(u) = \int_0^1 C(q(t), u(t)) dt$ avec $q = \Phi(u)$.

De plus pour une trajectoire $u \in L^2([0, 1], B_e)$, on peut définir les dérivées au sens de Gâteaux

$$\Phi'(u; \delta u) = \delta q \text{ et } c'(u; \delta u) = \delta c$$

où δq est la solution du linéarisé en u pour la perturbation δu et la donnée initiale $\delta q(0) = 0$ et $\delta c = \int_0^1 (\frac{\partial C}{\partial q}| \delta q) + (\frac{\partial C}{\partial u}| \delta u) dt$

Pour $p_1 \in B_e^*$ on considère l'application

$$\delta u \rightarrow c'(u; \delta u) - (p_1|\Phi'(u; \delta u)(1)) = \delta c - (p_1|\delta q(1))$$

THÉORÈME IV.3. *On suppose les hypothèses (H_f) et (H_C) vérifiées. Si p est l'unique solution de l'EDO rétrograde $p(1) = p_1$ et $\dot{p} = -\frac{\partial f^*}{\partial q}(q, u)p + \frac{\partial C}{\partial q}(q, u)$ (i.e. $p_1 = p(1) - \int_1^0 (\frac{\partial f^*}{\partial q}(q, u)p - \frac{\partial C}{\partial q}(q, u) ds)$ alors*

$$\delta c - (p_1|\delta q(1)) = \int_0^1 (\frac{\partial C}{\partial u}(q, u) - \frac{\partial f^*}{\partial u}(q, u)p|\delta u) dt.$$

Ainsi en considérant $H(q, p, u) = (p|f(q, u)) - C(q, u)$ on a donc pour p solution rétrograde de $\dot{p} = -\frac{\partial H}{\partial q}(q, p, u)$ et $\dot{q} = f(q, u) = \frac{\partial H}{\partial p}(q, p, u)$ que

$$\delta c - (p_1|\delta q(1)) = - \int_0^1 \left(\frac{\partial H}{\partial u}(q, p, u) | \delta u \right) dt.$$

Il faut noter que l'existence et l'unicité de p solution de l'équation rétrograde est garantie sous les hypothèses (H_f) (à q et u fixés). Noter qu'il s'agit d'une équation linéaire.

REMARQUE IV.2. Nous avons donc introduit ici une nouvelle fonction $H(q, p, u)$ qui va jouer le rôle du hamiltonien classique. Notons l'arrivée d'une variable supplémentaire u .

DÉMONSTRATION. Formellement si $t \rightarrow p(t)$ est dérivable avec $p(1) = p_1$ on peut écrire $(p_1|\delta q(1)) = (p(0)|\delta q(0)) + \int_0^1 (\dot{p}|\delta q) + (p|\dot{\delta q}) dt$ si bien que en écrivant $\dot{\delta q} = \frac{\partial f}{\partial q} \delta q + \frac{\partial f}{\partial u} \delta u$ on déduit par regroupement des termes que

$$\delta c - (p_1|\delta q(1)) = \int_0^1 \left(\frac{\partial C}{\partial q} - \frac{\partial f^*}{\partial q} p - \dot{p} | \delta q \right) + \left(\frac{\partial C}{\partial u} - \frac{\partial f^*}{\partial u} p | \delta u \right) dt - (p(0)|\delta q(0)).$$

On est amené à considérer $\frac{\partial C}{\partial q} - \frac{\partial f^*}{\partial q} p - \dot{p} = 0$ pour annuler la contribution en δq .

Rigoureusement on peut considérer la solution rétrograde $p(t) = p_1 - \int_t^1 (\frac{\partial f^*}{\partial q}(q, p) p) ds$ (qui existe) et l'équation $\delta q(t) = \delta q(0) + \int_0^1 \frac{\partial f}{\partial q} \delta q + \frac{\partial f}{\partial u} \delta u dt$ pour montrer en utilisant la formule d'intégration par partie pour les fonctions absolument continues (application de Fubini pour laquelle il n'est pas nécessaire de supposer que p et δq sont C^1) que

$$(p_1|\delta q(1)) = (p(0)|\delta q(0)) - \int_0^1 \left(\frac{\partial f^*}{\partial q} p | \delta q \right) + \left(p | \frac{\partial f}{\partial q} \delta q + \frac{\partial f}{\partial u} \delta u \right) dt$$

ce qui conduit au résultat. \square

Le théorème IV.3 s'applique dans deux situations d'importances que nous présentons maintenant.

3. Cas du problème (I')

On revient au problème (I') :

$$(I') \begin{cases} \operatorname{argmin} \int_0^1 C(q, u) dt + g(q(1)) \\ \text{avec } u \in L^2([0, 1], B_c), \quad q = \Phi(u) \\ q(0) \text{ fixé} \end{cases}$$

où $g : \mathcal{B} \rightarrow \mathbb{R}$ est supposé C^1 .

THÉORÈME IV.4. *On suppose les hypothèses (H_f) et (H_C) vérifiées et on définit*

$$J(u) \triangleq \int C(q, u) dt + g(q(1))$$

avec $q = \Phi(u)$ et $p \in \mathcal{C}([0, 1], B_e^*)$ la solution rétrograde de $p(1) + \frac{dq}{dq}(q(1)) = 0$ et $\dot{p} = -\frac{\partial H}{\partial q}(q, p, u)$. Alors,

(1) pour tout $\delta u \in L^2([0, 1], B_c)$, on a

$$J'(u; \delta u) = - \int_0^1 \left(\frac{\partial H}{\partial u}(q, p, u) | \delta u \right) dt$$

(2) si u atteint le minimum de J

alors on a

$$\frac{\partial H}{\partial u}(q, p, u) \equiv 0$$

En particulier les solutions vérifient

$$\begin{cases} \dot{q} = \frac{\partial H}{\partial p}(q, p, u) \\ \dot{p} = -\frac{\partial H}{\partial q}(q, p, u) \\ \frac{\partial H}{\partial u}(q, p, u) \equiv 0 \end{cases}$$

DÉMONSTRATION. Il suffit de remarquer que J est alors Gâteaux différentiable de dérivée $J'(u; \delta u) = c'(u; \delta u) + (\frac{dq}{dq}(q(1)) | \Phi'(u; \delta u)(1)) = \delta c + (\frac{dq}{dq}(q(1)) | \delta q(1))$ ce qui donne le résultat en posant $p_1 = -\frac{dq}{dq}(q(1))$ et en utilisant le théorème IV.3. \square

EXERCICE IV.5. Reprendre les exercices IV.2 et IV.3 et identifier dans chaque cas le hamiltonien $H(q, p, u)$ puis les équations hamiltoniennes (dans le deuxième cas, on se restreindra au cas où $p \in L^1(U, \mathbb{R}^d)$). Identifier les u tels que $\frac{\partial H}{\partial u} = 0$.

REMARQUE IV.3. On remarque facilement que dans le cas classique $u = \dot{q}$ et $L(q, \dot{q}) = C(q, u)$ que la condition $\frac{\partial H}{\partial u} = 0$ en u donne $p = \frac{\partial L}{\partial \dot{q}}$.

REMARQUE IV.4. Dans le cas $B_c = V$ qui est un cas central dans la suite, on peut écrire

$$\left(\frac{\partial H}{\partial u}(q, p, u)(t) | \delta u(t) \right) = \langle K[\frac{\partial H}{\partial u}(q(t), p(t), u(t))], \delta u(t) \rangle_V.$$

La partie (1) du théorème IV.4 donne alors la valeur du gradient de J calculée dans $L^2([0, 1], V)$ si bien que l'on pourrait ainsi construire une descente de gradient sur u par la mise à jour $u^{n+1} = u^n + \Delta t_n K[\frac{\partial H}{\partial u}(q^n(t), p^n(t), u^n(t))]$, $\dot{q}^{n+1} = \frac{\partial H}{\partial p}(q^n, p^n, u^{n+1})$ où $p^n(1) + \frac{dq}{dq}(q^n(1)) = 0$, $\dot{p}^n = -\frac{\partial H}{\partial q}(q^n, p^n, u^n)$.

4. Cas du problème (E')

Dans ce cas, pour retrouver la caractérisation du théorème IV.3, il faut avoir l'existence d'un multiplicateur de Lagrange $p(1)$ au temps 1. Pour cela, on veut appliquer le raisonnement suivant : on suppose que

$$(C1) \quad u \mapsto E(u) \triangleq (q_1(u), c(u)) \text{ est } \mathcal{C}^1$$

de $L^2([0, 1], B_c)$ dans $\mathcal{B} \times \mathbb{R}$ où $q_1(u) \triangleq \Phi(u)(1) = \Phi_1^u$. Supposons que le minimum de $c(u) = \int_0^1 C(q, u)dt$ soit atteint pour le problème sous contraintes (E') en un point u_* . Il est clair que E ne peut pas être surjective d'un voisinage ouvert U de u_* sur un voisinage ouvert V de $E(u_*)$ car sinon on pourrait trouver $u'_* \in U$ tel que $q_1(u'_*) = q_1(u_*)$ et $c(u'_*) < c(u_*)$ et contredit l'optimalité locale de u_* pour le problème (E') .

Or on sait qu'en dimension finie si E est \mathcal{C}^1 sur un voisinage de u_* de différentielle surjective en u_* , alors E est localement surjective. En dimension infinie, ce résultat perdure comme nous allons le voir.

THÉORÈME IV.5 (Inversion locale). *Soient E et F deux espaces de Banach, Ω un ouvert de E , $f \in \mathcal{C}^1(\Omega, F)$ et $a \in \Omega$. Supposons que $df(a)$ soit une bijection de E vers F . Alors il existe un ouvert V de E et un ouvert W de F contenant $b = f(a)$ et une application $g \in \mathcal{C}^1(W, V)$ telles que $g \circ f(x) = x$ pour tout $x \in V$ et $f \circ g(y) = y$ pour tout $y \in W$. Autrement dit $f|_V$ est un \mathcal{C}^1 -difféomorphisme de V sur W . On dit aussi que f induit un \mathcal{C}^1 -difféomorphisme de V sur W .*

DÉMONSTRATION. Voir Thm III.70 de [5] □

Du théorème d'inversion locale sur les espaces de Banach on tire le corollaire suivant.

COROLLAIRE IV.1 (Thm IV.9 de [5]). *Soit $f \in \mathcal{C}^1(E, F)$ telle que $df(a)$ est surjective en un point $a \in E$. Alors si $E = \ker df(a) \oplus E_1$ avec E_1 fermé, il existe un ouvert U de a , un ouvert U_0 de 0 dans $\ker df(a)$ et un ouvert V de $b = f(a)$ dans F tel que l'application $h : U \rightarrow V \times U_0$ qui définit par $u \rightarrow (f(u), \text{pr}(u - a))$ où pr est la projection sur $\ker df(a)$ parallèlement à E_1 est un \mathcal{C}^1 -difféomorphisme.*

DÉMONSTRATION. Il suffit de remarquer que $\tilde{h} : u \rightarrow (f(u), \text{pr}(u - a))$ est \mathcal{C}^1 de E dans $F \times \ker df(a)$ de différentielle $d\tilde{h} = (df, \text{pr})$ inversible en a puis d'appliquer le théorème d'inversion locale.² □

2. La condition E_1 fermé (et donc Banach) peut paraître superflue. Elle intervient en fait de façon cachée pour avoir la continuité de la projection pr qui entre dans la différentielle de dh . On voit bien alors que c'est une condition nécessaire puisqu'alors $E_1 = \ker(I - \text{pr})$ est fermé et une application du théorème du graphe fermé permet de montrer que c'est suffisant. Notons que si E est un Hilbert, alors comme $\ker df(a)$

REMARQUE IV.5. *On remarque en particulier que $f : U \rightarrow V$ est alors surjective d'un voisinage U de a dans E sur un voisinage V de b de différentielle df surjective en tout point de U . C'est donc une submersion surjective locale.*

En revenant à notre problème, si on considère une trajectoire minimisante contrôlée par $u_* \in L^2([0, 1], B_e)$ alors sous la condition (C1), une application du corollaire IV.1 permet d'affirmer que $\delta u \rightarrow dE(u_*)\delta u$ n'est pas *surjective* car sinon E serait une submersion surjective locale ce qui contredit comme nous l'avons vu l'optimalité de u_* . On dit parfois que u_* est un *contrôle singulier*. En notant $\delta q(1) = q'_1(u; \delta u)$ et $\delta c = c'(u; \delta u)$ comme précédemment pour $\delta u \in L^2([0, 1], B_c)$, on a $dE(u)\delta u = (\delta q(1), \delta c)$. La non surjectivité de $dE(u_*)$ se traduit donc par l'existence de $\lambda_0 \in \mathbb{R}$ et $\lambda \in B_e^*$ tel que pour tout δu on a

$$\lambda_0 \delta c + (\lambda | \delta q(1)) = 0$$

Si on suppose maintenant que $\delta u \rightarrow \delta q(1)$ est surjective (ce qui nous dit par application du corollaire que localement on peut amener la solution au temps 1 n'importe où dans un voisinage suffisamment petit de $q_1(u_*)$) alors $\lambda_0 \neq 0$ et on peut se ramener au cas $\lambda_0 = 1$ i.e.

$$(36) \quad \delta c + (\lambda | \delta q(1)) = 0 \quad \forall \delta u \in B_c$$

En utilisant maintenant le théorème IV.2, on déduit facilement le théorème suivant.

THÉORÈME IV.6. *On suppose que les hypothèse (H_f) et (H_C) sont vérifiées et que de plus*

(H1) B_c est un espace de Hilbert

(H2) E est \mathcal{C}^1

(H3) q_1 est une \mathcal{C}^1 -submersion (i.e. \mathcal{C}^1 de différentielle partout surjective).

alors si u_* est une solution de (E') , il existe $\lambda \in B_e^*$ tel que la trajectoire optimale $q = \Phi(u_*)$ vérifie

$$\begin{cases} \dot{q} = \frac{\partial H}{\partial q}(q, p, u_*) \\ \dot{p} = -\frac{\partial H}{\partial p}(q, p, u_*) \\ \frac{\partial H}{\partial u}(q, p, u_*) \equiv 0 \\ p(1) + \lambda = 0 \end{cases}$$

REMARQUE IV.6. *Les conditions (H2) et (H3) pourraient évidemment être un peu relâchées en prenant localisant les conditions dans un voisinage de u_* . Cependant comme généralement u_* n'est pas connu à priori, les conditions demandées restent assez naturelles.*

est fermé on peut prendre pour E_1 son orthogonal qui est toujours fermé. Dans le cas de E Hilbert, il est donc suffisant de vérifier que $df(a)$ est surjective.

Chapitre V

Flots géodésiques et equations hamiltoniennes

1. Hamiltonien réduit

Une question intéressante est de savoir si on peut construire le flot hamiltonien associé à l'un de nos problèmes.

DÉFINITION V.1. Soient E et F deux espaces de Banach. On dit que $f \in \mathcal{C}_{loc}^{1, lip}(E, F)$ si $f \in \mathcal{C}^1(E, F)$ et pour tout $R > 0$, il existe $K_R > 0$ tel que pour tous x, y dans la boule $B_E(0, R)$ sur E on a

$$|df(x) - df(y)| \leq K_R |x - y|.$$

PROPOSITION V.1. Si E, F et G sont trois espaces de Banach, $f \in \mathcal{C}_{loc}^{1, lip}(E, F)$ et $g \in \mathcal{C}_{loc}^{1, lip}(F, G)$ alors $g \circ f \in \mathcal{C}_{loc}^{1, lip}(E, G)$

EXERCICE V.1. Ecrire la preuve de la proposition.

Dans notre situation d'étude, on considère

$$H(q, p, u) = (p|f(q, u)) - C(q, u)$$

et par suite si f et C sont dans $\mathcal{C}_{loc}^{1, lip}(\mathcal{B} \times B_c, B_e)$, comme $(p, \delta q) \mapsto (p|\delta q)$ est bilinéaire et dans $\mathcal{C}_{loc}^{1, lip}(B_e^* \times B_e, \mathbb{R})$, on déduit de la proposition que

$$H(q, p, u) \in \mathcal{C}_{loc}^{1, lip}(B_e^* \times B_e \times B_c, \mathbb{R}).$$

Notons que lorsque f et C vérifient (H_f) et (H_C) , f et C sont automatiquement dans $\mathcal{C}_{loc}^{1, lip}$.

Le cadre géométrique des problèmes (E) et (I) sur lesquels nous nous concentrons maintenant est celui dans lequel $f(q, u) = u \cdot q$ (où $u \cdot q$ est l'action infinitésimale de u sur q obtenue en différenciant en l'identité l'action du groupe $(g, q) \mapsto g \cdot q$) et $C(q, u) = \frac{1}{2}|u|_{B_c}^2$ où B_e est un espace de Hilbert. Comme la définition de la différentielle de l'action peut poser problème, on suppose donc comme nous l'avons fait jusqu'ici que l'on se *donne* l'action infinitésimale $(u, q) \mapsto u \cdot q$ qui est linéaire en u (c'est formellement la différentielle en

$g = \text{Id}$). On introduit donc le cadre (C_{riem}) :

$$(C_{\text{riem}}) \left| \begin{array}{l} B_c \text{ Hilbert} \\ C(q, u) = \frac{1}{2}|u|_{B_c}^2 \\ f(q, u) = u \cdot q \end{array} \right.$$

On suppose de plus que pour tout $(q, p) \in \mathcal{B} \times B_e^*$

$$u \mapsto (p|u \cdot q) \in B_c^*$$

i.e. est une forme linéaire continue¹ si bien que l'on peut définir une application notée $j : \mathcal{B} \times B_e^* \rightarrow B_c^*$ définie par

$$(j(q, p)|u)_{B_c^* \times B_c} = (p|u \cdot q)_{B_e^* \times B_e}.$$

On appellera cette application *l'application moment* (momentum map). Elle intervient immédiatement dès que l'on considère des trajectoires critiques : En effet, en écrivant $|u|_{B_c}^2 = (Lu|u)$ où $L : B_c \rightarrow B_c^*$ est l'isométrie classique entre B_c et on a pour $K = L^{-1}$

$$\frac{\partial H}{\partial u} = 0 \Rightarrow j - Lu = 0 \Rightarrow u = Kj$$

et donc

$$H_r(q, p) \triangleq \max_{u \in B_c} H(q, p, u) = H(q, p, Kj(q, p)) = \frac{1}{2}(Kj|j).$$

Si on suppose maintenant que $j \in \mathcal{C}_{\text{loc}}^{1, \text{lip}}(\mathcal{B} \times B_e^*, B_c^*)$ alors

$$H_r \in \mathcal{C}_{\text{loc}}^{1, \text{lip}}(\mathcal{B} \times B_e^*, \mathbb{R})$$

et on a existence et unicité locale de la solution du système

$$\left| \begin{array}{l} \frac{dq}{dt} = \frac{\partial H_r}{\partial p} \\ \frac{dp}{dt} = -\frac{\partial H_r}{\partial q} \end{array} \right.$$

pour toute condition initiale (q_0, p_0) dans la classe $\mathcal{C}^1([0, t_0[, \mathcal{B} \times B_e^*)$ pour un $t_0 > 0$. En particulier *la solution est maintenant \mathcal{C}^1 en temps* ce qui n'était pas évident au départ.

2. Existence globale en tout temps du flot hamiltonien

En fait, on peut relativement facilement obtenir l'existence en tout temps du flot sous des hypothèses assez simples de régularité :

1. Il suffit pour cela que $u \rightarrow f(q, u) = u \cdot q$ soit continue de B_c dans B_e . On remarquera que si f vérifie les conditions (H_f) , alors en utilisant le fait que $u \mapsto f(q, u)$ est linéaire, $|f(q, u)| \leq K(|q - b| + 1)|u|$ et donc les condition (H_f) entraîne la continuité.

THÉORÈME V.1. *On se place dans le cadre (C_{riem}) et on suppose f vérifie les conditions (H_f) et que $j \in \mathcal{C}_{loc}^{1,lip}$. Alors le flot hamiltonien construit sur H_r est défini en tout temps pour toute donnée initiale (q_0, p_0) .*

DÉMONSTRATION. Sous les hypothèses du théorème, on a H_r défini et $\mathcal{C}_{loc}^{1,lip}$ et donc existence et unicité locale du flot. On va montrer que si $[0, t_0[$ est l'intervalle de définition d'une solution maximale, alors $t_0 = +\infty$. Par l'absurde, on suppose que $t_0 < \infty$.

- (1) On commence par vérifier que H_r est constante sur $[0, t_0[$ (il suffit de dériver en temps $H_r(q(t), p(t))$ et de remarquer que $\frac{dH_r}{dt} = 0$)
- (2) On obtient alors que $H_r(q, p) = \frac{1}{2}(Kj|j) = \frac{1}{2}|u|^2$ avec $u = Kj$ c'est à dire que si $u(t) \triangleq Kj(q(t), p(t))$ alors la norme reste constante sur $[0, t_0[$. De plus, on a $\frac{\partial H}{\partial u}(q, p, u)(t) = 0$, $\frac{dp}{dt} = -\frac{\partial H}{\partial q}(q, p, u)$ et $\frac{dq}{dt} = \frac{\partial H}{\partial p}(q, p, u) = f(q, u)$. En utilisant le lemme de Grönwall et le fait (via (H_f)) que $|f(q, u)| \leq |f(b, 0)| + K(|q - b| + 1)(|u| + 1)$, on déduit que $|q - b|$ reste borné sur $[0, t_0[$. De plus,

$$\frac{dp}{dt} = - \underbrace{\frac{\partial f^*}{\partial q}(q, u)}_{\text{borné}} p + \underbrace{\frac{\partial C}{\partial q}(q, u)}_{=0}$$

est une équation linéaire avec des coefficients bornés d'où p reste bornée sur $[0, t_0[$. Ainsi il existe $R > 0$ tel que $|q| + |p| \leq R$ pour tout $t \in [0, t_0[$.

- (3) La contradiction vient du fait que puisque H_r est $\mathcal{C}_{loc}^{1,lip}$, il existe un $\epsilon > 0$ ne dépendant que de R tel que si $|q|(t) + |p|(t) \leq R$, alors la solution existe sur $[t, t + \epsilon[$.

□

En mettant tout ensemble on déduit dans le cadre (C_{riem}) :

THÉORÈME V.2. *On se place dans le cadre (C_{riem}) et on suppose f vérifie les conditions (H_f) et que $j \in \mathcal{C}_{loc}^{1,lip}$. Alors les trajectoires critiques $q = \Phi(u)$ au sens où il existe $p_1 \in B_e^*$ tel pour tout $\delta u \in L^2([0, 1], B_c)$ on ait*

$$\delta c - (p_1|\delta q_1) \triangleq c'(u_*; \delta u) - (p_1|\Phi'(u_*; \delta u)(1)) = 0 \text{ avec } c(u) \triangleq \int |u|_{B_e}^2 dt$$

vérifient

$$\left| \begin{array}{l} u = Kj(q, p) \\ \frac{dq}{dt} = \frac{\partial H}{\partial p}(q, p, u_*) = \frac{\partial H_r}{\partial p}(q, p) \\ \frac{dp}{dt} = -\frac{\partial H}{\partial q}(q, p, u_*) = -\frac{\partial H_r}{\partial q}(q, p) \end{array} \right.$$

et par unicité du flot sur le hamiltonien réduit H_r , on déduit que la trajectoire (q, p, u) est \mathcal{C}^1 en temps et prolongeable en tout temps. De plus $|u| = Cte$.

Nous allons maintenant traiter trois exemples fondamentaux.

3. Cas des landmarks

Dans ce cas, on prend $q = (x_1, \dots, x_k) \in \mathbb{R}^{dk} = \mathcal{B}$ muni de la norme euclidienne et pour tout B_c un ENR admissible V et l'action infinitésimale est définie par

$$f(q, v) = v \cdot q = (v(x_1), \dots, v(x_k)).$$

On vérifie que $|(p|v \cdot q)| \leq |p||v|_\infty \leq K|p||v|_V$ si bien que $v \mapsto (p|v \cdot q)$ est une forme linéaire continue sur B_c et l'application moment est bien définie :

$$j(q, p) \triangleq \sum_{i=1}^k \delta_{x_i}^{p_i}$$

où $(\delta_{x_i}^{p_i}|v) = \langle v(x_i), p_i \rangle_{\mathbb{R}^d}$.

La régularité de l'application moment dépend en fait de la régularité des éléments de V .

PROPOSITION V.2. *Supposons que $V \xrightarrow{cont.} C_b^2(\mathbb{R}^d, \mathbb{R}^d)$ (injection continue). Alors*

$$(x, \alpha) \mapsto \eta(x, \alpha) \triangleq \delta_x^\alpha$$

est dans $\mathcal{C}_{loc}^{1, lip}(\mathbb{R}^d \times \mathbb{R}^d, V^)$.*

DÉMONSTRATION. Pour démontrer la proposition V.2 nous allons utiliser le lemme suivant.

LEMME V.1. *Soient E et F deux Banachs et $h : E \rightarrow F$. On suppose que pour tout $x \in E$, h est Gâteaux différentiable en x et qu'il existe une application linéaire continue $A(x) : E \rightarrow F$ telle que $h'(x; \delta x) = A(x)\delta x$ pour tout $\delta x \in E$. Alors, si $x \rightarrow A(x)$ est continue en x pour tout $x \in E$, $h \in \mathcal{C}^1(E, F)$ et $dh(x) = A(x)$.*

DÉMONSTRATION. Il suffit de vérifier que h est dérivable et que $h'(x) = A(x)$ pour tout $x \in E$. En effet, $h(x + \delta x) - h(x) - A(x)\delta x = \int_0^1 h'(x + t\delta x; \delta x) - A(x)\delta x dt = \int_0^1 (A(x + t\delta x) - A(x))\delta x dt$. Par suite $|h(x + \delta x) - h(x) - A(x)\delta x| \leq |\delta x| \int_0^1 |A(x + t\delta x) - A(x)| dt$. Comme $\int_0^1 |A(x + t\delta x) - A(x)| dt \rightarrow 0$ lorsque $\delta x \rightarrow 0$ puisque A est continue, on a le résultat annoncé. \square

Revenons à la preuve de la proposition. On commence par montrer que η est Gâteaux différentiable : En effet

$$t \mapsto (\eta(x + t\delta x, \alpha + t\delta\alpha)|v) = (\alpha + t\delta\alpha|v(x + t\delta x))$$

est dérivable en 0 de dérivée $(\alpha|dv(x).\delta x) + (\delta\alpha|v(x))$. Or

$$|(\alpha|dv(x).\delta x) + (\delta\alpha|v(x))| \leq K|\alpha||v|_V|\delta x| + K|\delta\alpha||v|_V$$

et donc

$$v \mapsto (\alpha|dv(x).\delta x) + (\delta\alpha|v(x)) \in V^*.$$

Notons $A(x, \alpha).(\delta x, \delta\alpha)$ l'unique élément de V^* tel que

$$(A(x, \alpha).(\delta x, \delta\alpha)|v) = (\alpha|dv(x).\delta x) + (\delta\alpha|v(x)).$$

On remarque que $(\delta x, \delta\alpha) \rightarrow A(x, \alpha).(\delta x, \delta\alpha)$ est linéaire continue puisque $|(A(x, \alpha).(\delta x, \delta\alpha)|v)| \leq K|v|_V(|\alpha||\delta x| + |\delta\alpha|)$ et donc $|A(x, \alpha).(\delta x, \delta\alpha)|_{V^*} \leq K(|\alpha||\delta x| + |\delta\alpha|)$. Il nous faut montrer maintenant que $\eta(x + t\delta x, \alpha + t\delta\alpha) - \eta(x, \alpha) - A(x, \alpha).(\delta x, \delta\alpha)|_{V^*} = o(|t|)$ pour montrer que η est Gâteaux différentiable et identifier sa dérivée avec A . Or pour tout $v \in V$

$$\begin{aligned} & |(\eta(x + t\delta x, \alpha + t\delta\alpha) - \eta(x, \alpha) - A(x, \alpha).(\delta x, \delta\alpha)|v)| \\ &= |(\alpha|v(x + t\delta x) - v(x) - t dv(x).\delta x) + t(\delta\alpha|v(x + t\delta x) - v(x))| \\ &\leq t^2|d^2v|_\infty|\delta x||\alpha| + t^2|dv|_\infty|\delta\alpha||\delta x| \\ &\leq Kt^2|v|_V(|\delta x||\alpha| + |\delta\alpha||\delta x|) \end{aligned}$$

où nous utilisons l'injection continue dans \mathcal{C}_b^2 pour la dernière inégalité. Nous avons donc montré à ce stade que η est Gâteaux différentiable de dérivée A . Il nous suffit de vérifier que A est localement lipschitzienne. Or

$$\begin{aligned} & |((A(x', \alpha') - A(x, \alpha)).(\delta x, \delta\alpha)|v)| \\ &\leq |\alpha' - \alpha||d^2v|_\infty|\delta x| + |\alpha||d^2v|_\infty|\delta x||x' - x| + |\delta\alpha||dv|_\infty|x' - x| \\ &\leq K|v|_V(1 + |\alpha|)(|\alpha' - \alpha| + |x' - x|)(|\delta\alpha| + |\delta x|). \end{aligned}$$

Par suite $|A(x', \alpha') - A(x, \alpha)| \leq \text{Cte}(1 + |\alpha|)(|\alpha' - \alpha| + |x' - x|)$ ce qui donne A localement lipschitzien. \square

EXERCICE V.2.

- (1) Montrer que si K est le noyau associé à l'ENR V alors le hamiltonien réduit $H_r : (\mathbb{R}^d)^k \times (\mathbb{R}^d)^k \rightarrow \mathbb{R}$ est défini par

$$H_r(q, p) = \frac{1}{2} \sum_{i,j=1}^k p_i^T K(x_i, x_j) p_j.$$

- (2) Que peut-on dire de la régularité de K si $V \xrightarrow{\text{cont.}} C_b^2(\mathbb{R}^d, \mathbb{R}^d)$.
- (3) Ecrire les équations hamiltoniennes.

4. Cas des courbes

Supposons ici que q est une courbe identifiée à un élément de $\mathcal{C}([0, 1], \mathbb{R}^d)$ si bien que l'on peut prendre ici $\mathcal{B} = B_e = \mathcal{C}([0, 1], \mathbb{R}^d)$ et pour B_e^* les mesures boréliennes finies vectorielles. On a alors $f(q, v) \triangleq v \cdot q = v \circ q$ et $|v \circ q| \leq K|v|_V$ d'où

$$|(p|v \cdot q)| \leq |p|_{B_e^*} K|v|_V$$

et l'application moment $j : \mathcal{B} \times B_e^* \rightarrow V^*$ est bien définie.

Commençons par vérifier que $f \in \mathcal{C}_{\text{loc}}^{1, \text{lip}}(\mathcal{B} \times V, B_e)$. En effet pour tout $q, q' \in \mathcal{B}$, on a

$$\begin{aligned} & |v \circ q' - v \circ q - dv(q) \cdot (q' - q)| \\ & \leq \int_0^1 |(dv((1-h)q + hq')) - dv(q))(q' - q)| dh \\ & \leq |d^2 v|_{\infty} |q' - q|^2 \leq K|v|_V |q' - q|^2, \end{aligned}$$

d'où l'on déduit que $q \rightarrow v \circ q$ est différentiable de différentielle $dv \circ q$ que l'on reconnaît comme $\partial_q f(q, v)$ ². De plus $|\partial_q f(q', v') - \partial_q f(q, v)| \leq K(|v' - v| + |v||q' - q|)$. Comme par ailleurs $v \rightarrow v \circ q$ est linéaire que $|\delta v \circ q - \delta v \circ q'| \leq K|\delta v|_V |q' - q|$ on déduit aisément que $\partial_v f$ existe et est localement lipschitzienne ce qui donne au final $f \in \mathcal{C}_{\text{loc}}^{1, \text{lip}}(\mathcal{B} \times V, B_e)$.

Montrons maintenant que $j \in \mathcal{C}_{\text{loc}}^{1, \text{lip}}(\mathcal{B} \times B_e^*, V^*)$. On a $|(j(q', p') - j(q, p)|v) - (p|dv(q) \cdot q' - q) - (p' - p|v(q))| = |(p'|v(q')) - (p|v(q)) - (p|dv(q) \cdot q' - q) - (p' - p|v(q))| = |(p' - p|v(q') - v(q)) + (p|v(q') - v(q) - dv(q) \cdot (q' - q))| \leq K|v|_V |p' - p| |q' - q| + K|p| |v|_V |q' - q|^2$. Par suite j est différentiable de différentielle dj définie par

$$(dj(q, p) \cdot (\delta q, \delta p)|v) = (\delta p, v \circ q) + (p|dv \circ q \cdot \delta q)$$

EXERCICE V.3. Vérifier que dj est localement lipschitzienne puis conclure que $j \in \mathcal{C}_{\text{loc}}^{1, \text{lip}}(\mathcal{B} \times B_e^*, V^*)$.

Nous pouvons donc utiliser le théorème V.2 et déduit l'existence du flot hamiltonien en tout temps pour toute données initiale (q_0, p_0) . La courbe $q(t)$ est alors une courbe dans $\mathcal{C}([0, 1], \mathcal{B})$ (courbe de courbes!) et la variable adjointe $p(t)$, souvent appelé dans notre contexte *le moment*, est une mesure vectorielle finie sur $[0, 1]$ à valeur dans \mathbb{R}^d . On peut étendre sans difficulté ce résultat aux surfaces.

2. On utilisera parfois la notation $\partial_q f$ au lieu de $\frac{\partial f}{\partial q}$ pour désigner une dérivée partielle

4.1. Régularité de la variable adjointe. A priori $p(t)$ est une mesure vectorielle. Cela arrive par exemple avec de problème de matching inexact d'une courbe initiale sur un nombre fini de points cibles avec une attache aux données g du type :

$$g(q_1) \triangleq \sum_{i=1}^n |q_1(s_i) - y_i|^2$$

qui est \mathcal{C}^1 avec

$$g'(q_1)|\delta q_1 = \sum_{i=1}^n \langle q(s_i) - y_i, \delta q_1(s_i) \rangle = - \int \langle \delta q_1(s), dp_1(s) \rangle$$

où

$$p_1 = \sum_{i=1}^n \delta_{t_i}^{q_1(s_i) - y_1}$$

est la mesure vectorielle discrète que nous avons vue comme un élément de V^* mais qui s'interprète ici comme un élément de $\mathcal{C}([0, 1], \mathbb{R}^d)^*$ somme de masses de Dirac vectorielle. On s'est donc ramené au cas des landmarks.

Cependant de façon plus générale, si (q, p) est solution du système hamiltonien réduit et $u = Kj$ alors p est solution de l'EDO linéaire $\dot{p} = -\partial_q f^*(q, u)p$. Or on vérifie facilement que $p \mapsto -\partial_q f^*(q, u)p$ est un opérateur linéaire bornée dans L^1 puisque

$$(\partial_q f^*(q, v)p|\delta q) = (p|dv \circ q.\delta q) = \int \langle p(s), dv \circ q(s).\delta q(s) \rangle ds = \int \langle dv^* \circ q.p, \delta q(s) \rangle ds$$

et

$$|\partial_q f^*(q, v)p|_{L^1} \leq K|v|_V|p|_{L^1}$$

i.e. $\partial_q f^* \in \text{End}(L^1)$ et donc la solution reste dans L^1 . La même propriété reste vraie pour tout les L^p ou encore dans $\mathcal{C}([0, 1], \mathbb{R}^d)$. En particulier, si la cible est une courbe $y(s)$ continue et que

$$g(q_1) \triangleq \frac{1}{2} \int_0^1 |q_1(s) - y(s)|^2 ds$$

alors $g'(q_1; \delta q_1) = \int \langle q_1(s) - y(s), \delta q_1(s) \rangle ds$

EXERCICE V.4. Montrer que g est \mathcal{C}^1 sur $\mathcal{C}([0, 1], \mathbb{R}^d)$ de dérivée un élément de $\mathcal{C}([0, 1], \mathbb{R}^d)$.

En utilisant l'exercice, on déduit que $p_1 \in \mathcal{C}([0, 1], \mathbb{R}^d)$ et le moment reste continu. En particulier le moment initial est continu. On aura en tout temps

$$v_t = Kj(q_t, p_t)$$

si bien que

$$\langle v_t(x), \alpha \rangle = (\delta_x^\alpha | v) = (\delta_x^\alpha | K j) = (j | K \delta_x^\alpha) = (p_t | K \delta_x^\alpha \circ q_t) = \int \langle p_t(s), K(q_t(s), x) \alpha \rangle ds$$

et

$$v_t(x) = \int p_t(s)^T K(q_t(s), x) ds.$$

On a donc au final une expression explicite de $v_t = K j$ dès que p et q sont connus.

EXERCICE V.5. Vérifier que si $q, p \in \mathcal{C}([0, 1], \mathbb{R}^d)$, alors le hamiltonien réduit est donné par

$$H_r(q, p) = \frac{1}{2} \int_0^1 \int_0^1 p^T(r) K(q(r), q(s)) p(s) dr ds$$

puis écrire les équations hamiltoniennes.

5. Cas des difféomorphismes

Ici on considère que $q = \varphi \in \mathcal{C}_{\text{Id}}(\mathbb{R}^d, \mathbb{R}^d) = \mathcal{B}$, $B_e = \mathcal{C}_b(\mathbb{R}^d, \mathbb{R}^d)$ et que $V \xrightarrow{\text{cont.}} C_b^2(\mathbb{R}^d, \mathbb{R}^d)$.

Jusqu'à présent, nous avons seulement montré que les éléments de G_V sont des homéomorphismes. Nous allons ici commencé par montrer qu'il s'agit en fait de difféomorphismes.

THÉORÈME V.3. Si $V \xrightarrow{\text{cont.}} C_b^1(\mathbb{R}^d, \mathbb{R}^d)$ alors pour tout $v \in L^1([0, 1], V)$ la solution de

$$q_t = \text{Id} + \int_0^t v_s \circ q_s ds$$

définit un \mathcal{C}^1 difféomorphisme pour lequel on a

$$dq_t = \text{Id} + \int_0^t dv_s \circ q_s dq_s ds.$$

DÉMONSTRATION. On commence par remarquer que

$$q_t(x) = x + \int_0^t v_s(q_s(x)) ds$$

et que si on considère maintenant les perturbations en la condition initiale de cette EDO sur \mathbb{R}^d , alors en posant $f(x, v) = v(x)$ on a $|\partial_x f| \leq K|v|_V$ et reprenant la preuve du théorème IV.1 dans le cas où on ne considère que des perturbations en la condition initiale (ici x) et non en le contrôle (ici v) on obtient en considérant la perturbation $x + sh$ que q_t admet une dérivée directionnelle dans la direction h et que

$$q'_t(x; h) = h + \int_0^t dv_s(q_s(x)) \cdot q'_s(x; h) dx.$$

Par suite, si on considère l'équation différentielle linéaire à valeurs dans $\mathcal{M}_d(\mathbb{R})$

$$A_t(x) = \text{Id} + \int_0^t dv_s(q_s(x))A_s(x)ds$$

on montre facilement l'existence et l'unicité de la solution (par la méthode itérative classique) puis que $x \mapsto A_t(x)$ est continue en x . En effet en écrivant

$$A_t(x') - A_t(x) = \int_0^t dv_s(q_s(x'))(A_s(x') - A_s(x))ds + \int_0^t (dv_s(q_s(x')) - dv_s(q_s(x)))A_s(x)dx$$

on obtient par une application du lemme de Grönwall

$$\sup_t |A_t(x') - A_t(x)| \leq \exp(K \int_0^1 |v_s|_V ds) M(x, x')$$

où $M(x, x') = \int_0^1 |dv_s(q_s(x')) - dv_s(q_s(x))| dx$. Or par convergence dominée, $M(x, x') \rightarrow 0$ lorsque $x' \rightarrow x$. En utilisant le lemme V.1, on déduit que q_t est \mathcal{C}^1 et donc un \mathcal{C}^1 difféomorphisme (puisque son inverse est aussi \mathcal{C}^1 comme solution du flot retourné). De plus on a

$$dq_t(x) = \text{Id} + \int_0^t dv_s(q_s(x))dq_s(x)ds$$

□

Revenons à notre discussion. Puisque $B_e = \mathcal{C}_b(\mathbb{R}^d, \mathbb{R}^d)$, l'espace dual $B_e^* = \mathcal{M}_f(\mathbb{R}^b, \mathbb{R}^d)$ est l'ensemble des mesures boréliennes vectorielles finies sur \mathbb{R}^d . On a $f(q, v) \triangleq v \circ q$, et comme dans le cas précédent on montre que f , j et C sont $\mathcal{C}_{\text{loc}}^{1, \text{lip}}$ si bien que l'on a existence des géodésiques en tout temps pour toute condition initiale (q_0, p_0) avec p_0 mesure vectorielle finie.

5.1. Régularité du moment. Supposons que l'on considère le problème du matching inexact et une attache aux données g définie par :

$$g(q_1) \triangleq \int |I_{\text{targ}}(x) - I_0 \circ q_1^{-1}(x)|^2 dx$$

où I_{targ} est une image continue à support compact, I_0 une image \mathcal{C}_b^1 à support compact.

On vérifie que si \tilde{q}_s est une perturbation d'un \mathcal{C}^1 difféomorphisme $q \in \mathcal{B}$ tangente à δq et telle que \tilde{q}_s est inversible pour tout s , alors pour tout compact K , on a

$$\sup_K |\tilde{q}_s^{-1} - q^{-1} + s dq^{-1} \delta q \circ q^{-1}|(x) = o(|s|).$$

et donc \tilde{q}_s^{-1} est une perturbation de q^{-1} tangente à $-dq^{-1} \delta q \circ q^{-1}$ pour la norme de la convergence uniforme sur tout compact. Par suite, on a par convergence dominée que

$s \mapsto h(s) \triangleq g(\tilde{q}_s)$ est dérivable en 0, de dérivée

$$\begin{aligned} h'(0) &= \int \langle I_{\text{targ}}(y) - I_0 \circ q^{-1}(y), dI_0 \circ q^{-1} dq^{-1} \delta q \circ q^{-1}(y) \rangle dy \\ &= \int \langle I_{\text{targ}} \circ q - I_0 \rangle \nabla(I_0 \circ q^{-1}) \circ q |Jq|(x), \delta q(x) dx \end{aligned}$$

où nous avons fait le changement de variable $y = q(x)$. Ainsi on vérifie aisément que les solutions du problème de matching inexact correspondant au temps 1 à un multiplicateur de Lagrange p_1 qui est dans L^1 et que l'on peut identifier avec la fonction

$$p_1(x) = -(I_{\text{targ}} \circ q_1 - I_0) \nabla(I_0 \circ q_1^{-1}) \circ q_1 |Jq_1|(x)$$

Comme précédemment, puisque $p \rightarrow -\partial_q f^*(q, u)p$ est un opérateur bornée dans L^1 , p_t reste dans L^1 et en particulier p_0 . On obtient donc ici que les géodésiques qui sortent des problèmes de matching avec une template régulière sont en fait non pas des mesures mais de « bonnes » fonctions L^1 . Ces géodésiques correspondent au relèvement « horizontal » dans G_V de géodésiques pour la métrique induite dans l'orbite $G_V \cdot I_0$.

6. Méthode de gradient trajectoriel et de tir pour les landmarks

Dans les applications actuelles, le problème des appariements de landmarks joue un rôle central car beaucoup de problèmes peuvent s'y ramener via une discrétisation adéquate de l'attache aux données. Dans cette partie nous allons présenter deux méthodes de construction d'un appariement optimal : l'une introduite dans [9] qui est une descente de gradient sur la trajectoire temporelle des variables adjointes, l'autre est construite sur une méthode dite de tir et utilise l'équation géodésique pour ajuster le choix des moments initiaux via une descente de gradient [14].

6.1. Descente de gradient sur la trajectoire. En introduisant le hamiltonien réduite H_r dans le cas des landmarks, les trajectoires géodésiques sont solutions du système

$$\dot{q} = \frac{\partial H_r}{\partial p}, \quad \dot{p} = -\frac{\partial H_r}{\partial q}$$

où $H_r(q, p) = \sum_{i,j} p_i^T K(q_i, q_j) p_j$. On sait également que la vitesse instantanée pour la reconstruction de l'appariement dense vérifie $v = Kj(q, p)$ ou de façon plus explicite $v(x) = \sum_i K(x, x_i) p_i$ et qu'enfin $\frac{1}{2} |v|_V^2 = H(q, p)$.

On pourrait donc essayer d'utiliser la réduction du nombre de paramètre pour la représentation de v via l'égalité $v = Kj(q, p)$ pour construire une méthode de gradient non plus sur la trajectoire $t \mapsto v_t$ mais sur celle de $t \mapsto \alpha_t$ en imposant $v_t = Kj(q, \alpha)$ ce qui permet une réduction importante de la complexité.

On considère donc sur $\mathcal{C}([0, 1], (\mathbb{R}^d)^k)$ le problème de minimisation

$$(37) \quad \left| \begin{array}{lcl} \min J(\alpha) & = & \int H_r(q, \alpha) dt + g(q_1) \\ \dot{q} & = & \frac{\partial H_r}{\partial \alpha}(q, \alpha) \\ q(0) & = & q_0 \end{array} \right.$$

Si on suppose que $V \xrightarrow{\text{cont.}} C_b^2(\mathbb{R}^d, \mathbb{R}^d)$ alors H_r est \mathcal{C}^2 et

$$J'(\alpha; \delta\alpha) = - \int_0^1 \left(\frac{\partial \tilde{H}}{\partial \alpha}(q_t, \alpha_t, p_t) | \delta\alpha \right) dt,$$

où $\tilde{H}(q, p, \alpha) = (p | \frac{\partial H_r}{\partial \alpha}(q, \alpha)) - H_r(q, \alpha)$, $\frac{dq}{dq}(q_1) + p_1 = 0$ et $\dot{p} = -\frac{\partial \tilde{H}}{\partial q}(q, p, \alpha) = -(dv_{j(q,p)}^*(q) \cdot \alpha + dv_{j(q,\alpha)}^*(q) \cdot (p - \alpha))$ avec $v_j \triangleq K j$ ³.

En effet il suffit d'appliquer le théorème IV.3 (dont on vérifie les hypothèses) dans la nouvelle situation où le contrôle u est α , $f(q, \alpha) = \frac{\partial H_r}{\partial \alpha}(q, \alpha)$.

Or on obtient facilement que

$$\frac{\partial \tilde{H}}{\partial \alpha}(q, p, \alpha) = \mathbb{K}(q, q)(p - \alpha)$$

où $\mathbb{K}(q, q)$ est la matrice blocs $(K(q_i, q_j))_{i,j}$. Comme $\mathbb{K}(q, q)$ est une matrice définie positive, on déduit que $p - \alpha$ est une direction de descente puisque

$$J'(\alpha; p - \alpha) = - \int_0^1 (p_t - \alpha_t)^T \mathbb{K}(q_t, q_t) (p_t - \alpha_t) dt = - \int_0^1 |v_{j(q_t, p_t - \alpha_t)}|_V^2 dt \leq 0.$$

6.1.1. *Algorithme de descente.* On peut donc construire un algorithme avant-arrière de la façon suivante :

- (1) (Intégration avant) Partant d'une trajectoire $t \mapsto \alpha_t^n$, on construit une trajectoire q_t^n en intégrant $\dot{q}^n = \mathbb{K}(q_t^n, q_t^n) \alpha_t^n$ pour obtenir la position finale q_1^n des landmarks.
- (2) (Calcul du critère à l'extrémité) On calcule la dérivée de g en q_1^n et on initialise p_1^n par $p_1^n + \frac{dg}{dq}(q_1^n) = 0$.
- (3) (Intégration arrière) On calcule la trajectoire p_t^n en intégrant en arrière l'équation $\dot{p}_t^n = -\frac{\partial \tilde{H}}{\partial q}(q_t^n, p_t^n, \alpha_t^n)$.
- (4) (Pas de gradient) On met à jour la trajectoire des α par $\alpha^{n+1} = \alpha^n + \Delta t_n (p^n - \alpha^n)$.

6.2. Méthode de tir. Dans la méthode précédente, il n'y a pas de garantie tant que l'algorithme n'a pas convergé que la trajectoire courante soit une géodésique : En effet, la condition pour être une géodésique est en fait que l'on est $p = \alpha^4$ ce qui n'est atteint qu'à convergence. En particulier, si on reconstruit la géodésique partant de la condition initiale α_0 en intégrant le hamiltonien réduit $H_r(q, \alpha)$ i.e. en résolvant $\dot{q} = \partial_\alpha H_r(q, \alpha)$,

3. En effet, si

4. On vérifie par exemple que $\partial_p \tilde{H}(q, \alpha, \alpha) = \partial_\alpha H_r(q, \alpha)$

$\dot{\alpha} = -\partial_q H_r(q, \alpha)$, alors on peut trouver une différence significative avant convergence avec la trajectoire courante de l'algorithme de descente sur la trajectoire⁵. L'idée alternative est de contraindre la recherche sur l'espace des géodésiques et en optimisant la vitesse initiale ce qui revient dans le langage de la géométrie riemannienne à s'intéresser aux champs de Jacobi. On considère donc la minimisation de

$$(38) \quad \left| \begin{array}{lcl} \min J(\alpha_0) & = & H_r(q_0, \alpha_0) + g(q_1) \\ \dot{q} & = & \frac{\partial H_r}{\partial \alpha}(q, \alpha) \\ \dot{\alpha} & = & -\frac{\partial H_r}{\partial q}(q, \alpha) \\ q(0) & = & q_0 \end{array} \right.$$

On voit bien ici que nous retrouvons une situation particulièrement simple de contrôle optimal dans lequel... *il n'y a pas de contrôle* du type

$$(39) \quad \left| \begin{array}{lcl} \min \mathcal{J}(s_0) & = & g_0(s_0) + g_1(s_1) \\ \dot{s} & = & F(s) \end{array} \right.$$

pour laquelle en introduisant le hamiltonien $\mathcal{H}(s, \nu) = (\nu | F(s))$, on obtient par application du théorème IV.3 que

$$J'(s_0; \delta s_0) = \left(\frac{dg_0}{ds}(s_0) - \nu_0 | \delta s_0 \right)$$

où $\nu_1 + \frac{dg_1}{ds}(s_1) = 0$ et ν_0 est obtenu par l'intégration en arrière de l'équation $\dot{\nu} = -\partial_s \mathcal{H}(s, \nu) = -dF(s)^* \nu$. Pour le cas précédent, il suffit de poser $s = (q, \alpha)$,

$$s = (q, \alpha), \quad F(s) = \begin{pmatrix} \partial_\alpha H_r(q, \alpha) \\ -\partial_q H_r(q, \alpha) \end{pmatrix}, \quad g_1(s_1) = g(q_1), \quad g_0(s_0) = H_r(q_0, \alpha_0).$$

En particulier comme g_1 ne dépend que de q_1 et non de α_1 , on aura donc pour $\nu = (\nu^q, \nu^\alpha)$ que $\nu_1^\alpha = 0$, $\nu_1^q = \frac{dg}{dq}(q_1)$ et $J'(\alpha_0; \delta \alpha_0) = (\mathbb{K}(q_0, q_0) \alpha_0 - \nu_0^\alpha | \delta \alpha_0)$.

Le coût de cette approche est que l'on doit augmenter la taille de système et calculer le linéarisé i.e. construire la matrice dF . Cependant ceci fait, on peut travailler complètement avec les moments initiaux α_0 en garantissant à chaque étape de travailler sur des géodésiques. On peut de la même manière regarder des situations plus riche comme celle présentée par Stanley Durrleman [7] en introduisant des marqueurs passifs $x = (x_j)$ dans la dynamique : on pose alors $s = (q, x, \alpha)$ avec $\dot{q} = \partial_\alpha H_r(q, \alpha) = \mathbb{K}(q, q)\alpha$, $\dot{x} = v_{j(q, \alpha)} \cdot x = \mathbb{K}(x, q)\alpha$ ⁶, $\dot{\alpha} = -\partial_q H_r(q, \alpha)$ et procéder de façon similaire.

5. Cela pose des problème si l'on veut faire ensuite des statistiques dans l'espace tangent du type "tangent PCA" construites sur la distribution des α_0

6. Les marqueurs passifs sont entraînés par le flot de v

Chapitre VI

Continuité faible du flot et théorèmes d'existence

1. Complétude de G_V

Nous avons vu dans le chapitre III un certain nombre de propriétés du groupe G_V . Tout d'abord sa construction par l'intégration des flots des champs de vecteurs $u \in L^1([0, 1], V)$ via l'application $\Phi : L^1([0, 1], V) \rightarrow \mathcal{C}([0, 1], \mathcal{C}_{\text{Id}}(\mathbb{R}^d, \mathbb{R}^d))$:

$$G_V = \{ \Phi_1^v \mid v \in L^1([0, 1], V) \}$$

pour un espace V admissible ($V \hookrightarrow \mathcal{C}_b^1(\mathbb{R}^d, \mathbb{R}^d)$ cf Définition III.2). Nous avons montré que G_V est un groupe, puis que

$$d_G(\text{Id}, \varphi) \triangleq \{ \|v\|_1 \mid v \in L^1([0, 1], V), \Phi_1^v = \varphi \}$$

définit une distance invariante à droite sur G_V . Nous allons prouver que le groupe G_V est *complet* pour la distance d_G .

THÉORÈME VI.1. *Le groupe G_V est complet pour d_V .*

DÉMONSTRATION. Soit donc $(\varphi_n)_{n \in \mathbb{N}}$ une suite de Cauchy. Nous pouvons quitte à extraire une sous-suite supposer que $\sum_{n \geq 0} d_G(\varphi_n, \varphi_{n+1}) < \infty$. On choisit maintenant pour tout $n \geq 0$ un $v_n \in L_V^1$ tel que $\varphi^{v_n} \circ \varphi_n = \varphi_{n+1}$ et $\|v_n\|_1 \leq 2d_G(\varphi_n, \varphi_{n+1})$.

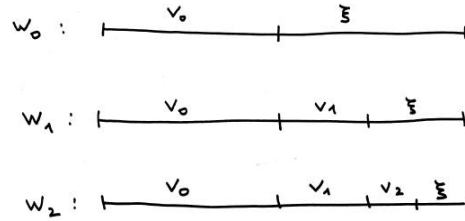


FIGURE 1. Définition des w_n

On note $\xi \equiv 0$ le champ de L_V^1 contenant égal à 0, puis $w_n = (\cdots (\xi \star v_n) \star v_{n-1}) \cdots \star v_0$ pour tout $n \geq 0$. On vérifie que $\varphi_{n+1} = \varphi_1^{w_n} \circ \varphi_0$ et $\|w_n\|_1 = \sum_{k=0}^n \|v_k\|_1 < 2 \sum_{k=0}^{\infty} d_G(\varphi_k, \varphi_{k+1}) < \infty$.

∞ . Comme par ailleurs $\|w_{n+1} - w_n\|_1 = \|v_n\|_1$ on déduit que $(w_n)_{n \geq 0}$ est de Cauchy dans L_V^1 qui est un Banach et donc converge vers $w_\infty \in L_V^1$. En notant $\varphi_\infty = \varphi_1^{w_\infty} \circ \varphi_0 \in G_V$ on vérifie que $\varphi_\infty = \varphi^{w_\infty - w_{n-1}} \circ \varphi_n$ d'où $d_G(\varphi_\infty, \varphi_n) \leq \|w_\infty - w_n\|_1 \rightarrow 0$ lorsque $n \rightarrow \infty$. \square

2. Continuité faible du flot

La complétude de G_V assure déjà que l'on retrouve en partie pour le groupe, une propriété importante de l'espace V qui sert de modèle à l'espace tangent en l'identité : la complétude. Nous allons plus loin maintenant en montrant que l'on peut retrouver de la compacité pour une topologie faible.

DÉFINITION VI.1. On définit $L_V^2 \triangleq \{ v \in L_V^1 \mid \int_0^1 |v_s|_V ds < \infty \}$. Sur L_V^2 on définit $\|v\|_2 \triangleq (\int_0^1 |v_s|_V^2 ds)^{1/2}$ qui fait de L_V^2 un Hilbert.

REMARQUE VI.1. L_V^2 est séparable dès que V l'est. On se placera dorénavant dans ce cas.

On rappelle ici un résultat démontré dans le chapitre III (Proposition III.4)

PROPOSITION VI.1. On a $G_V = \{ \Phi_1^v \mid v \in L^2([0, 1], V) \}$ et

$$d_G(\text{Id}, \varphi) = \inf \{ \|v\|_2 \mid v \in L^2([0, 1], V), \Phi_1^v = \varphi \}$$

$$\text{où } \|v\|_2 = \left(\int_0^1 |v_t|_V^2 dt \right)^{1/2}.$$

REMARQUE VI.2. On remarque au passage qu'un corollaire trivial de la proposition dit que pour tout φ et $\varphi' \in G_V$, il existe un champ de vitesse dépendant du temps $v \in L_V^2$ (i.e. d'énergie finie) tel que $\varphi' = \varphi_1^v \circ \varphi$.

THÉORÈME VI.2. L'application flot $\varphi : L_V^2 \rightarrow \mathcal{C}_{\text{Id}}([0, 1] \times \mathbb{R}^d, \mathbb{R}^d)$ restreinte à L_V^2 est continue pour la topologie faible sur L_V^2 et la norme de la convergence uniforme sur les compacts de $[0, 1] \times \mathbb{R}^d$ sur $\mathcal{C}_{\text{Id}}([0, 1] \times \mathbb{R}^d, \mathbb{R}^d)$.

DÉMONSTRATION. Soit $(v_n)_{n \geq 0}$ une suite de L_V^2 convergeant faiblement vers v . Rappelons que par le théorème de Banach-Steinhaus, on a $\sup_n \|v_n\| < \infty$. De plus, on a

$$(40) \quad \varphi^{v_n}(t, x) - \varphi^v(t, x) = \int_0^t (v_n - v)(s, \varphi^v(s, x)) ds + \int_0^t v_n(s, \varphi^{v_n}(s, x)) - v_n(s, \varphi^v(s, x)) ds.$$

Fixons un instant x et notons $\epsilon_n(t) \triangleq \int_0^t (v_n - v)(s, \varphi^v(s, x)) ds$. On a pour tous $0 \leq t \leq t' \leq 1$,

$$|\epsilon_n(t) - \epsilon_n(t')| \leq \int_t^{t'} |(v_n - v)(s, \cdot)|_\infty ds \stackrel{\text{C.S.}}{\leq} C(t' - t)^{1/2} \|v_n - v\|_2 \leq C(t' - t)^{1/2} \sup_{m \geq 0} \|v_m - v\|_2,$$

d'où la suite ϵ_n est équicontinue en t . Or comme $w \rightarrow \int_0^t w(s, \varphi^v(s, x)) ds$ est une forme linéaire continue sur L_V^2 , on déduit que ϵ_n tend simplement vers 0 en n et donc par le théorème d'Ascoli, elle tend uniformément vers 0 pour $t \in [0, 1]$.

En utilisant le lemme de Grönwall à l'inégalité $|\varphi^{v_n}(t, x) - \varphi^v(t, x)| \leq |\epsilon_n|_\infty + \int_0^t |\partial_2 v_n(s, \cdot)|_\infty |\varphi^{v_n}(s, x) - \varphi^v(s, x)| ds$ tirée de (40), on obtient

$$|\varphi^{v_n}(t, x) - \varphi^v(t, x)| \leq |\epsilon_n|_\infty \exp\left(\int_0^1 |\partial_2 v_n(s, \cdot)|_\infty\right) \leq |\epsilon_n|_\infty \exp(K\|v_n\|_2) \rightarrow 0$$

car $\sup_{n \geq 0} \|v_n\|_2 < \infty$ et $|\epsilon_n|_\infty \rightarrow 0$ lorsque $n \rightarrow \infty$ par convergence uniforme.

Nous avons donc montré la convergence simple en (t, x) de $\varphi^{v_n}(t, x)$ vers $\varphi^v(t, x)$. Pour montrer la convergence uniforme sur tout compact de $[0, 1] \times \mathbb{R}^d$, il suffit de montrer l'équicontinuité de la suite $\varphi^{v_n}(t, x)$ en tout point (t, x) . Or

$$\begin{aligned} |\varphi^{v_n}(t, x) - \varphi^{v_n}(t, y)| &\leq |x - y| + \int_0^t |v_n(s, \varphi^{v_n}(s, x)) - v_n(s, \varphi^{v_n}(s, y))| ds \\ &\leq |x - y| + \int_0^t |\partial_2 v_n(s, \cdot)|_\infty |\varphi^{v_n}(s, x) - \varphi^{v_n}(s, y)| ds \end{aligned}$$

qui donne par Grönwall

$$|\varphi^{v_n}(t, x) - \varphi^{v_n}(t, y)| \leq |x - y| \exp\left(C \sup_{n \geq 0} \|v_n\|_2\right),$$

et

$$|\varphi^{v_n}(t, x) - \varphi^{v_n}(t + h, x)| \leq C\sqrt{h} \sup_{n \geq 0} \|v_n\|_2.$$

□

3. Existence de géodésiques minimisantes et solution de problèmes variationnels

Le théorème de continuité faible du flot permet d'obtenir de nombreux résultats d'existence. Le premier est celui des géodésiques minimisante entre deux difféos quelconques de G_V .

THÉORÈME VI.3. *Pour tous $\varphi, \varphi' \in G_V$ il existe $u \in L_V^2$ tel que*

$$d_G(\varphi, \varphi') = \|u\|_1 = \|u\|_2$$

et $\varphi_1 = \varphi_1^u \circ \varphi$.

REMARQUE VI.3. *La courbe $t \rightarrow \varphi_t \triangleq \varphi_t^v \circ \varphi$ est une courbe joignant $\varphi = \varphi_0$ à $\varphi' = \varphi_1$ de longueur $d_G(\varphi, \varphi')$. L'égalité $\|u\|_1 = \|u\|_2$ assure que $|u_t|_V \equiv C$ si bien que $t \rightarrow \varphi_t$ est une géodésique à vitesse constante entre φ et φ' .*

DÉMONSTRATION. Soit $F \triangleq \{ u \in L_V^2 \mid \varphi' = \varphi_1^u \circ \varphi \}$. Le résultat de continuité faible nous donne que F est fermé pour la topologie faible. De plus on déduit de la remarque VI.2 que F est non vide. Il faut montrer que l'infimum de l'énergie $\|u\|_2$ pour $u \in F$ est atteint. Soit donc $(u_n)_{n \geq 0}$ une suite minimisante. Comme $\|u_n\|_2$ est borné on peut supposer quitte à extraire une sous-suite, que $u_n \rightharpoonup u_\infty \in F$. Or $\|u_\infty\|_2 \leq \liminf \|u_n\|_2 = \inf_{u \in F} \|u\|_2$ (car la norme forte est s.c.i. pour la topologie faible) donc $d_G(\varphi, \varphi') \leq \|u_\infty\|_1 \leq \|u_\infty\|_2 = d_G(\varphi, \varphi')$ d'après la proposition VI.1. \square

Comme pour les petites déformations, nous tirons également de la continuité faible du flot un résultat d'existence assez général pour les problèmes variationnels liés aux problèmes d'appariements.

THÉORÈME VI.4. *Soit $E : G_V \rightarrow \mathbb{R}^+$ une application semi-continue inférieurement pour la convergence uniforme sur tous les compacts de \mathbb{R}^d et $R : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ une application croissante telle que $R(\rho) \rightarrow +\infty$ lorsque $\rho \rightarrow +\infty$. On pose*

$$J(\varphi) = R(d_G(\text{Id}, \varphi)) + E(\varphi).$$

Alors, il existe $\varphi_ \in G_V$ tel que*

$$J(\varphi_*) = \inf_{\varphi \in G_V} J(\varphi).$$

EXERCICE VI.1. Montrer le théorème VI.4.

Le théorème précédent s'applique à la majorité des situations que nous avons vues : matching de landmarks, de mesures, de courants (voir chapitre suivant). Dans la cas des images, si Ω est borné et $E(\varphi) = \int_\Omega |I_1 - I_0 \circ \varphi^{-1}|^2(x) dx$ avec I_0 bornée et presque partout continue, alors le résultat s'applique.

Chapitre VII

Normes duales et appariements de sous-variétés

Pour faire de l'appariement de lignes ou de surfaces, il nous manque une attache aux données spécifique qui ne dépende pas du positionnement de landmarks en correspondance deux à deux. Il existe une notion naturelle pour faire de tels appariements qui est celle de courants qui en quelque sorte sont des mesures à valeurs éléments de ligne, éléments de surface, etc (ceci sera bien sûr éclairci dans la suite). C'est une notion qui fait appel à la notion de forme différentielle. Pour les introduire, nous allons tout d'abord parler du produit extérieur. C'est une notion un peu abstraite, mais tout à fait effective pour calculer avec des *éléments de surfaces ou de lignes*.

1. p-vecteurs et produit extérieur

1.1. Définition. Soient E un \mathbb{R} espace vectoriel de dimension fini d et p un entier $p \leq d$. Nous donnons ici une définition naïve du produit extérieur $\Lambda^p E$ à partir d'une base e_1, \dots, e_d de E on disant que $(e_{i_1} \wedge \dots \wedge e_{i_p})_{i_1 < \dots < i_p}$ est une base de $\Lambda^p E$ et qu'il existe une application *multilinéaire alternée* $E^p \xrightarrow{w} \Lambda^p E$ définie par $(u_1, \dots, u_p) \mapsto u_1 \wedge \dots \wedge u_p$.

Le produit extérieur $u_1 \wedge \dots \wedge u_p$ "code" en fait un élément de surface de dimension p porté par le sous-espace $F = \text{Vect}\{u_1, \dots, u_p\}$, dont l'orientation est donnée par celle de (u_1, \dots, u_p) . Ceci est précisé par l'exercice suivant :

EXERCICE VII.1. Si $v_i = \sum_{1 \leq j \leq p} A_{ij} u_j$ pour tout $1 \leq i \leq p$ est un changement de base de F , vérifier en utilisant la multilinéarité de w , que $v_1 \wedge \dots \wedge v_p = \det(A) u_1 \wedge \dots \wedge u_p$. En particulier, lorsque $\det(A) = 1$ (ce qui vaut dire que les aires orientées des parallélogrammes définis par (u_1, \dots, u_p) et (v_1, \dots, v_p) sont identiques) alors $v_1 \wedge \dots \wedge v_p = u_1 \wedge \dots \wedge u_p$.

On a immédiatement que $\dim(\Lambda^p E) = \binom{d}{p}$ et les éléments génériques s'écrivent $\sum_{i_1 < \dots < i_p} \alpha_{i_1, \dots, i_p} e_{i_1} \wedge \dots \wedge e_{i_p}$. On prendra la convention $\Lambda^0 E = \mathbb{R}$ et on remarquera que $\Lambda^1 E = E$.

EXERCICE VII.2. En écrivant $u_i = \sum_{j=1}^d u_{ij} e_j$, vérifier que en posant

$$\alpha_{i_1 \dots i_p} = \det \begin{bmatrix} u_{1i_1} & \cdots & u_{1i_p} \\ \vdots & \vdots & \vdots \\ u_{pi_1} & \cdots & u_{pi_p} \end{bmatrix}$$

on a $u_1 \wedge \cdots \wedge u_p = \sum_{i_1 < \dots < i_p} \alpha_{i_1, \dots, i_p} e_{i_1} \wedge \cdots \wedge e_{i_p}$

1.2. Produit scalaire sur $\Lambda^p E$. Si E muni d'un produit scalaire \langle, \rangle , alors on peut munir $\Lambda^p E$ d'un produit scalaire vérifiant $\langle u_1 \wedge \cdots \wedge u_p, v_1 \wedge \cdots \wedge v_p \rangle \triangleq \det(\langle u_i, v_j \rangle_{ij})$.

EXERCICE VII.3. Si (e_1, \dots, e_p) est une b.o.n de E , vérifier alors que $(e_{i_1} \wedge \cdots \wedge e_{i_p})_{i_1 < \dots < i_p}$ en est une de $\Lambda^p E$.

1.3. Lien avec le produit vectoriel. Si E est orienté (on choisit une base connue comme positivement orientée et l'orientation de toute autre base s'obtient en considérant le signe du déterminant de la matrice de changement de base).

Ici on suppose e_1, \dots, e_d positivement orientée. On peut alors considérer l'isomorphisme (souvent appelé opérateur de Hodge) : $*$: $\Lambda^p E \mapsto \Lambda^{d-p} E$ défini par $*(e_{i_1} \wedge \cdots \wedge e_{i_p}) \mapsto e_{j_1} \wedge \cdots \wedge e_{j_{d-p}}$ telle que $(e_{i_1}, \dots, e_{i_p}, e_{j_1}, \dots, e_{j_{d-p}})$ est une base positivement orientée. Dans le cas $p = d - 1$, $\Lambda^{d-1} E \xrightarrow{w} \Lambda^1 E = E$.

REMARQUE VII.1. — En particulier, dans le cas de la géométrie 3D ($E = \mathbb{R}^3$), alors pour $a, b \in E$, $*(a \wedge b)$ est exactement le vecteur $c \in \mathbb{R}^3$ tel que (a, b, c) est une base orthogonale positivement orientée et tel que $|c|$ est l'aire du parallélogramme défini par a et b . Ainsi, $*(a \wedge b)$ n'est rien d'autre que le produit vectoriel classique

— Cette identification pour $d = 3$ entre 2-vecteur et vecteur est cependant source de confusion. Elle n'apporte dans les faits qu'une simplification "apparente" et ne se généralise pas à $d > 3$.

EXERCICE VII.4. Vérifier que $\langle u_1 \wedge \cdots \wedge u_p, v_1 \wedge \cdots \wedge v_p \rangle = \langle *(u_1 \wedge \cdots \wedge u_p), *(v_1 \wedge \cdots \wedge v_p) \rangle$ ie que $*$ est une isométrie de $\Lambda^p E$ dans $\Lambda^{d-p} E$.

2. Formes différentielles

2.1. Dual de $\Lambda^p E$. On définit $(\Lambda^p E)^*$ le dual de $\Lambda^p E$ comme l'espace des formes lineaires sur $\Lambda^p E$. Pour tout $\omega \in (\Lambda^p E)^*$, ω est définie par ses valeurs sur les p -vecteurs $u_1 \wedge \cdots \wedge u_p$.

En considérant la base $(e_{i_1} \wedge \cdots \wedge e_{i_p})_{i_1 < \dots < i_p}$ de $\Lambda^p E$, il suffit de connaître $\omega(e_{i_1} \wedge \cdots \wedge e_{i_p})$.

On note $dx^{i_1} \wedge \cdots \wedge dx^{i_p} \in (\Lambda^p E)^*$ la base duale telle que :

$$(dx^{i_1} \wedge \cdots \wedge dx^{i_p})(e_{j_1} \wedge \cdots \wedge e_{j_p}) = \begin{cases} 1, & \text{si } i_k = j_k, \forall k \\ 0, & \text{sinon} \end{cases}$$

avec $i_1 < \cdots < i_p, j_1 < \cdots < j_p$.

REMARQUE VII.2. On identifie ainsi $(\Lambda^p E)^*$ et $\Lambda^p E^*$

EXERCICE VII.5. En écrivant $u_1 \wedge \cdots \wedge u_p = \sum_{i_1 < \cdots < i_p} dx^{i_1} \wedge \cdots \wedge dx^{i_p} (u_1 \wedge \cdots \wedge u_p) e_{i_1} \wedge \cdots \wedge e_{i_p}$, déduire de l'exercice VII.2 que

$$dx^{i_1} \wedge \cdots \wedge dx^{i_p} (u_1 \wedge \cdots \wedge u_p) = \det \begin{bmatrix} u_{1i_1} & \cdots & u_{pi_1} \\ \vdots & \ddots & \vdots \\ u_{1i_p} & \cdots & u_{pi_p} \end{bmatrix}$$

2.2. Exemples.

- Pour $p = d$, $\dim(\Lambda^d E)^* = 1$, $\omega(u_1 \wedge \cdots \wedge u_d) = \det(u_1, \dots, u_d) \omega(e_1 \wedge \cdots \wedge e_d)$. Toutes les formes linéaires sont donc dans ce cas proportionnelles au déterminant (forme volume) qui lui-même s'identifie avec $dx^1 \wedge \cdots \wedge dx^n$.
- Dans le cas $d = 3$, pour $p = 2$, $(\Lambda^d E)^*$ peut s'interpréter comme une mesure de *flux*,
- Dans le cas $p = d - 1$, pour toute forme linéaire $\omega \in \Lambda^{p-1} E$, il existe $\eta \in \Lambda^{p-1} E$ telle que $\omega(u_1 \wedge \cdots \wedge u_p) = \langle \eta, (u_1 \wedge \cdots \wedge u_p) \rangle$. En utilisant l'isométrie $*$ (cf exercice VII.4), on a $\omega(u_1 \wedge \cdots \wedge u_p) = \langle *\eta, *(u_1 \wedge \cdots \wedge u_p) \rangle$ qui s'interprète géométriquement comme le flux du vecteur $*\eta \in E$ à travers l'élément d'hypersurface définie par $u_1 \wedge \cdots \wedge u_p$.
- Dans le cas $p = 1$, alors de façon cette fois immédiate, en remarquant que $(\Lambda^1 E)^* = E^*$, on a pour tout $\omega \in E^*$, il existe $v \in E$ telle que $\omega(u) = \langle v, u \rangle$ pour tout $u \in \Lambda^1 E = E$.
- Enfin pour $p = 0$, on identifie $(\Lambda^0 E)^* = \mathbb{R}$

DÉFINITION VII.1. Pour tout $\omega \in (\Lambda^p E)^*$, on définit la norme duale

$$|\omega|_{(\Lambda^p E)^*} = \sup_{\eta \in \Lambda^p E, |\eta|=1} \omega(\eta).$$

REMARQUE VII.3. La norme duale est issue d'un produit scalaire sur $(\Lambda^p(E))^*$ pour lequel $(dx^{i_1} \wedge \cdots \wedge dx^{i_p})_{i_1 < \cdots < i_p}$ est une b.o.n.

REMARQUE VII.4. Pour tout $\omega \in \Lambda^p E$, si on définit $\tilde{\omega}(u_1, \dots, u_p) \triangleq \omega(u_1 \wedge \cdots \wedge u_p)$ ie $\tilde{\omega} \triangleq \omega \circ w$, alors $\tilde{\omega}$ est forme multilinéaire alternée. Il est d'usage d'identifier $\tilde{\omega}$ et ω en écrivant parfois $\omega(u_1, \dots, u_p)$ au lieu de $\omega(u_1 \wedge \cdots \wedge u_p)$.

On considère maintenant $\Omega_0^p(\mathbb{R}^d) \triangleq C^\infty(\mathbb{R}^d, (\Lambda^p \mathbb{R}^d)^*)$ les p -formes différentielles sur \mathbb{R}^d continues nulles à l'infini munies de la norme

$$|\omega|_\infty = \sup_x |\omega_x|_{(\Lambda^p \mathbb{R}^d)^*}$$

qui fait de $\Omega_0^p(\mathbb{R}^d)$ un espace de Banach

- REMARQUE VII.5. — Pour $p = 0$, $\Omega_0^p(\mathbb{R}^d)$ s'identifie avec $C_0(\mathbb{R}^d, \mathbb{R})$ et nous retrouvons l'espace de fonctions à partir duquel nous avons par dualité les mesures signées ;
- pour $p = 1$, $\Omega_0^1(\mathbb{R}^d)$ s'identifie avec les champs de vecteurs ;
- nouvel objet pour $p \geq 2$, (cependant si $d = 3$, $\Omega_0^2(\mathbb{R})$ peut s'identifier encore aux champs de vecteurs).

REMARQUE VII.6. Dans le cas d'une variété M , on peut étendre la notion de formes différentielles en considérant les sections continues du fibré sur M de fibre $\Lambda^p T_x^*$ en $x \in M$.

3. Courants

Ce qui nous intéresse ici est de considérer le dual $(\Omega_0^p(\mathbb{R}^d))^*$ de $\Omega_0^p(\mathbb{R})$.

- dans le cas $p = 0 \mapsto$ mesures signées
- dans le cas $p = 1 \mapsto$ mesures à valeurs vecteurs (on peut les définir en regardant les choses coordonnées par coordonnées).

La question est de voir comment une surface classique peut être vue comme un courant.

3.1. Cas des lignes. Prenons le cas d'une ligne définie par la courbe paramétrée $\gamma : [0, 1] \mapsto \mathbb{R}^d$ qui soit disons C^1 . On peut associer à γ une forme linéaire sur $\Omega_0^1(\mathbb{R}^d)$:

$$\omega \mapsto \int_0^1 \omega_{\gamma(t)}(\gamma'(t)) dt.$$

Évidemment le point principal est de vérifier que cette forme linéaire ne dépend pas de la paramétrisation mais de la ligne géométrique elle-même et d'une orientation.

Pour cela, prenons $\psi : [0, 1] \rightarrow [0, 1]$ qui soit C^1 telle que $\psi'(x) > 0$ pour tout $x \in [0, 1]$. La nouvelle courbe paramétrée $\tilde{\gamma} \triangleq \gamma \circ \psi$ définit la même courbe géométrique avec le même sens de parcours (en d'autres termes on considère la même variété orientée (ici de dimension 1) à travers une autre carte). Alors on a

$$\int_0^1 \omega_{\tilde{\gamma}(s)}(\tilde{\gamma}'(s)) ds = \int_0^1 \omega_{\gamma \circ \psi(s)}(\gamma' \circ \psi(s) \psi'(s)) ds = \int_0^1 \omega_{\gamma(t)}(\gamma'(t)) dt$$

où la dernière inégalité est obtenue en utilisant la linéarité de ω et le changement de variable $t = \psi(s)$. Par suite, si L est la courbe orientée définie de façon formelle par la classe d'équivalence de γ dans $C^1([0, 1], \mathbb{R}^d)/\text{Diff}_+^1([0, 1])$, on peut définir

$$T_L(\omega) = \int_L \omega \triangleq \int_0^1 \omega_{\gamma(t)}(\gamma'(t)) dt.$$

Comme $|T_L(\omega)| \leq |\omega|_\infty \int_0^1 |\gamma'(t)| dt \leq |\omega|_\infty |L|$ où $|L|$ est la longueur de L , T_L est bien une forme linéaire continue pour la norme $|\omega|_\infty$ et donc un élément de $\Omega_0^1(\mathbb{R}^d)$.

Au passage on a transformé une ligne géométrique en un élément d'un Banach sur lequel on peut faire des additions. Par exemple, $T_{L_1} + T_{L_2}$ code pour deux lignes L_1 et L_2 prises simultanément. Comme dans le cas des mesures signées, cet espace contient bien d'autres choses et notamment des objets discrets.

En effet, prenons $x \in \mathbb{R}^d$ et $\alpha \in \mathbb{R}^d$, si $\delta_x^\alpha(\omega) = \omega_x(\alpha)$, alors δ_x^α peut être comme un courant correspondant en un élément "ponctuel" de ligne positionné en $x \in \mathbb{R}^d$ orienté selon la direction α et de longueur $|\alpha|$. On voit dans l'exercice suivant que c'est un élément de base qui permet d'approximer des courants 1D plus généraux.

EXERCICE VII.6.

- (1) Montrer que $(\sum_{i=1}^n \delta_{\gamma(i/n)}^{\gamma'(i/n)/n})(\omega) \rightarrow T_\gamma(\omega)$ lorsque $n \rightarrow \infty$ pour tout $\omega \in \Omega_0^1(\mathbb{R}^d)$.
- (2) Montrer que si $x_i = \gamma(i/n)$ pour $0 \leq i \leq n$ alors $(\sum_{i=1}^n \delta_{x_i}^{x_{i+1}-x_i})(\omega) \rightarrow T_\gamma(\omega)$ lorsque $n \rightarrow \infty$ pour tout $\omega \in \Omega_0^1(\mathbb{R}^d)$.

3.2. Cas des surfaces. Soit $\gamma : U \subset \mathbb{R}^2 \mapsto \mathbb{R}^d$ une paramétrisation C^1 d'une surface S . On suppose que

$$(41) \quad |S| \triangleq \int_U \left| \frac{\partial \gamma}{\partial s} \wedge \frac{\partial \gamma}{\partial t} \right| ds dt < \infty$$

où on rappelle que $|\frac{\partial \gamma}{\partial s} \wedge \frac{\partial \gamma}{\partial t}| = \sqrt{\det(\text{Gram}(\frac{\partial \gamma}{\partial s}, \frac{\partial \gamma}{\partial t}))}$ avec $\text{Gram}(u_1, \dots, u_p) \triangleq (\langle u_i, u_j \rangle)_{1 \leq i, j \leq p}$ la matrice de Gram¹.

1. La valeur de $|S|$ peut être prise comme l'aire de S . En effet, en découpant U sur un quadrillage en petits carrés C_i de côté $\Delta s \times \Delta t$, puis en approximant les aires élémentaires des morceaux $\gamma(C_i)$ au premier ordre par les aires des parallélogrammes P_i définis par les vecteurs $\frac{\partial \gamma}{\partial s} \Delta s(s_i, t_i)$ et $\frac{\partial \gamma}{\partial t} \Delta t(s_i, t_i)$ où $(s_i, t_i) \in C_i$ on a

$$\text{Aire}(S) \simeq \sum_{i \in I} \text{Aire}(P_i) = \sum_{i \in I} \left| \frac{\partial \gamma}{\partial s} \Delta s \wedge \frac{\partial \gamma}{\partial t} \Delta t \right|(s_i, t_i) = \sum_{i \in I} \left| \frac{\partial \gamma}{\partial s} \wedge \frac{\partial \gamma}{\partial t} \right|(s_i, t_i) \Delta s \Delta t.$$

où l'on reconnaît une somme de Riemann qui converge vers (41).

Si $\omega \in \Omega_0^2(\mathbb{R}^d)$ on peut définir

$$T_\gamma(\omega) \triangleq \int_U \omega \left(\frac{\partial \gamma}{\partial s} \wedge \frac{\partial \gamma}{\partial t} \right) ds dt$$

qui vérifie $|T_\gamma(\omega)| \leq |\omega|_\infty |S|$ et qui donc est une forme linéaire continue. Or si $\psi : V \mapsto U$, $(a, b) \mapsto \psi(a, b)$ est dans $C^1(U, V)$ tel que $J\psi \triangleq \det(\frac{\partial \psi}{\partial a}, \frac{\partial \psi}{\partial b}) > 0$, alors pour $\tilde{\gamma} = \gamma \circ \psi$ on a

$$(42) \quad T_{\tilde{\gamma}} = \int_V \omega \left(\frac{\partial \tilde{\gamma}}{\partial a} \wedge \frac{\partial \tilde{\gamma}}{\partial b} \right) da db = \int_U \omega \left(\frac{\partial \gamma}{\partial s} \wedge \frac{\partial \gamma}{\partial t} \right) ds dt = T_\gamma(\omega)$$

EXERCICE VII.7. Montrer l'égalité (42) en montrant que $\frac{\partial \tilde{\gamma}}{\partial a} \wedge \frac{\partial \tilde{\gamma}}{\partial b} = (\frac{\partial \gamma}{\partial s} \wedge \frac{\partial \gamma}{\partial t}) \circ \psi \underbrace{J\psi}_{>0} = (\frac{\partial \gamma}{\partial s} \wedge \frac{\partial \gamma}{\partial t}) \circ \psi |J\psi|$ puis en effectuant un changement de variable.

Une fois encore, T_γ ne dépend que de la classe de γ modulo une reparamétrisation qui ne change pas l'orientation. En particulier T_γ ne dépend que de la structure de sous-variété orientée de S^2 . On peut donc écrire

$$(43) \quad T_S(\omega) = \int_S \omega \triangleq T_\gamma(\omega).$$

Ainsi, les 2-surfaces orientées s'identifient avec des éléments de $(\Omega_0^2(\mathbb{R}^d))^*$. Ici aussi beaucoup d'autres choses sont des 2-courants, en particulier le courant élémentaire $\delta_x^{u \wedge v} : \omega \mapsto \omega_x(u \wedge v)$ qui représente un élément "ponctuel" de surface orientée positionné en x d'aire $|u \wedge v|$ dans le plan engendré par (u, v) . Plus généralement, on peut voir un 2-courant comme une mesure à valeurs éléments de surface.

EXERCICE VII.8. Si $(X_i)_{i \geq 1}$ est une suite iid de loi uniforme sur U et pour tout i , $\nu_i = (\frac{\partial \gamma}{\partial s} \wedge \frac{\partial \gamma}{\partial t})(X_i)$, vérifier par la loi forte des grands nombres que pour tout $\omega \in \Omega_0^2(\mathbb{R}^d)^*$, on a

$$\frac{1}{n} \sum_{i=1}^n \delta_{X_i}^{\nu_i}(\omega) \rightarrow T_S(\omega),$$

p.s. lorsque $n \rightarrow \infty$.

4. Transport de courants

Si $\varphi : \mathbb{R}^d \mapsto \mathbb{R}^d$ et $\omega \in \Omega_0^p(\mathbb{R}^d)$, on peut définir $\varphi(\omega)$ par

$$(44) \quad \varphi(\omega)_x(u_1 \wedge \cdots \wedge u_p) \triangleq \omega_{\varphi(x)}(d\varphi_x(u_1) \wedge \cdots \wedge d\varphi_x(u_p)).$$

2. En fait, ceci n'est pas tout à fait exact car S peut ne pas être une sous-variété puisque la surface définie par γ pourrait éventuellement se recouper elle-même. Il faudrait plutôt comme pour le cas des lignes parler d'immersion à reparamétrisation près et définir la structure quotient adéquate.

EXERCICE VII.9.

(1) Montrer que l'on a $\varphi_1(\varphi_2(\omega)) = (\varphi_2 \circ \varphi_1)(\omega)$.

(2) Lorsque $\varphi = \text{Id} + u$ pour $u \in C_0^1(\mathbb{R}^d, \mathbb{R}^d)$ vérifier que $\varphi(\omega) \in \Omega_0^p(\mathbb{R}^d)$.

Si $T \in \Omega_0^p(\mathbb{R}^d)^*$, on peut définir

$$(45) \quad \varphi(T)(\omega) \triangleq T(\varphi(\omega))$$

si bien que l'on déduit de l'exercice précédent que

$$(46) \quad \varphi_1(\varphi_2(T)) = (\varphi_1 \circ \varphi_2)(T).$$

4.1. Cas des lignes. Que se passe-t-il pour les lignes ? On a

$$\int_0^1 \omega_{\varphi \circ \gamma(t)}(d\varphi_{\gamma(t)}(\gamma'(t))) dt = \int_{\varphi(L)} \omega$$

ce qui correspond au transport de la ligne par φ .

4.2. Cas des surfaces. Que se passe-t-il pour les surfaces ? On a

$$(47) \quad \begin{aligned} \varphi(T_\gamma)(\omega) &= T_\gamma(\varphi(\omega)) = \int_U \omega_{\varphi \circ \gamma(s,t)}(d\varphi_\gamma(\frac{\partial \gamma}{\partial s}) \wedge d\varphi_\gamma(\frac{\partial \gamma}{\partial t})) ds dt \\ &= \int_U \omega_{\varphi \circ \gamma(s,t)}(\frac{\partial \varphi \circ \gamma}{\partial s} \wedge \frac{\partial \varphi \circ \gamma}{\partial t}) ds dt = T_{\varphi \circ \gamma}(\omega) \end{aligned}$$

ce qui correspond encore au transport des 2-surfaces au sens géométrique.

5. ENR sur les courants

Si $W \hookrightarrow \Omega_0^p(\mathbb{R}^d)$ est une injection continue, W sev dense, alors W est un ENR sur les p -formes. On a de plus $\Omega_0^p(\mathbb{R}^d)^* \hookrightarrow W^*$ si bien que tout courant $T \in \Omega_0^p(\mathbb{R}^d)^*$ peut être examiné dans W^* . Or pour tout $T \in W^*$ on dispose de la norme duale :

$$(48) \quad |T|_{W^*} = \sup_{|\omega|_W=1} T(\omega).$$

et des techniques de calcul des normes à partir du noyau K associée à W .

5.1. Cas des lignes. On a $T_L(\omega) = \int_L \omega = \int_0^1 \omega_\gamma(\gamma'(t)) dt = \int_0^1 \langle K \delta_\gamma', \omega \rangle_W dt$ et

$$(49) \quad \begin{aligned} \langle T_L, T_L \rangle_{W^*} &\stackrel{K_{\text{isom}}}{=} \langle K T_L, K T_L \rangle_W = T_L(K T_L) \\ &\stackrel{\text{Def de } K}{=} \int_0^1 \langle K \delta_\gamma', K T_L \rangle_W dt = \int_0^1 \int_0^1 \langle K \delta_{\gamma(s)}^{\gamma'(s)}, K \delta_{\gamma(t)}^{\gamma'(t)} \rangle_W ds dt \\ &= \int_0^1 \int_0^1 \gamma'(s)^T K(\gamma(s), \gamma(t)) \gamma'(t) ds dt = \int_{L \times L} u(q)^T K(q, q') u(q') dq dq', \end{aligned}$$

où en tout point de q de L , $u(q)$ est le vecteur tangent orienté unitaire et dq désigne la variation de l'abscisse curviligne. En peut donc calculer la distance entre deux lignes géométriques L_0 et L_1 en calculant $|T_{L_1} - T_{L_2}|_{W^*}$ qui se ramène à un calcul d'intégrales doubles.

5.2. Cas des 2-surfaces. Dans ce cas, on a $T_S(\omega) = \int_S \omega = \int_U \omega_\gamma (\frac{\partial \gamma}{\partial s} \wedge \frac{\partial \gamma}{\partial t}) ds dt$. Par suite, par un calcul similaire à celui des lignes, on a

$$(50) \quad |T_S|_{W^*}^2 = \int_U \int_U (\frac{\partial \gamma}{\partial s} \wedge \frac{\partial \gamma}{\partial t})^T K(\gamma(s, t), \gamma(s', t')) \frac{\partial \gamma}{\partial s'} \wedge \frac{\partial \gamma}{\partial t'} ds dt ds' dt'$$

Encore, une fois il faut bien remarquer que si l'expression de droite dans (50) fait intervenir une paramétrisation γ , la norme est intrinsèque à S et ne dépend pas de cette paramétrisation (à condition que l'orientation soit préservée).

Dans le cas $d = 3$, alors on peut identifier le produit extérieur de deux vecteurs avec le vecteur donnée par le produit vectoriel (en utilisant l'opérateur $*$). Dans ce cas, si en tous points de $q \in S$, on définit la normale orientée $n(q)$, on a

$$(51) \quad |T_S|_{W^*}^2 = \int_{S \times S} n(q)^T K(q, q') n(q') dq dq'.$$

où $K(q, q')$ peut-être alors considéré à valeurs dans $\mathcal{M}_d(\mathbb{R})$ et dq désigne la mesure de surface sur q . Ainsi, pour comparer deux surfaces S et S' , l'approche par ENR opère par comparaisons des normales :

$$|T_{S_0} - T_{S_1}|_{W^*}^2 = \sum_{0 \leq i, j \leq 1} (-1)^{(i-j)} \int_{S_i \times S_j} n_i(q)^T K(q, q') n_j(q') dq dq'.$$

6. Calcul des appariements

C'est une situation finalement très similaire à celle des mesures. Dans le cas des lignes, on peut représenter chaque ligne L par des éléments ponctuels de lignes sur la forme $\sum_{i=1}^n \delta_{x_i}^{x_{i+1}-x_i}$ où (x_i) sont distribués le long de L .

REMARQUE VII.7. (1) *La régularité de la distribution de point n'est pas primordiale, l'approximation sera de bonne qualité tant que la distance entre points restera petite par rapport à dynamique du noyau*

(2) *Cette approximation discrète ne code pas une paramétrisation spécifique de la ligne. Elle approxime comme nous l'avons dit une quantité intrinsèque liée à la structure différentielle et à orientation de la ligne.*

Dans cas des 2-surfaces, si on dispose d'une triangulation sous la forme d'un découpage en triangles $T_i = (a_i, b_i, c_i)$ pour $i \in I$ de la surface, alors si pour tout $i \in I$, x_i est un point du triangle (par exemple le barycentre), on peut approximer le courant T_S par

$$\tilde{T}_S = \sum_{i \in I} \delta_{x_i}^{\overrightarrow{a_i b_i} \wedge \overrightarrow{a_i c_i}}$$

Les remarques précédentes s'appliquent ici : cette approximation est relativement insensible au découpage en triangles pourvu qu'ils soient de tailles raisonnables par rapport à la dynamique du noyau. Notons également que nous n'avons pas besoin de connaître les relations de voisinages entre triangles. Cependant, n'oublions pas le problème de l'orientation :

EXERCICE VII.10. Vérifier que $\overrightarrow{a_i b_i} \wedge \overrightarrow{a_i c_i} = \overrightarrow{b_i c_i} \wedge \overrightarrow{b_i a_i} = \overrightarrow{c_i a_i} \wedge \overrightarrow{c_i b_i}$. En déduire que l'on peut échanger les rôles de a_i , b_i et c_i par permutation circulaire.

Dans le cas de la triangulation d'une 2-surface orientable, une orientation de chaque triangle correspond à une façon de circuler le long du bord du triangle (dans notre cas, l'ordre d'apparition des sommets (a_i, b_i, c_i) dans le triangle T_i décrit cette orientation). Les orientations sont consistantes si deux triangles adjacents ont des sens de parcours *opposés* lorsqu'ils partagent une arête commune. Pour cela, il peut être utile d'avoir une information sur les triangles partageant une arête commune.

L'appariement entre deux 2-surfaces S_0 et S_1 se fera du point de vue pratique sur leurs approximations \tilde{T}_{S_i} en minimisant sur $u \in V$ une fonctionnelle du type :

$$(52) \quad J(u) \triangleq \frac{1}{2} |u|_V^2 + \frac{\gamma}{2} |\varphi^u(\tilde{T}_{S_0}) - \tilde{T}_{S_1}|_{W^*}^2,$$

avec $\varphi^u(x) = x + u(x)$. L'attache aux données ne dépend en fait que de $(u(x_i), du_{a_i}, du_{b_i}, du_{c_i})_{i \in I}$. On peut montrer alors que les solutions sont à chercher dans un espace de dimension fini un peu plus compliqué que dans le cas de l'interpolation simple. La solution la plus simple est de faire l'approximation naturelle proposée par J. Glaunès [9] :

$$\varphi(\delta_{x_i}^{\overrightarrow{a_i b_i} \wedge \overrightarrow{a_i c_i}}) \simeq \delta_{\varphi(x)}^{\overrightarrow{\varphi(a_i) \varphi(b_i)} \wedge \overrightarrow{\varphi(a_i) \varphi(c_i)}}$$

L'attache aux données ne dépend plus que des valeurs de u sur les x_i et les sommets des triangles, si bien que les solutions s'écrivent $u_*(x) = \sum_{i \in I} (K(x, x_i) \xi_i + (K(x, a_i) \alpha_i + K(x, b_i) \beta_i + K(x, c_i) \gamma_i))$. On se ramène à un problème de minimisation de dimension finie (non linéaire).

Présentation orale

Les papiers qui peuvent faire l'objet de la présentation orale sont accessibles sur la page moodle du cours que vous pouvez atteindre à l'adresse <https://elearn.ens-cachan.fr/course/view.php?id=1683>. Le clé d'accès pour le cours est Thompson.

Ce sont tous des papiers assez récents. On attend une exposition argumentée du papier en 20 min pour laisser place à une discussion libre avec le jury (durée totale de la présentation 45min). Le jury pourra poser des question sur le contenu du papier ou sur ses liens avec le cours. Pour certains papier, on propose une grille de lecture particulière. Merci de choisir assez vite l'un d'entre eux. Vous pourrez disposer d'un vidéo projecteur.

Morphometry of anatomical shape complexes, S. Durrleman et al.: Le papier, paru dans Neuroimage en 2014, presente une étude de la variabilité de complexe de formes par des outils de grandes déformations difféomorphiques illustrant les techniques vues en cours. On attend une compréhension assez fine du principe de la méthode et de sa mise en oeuvre. La partie statistique pourra être abordée mais ne doit constituer le centre de l'exposé.

A Bayesian Framework for Joint Morphometry of Surface and Curve Meshes in Multi-Object Complexes, P. Gori et al.: Ce papier paru dans Med Image Analysis en 2017 est assez dans l'esprit du papier précédent mais regarde le probleme considère des mélanges de surface et de structures en lignes pour une application à l'étude de la variabilité de la matière blanche. Il plonge aussi le recalage dans un cadre statistique permettant de dégager des heuristique pour la calibration des paramètres. Ce papier est particulièrement adapté à ceux qui suivent par ailleurs le cours de S. Allasonnière au MVA. Il est attendu une présentation du cadre statistique et numérique du papeir ainsi qu'une analyse d ela pertinence de la modélisation dans le cadre de l'application.

Higher order kernels, S. Sommer et al.: Le papier construit l'analyse de la situation en introduisant l'equation EPdiff. Nous avons vu dans le cours une approche s'appuyant sur le controle optimal et l'approche hamiltonienne. Le lecteur essaiera de

voir comment cette approche permet une relecture plus simple du papier. Donner votre appréciation sur la portée de l'introduction d'un tel modèle.

Constrained diffeomorphic Shape Evolution, L. Younes: Exposer l'idée principale et en particulier la dérivation des équations d'évolution dans le contexte du contrôle optimal. Quelles sont pour vous les forces et les limitations d'une telle approche ?

Shape Splines, A. Trounev et al.: Expliquer le principe de la construction des splines de formes. La partie stochastique n'est pas directement accessible dans le contexte du cours et pourra être sautée. Donnez votre avis sur la possibilité d'utiliser un tel modèle pour le suivi temporel de formes.

Diffeomorphic 3D Images Registration, F.-X. Vialard et al.: Expliquer comment on peut obtenir (du moins formellement) les équations de shooting dans le cadre des images. On essaiera de comprendre aussi pourquoi dans le cas des images, les équations d'évolutions sur le hamiltonien réduit ne peut pas être vu comme des EDO sur un Banach comme fait dans le cours. Le papier contient une problématique d'analyse numérique que l'on ne demande pas de regarder en profondeur. On s'attachera plutôt à replacer le papier dans le cadre du cours.

Diffeomorphic Longitudinal Shape Analysis, N. Singh et al.: Ce papier paru dans la conférence IPMI 2013 (Image Processing in Medical Images), explore le problème de l'analyse de l'évolution d'une forme dans le temps. Le modèle de base est celui d'une évolution géodésique mais on considère ici une population comprenant l'évolution de plusieurs sujets dont on veut reconstruire une évolution moyenne. Le papier est court mais il pourra être utile de consulter la biblio. On s'attachera à exposer clairement le principe du modèle et de sa mise en œuvre.

The varifold representation of non-oriented shapes, N. Charon et al.: Ce papier paru cette année à SIAM Journal of Imaging Sciences introduit pour la première fois des attaches aux données de type varifolds entre surfaces. La théorie des varifolds, issue de la théorie de la mesure géométrique, permet de construire des représentations non-orientées des surfaces et le papier dérive des espaces RKHS dans lesquels ces types de représentation conduisent à de nouveaux termes d'attache. Contrairement aux attaches aux données dans la littérature entre surface, ceux-ci ne dépendent pas du choix d'une paramétrisation des surfaces, sont différentiables sous l'action des déformations, ne souffrent pas des phénomènes d'annulation des orientations opposées comme dans le cas des courants et intègrent des informations sur les distributions des espaces tangents à la surface qui leurs

donnent des comportements complètement nouveaux. Ceci en pour toutes les dimensions. On s'attachera à exposer clairement le principe des varifolds et la façon dont ils peuvent s'intégrer dans un contexte d'espaces de formes. Attention le papier est un peu long. Il faudra faire quelques choix.

Atlas-based Shape Analysis using the Functional Shape Framework: (S. Lee et al) Ce papier paru à Medical Image Analysis en 2017 propose une extension de l'analyse de la variabilité des shapes géométriques à celle de la variabilité des formes fonctionnelles (fshapes) c'est à dire des surfaces texturée. Le travail présente une application à l'étude de données de rétines par tomographie en cohérence optique pour la détection du glaucome. Il s'agira de présenter l'approche par espaces de formes dans le cas des fshapes, en particulier la nouveauté par rapport au cadre que nous avons vu en cours et de développer l'application à l'analyse des épaisseurs des fibres nerveuses autour de la papille pour le diagnostique du glaucome.

Multiple Shape Registration using Constrained Optimal Control: (S. Arguillière et al.) Ce papier paru dans SIAM J. Imaging Science. Il propose de regarder si c'est possible lorsque l'on considère plusieurs formes simultanément, de construire un espace de formes en faisant agir un difféomorphisme par objet + 1 pour le fond. Evidement il faut pouvoir recoudre les difféos le long des frontières des objets. On obtient une évolution simultanée de plusieurs difféos sous contraintes mais qui permet de choisir des métriques différentes (et donc des régularités différentes) pour le transport de chaque objet. On demande de présenter le modèle dans le cas des contraintes d'identités (recolllements sur les bords) et pour des modèles discrets (le cas des contraintes de glissement, plus technique n'est pas demandé). Ce papier, bien dans l'esprit du cours considère une extension importante avec l'arrivée de contraintes. On s'attachera à en comprendre les éléments nouveaux et le potentiel impact sur les applications à la neuro-imagerie.

Shape deformation analysis from the optimal control viewpoint: (S. Arguillière et al). Ce papier paru au Journal de Mathématiques Pures et Appliquées en 2015 est une présentation d'un cadre assez général pour les espaces de formes et qui interressera les plus mathématiciens d'entre vous qui ont envie de creuser des aspects qui n'ont pas été complètement approfondis dans le cours. Le papier est assez touffu et il est demandé seulement d'étudier de près et de présenter les deux premières sections du papier et notamment de comprendre le cadre et la preuve du théorème 1 (20 premières pages).

Bibliographie

- [1] N. Aronszajn. La théorie des noyaux reproduisants et ses applications. *Proc. Cambridge Philos. Soc.*, pages 133–153, 1943.
- [2] S. Bochner. Integration von funktionen, deren werte die elemente eines vektorraumes sind. *Fundamenta Mathematicae*, 20 :262–276, 1933.
- [3] L. Bookstein, F. Principal warps : Thin plate splines and the decomposition of deformations. *IEEE trans. PAMI*, 11(6) :567–585, 1989.
- [4] L. Bookstein, F. *Morphometric tools for landmark data ; geometry and biology*. Cambridge University press, 1991.
- [5] J. Bourguignon. *Calcul variationnel*. Ecole polytechnique, 2007.
- [6] H. Brezis. *Analyse fonctionnelle, théorie et applications*. Masson, Paris, 1983. English translation : Springer Verlag.
- [7] S. Durrleman, M. Prastawa, G. Gerig, and S. Joshi. Optimal data-driven sparse parametrization of diffeomorphisms for population analysis. In *Information Processing in Medical Imaging (IPMI)*, volume Lecture Notes in Computer Science (LNCS) 6801, pages 123–134, 2011.
- [8] S. Gallot, D. Hullin, and J. Lafontaine. *Riemannian Geometry*. Springer, 1993.
- [9] J. Glaunes. *Transports par difféomorphismes de points, de mesures et de courants pour la comparaison de formes et l'anatomie numérique*. PhD thesis, Université Paris 13, 2005.
- [10] M. Hein and O. Bousquet. Kernels, associated structures and generalizations. Technical report, Max Plack Institute for Biological Cybernetics, July 2004.
- [11] D. G. Kendall. A survey of the statistical theory of shape. *Statistical Science*, pages 87–99, 1989.
- [12] Š. Schwabik and G. Ye. *Topics in Banach space integration*, volume 10. World Scientific Pub Co Inc, 2005.
- [13] D. W. Thompson. *On growth and form*. Cambridge University Press, 1945.
- [14] M. Vaillant, M. Miller, L. Younes, and A. Trouvé. Statistics on diffeomorphisms via tangent space representations. *NeuroImage*, 23 :S161–S169, 2004.
- [15] L. Younes. *Shapes and Diffeomorphisms*. Springer, 2010.