

MASTER MVA

Géométrie et espaces de formes

Alain Trouvé
2009-2010
Version 0.7

Ce texte ne constitue pas à vrai dire des notes de cours mais reprend en partie certains points évoqués dans le cours parfois avec plus de détails. A ce titre il peut être utile. Il s'agit cependant d'un texte en cours d'écriture et comportant encore pas mal de coquilles. Merci de me communiquer vos remarques à trouve@cmla.ens-cachan.fr¹

Les mises à jour seront postées sur la page :
<http://www.cmla.ens-cachan.fr/fileadmin/Membres/trouve/GEF/>
ainsi que d'autres informations utiles.

¹Je remercie Philipp Harms pour une première relecture attentive

Table des matières

Chapitre 1. Modèles de petites déformations	5
1. Interpolation et ENR	5
2. Appariement dense de points indexés	7
3. Appariement de nuages de points, appariement de mesures	8
4. Appariements de courants	13
5. Appariement d'images, cas général	22
6. Les ENR en brefs	24
Chapitre 2. Actions de groupes, distances invariantes	29
1. Actions de groupes, espaces homogènes	29
2. Distances, distances quotient sur les espaces homogènes	31
3. Approche différentielle	34
4. Métriques riemanniennes	40
Chapitre 3. Construction de groupes de déformations	43
1. Flots	43
2. Définition et premières propriétés du groupe G_V	48
3. Continuité faible du flot	50
4. Existence de géodésiques minimisantes et solution de problèmes variationnels	51
Bibliographie	53

CHAPITRE 1

Modèles de petites déformations

1. Interpolation et ENR

On se place sur \mathbb{R}^d et on considère une famille finie $\mathbf{x}_I = (\mathbf{x}_i)_{i \in I}$ de points sur \mathbb{R}^d et $\mathbf{a}_I = (\mathbf{a}_i)_{i \in I}$ une famille de vecteurs dans \mathbb{R}^m .

Problème de l'interpolation: .

Trouver $f : \mathbb{R}^d \rightarrow \mathbb{R}^m$ tel que $f(\mathbf{x}_i) = \mathbf{a}_i$ pour tout $i \in I$.

Les valeurs en dehors de la famille \mathbf{x}_I sont indéfinies : *on a donc un problème mal posé.*

Prenons un espace de Hilbert \mathcal{W} (e.v.n. complet muni d'un produit scalaire $(f, g) \rightarrow \langle f, g \rangle_{\mathcal{W}}$) de fonctions de \mathbb{R}^d dans \mathbb{R}^m . La norme $\|f\|_{\mathcal{W}}^2$ de f est une mesure de la régularité de f . On transforme le problème en un nouveau problème régularisé :

$$(1) \quad (P) \begin{cases} \min_{\mathcal{W}} \frac{1}{2} \|f\|_{\mathcal{W}}^2 \\ \text{avec} \\ f(\mathbf{x}_i) = \mathbf{a}_i, \quad i \in I \end{cases}$$

C'est un problème de projection sur un sous-espace affine $\mathcal{W}_{\mathbf{x}_I, \mathbf{a}_I} = \{f \in \mathcal{W} \mid f(\mathbf{x}_i) = \mathbf{a}_i, \quad i \in I\}$. Lorsque le sous-espace est fermé, on a existence et unicité de la solution f_* solution de (P). Pour cela, il suffit que pour tout $\mathbf{x} \in \mathbb{R}^d$, l'application $\mathbf{x} \rightarrow f(\mathbf{x})$ soit continue.

DÉFINITION 1. *Lorsque l'application $\delta_{\mathbf{x}} : \mathcal{W} \rightarrow \mathbb{R}^m$ définie par $\delta_{\mathbf{x}}(f) = f(\mathbf{x})$ est continue, on dit que \mathcal{W} est un espace à noyau reproduisant (ENR).*

On suppose donc maintenant que \mathcal{W} est un ENR. On notant, pour tout $\mathbf{x} \in \mathbb{R}^d$ et $\boldsymbol{\alpha} \in \mathbb{R}^m$, $\delta_{\mathbf{x}}^{\boldsymbol{\alpha}} : \mathcal{W} \rightarrow \mathbb{R}$ définie par

$$(2) \quad \delta_{\mathbf{x}}^{\boldsymbol{\alpha}}(f) = \langle \boldsymbol{\alpha}, f(\mathbf{x}) \rangle_{\mathbb{R}^m}$$

on a que $\delta_{\mathbf{x}}^{\boldsymbol{\alpha}} \in \mathcal{W}^*$ où \mathcal{W}^* est l'espace des applications linéaires continues de \mathcal{W} dans \mathbb{R} (ie des formes linéaires continues).

Un résultat central est le théorème de représentation de Riesz qui dit que pour tout $\mathbf{l} \in \mathcal{W}^*$ il existe unique $\mathbf{K}(\mathbf{l}) \in \mathcal{W}$ tel que

$$(3) \quad \langle \mathbf{K}(\mathbf{l}), f \rangle_{\mathcal{W}} = \mathbf{l}(f) \quad \forall f \in \mathcal{W}$$

ie la forme linéaire \mathbf{l} est représentée par le vecteur $\mathbf{K}(\mathbf{l}) \in \mathcal{W}$.

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

On remarque immédiatement que $(\alpha, \beta) \rightarrow \langle \mathbf{K}(\delta_x^\alpha), \mathbf{K}(\delta_y^\beta) \rangle_{\mathcal{W}}$ est une forme linéaire, et donc il existe $\mathbf{K}(x, y) \in \mathcal{M}_m(\mathbb{R})$ telle que

$$(4) \quad \langle \mathbf{K}(\delta_x^\alpha), \mathbf{K}(\delta_y^\beta) \rangle_{\mathcal{W}} = \alpha^T \mathbf{K}(x, y) \beta = \langle \alpha, \mathbf{K}(x, y) \beta \rangle_{\mathbb{R}^d}.$$

Comme $\langle \mathbf{K}(\delta_x^\alpha), \mathbf{K}(\delta_y^\beta) \rangle_{\mathcal{W}} = \delta_x^\alpha(\mathbf{K}(\delta_y^\beta))$ on tire de (4) que

$$(5) \quad \mathbf{K}(\delta_y^\beta)(x) = \mathbf{K}(x, y) \beta.$$

EXERCICE 1. Montrer que si $f = \mathbf{K}(\delta_y^\beta)$ alors $|f|_{\mathcal{W}}^2 = \beta^T \mathbf{K}(x, x) \beta$ et plus généralement, que si $f(x) = \sum_{i \in I} \mathbf{K}(x, x_i) \alpha_i$, alors $|f|_{\mathcal{W}}^2 = \sum_{i, j \in I} \alpha_i^T \mathbf{K}(x_i, x_j) \alpha_j$.

THÉORÈME 1. On suppose que \mathcal{W} est un ENR. Si (P) admet une solution, alors elle est unique et s'écrit $f_* = \mathbf{K}(\sum_{i \in I} \delta_{x_i}^{\alpha_i})$ pour une famille $\alpha = (\alpha_i)_{i \in I}$ de coefficients dans \mathbb{R}^d solution du système linéaire $|\mathbf{I}|m \times |\mathbf{I}|m$:

$$(6) \quad \mathbb{K} \alpha_{\mathbf{I}} = \mathbf{a}_{\mathbf{I}}$$

où \mathbb{K} est la matrice bloc $\mathbb{K} = (\mathbf{K}(x_i, x_j))_{1 \leq i, j \leq n}$, $\alpha_{\mathbf{I}} = (\alpha_i)_{i \in I}$ et $\mathbf{a}_{\mathbf{I}} = (\mathbf{a}_i)_{i \in I}$.

DÉMONSTRATION. Si (P) admet une solution, alors $\mathcal{W}_{x_i, \alpha_i}$ est un fermé non vide et par le théorème de projection, 0 se projette en un point unique f_* . Comme $\mathcal{W}_{x_i, \alpha_i} = f_* + \mathcal{W}_{x_i, 0}$ où $\mathcal{W}_{x_i, 0} = \{f \in \mathcal{W} \mid f(x_i) = 0, i \in I\}$. Il est clair que $\mathcal{W}_{x_i, 0} = \{f \in \mathcal{W} \mid \langle f(x_i), \alpha_i \rangle_{\mathbb{R}^m} = 0 \forall \alpha_i \in \mathbb{R}^m \text{ et } i \in I\}$. D'après (2), on déduit que $\mathcal{W}_{x_i, 0}$ est l'orthogonal de $\text{Vect}\{\mathbf{K}(\delta_{x_i}^{\alpha_i}), i \in I, \alpha_i \in \mathbb{R}^m\}$ et donc $\mathcal{W}_{x_i, 0}^\perp = \text{Vect}\{\mathbf{K}(\delta_{x_i}^{\alpha_i}), i \in I, \alpha_i \in \mathbb{R}^m\}$.

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

$$(7) \quad \sum_{j \in I} \mathbf{K}(x_i, x_j) \alpha_j = \mathbf{a}_i$$

ce qui conduit au système (6). □

REMARQUE 1. (1) La taille du système ne dépend pas de la dimension d mais seulement de m et du nombre de points. Potentiellement d pourrait être de dimension très grande (ce qui n'est pas intéressant pour nous, mais fondamental pour les applications des "méthodes à noyaux" en reconnaissance de formes).

(2) L'inversibilité de la matrice \mathbb{K} est vraie dès que \mathcal{W} contient suffisamment de fonction ie $\Pi : f \rightarrow (f(x_i))_{i \in I}$ est surjective. En effet, si $\mathbb{K} \alpha_{\mathbf{I}} = 0$ alors $\alpha_{\mathbf{I}}^T \mathbb{K} \alpha_{\mathbf{I}} = \sum_{i, j} \alpha_i^T \mathbf{K}(x_i, x_j) \alpha_j = |\mathbf{K}(\sum_{i \in I} \delta_{x_i}^{\alpha_i})|_{\mathcal{W}}^2 = |\sum_{i \in I} \delta_{x_i}^{\alpha_i}|_{\mathcal{W}^*}^2 = 0$. Par suite $\sum_{i \in I} \delta_{x_i}^{\alpha_i}(f) = \sum_{i \in I} \langle \alpha_i, f(x_i) \rangle_{\mathbb{R}^m} = 0$ et donc la surjectivité de Π , on a le résultat.

On verra plus tard, qu'on peut construire W directement en se donnant un noyau explicite $K(x, y)$ pour tout $x, y \in \mathbb{R}^d$. Le théorème précédent montre qu'alors le problème d'interpolation devient un problème linéaire complètement explicite.

EXERCICE 2. Un noyau très courant est le noyau (dit "gaussien") $K_\lambda(x, y) = \exp(-|x-y|^2/\lambda^2)\text{Id}$. Les solutions sont alors du type RBF (Radial Basis Functions). Mettre en oeuvre numériquement la méthode d'interpolation sur quelques exemples en 1D, 2D.

EXERCICE 3. [Interpolation inexacte] On cherche un appariement qui tiennent compte du bruit sur les données en minimisant

$$(8) \quad J(f) \triangleq \frac{1}{2} \|f\|_W^2 + \frac{\gamma}{2} \sum_{i \in I} |f(x_i) - \alpha_i|^2.$$

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

2. Appariement dense de points indexés

C'est le problème que nous avons introduit dans le Chap 0. On a deux familles $\mathbf{x}_I = (x_i)_{i \in I}$ et $\mathbf{y}_I = (y_i)_{i \in I}$ de points dans \mathbb{R}^d que l'on cherche à appairer (lorsque, et c'est un point crucial, les correspondance point à point sont *connues*, cf Fig 2.2.2). On cherche donc $\phi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ telle que $\phi(x_i) = y_i$.

On peut alors en considérant $\mathbf{m} = \mathbf{d}$ et $\mathbf{a}_i = y_i - x_i$ et $\mathbf{u}(x) = \phi(x) - x$ se ramener à un problème d'interpolation. L'espace W est dans ce cas un Hilbert de *champs de vecteurs*. Pour éviter malgré tout la confusion avec le problème d'interpolation sur les fonctions, nous noterons dans ce cas V l'espace de Hilbert sous-jacent (les éléments de V seront notés \mathbf{u} , ou \mathbf{v}). La façon de procéder reste à ce changement de notations près identique.

Bien sûr, rien ne garanti a priori l'inversibilité de $\phi(x) = x + \mathbf{u}(x)$ sauf si $\nabla \mathbf{u}$ reste suffisamment petit.

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

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

$$(9) \quad \|\mathbf{u}\|_V^2 = \alpha_I^T \mathbb{K} \alpha_I = \mathbf{u}_I^T \mathbb{K}^{-1} \mathbf{u}_I.$$

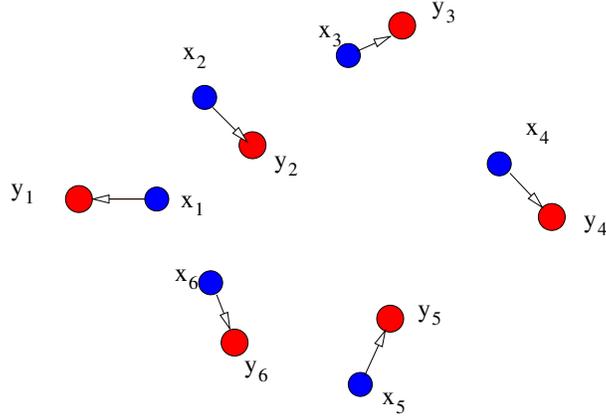


FIG. 1. Exemple de problème d'appariement de points

La matrice \mathbb{K}^{-1} définit donc une métrique locale sur les \mathbf{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^n 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) $(\mathbf{u}_I^k, \lambda^k)$ (resp. (\mathbf{a}_I^k, μ^k)) orthonormés pour la métrique euclidienne. On a alors $\lambda^k = 1/\mu^k$ et $\mathbf{u}_I^k = \mathbf{a}_I^k = \lambda^k \mathbb{K} \mathbf{a}_I^k$.

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

EXERCICE 4. En choisissant quelques configurations de références, afficher les modes propres pour un noyau gaussien et différentes valeurs du paramètre λ .

3. Appariement de nuages de points, appariement de mesures

Référence ([8]).

L'objectif est ici d'enlever la contrainte de correspondances bi-univoque entre les points $\mathbf{x}_i \leftrightarrow \mathbf{y}_i$. Une méthode classique et assez populaire pour traiter ce problème est la méthode du plus proche voisin itéré (ICP pour Iterated Closest Point). Comme nous avons vu en cours, cette méthode souffre de plusieurs défauts partiellement corrigés par les variantes proposées à la version de base (voir [12] pour une comparaison de ces variantes). Nous nous proposons d'introduire ici un point de vue un peu nouveau qui va nous permettre de découvrir toutes les ressources des ENR dans un contexte un peu inattendu.

3.1. Nuages de points comme mesures. On considère pour commencer le cas de deux ensembles de points $\mathbf{x}_I = (\mathbf{x}_i)_{i \in I}$ et $\mathbf{y}_J = (\mathbf{y}_j)_{j \in J}$ de cardinaux éventuellement différents

qu'on imagine avoir été grossièrement recalés par des transformations rigides.

Idée de base : Interpréter les nuages de points comme des mesures ponctuelles

$$(10) \quad \mu_{x_i} = \frac{1}{|I|} \sum_I \delta_{x_i}, \quad \mu_{y_j} = \frac{1}{|J|} \sum_J \delta_{y_j}$$

et construire le problème d'appariement comme le problème plus général d'appariement de mesures.

Comment définir un critère pour quantifier la proximité de deux mesures? Ici les mesures sont de masse 1 et on peut penser à utiliser la distance de Kullback ($K(\mu, \nu) = \int \log(\frac{d\mu}{d\nu}) d\mu$) qui n'est cependant pas adaptée pour des mesures étrangères (il faudrait passer via l'estimation de densité).

3.2. Intégration de fonctions tests et normes duales. Les mesures sont des objets qui permettent d'intégrer ... des fonctions :

$$(11) \quad f \rightarrow (\mu, f) = \int_{\mathbb{R}^d} f d\mu$$

et si $\mu \simeq \nu$, on doit avoir $(\mu, f) \simeq (\nu, f)$. Bien sûr, si $\mu = \mu_{x_i}$, $\nu = \mu_{y_j}$ et si f vaut M sur les x_i et $-M$ sur les y_j , alors $|(\mu, f) - (\nu, f)| = 2M$ qui peut être très grand. Cependant si les nuages sont proches, la fonction f résultante va être très irrégulière. En contraignant la régularité, on obtient une notion plus intéressante qui peut être interprétée comme une norme duale :

Partons donc d'un espace W de fonctions de $C_0(\mathbb{R}^d, \mathbb{R})$ (espace des fonctions continues qui tendent vers 0 en l'infini) qui soit un Banach pour une norme $|\cdot|_W$. On pourrait prendre $W = C_0(\mathbb{R}^d, \mathbb{R})$ lui-même muni de la norme uniforme $|f|_\infty = \sup_{\mathbb{R}^d} |f(x)|$. Considérons alors

$$(12) \quad |\mu - \nu|_{W^*} \triangleq \sup_{f \in W, |f|_W \leq 1} |(\mu, f) - (\nu, f)|$$

Ici, la contrainte $|f|_W \leq 1$ donne un contrôle de la régularité de f et $|\cdot|_{W^*}$ est la norme duale dont nous avons déjà parlé.

EXERCICE 5. Dans le cas où $W = C_0(\mathbb{R}^d, \mathbb{R})$ le dual de $C_0(\mathbb{R}^d, \mathbb{R})$ est l'ensemble des mesures boreliennes signées noté $\mathcal{M}_s(\mathbb{R}^d)$ sur \mathbb{R}^d qui peuvent toutes se décomposer comme différences de 2 mesures *positives finies* étrangères entre elles

$$(13) \quad \mu = \mu_+ - \mu_-$$

Montrer alors que l'on a

$$(14) \quad |\mu|_{\mathcal{M}_s(\mathbb{R}^d)} = \mu_+(\mathbb{R}^d) + \mu_-(\mathbb{R}^d)$$

et montrer que $|\mu_{x_i} - \mu_{y_j}|_{\mathcal{M}_s(\mathbb{R}^d)} = 2$ dès que les nuages sont disjoints.

L'exercice 5 montre que le contrôle de la régularité par une norme uniforme est trop faible ce qui donne une norme sur le dual trop forte (trop facile de séparer les mesures).

Supposons donc maintenant que W est un Hilbert ($\|\cdot\|_W$ est une norme hilbertienne) et que W s'injecte continûment dans $C_0(\mathbb{R}^d, \mathbb{R})$ ie il existe $C > 0$ tel que

$$(15) \quad \|f\|_\infty \leq C\|f\|_W.$$

Notons $i : W \rightarrow C_0(\mathbb{R}^d, \mathbb{R})$ cette injection. Alors, pour les duals on a $i^* : \mathcal{M}_s(\mathbb{R}^d) \rightarrow W^*$, et :

$$(16) \quad \|i^*(\mu)\|_{W^*} \leq C\|\mu\|_{\mathcal{M}_s(\mathbb{R}^d)}$$

EXERCICE 6. (Facile) Montrer l'inégalité (16).

EXERCICE 7. Montrer que si W est un sous-espace dense de $C_0(\mathbb{R}^d, \mathbb{R})$ alors i^* est injective. On identifie alors $\mathcal{M}_s(\mathbb{R}^d)$ à un sous-espace de W^* .

L'inégalité (15) entraîne que si W est un ENR pour un certain noyau K^W , en utilisant l'isométrie $W^* \rightarrow W$ donnée par K^W , on a pour une mesure signée discrète de la forme $\nu = \sum_{h=1}^n \alpha_h \delta_{z_h}$

$$(17) \quad \|\nu\|_{W^*}^2 = \left\| \sum_{h=1}^n \alpha_h \delta_{z_h} \right\|_{W^*}^2 = \|K^W \sum_{h=1}^n \alpha_h \delta_{z_h}\|_W^2 = \sum_{h,h'} \alpha_h \alpha_{h'} K^W(z_h, z_{h'}).$$

Ainsi, en prenant $\nu = \mu_{x_1} - \mu_{y_1}$, on peut calculer explicitement $\|\mu_{x_1} - \mu_{y_1}\|$ en fonction du noyau K^W .

EXERCICE 8. Montrer $\|\mu\|_{W^*}^2 = \int \int g(x)g(y)K^W(x,y)dx dy$ lorsque $d\mu(x) = g(x)dx$ et g est une fonction intégrable sur \mathbb{R}^d pour la mesure de Lebesgue. Donner l'expression générale lorsque $\mu \in \mathcal{M}_s(\mathbb{R}^d)$

REMARQUE 2. Concrètement, on peut tout à fait utiliser un noyau gaussien du type $K^W(x,y) = \exp(-|x-y|/\lambda)^2$ où λ est un paramètre d'échelle.

3.3. Consistance. Un point intéressant dans cette approche utilisant la norme duale d'un ENR W est qu'on peut avoir des résultats d'approximations des mesures générales par des mesures discrètes.

On a le résultat suivant¹ :

Si μ est une mesure de probabilité et si (x_1, \dots, x_n) est obtenue par tirage aléatoire selon

¹ On ne résiste pas au plaisir de donner la preuve qui utilise de façon déguisée des techniques venant tout droit de la partie Apprentissage du Master...

la mesure μ alors, on montre que

$$(19) \quad \mathbb{E} \left(\left| \mu - \frac{1}{n} \sum_{i=1}^n \delta_{x_i} \right|_{W^*} \right) \leq 2 \sqrt{\frac{M}{n}}$$

où $M = \sup_{x \in \mathbb{R}^d} K^W(x, x)$. On peut s'étonner que si K^W est un noyau gaussien de la forme $K^W(x, y) = \exp(-|(x - y)/\lambda|^2)$, la borne ne dépende pas de l'échelle λ . Ceci n'est en fait qu'apparent et apparaît dès que l'on regarde l'erreur *relative* d'approximation :

EXERCICE 9. (suite ex 8) Montrer que si $d\mu(x) = g(x)dx$ avec $g \geq 0$ et $\int g(x)dx = 1$, alors $|\mu|_{W_\lambda^*} \sim (\pi\lambda^2)^d \int g(x)^2 dx$ pour λ au voisinage de 0. Quelles sont les mesures faciles à estimer ?

3.4. Transport de mesures. Puisque nous savons mesurer les écarts entre nuages de points, comment définir alors une déformation de mesures ? Il suffit de transporter la mesure sous l'effet de la déformation de la façon suivante : on définit d'abord la transformation φ associée à un champ de déplacement u par $\varphi(x) = x + u(x)$. On définit alors le transport de $\varphi\nu$ de $\nu \in \mathcal{M}_s(\mathbb{R}^d)$ par φ par

$$(20) \quad (\varphi\nu, f) = (\nu, f \circ \varphi)$$

REMARQUE 3. *Il s'agit simplement de la mesure image de ν par φ et $\varphi\nu \in \mathcal{M}_s(\mathbb{R}^d)$*

DÉMONSTRATION. Si x'_1, \dots, x'_n est un n -échantillon de loi μ indépendant du premier échantillon, alors on a pour tout $f \in W$ tel que $|f|_W \leq 1$,

$$\mathbb{E}_{x'_1, \dots, x'_n} \left(\left(\frac{1}{n} \sum_{i=1}^n f(x'_i) - \frac{1}{n} \sum_{i=1}^n f(x_i) \right)^2 \right) = ((\mu, f) - \frac{1}{n} \sum_{i=1}^n f(x_i))^2 + \mathbb{E} \left(\left(\frac{1}{n} \sum_{i=1}^n f(x'_i) - (\mu, f) \right)^2 \right)$$

d'où $((\mu, f) - \frac{1}{n} \sum_{i=1}^n f(x_i))^2 \leq \mathbb{E}_{x'_1, \dots, x'_n} (|\frac{1}{n} \sum_{i=1}^n \delta_{x'_i} - \frac{1}{n} \sum_{i=1}^n \delta_{x_i}|_{W^*}^2)$ puis en prenant le sup sur de tels f et en prenant l'espérance par rapport aux x_i , on obtient

$$(18) \quad \mathbb{E} \left(\left| \mu - \frac{1}{n} \sum_{i=1}^n \delta_{x_i} \right|_{W^*}^2 \right) \leq \mathbb{E} \left(\left| \frac{1}{n} \sum_{i=1}^n (\delta_{x'_i} - \delta_{x_i}) \right|_{W^*}^2 \right)$$

Or comme la loi jointe des x et x' est invariante sous l'action des échanges d'une variable x_i avec x'_i , on peut introduire une suite de variables η_i iid indépendantes des x et x' telles que $P(\eta_i = 1) = P(\eta_i = -1) = 1/2$. On a ainsi $\mathbb{E} \left(\left| \frac{1}{n} \sum_{i=1}^n (\delta_{x'_i} - \delta_{x_i}) \right|_{W^*}^2 \right) = \mathbb{E} \left(\left| \frac{1}{n} \sum_{i=1}^n \eta_i (\delta_{x'_i} - \delta_{x_i}) \right|_{W^*}^2 \right)$. Par ailleurs lorsque les x et les x' sont fixés, on a

$$\mathbb{E}_{\eta_1, \dots, \eta_n} \left(\left| \frac{1}{n} \sum_{i=1}^n \eta_i (\delta_{x'_i} - \delta_{x_i}) \right|_{W^*}^2 \right) = \frac{1}{n^2} \sum_{i=1}^n |\delta_{x_i} - \delta_{x'_i}|_{W^*}^2 \leq \frac{2}{n^2} \sum_{i=1}^n (|\delta_{x_i}|_{W^*}^2 + |\delta_{x'_i}|_{W^*}^2)$$

Il suffit pour terminer de remarquer que $|\delta_x|_{W^*}^2 = K^V(x, x) \leq M$ pour avoir le résultat. \square

Regardons le cas particulier important $\nu = \sum_{h=1}^n \alpha_h \delta_{z_h}$, on a alors

$$(21) \quad (\varphi\nu, f) = \sum_{h=1}^n \alpha_h f(\varphi(z_h))$$

d'où

$$(22) \quad \varphi\nu = \sum_{h=1}^n \alpha_h \delta_{\varphi(z_h)}$$

ce qui correspond au déplacement des points avec leurs poids qui restent inchangés.

EXERCICE 10. Vérifier la propriété de *transport du support* : $\text{supp}(\varphi\nu) = \varphi(\text{supp}(\nu))$ lorsque φ est un homéomorphisme² (continu d'inverse continue). Vérifier qu'alors $\nu \rightarrow \varphi\nu$ conserve la norme dans $\mathcal{M}_s(\mathbb{R}^d)$ (en particulier la masse est conservée lorsque ν est une mesure positive).

Par suite si μ est une mesure concentrée sur une sous-variété M , $\varphi\mu$ est une mesure concentrée sur l'image $\varphi(M)$. *L'approche est donc bien plus large que l'appariement de nuages de points*

Cependant, dans le cas d'une mesure diffuse $d\nu(x) = g(x)dx$, lorsque φ est un C^1 difféomorphisme, on calcule par changement de variables

$$d(\varphi\nu)(y) = g(\varphi^{-1}(y)) |d_y \varphi^{-1}| dy,$$

où $|d_y \varphi^{-1}|$ est le jacobien de φ^{-1} . Ainsi comme dans tout transport de masse, la densité en $y = \varphi(x)$ est le produit de la densité transportée et d'un terme de compression venant de φ .

En conséquence, si $g = \mathbf{1}_U$ ou U est un domaine borné, alors $\varphi\nu$ est supportée par $\varphi(U)$ mais *n'a pas de raison d'être une mesure uniforme sur $\varphi(U)$* . Ceci se généralise pour le cas des sous-variétés et en particulier des courbes :

EXERCICE 11. Calculer le transport de masse pour une mesure uniforme définie sur une courbe C^1 par morceau et de longueur finie.

En particulier, pour la comparaison de courbes, si μ_γ est la mesure uniforme sur une courbe $\gamma : [0, 1] \rightarrow \mathbb{R}^d$, on peut construire évidemment deux transformations φ et φ' telles que $\varphi(\gamma([0, 1])) = \varphi'(\gamma([0, 1]))$ mais $\varphi\mu_\gamma \neq \varphi'\mu_\gamma$ (il suffit que $\varphi \circ \gamma$ et $\varphi' \circ \gamma$ génèrent deux paramétrages différents de la courbe image).

²Notons que ceci ne peut être garanti dans le cadre des petites déformations que si u varie suffisamment lentement

REMARQUE 4. *Le transport de mesures ne traite en particulier pas intrinsèquement les surfaces comme des ensembles géométriques indépendamment de leur paramétrage. Il faut le savoir pour bien comprendre comment utiliser cependant un tel algorithme pour l'appariement de sous-variétés. La déformation doit être capable non seulement d'amener le support de la première surface au voisinage de la seconde mais aussi de reproduire localement la densité de masses de la mesure but décrite souvent comme une approximation d'une mesure uniforme sur la surface but (voir ci-dessous). Il ne faut pas que les déformations soient trop fortes.*

3.5. Calcul de la solution. Dans le cas de deux mesures discrètes, on a en exprimant φ en fonction du déplacement \mathbf{u} :

$$(23) \quad \gamma |\varphi \mu_{x_i} - \mu_{y_i}|_W^2 = g(\mathbf{u}_I) \text{ avec } \mathbf{u}_I = (\mathbf{u}(x_i))_{i \in I}$$

et l'approche variationnelle consiste à minimiser

$$(24) \quad J(\mathbf{u}) = \frac{1}{2} |\mathbf{u}|_V^2 + g(\mathbf{u}_I).$$

En utilisant le principe de réduction du cours précédent, le problème se ramène à résoudre d'abord un problème de dimension finie :

$$(25) \quad \min_{\alpha_I} J_I(\alpha_I) \text{ où } J_I(\alpha_I) \triangleq \frac{1}{2} \alpha_I^* K_{x_i, x_i}^V \alpha_I + g(K_{x_i, x_i}^V \alpha_I)$$

où K^V est le noyau associé à V et K_{x_i, x_i}^V est la matrice bloc

$$(26) \quad K_{x_i, x_i}^V = (K^V(x_i, x_j))_{i, j \in I} \text{ où les blocs } K^V(x, y) \in \mathcal{M}_d(\mathbb{R}).$$

A partir de la solution α_{I*} de (25), on construit une solution minimisante globale \mathbf{u}_* de J par

$$(27) \quad \mathbf{u}_*(x) = \sum_{i \in I} K^V(x, x_i) \alpha_{i,*}$$

qui est définie *partout*.

REMARQUE 5. *Le problème bien que de dimension finie est non-linéaire. La minimisation doit être faite par un algorithme de gradient sur α_I . On obtient facilement (le vérifier tout de même) que $\nabla J_I(\alpha_I) = K_{x_i, x_i}^V (\alpha_I + \frac{\partial g}{\partial \mathbf{u}_I})$.*

4. Appariements de courants

Comme nous l'avons vu, l'appariement de mesures n'est pas tout à fait adapté à un appariement géométrique de courbes, de surfaces ou plus généralement de sous-variétés. Il existe cependant 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".

4.1. p-vecteurs et produit extérieur.

4.1.1. *Définition.* Soit E un \mathbb{R} espace vectoriel de dimension fini 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 12. 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 parallélotopes 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 13. En écrivant $u_i = \sum_{j=1}^d u_{ij} e_j$, vérifier que en posant

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

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

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

EXERCICE 14. Si (e_1, \dots, e_p) est une b.o.n de E , vérifier alors que $(e_{i_1} \wedge \dots \wedge e_{i_p})_{i_1 < \dots < i_p}$ est une de $\Lambda^p E$.

4.1.3. *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 $*(e_{i_1} \wedge \dots \wedge e_{i_p}) \mapsto e_{j_1} \wedge \dots \wedge e_{j_{d-p}}$ telle que $(e_{i_1}, \dots, e_{i_p}, e_{j_1}, \dots, e_{j_{d-p}})$ est une base positivement orientée. Dans le case $p = d - 1$, $\Lambda^{d-1} E \xrightarrow{w} \Lambda^1 E = E$.

REMARQUE 6. – En particulier, dans le cas de la géométrie 3D ($E = \mathbb{R}^3$), alors pour $\mathbf{a}, \mathbf{b} \in E$, $*(\mathbf{a} \wedge \mathbf{b})$ est exactement le vecteur $\mathbf{c} \in \mathbb{R}^3$ tel que $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ est une base orthogonale positivement orientée et tel que $|\mathbf{c}|$ est l'aire du parallélogramme défini par \mathbf{a} et \mathbf{b} . Ainsi, $*(\mathbf{a} \wedge \mathbf{b})$ 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 15. Vérifier que $\langle u_1 \wedge \dots \wedge u_p, v_1 \wedge \dots \wedge v_p \rangle = \langle *(u_1 \wedge \dots \wedge u_p), *(v_1 \wedge \dots \wedge v_p) \rangle$ ie que $*$ est une isométrie de $\Lambda^p E$ dans $\Lambda^{d-p} E$.

4.2. Formes différentielles.

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

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

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

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

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

REMARQUE 7. On identifie ainsi $(\Lambda^p E)^*$ et $\Lambda^p E^*$

EXERCICE 16. En écrivant $u_1 \wedge \dots \wedge u_p = \sum_{i_1 < \dots < i_p} dx^{i_1} \wedge \dots \wedge dx^{i_p} (u_1 \wedge \dots \wedge u_p) e_{i_1} \wedge \dots \wedge e_{i_p}$, déduire de l'exercice 13 que

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

4.2.2. *Exemples.*

– Pour $p = d$, $\dim(\Lambda^d E)^* = 1$, $\omega(u_1 \wedge \dots \wedge u_d) = \det(u_1, \dots, u_d) \omega(e_1 \wedge \dots \wedge e_d)$. Toutes les formes linéaires sont donc dans ce cas proportionnelles au déterminant (forme volume) qui lui-même s'identifie avec $dx^1 \wedge \dots \wedge dx^n$.

- Dans le cas $d = 3$, pour $p = 2$, $(\wedge^d E)^*$ peut s'interpréter comme une mesure de *flux*,
- Dans le cas $p = d - 1$, pour toute forme linéaire $\omega \in \wedge^{p-1} E$, il existe $\eta \in \wedge^{p-1} E$ telle que $\omega(\mathbf{u}_1 \wedge \cdots \wedge \mathbf{u}_p) = \langle \eta, (\mathbf{u}_1 \wedge \cdots \wedge \mathbf{u}_p) \rangle$. En utilisant l'isométrie $*$ (cf exercice 15), on a $\omega(\mathbf{u}_1 \wedge \cdots \wedge \mathbf{u}_p) = \langle *\eta, *(\mathbf{u}_1 \wedge \cdots \wedge \mathbf{u}_p) \rangle$ qui s'interprète géométriquement comme le flux du vecteur $*\eta \in E$ à travers l'élément d'hypersurface définie par $\mathbf{u}_1 \wedge \cdots \wedge \mathbf{u}_p$.
- Dans le cas $p = 1$, alors de façon cette fois immédiate, en remarquant que $(\wedge^1 E)^* = E^*$, on a pour tout $\omega \in E^*$, il existe $\mathbf{v} \in E$ telle que $\omega(\mathbf{u}) = \langle \mathbf{v}, \mathbf{u} \rangle$ pour tout $\mathbf{u} \in \wedge^1 E = E$.
- Enfin pour $p = 0$, on identifie $(\wedge^0 E)^* = \mathbb{R}$

DÉFINITION 2. Pour tout $\omega \in (\wedge^p E)^*$, on définit la norme duale

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

REMARQUE 8. La norme duale est issue d'un produit scalaire sur $(\wedge^p(E))^*$ pour lequel $(dx^{i_1} \wedge \cdots \wedge dx^{i_p})_{i_1 < \cdots < i_p}$ est une b.o.n.

REMARQUE 9. Pour tout $\omega \in \wedge^p E$, si on définit $\tilde{\omega}(\mathbf{u}_1, \dots, \mathbf{u}_p) \triangleq \omega(\mathbf{u}_1 \wedge \cdots \wedge \mathbf{u}_p)$ ie $\tilde{\omega} \triangleq \omega \circ \omega$, alors $\tilde{\omega}$ est forme multilinéaire alternée. Il est d'usage d'identifier $\tilde{\omega}$ et ω en écrivant parfois $\omega(\mathbf{u}_1, \dots, \mathbf{u}_p)$ au lieu de $\omega(\mathbf{u}_1 \wedge \cdots \wedge \mathbf{u}_p)$.

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

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

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

- REMARQUE 10. – 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 11. 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 $\wedge^p T_x^*$ en $x \in M$.

4.3. 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.

4.3.1. *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 $\Omega_0^1(\mathbb{R}^d)$:

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

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 égalité 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 17.

- (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 $\omega \in \Omega_0^1(\mathbb{R}^d)$.

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

$$(28) \quad |S| \triangleq \int_{\mathbf{U}} \left| \frac{\partial \gamma}{\partial s} \wedge \frac{\partial \gamma}{\partial t} \right| ds dt < \infty$$

où on rappelle que $\left| \frac{\partial \gamma}{\partial s} \wedge \frac{\partial \gamma}{\partial t} \right| = \sqrt{\det(\text{Gram}(\frac{\partial \gamma}{\partial s}, \frac{\partial \gamma}{\partial t}))}$ avec $\text{Gram}(\mathbf{u}_1, \dots, \mathbf{u}_p) \triangleq (\langle \mathbf{u}_i, \mathbf{u}_j \rangle)_{1 \leq i, j \leq p}$ la matrice de Gram³.

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

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

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

$$(29) \quad T_{\tilde{\gamma}} = \int_{\mathbf{V}} \omega \left(\frac{\partial \tilde{\gamma}}{\partial \mathbf{a}} \wedge \frac{\partial \tilde{\gamma}}{\partial \mathbf{b}} \right) d\mathbf{a} d\mathbf{b} = \int_{\mathbf{U}} \omega \left(\frac{\partial \gamma}{\partial s} \wedge \frac{\partial \gamma}{\partial t} \right) ds dt = T_\gamma(\omega)$$

EXERCICE 18. Montrer l'égalité (29) en montrant que $\frac{\partial \tilde{\gamma}}{\partial \mathbf{a}} \wedge \frac{\partial \tilde{\gamma}}{\partial \mathbf{b}} = \left(\frac{\partial \gamma}{\partial s} \wedge \frac{\partial \gamma}{\partial t} \right) \circ \psi \underbrace{J\psi}_{>0} = \left(\frac{\partial \gamma}{\partial s} \wedge \frac{\partial \gamma}{\partial t} \right) \circ \psi |J\psi|$ 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^4 . On peut donc écrire

$$(30) \quad T_S(\omega) = \int_S \omega \triangleq T_\gamma(\omega).$$

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 $\delta_x^{\mathbf{u} \wedge \mathbf{v}} : \omega \mapsto \omega_x(\mathbf{u} \wedge \mathbf{v})$ qui représente un élément "ponctuel" de surface orientée positionné en x

³La valeur de $|S|$ peut être prise comme l'aire de S . En effet, en découpant \mathbf{U} sur un quadrillage en petits carrés C_i de coté $\Delta s \times \Delta t$, 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 (28).

⁴En 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.

d'aire $|\mathbf{u} \wedge \mathbf{v}|$ dans le plan engendré par (\mathbf{u}, \mathbf{v}) . Plus généralement, on peut voir un 2-courant comme une mesure à valeurs éléments de surface.

EXERCICE 19. Si $(X_i)_{i \geq 1}$ est une suite iid de loi uniforme sur \mathbf{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 $n \rightarrow \infty$.

4.4. Transport de courants. Si $\phi : \mathbb{R}^d \mapsto \mathbb{R}^d$ et $\omega \in \Omega_0^p(\mathbb{R}^d)$, on peut définir $\phi(\omega)$ par

$$(31) \quad \phi(\omega)_x(\mathbf{u}_1 \wedge \cdots \wedge \mathbf{u}_p) \triangleq \omega_{\phi(x)}(d\phi_x(\mathbf{u}_1) \wedge \cdots \wedge d\phi_x(\mathbf{u}_p)).$$

EXERCICE 20.

(1) Montrer que l'on a $\phi_1(\phi_2(\omega)) = (\phi_2 \circ \phi_1)(\omega)$.

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

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

$$(32) \quad \phi(T)(\omega) \triangleq T(\phi(\omega))$$

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

$$(33) \quad \phi_1(\phi_2(T)) = (\phi_1 \circ \phi_2)(T).$$

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

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

ce qui correspond au transport de la ligne par ϕ .

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

$$(34) \quad \begin{aligned} \phi(T_\gamma)(\omega) = T_\gamma(\phi(\omega)) &= \int_{\mathbf{U}} \omega_{\phi \circ \gamma(s,t)}(d\phi_\gamma(\frac{\partial \gamma}{\partial s}) \wedge d\phi_\gamma(\frac{\partial \gamma}{\partial t})) ds dt \\ &= \int_{\mathbf{U}} \omega_{\phi \circ \gamma(s,t)}(\frac{\partial \phi \circ \gamma}{\partial s} \wedge \frac{\partial \phi \circ \gamma}{\partial t}) ds dt = T_{\phi \circ \gamma}(\omega) \end{aligned}$$

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

4.5. ENR sur les courants. Si $W \hookrightarrow \Omega_0^p(\mathbb{R}^d)$ est une injection continue, W se 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 :

$$(35) \quad |T|_{W^*} = \sup_{|\omega|_W=1} T(\omega).$$

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

4.5.1. *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

$$(36) \quad \begin{aligned} \langle T_L, T_L \rangle_{W^*} &\stackrel{\text{Kisom}}{=} \langle KT_L, KT_L \rangle_W = T_L(KT_L) \\ &\stackrel{\text{Def de } K}{=} \int_0^1 \langle K\delta_\gamma', KT_L \rangle_W dt = \int_0^1 \int_0^1 \langle K\delta_{\gamma(s)}^{\gamma'(s)}, K\delta_{\gamma(t)}^{\gamma'(t)} \rangle_W ds dt \\ &= \int_0^1 \int_0^1 \gamma'(s)^T K(\gamma(s), \gamma(t)) \gamma'(t) ds dt = \int_{L \times L} \mathbf{u}(\mathbf{q})^T K(\mathbf{q}, \mathbf{q}') \mathbf{u}(\mathbf{q}') d\mathbf{q} d\mathbf{q}', \end{aligned}$$

où en tout point de \mathbf{q} de L , $\mathbf{u}(\mathbf{q})$ est le vecteur tangent orienté unitaire et $d\mathbf{q}$ 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.

4.5.2. *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

$$(37) \quad |T_S|_{W^*}^2 = \int_U \int_U (\frac{\partial \gamma}{\partial s} \wedge \frac{\partial \gamma}{\partial t})^T K(\gamma(s, t), \gamma(s', t')) \frac{\partial \gamma}{\partial s'} \wedge \frac{\partial \gamma}{\partial t'} ds dt ds' dt'$$

Encore, une fois il faut bien remarquer que si l'expression de droite dans (37) 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 $\mathbf{q} \in S$, on définit la normale orientée $\mathbf{n}(\mathbf{q})$, on a

$$(38) \quad |T_S|_{W^*}^2 = \int_{S \times S} \mathbf{n}(\mathbf{q})^T K(\mathbf{q}, \mathbf{q}') \mathbf{n}(\mathbf{q}') d\mathbf{q} d\mathbf{q}'.$$

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

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

4.6. 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 12. (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 $T_i = (\mathbf{a}_i, \mathbf{b}_i, \mathbf{c}_i)$ 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

$$\tilde{T}_S = \sum_{i \in I} \delta_{x_i}^{\overrightarrow{\mathbf{a}_i \mathbf{b}_i} \wedge \overrightarrow{\mathbf{a}_i \mathbf{c}_i}}$$

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 21. Vérifier que $\overrightarrow{\mathbf{a}_i \mathbf{b}_i} \wedge \overrightarrow{\mathbf{a}_i \mathbf{c}_i} = \overrightarrow{\mathbf{b}_i \mathbf{c}_i} \wedge \overrightarrow{\mathbf{b}_i \mathbf{a}_i} = \overrightarrow{\mathbf{c}_i \mathbf{a}_i} \wedge \overrightarrow{\mathbf{c}_i \mathbf{b}_i}$. En déduire que l'on peut échanger les rôles de \mathbf{a}_i , \mathbf{b}_i et \mathbf{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 $(\mathbf{a}_i, \mathbf{b}_i, \mathbf{c}_i)$ 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 $\mathbf{u} \in V$ une fonctionnelle du type :

$$(39) \quad J(\mathbf{u}) \triangleq \frac{1}{2} |\mathbf{u}|_V^2 + \frac{\gamma}{2} |\phi^{\mathbf{u}}(\tilde{T}_{S_0}) - \tilde{T}_{S_1}|_{W^*}^2,$$

avec $\phi^{\mathbf{u}}(\mathbf{x}) = \mathbf{x} + \mathbf{u}(\mathbf{x})$. L'attache au données ne dépend en fait que de $(\mathbf{u}(x_i), d\mathbf{u}_{\mathbf{a}_i}, d\mathbf{u}_{\mathbf{b}_i}, d\mathbf{u}_{\mathbf{c}_i})_{i \in I}$. On peut montrer alors que les solutions sont à chercher dans un espace de dimension fini un peut 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 [7] :

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

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

EXERCICE 22. Mettre en place l'approche précédente pour la comparaison de deux lignes en choisissant un noyau pour les espaces V et W . Comparer une approche par minimisation de landmarks, par l'approche mesure et enfin par l'approche courant. On pourra se restreindre à $d = 2$.

5. Appariement d'images, cas général

Nous allons séparer deux problèmes qui ne sont pas tout à fait équivalents.

5.1. Appariement image-modèle. C'est le problème de l'appariement d'une observation *digitale* I_{obs} définie sur une grille de pixels Λ avec un *modèle numérique continu* $\mathbf{m}_0 : \mathbb{R}^d \rightarrow \mathbf{R}$ (qui est une image en niveau de gris définie potentiellement sur *l'espace entier*, en utilisant par exemple une forme paramétrique sur une base de fonctions). On note $(x_s)_{s \in \Lambda}$ la famille donnant pour chaque pixel $s \in \Lambda$ sa position x_s dans \mathbb{R}^d .

Dans ce contexte, une solution variationnelle pour le problème d'appariement peut être obtenu en minimisant une fonctionnelle du type :

$$(40) \quad J(\mathbf{u}) = \frac{1}{2} |\mathbf{u}|_V^2 + \gamma \sum_{s \in \Lambda} (\mathbf{m}_0(x_s - \mathbf{u}(x_s)) - I_{\text{obs}}(s))^2$$

où plus généralement

$$(41) \quad J(\mathbf{u}) = \frac{1}{2} |\mathbf{u}|_V^2 + \gamma \sum_{s \in \Lambda} \rho_s(I_{\text{obs}}(s), \mathbf{m}_0(x_s - \mathbf{u}(x_s))).$$

On remarque que dans ce cas, comme \mathbf{m}_0 est défini sur l'espace entier, il n'y a pas de problème pour parler de $\mathbf{m}_0(x_s - \mathbf{u}(x_s))$ *quelque soit la valeur de $\mathbf{u}(x_s)$* et J est bien définie sur tout $\mathbf{u} \in V^5$

⁵Evidemment, il n'en est pas moins vrai que $\psi(x) \triangleq x - \mathbf{u}(x)$ peut ne pas être injective et deux régions différentes de l'images peuvent être appariées au même endroit dans le modèle numérique (problème des repliement de l'appariement).

En notant $\mathbf{u}_\Lambda = (\mathbf{u}(x_s))_{s \in \Lambda}$ et $\mathbf{g}(\mathbf{u}_\Lambda) \triangleq \gamma \sum_{s \in \Lambda} \rho_s(I_{\text{obs}}(s), \mathbf{m}_0(x_s - \mathbf{u}(x_s)))$, on retrouve un problème du même type que précédemment :

$$(42) \quad J(\mathbf{u}) = \frac{1}{2} |\mathbf{u}|_V^2 + \mathbf{g}(\mathbf{u}_\Lambda)$$

qui peut être traité de la même façon, en se ramenant à la dimension finie. Notons encore qu'ici \mathbf{g} est non-linéaire et qu'il faut procéder par descente de gradient.

On peut se poser cependant la question d'une formulation plus générale où toutes les valeurs de \mathbf{u} interviendraient dans \mathbf{g} :

$$(43) \quad J(\mathbf{u}) = \frac{1}{2} |\mathbf{u}|_V^2 + \mathbf{g}(\mathbf{u}).$$

Dans ce cas, on ne peut se ramener à la dimension finie pour montrer l'existence de solution mais on peut utiliser le résultat classique suivant d'existence d'un minimum en dimension infinie :

THÉORÈME 2. *Si $\mathbf{g} \geq 0$ et semi-continue inférieurement pour la topologie faible⁶ sur \mathbf{g} , alors J donné par (43) atteint son minimum.*

DÉMONSTRATION. La preuve repose sur les deux faits suivants : dans un Hilbert séparable, de toute suite bornée on peut extraire une sous suite faiblement convergente ; l'application $\mathbf{u} \rightarrow |\mathbf{u}|_V$ est sci pour la topologie faible (cf [5], ou vérification facile à la main dans notre cas en prenant une base hilbertienne). Par suite J est sci, et si (\mathbf{u}_n) est une suite minimisante, $\frac{|\mathbf{u}_n|_V^2}{2} \leq J(\mathbf{u}_0)$, donc (\mathbf{u}_n) est bornée. Si (\mathbf{v}_n) est une sous-suite extraite convergente pour la topologie faible vers \mathbf{u}_* , on a $J(\mathbf{u}_*) \leq \liminf J(\mathbf{v}_n) = \inf J$. \square

5.2. Appariement image-image. C'est le problème de l'appariement cette fois-ci de *deux images digitales* I_0 et I_1 définies sur Λ et qui jouent des rôles semblables. C'est un cas pratique important. On peut se replacer dans le cadre précédent en décidant que l'une joue le rôle de l'observation I_{obs} et l'autre de modèle \mathbf{m}_0 . Le passage $I_0 \rightarrow \mathbf{m}_0$ n'est pas trivial à cause des problèmes sous-jacents d'*interpolation* entre les pixels (en particulier en présence de bruit) et d'*extrapolation* en dehors de la grille. Lorsque les bords de I_0 sont homogènes, le problème d'extrapolation est résolu facilement par prolongement constant, dans les autres cas, il faut se rabattre sur des solutions partielles en imposant des conditions de Neumann sur les pixels du bords de I_0 ⁷.

⁶ On rappelle ici que (\mathbf{u}_n) conserve faiblement vers \mathbf{u} dans V (noté $\mathbf{u}_n \rightharpoonup \mathbf{u}$) si $(\eta, \mathbf{u}_n) \rightarrow (\eta, \mathbf{u})$ pour toute forme linéaire continue η de V^* . Dire que \mathbf{g} est sci pour la topologie faible revient alors à dire que $\mathbf{g}(\mathbf{u}) \leq \liminf \mathbf{g}(\mathbf{u}_n)$ dès que $\mathbf{u}_n \rightharpoonup \mathbf{u}$

⁷Notons que de telles conditions ne permettent pas d'imposer que $x_s + \mathbf{u}(x_s)$ reste dans le domaine de l'image lorsque les déformations deviennent importantes

6. Les ENR en brefs

Les paragraphes précédents nous ont permis de nous rendre compte que les ENR peuvent jouer un rôle important sur de multiples aspects dans modèles de petites déformations. Il joueront aussi un rôle pour les grandes déformations et il est temps maintenant d'en explorer un peu systématiquement les propriétés. Cette section regroupe celles qui nous seront le plus utiles.

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.⁸

6.1. Qu'est-ce qu'un ENR ?. Un ENR H sera pour nous un espace de Hilbert de fonctions $\mathbf{h} : \mathcal{X} \rightarrow \mathbb{R}^m$ tel que pour tout $\mathbf{x} \in \mathcal{X}$ et tout $\boldsymbol{\alpha} \in \mathbb{R}^m$, la fonction d'évaluation $\delta_{\mathbf{x}}^{\boldsymbol{\alpha}} : \mathbf{h} \rightarrow \langle \mathbf{h}(\mathbf{x}), \boldsymbol{\alpha} \rangle_{\mathbb{R}^m}$ soit une forme linéaire continue pour la norme de H . En considérant l'isométrie $K : H^* \rightarrow H$ donnée par le théorème de représentation de Riesz, on définit pour tous $\mathbf{x}, \mathbf{y} \in \mathcal{X}$, une matrice $K(\mathbf{x}, \mathbf{y}) \in \mathcal{M}_m(\mathbb{R})$ unique telle que

$$(44) \quad \langle K\delta_{\mathbf{x}}^{\boldsymbol{\alpha}}, K\delta_{\mathbf{y}}^{\boldsymbol{\beta}} \rangle_H = \boldsymbol{\alpha}^* K(\mathbf{x}, \mathbf{y}) \boldsymbol{\beta}$$

pour tous $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}^m$.⁹ On appelle alors noyau associé à H , la fonction $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{M}_m(\mathbb{R})$ définie par (44).¹⁰

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

$$(45) \quad K(\mathbf{x}, \mathbf{y}) = K(\mathbf{y}, \mathbf{x})^*$$

et pour toute famille finie $(\mathbf{x}_i)_{1 \leq i \leq n}$ de \mathcal{X} et toute famille $(\boldsymbol{\alpha}_i)_{1 \leq i \leq n}$ de vecteurs de \mathbb{R}^m :

$$(46) \quad \sum_{1 \leq i, j \leq n} \boldsymbol{\alpha}_i^* K(\mathbf{x}_i, \mathbf{x}_j) \boldsymbol{\alpha}_j \geq 0.$$

Lorsqu'un noyau $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{M}_m(\mathbb{R})$ vérifie (46) et (45), on dit que c'est un *noyau positif*¹¹.

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

⁸On pourra consulter la thèse de J. Glaunes [7] pour un traitement très complet des ENR dans le cadre qui nous intéresse ici, ou encore le papier assez complet [9] consultable ici

⁹Il suffit de vérifier que $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \rightarrow \langle K\delta_{\mathbf{x}}^{\boldsymbol{\alpha}}, K\delta_{\mathbf{y}}^{\boldsymbol{\beta}} \rangle_H$ est une application bilinéaire

¹⁰Notons ici un l'abus de notation consistant à identifier opérateur et noyau

¹¹ou plutôt un noyau vectoriel positif

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.¹²

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 (48)).

6.3. Comment construire des noyaux ? Il y a de nombreuses façons et cela constitue un sujet de recherche actif en particulier lorsque l'espace \mathcal{X} est un espace de données non conventionnel (séquences de longueurs variables, arbres, etc... (nous renvoyons au cours dans ce Master de J.-P. Vert ou plus généralement au site <http://www.kernel-machines.org/>)).

La première façon de construire des noyaux est de partir de noyaux connus et d'utiliser les recettes suivantes

- (1) $K(x, y) = f(x)^*f(y)$ 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 (ie le produit coordonnée par coordonnée des matrices¹³) de $K_1(x, y)$ avec $K_2(x, y)$ ie

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

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

¹²La preuve consiste à considérer l'év H_f des $h : \mathcal{X} \rightarrow \mathbb{R}^m$ qui se représentent sous la forme d'une somme finie du type :

$$(47) \quad h(x) = \sum_{1 \leq i \leq n} K(x, x_i)\alpha_i$$

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

$$(48) \quad \langle g, h \rangle_H \triangleq \sum_{i,j} \alpha_i^* K(x_i, y_j)\beta_j = \sum_j g(y_j)^* \beta_j = \sum_i \alpha_i^* h(x_i)$$

définit un produit scalaire sur H_f (notons que les deux dernières égalités de (48) 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 .

¹³avec des notations "Scilab" ou "Matlab", on a $A \circ B = A .* B$

- (5) Si $K : \mathcal{Y} \times \mathcal{Y} \rightarrow \mathcal{M}_m(\mathbb{R})$ est un noyau et $\phi : \mathcal{X} \rightarrow \mathcal{Y}$, alors $K_\phi : \mathcal{X} \times \mathcal{X}$ défini par $K_\phi(\mathbf{x}, \mathbf{y}) = K(\phi(\mathbf{x}), \phi(\mathbf{y}))$ 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 23. Montrer le point (1). Que vaut l'ENR associé à K ?

EXERCICE 24. 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 $A = \mathbf{u}^* \mathbf{u}$ et $B = \mathbf{v}^* \mathbf{v}$ 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, $K(\mathbf{x}, \mathbf{y}) = k(\mathbf{x}, \mathbf{y}) \text{Id}_d$, il suffit de partir d'un noyau fonctionnel $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$.

6.3.1. *Exemple du noyau gaussien.* Vérifions par exemple que

$$(49) \quad K(\mathbf{x}, \mathbf{y}) = \exp(-|\mathbf{x} - \mathbf{y}|^2) \text{Id}_d$$

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

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

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

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

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

$$(50) \quad K(\mathbf{x}, \mathbf{y}) = \rho(\mathbf{y} - \mathbf{x})$$

engendrent des ENR pour lesquels les opérateurs de translations $\tau : H \rightarrow H$ définis par $\tau \mathbf{h}(\mathbf{x}) = \mathbf{h}(\mathbf{x} + \tau)$ sont des isométries¹⁴ :

$$\langle \tau \mathbf{g}, \tau \mathbf{h} \rangle_H = \langle \mathbf{g}, \mathbf{h} \rangle_H$$

¹⁴cette propriété est intéressante pour construire des méthodes d'appariement qui sont invariante par translation

EXERCICE 25. Vérifier que réciproquement, si les translations sont des isométries, alors K est invariant par translation.

La question est alors du choix de ρ dans (50). 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 $\rho : \mathbb{R}^d \rightarrow \mathcal{M}_m(\mathbb{R})$ ¹⁵ définie un noyau positif invariant par translation ssi sa transformée de Fourier $\widehat{\rho}$ est telle que $\widehat{\rho}(\omega)$ est hermitienne positive pour tout $\omega \in \mathbb{R}^m$.

En prenant $\rho(\mathbf{u}) = \exp(-|\mathbf{u}|^2/2)\text{Id}_m$, on obtient $\widehat{\rho}(\xi) = \sqrt{2\pi}^m \exp(-|\xi|^2/2)\text{Id}_m$ et le théorème de Bochner nous permet de retrouver la positivité du noyau gaussien.

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

$$\mathbf{R}h(\mathbf{x}) = h(\mathbf{R}\mathbf{x})$$

est une isométrie, alors on montre que $K(\mathbf{x}, \mathbf{y}) = \rho(|\mathbf{y} - \mathbf{x}|)$ ie ne dépend que de la norme de $\mathbf{x} - \mathbf{y}$. 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 $\rho(r) = \int \exp(-r^2\mathbf{u}^2)d\mu(\mathbf{u})$ où μ est une mesure borélienne positive finie sur \mathbb{R}_+ , alors $K(\mathbf{x}, \mathbf{y}) = \rho(|\mathbf{x} - \mathbf{y}|)$ 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 26. Montrer que $\rho(r) = 1/(1+r^2)$ définit un noyau positif (bonne alternative au noyau gaussien en pratique)

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

6.4.1. *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 $C_0(\Omega, \mathbb{R}^m)$ l'espace des applications h continues de Ω dans \mathbb{R}^m qui tendent vers 0 lorsque

¹⁵on suppose que ρ est intégrable, ainsi que sa transformée de Fourier et que ρ s'obtient par transformée de Fourier inverse.

$x \rightarrow \partial\Omega$ ¹⁶. Muni de la norme uniforme $\| \cdot \|_\infty$ sur Ω , c'est un Banach. On a le résultat suivant¹⁷ :

Si le noyau K est continu borné sur $\Omega \times \Omega$ et que pour tout x , $K(x, \cdot) \in C_0(\Omega, \mathcal{M}_m(\mathbb{R}))$ ¹⁸, alors H s'injecte continûment dans $C_0(\Omega, \mathbb{R}^m)$.

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 $k \in C_0^{2p}(\Omega, \mathcal{M}_m(\mathbb{R}))$ ¹⁹ alors H s'injecte continûment dans $C_0^p(\mathbb{R}^d, \mathbb{R}^m)$.

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

6.4.2. *Séparabilité.* Une dernière chose mérite d'être contrôlée? Est-ce que l'ENR construit avec un noyau K est séparable (ie 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 $C(\mathcal{X} \times \mathcal{X}, \mathcal{M}_m(\mathbb{R}))$, alors H est séparable dès que \mathcal{X} l'est.

EXERCICE 27. Montrer si K est continue, alors $x \rightarrow K\delta_x\alpha$ est continue sur H . En déduire le résultat.

¹⁶on considère que si Ω est non borné, $\infty \in \partial\Omega$ et $h(x)$ tends vers 0 lorsque $|x| \rightarrow \infty$. En particulier $C_0(\mathbb{R}^d, \mathbb{R}^m)$ est l'espace des fonctions continues qui tendent vers 0 à l'infini

¹⁷la 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 $H_f \subset C_0(\Omega, \mathbb{R}^m)$. 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 $C_0(\Omega, \mathbb{R}^m)$), on obtient le résultat

¹⁸ $K(x, \cdot)$ désigne ici l'application $y \rightarrow K(x, y)$

¹⁹Cette fois ci, on suppose que les dérivées partielles jusqu'à l'ordre $2p$ sont dans $C_0(\Omega, \mathcal{M}_m(\mathbb{R}))$

CHAPITRE 2

Actions de groupes, distances invariantes

1. Actions de groupes, espaces homogènes

1.1. Définition d'une action de groupe. Dans la suite, on considère un groupe G agissant sur un ensemble. L'élément neutre de G sera noté e et l'inverse d'un élément $g \in G$ sera noté g^{-1} .

DÉFINITION 3. Soit M un ensemble et G un groupe. Une action à gauche de G sur M est une application $A : G \times M \rightarrow M$ telle que

- (1) $A(e, m) = m$ pour tout $m \in M$,
- (2) $A(g', A(g, m)) = A(g'g, m)$ pour tous $g, g' \in G$ et $m \in M$.

REMARQUE 13. Nous pouvons déjà faire les trois remarques suivantes :

– On définit de même une action à droite, en remplaçant 2) par 2') :

$$A(g', A(g, m)) = A(gg', m).$$

- Si A est une action à gauche, alors $\tilde{A}(g, m) = A(g^{-1}, m)$ définit une action à droite.
– On simplifie souvent la notation $A(g, m)$ par gm pour une action à gauche et mg pour une action à droite. La propriété 2) s'écrit donc alors plus simplement comme

$$g'(gm) = (g'g)m.$$

1.2. Exemples.

1.2.1. *Action sur les n-uplets de points.* Soit $M \triangleq \{ x_I = (x_i)_{i \in I} \mid x_i \in \mathbb{R}^d \text{ pour tout } i \in I \}$ et $G = \text{Hom}(\mathbb{R}^d)$ l'espace de homéomorphismes sur \mathbb{R}^d c'est à dire des applications inversibles continues d'inverses continus. On dispose de l'action :

$$\phi x_I \triangleq (\phi(x_i))_{i \in I}.$$

1.2.2. *Action sur les images.* Soit $M = \mathcal{F}(\mathbb{R}^d, \mathbb{R})$ et $G = \text{Hom}(\mathbb{R}^d)$. On a l'action de transport des images sur \mathbb{R}^d définie par

$$\phi f \triangleq f \circ \phi^{-1}.$$

1.2.3. *Action sur les mesures.* L'action précédente induit une action duale sur le dual $M = \mathcal{M}_s(\mathbb{R}^d)$ définie par :

$$\phi\mu(f) \triangleq \mu(f \circ \phi^{-1})$$

pour tout $f \in C_0(\mathbb{R}^d, \mathbb{R})$.

1.2.4. *Action sur les p-formes.* Soit $M = \mathcal{F}(\mathbb{R}^d, (\Lambda^p \mathbb{R}^d)^*)$ et $G = \text{Diff}^1(\mathbb{R}^d)$ (C^1 d'inverse C^1), on définit le transport d'une p-forme ω par

$$(\phi\omega)_x(u_1 \wedge \cdots \wedge u_p) \triangleq \omega_{\phi^{-1}(x)}(d\phi_x^{-1}(u_1) \wedge \cdots \wedge d\phi_x^{-1}(u_p)).$$

1.2.5. *Action sur les courants.* Comme précédemment, on transforme cette action sur nos ensemble de courants $\Omega_0^p(\mathbb{R}^d)^*$ par

$$\phi T(\omega) = T(\phi^{-1}\omega)$$

1.2.6. *Action sur les pré-formes de Kendall.* On pose $M = S_d^n = \{m = (x_1, \dots, x_n) \mid (x_1, \dots, x_n) \in \mathbb{R}^{dn}, \sum_{i=1}^n x_i = 0 \text{ et } \sum_{i=1}^n |x_i|^2 = 1\}$ et $G = \text{SO}(d)$ le groupe des rotations de déterminant 1 sur \mathbb{R}^d . On a alors l'action

$$Rx \triangleq (R(x_1), \dots, R(x_n)).$$

1.3. Orbite et stabilisateur d'un élément. On suppose que $G \times M \rightarrow M$ est une action à gauche de G sur M (évidemment toutes les notions qui suivent ont des versions équivalentes pour les actions à droite).

DÉFINITION 4. Pour tout $m \in M$, on note Gm , appelée orbite de m dans M , l'ensemble

$$Gm \triangleq \{ gm \mid g \in G \}.$$

EXERCICE 28. Montrer que la relation $m \sim m'$ ssi $m' \in Gm$ est une relation d'équivalence.

L'ensemble des orbites forme donc une partition de M .

DÉFINITION 5. On note $G \backslash M$ l'ensemble des orbites (lire "M sur G"; G est placé à gauche pour rappeler que l'action est à gauche).

REMARQUE 14. On définit de même pour une action à droite les orbites à droite mG et l'ensemble quotient M/G (ici G est à droite car l'action est à droite).

DÉFINITION 6. On note $G_m = \{g \in G \mid gm = m\}$ le stabilisateur de m sous l'action de G .

EXERCICE 29. Montrer que G_m est un groupe appelé aussi groupe isotropie de m . Montrer que $G_{gm} = gG_m g^{-1}$.

1.4. Action transitive et espaces homogènes.

DÉFINITION 7. On dit que l'action est transitive si $G\mathfrak{m} = M$ pour tout $\mathfrak{m} \in M$. L'ensemble M muni de l'action de G sur M est alors appelé espace homogène¹

EXERCICE 30. Vérifier que si il existe \mathfrak{m}_0 tel que $G\mathfrak{m}_0 = M$, alors M est homogène.

EXERCICE 31.

- (1) Soit H un sous-groupe de G . Alors $(h, g) \rightarrow gh$ définit une action à droite de H sur G (le vérifier). On note G/H l'ensemble des classes d'équivalences ie $G/H \triangleq \{gH \mid g \in G\}$.
- (2) Maintenant, vérifier que pour tous $g, g_1, g'_1 \in G$, si $g_1 \sim g'_1$ alors $gg_1 \sim gg'_1$ si bien que l'on peut définir une application de $G \times G/H \rightarrow G/H$ par $(g, g_1H) \rightarrow gg_1H$ qui est une action à gauche de G sur G/H . Par suite, G/H est un espace homogène. Quel est le stabilisateur de $[e]$?

En reprenant l'exercice précédent, si on part d'un élément $\mathfrak{m}_0 \in M$ que nous appellerons *origine* et que nous considérons $H = G_{\mathfrak{m}_0}$ alors on vérifie que $g \sim g'$ ie $gG_{\mathfrak{m}_0} = g'G_{\mathfrak{m}_0}$ ssi $g\mathfrak{m}_0 = g'\mathfrak{m}_0$. Par suite l'application $\pi_{\mathfrak{m}_0} : G \rightarrow G_{\mathfrak{m}_0}$ définie par $\pi_{\mathfrak{m}_0}(g) = g\mathfrak{m}_0$ se factorise en $\tilde{\pi}_{\mathfrak{m}_0} : G/G_{\mathfrak{m}_0} \rightarrow G_{\mathfrak{m}_0}$. C'est de plus une bijection et on a la relation remarquable :

$$\tilde{\pi}_{\mathfrak{m}_0}(g[g']) = g\tilde{\pi}_{\mathfrak{m}_0}([g']) \text{ pour tout } g \in G, [g'] \in G/G_{\mathfrak{m}_0},$$

La morale de cette histoire (cf figure 1.4) est que l'action de G sur l'orbite d'un élément \mathfrak{m}_0 est isomorphe à l'action de G sur le quotient $G/G_{\mathfrak{m}_0}$. En particulier, toutes les actions à gauche de G sur des ensembles homogènes peuvent être ramenées à des actions à gauche de G sur des quotients à droite G/H pour un sous-groupe H de G . Il suffit pour cela de choisir une origine \mathfrak{m}_0 et de prendre $H = G_{\mathfrak{m}_0}$. La classe $[e] = G_{\mathfrak{m}_0}$ de $[e]$ correspond alors à l'origine \mathfrak{m}_0 dans M .

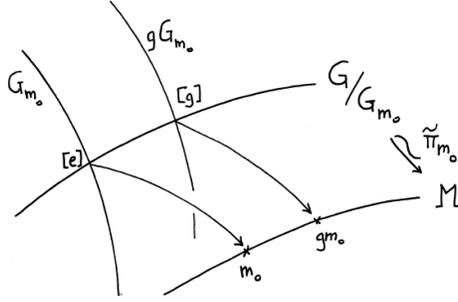
2. Distances, distances quotient sur les espaces homogènes

2.1. Distances et pseudo-distances.

DÉFINITION 8. Soient N un ensemble et $d_N : N \times N \rightarrow [0, +\infty]$. On dit que d_N est une pseudo-distance si

- (1) $d_N(\mathfrak{n}, \mathfrak{n}') = d_N(\mathfrak{n}', \mathfrak{n})$ pour tout $\mathfrak{n}, \mathfrak{n}' \in N$ (symétrie);

¹Cette définition d'un espace homogène n'est pas la définition complète qui normalement demande de considérer une structure de groupe de Lie sur G et une structure de variété sur M . Nous garderons cette définition sans structure car nous souhaitons dans cette première partie garder un point de vue purement algébrique

FIG. 1. Bijection entre G/G_{m_0} et G_{m_0}

(2) $d_N(n, n'') \leq d_N(n, n') + d_N(n', n'')$ pour tout $n, n', n'' \in N$ (inégalité triangulaire);

si de plus d_N vérifie

(3) $d(n, n') = 0 \Rightarrow n = n'$ pour tout $n, n' \in N$

alors d_N est appelée une distance sur N .

DÉFINITION 9. si $H \times N \rightarrow N$ est une action à droite et d_N est une distance sur N , on dit que d_N est H invariante à droite (on dit parfois H équivariante à droite) si

$$d_N(nh, nh') = d_N(n, n') \text{ pour tout } (h, n, n') \in H \times N \times N.$$

REMARQUE 15. On a bien évidemment une définition similaire pour une distance invariante à gauche.

EXERCICE 32. Montrer que l'ensemble des $h \in G$ tels que $d_N(nh, n'h) = d_N(n, n')$ pour tout $n, n' \in N$ forme un groupe; Vérifier que pour un tel h , la multiplication à droite par h est une isométrie pour la distance d_N .

PROPOSITION 2. Si d_N est H i. à d. alors

$$d_{N/H}([n], [n']) \triangleq \inf\{d_N(nh, n'h') \mid h, h' \in H\}$$

est une pseudo-distance sur N/H .

REMARQUE 16. On remarque que par invariance à droite sous l'action de H on a $d_N(nh, n'h') = d_N(n, n'h'h^{-1})$ d'où $d_{N/H}([n], [n']) = \inf\{d_N(n, n'h') \mid h' \in H\} = \inf\{d_N(nh, n') \mid h \in H\}$

DÉMONSTRATION. La propriété 1) est immédiate. Pour la propriété 2), on remarque que $d_{N/H}([n], [n'']) \leq d_N(nh, n') + d_N(n', n''h)$. En prenant successivement l'infimum sur h puis $h'' \in H$, on déduit de la remarque précédente le résultat. \square

2.2. Application à la construction d'une distance sur G/H . Une application directe de la proposition 2 est la suivante : si on considère l'action à droite d'un sous-groupe H sur G , alors en prenant $N = G$, on déduit que si l'on dispose d'une distance sur G qui est H invariante à droite, on obtient une distance sur G/H . Bien sur, dans le cas où $H = G_{\mathfrak{m}_0}$, de part l'isomorphisme entre $G/G_{\mathfrak{m}_0}$ et $G\mathfrak{m}_0$ on déduit immédiatement une distance $d_{G\mathfrak{m}_0}$ sur l'orbite $G\mathfrak{m}_0$. Dans ce cas, l'application $\tilde{\pi}_{\mathfrak{m}_0}([g]) = g\mathfrak{m}_0$ devient une isométrie. Enfin, si M est homogène, alors $G\mathfrak{m}_0 = M$ et on obtient ainsi une distance sur M qui définit par :

$$d_M(\mathfrak{m}, \mathfrak{m}') = \inf\{d_G(g, g') \mid g\mathfrak{m}_0 = \mathfrak{m}, g'\mathfrak{m}_0 = \mathfrak{m}'\}.$$

2.2.1. *Cas d'une distance d_G G invariante à droite.* Un cas particulièrement agréable est celui d'une distance d_G qui est G invariante à droite. Dans ce cas, pour toute origine \mathfrak{m}_0 , $G\mathfrak{m}_0 \subset G$, et donc la distance est $G_{\mathfrak{m}_0}$ invariante à droite. On peut donc toujours induire une distance sur l'orbite $G\mathfrak{m}_0$. Plus généralement, tout espace homogène M peut être équipé d'une métrique induite par celle de d_G . C'est dans ce cadre à la fois riche et simple que nous nous placerons presque exclusivement dans la suite.

EXERCICE 33. Vérifier que dans le cas où d_G est invariante à droite et M homogène, alors

$$(51) \quad d_M(\mathfrak{m}, \mathfrak{m}') = \inf\{d_G(e, g) \mid g\mathfrak{m} = \mathfrak{m}'\}.$$

Il ne faut cependant pas oublier que la G invariance à droite n'est pas une condition nécessaire. Travailler avec des distances seulement $G_{\mathfrak{m}_0}$ invariante est sans doute un champ intéressant trop peu exploré.

2.2.2. *Changement d'origine.* Supposons que l'on dispose d'une distance d_G qui soit à la fois $G_{\mathfrak{m}_0}$ et $G_{\mathfrak{m}'_0}$ invariante sur un espace homogène M . On peut alors construire deux distances sur M à partir des deux quotients $G/G_{\mathfrak{m}_0}$ et $G/G_{\mathfrak{m}'_0}$. Les distances induites sur M sont-elles différentes ?

On remarque que l'on a une bijection entre $G/G_{\mathfrak{m}_0}$ et $G/G_{\mathfrak{m}'_0}$ via les bijections $\tilde{\pi}_{\mathfrak{m}_0}$ et $\tilde{\pi}_{\mathfrak{m}'_0}$ sur M .

Supposons que $\mathfrak{m}'_0 = h\mathfrak{m}_0$ avec $h \in H$. En remarquant que pour tout $g \in G$, $gG_{\mathfrak{m}_0}h^{-1} = gh^{-1}hG_{\mathfrak{m}_0}h^{-1} = gh^{-1}G_{\mathfrak{m}'_0}$ et que $g\mathfrak{m}_0 = gh^{-1}\mathfrak{m}'_0$, on voit que la bijection entre $G/G_{\mathfrak{m}_0}$ dans $G/G_{\mathfrak{m}'_0}$ est donnée par l'application $\psi_{\mathfrak{m}_0, \mathfrak{m}'_0}$:

$$gG_{\mathfrak{m}_0} \rightarrow gG_{\mathfrak{m}_0}h^{-1} = gh^{-1}G_{\mathfrak{m}'_0}$$

On vérifie immédiatement que ψ est une isométrie dès que la multiplication à droite par h^{-1} (ou par h) est une isométrie. On a donc la proposition suivante :

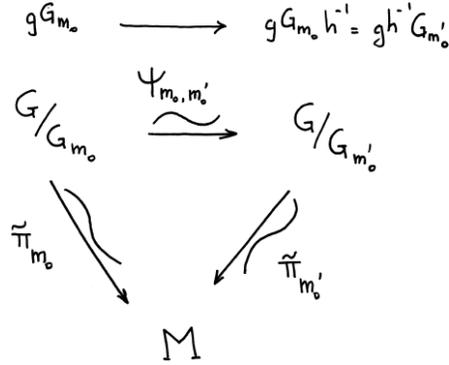


FIG. 2. Bijection entre G/G_{m_0} et $G/G_{m'_0}$ via les bijections $\tilde{\pi}_{m_0}$ et $\tilde{\pi}_{m'_0}$ sur M

PROPOSITION 3. Si d_G est H invariante à droite, alors pour tout $h \in H$, m_0 et hm_0 sont deux origines équivalentes au sens où G/G_{m_0} et G/G_{hm_0} sont isométriques et induisent la même distance sur M (on suppose M homogène).

REMARQUE 17. Dans le cas d'une invariance à droite, tous les points peuvent servir d'origine et pour deux points m et $m' \in M$, en prenant m comme origine, on retrouve immédiatement l'expression (51).

3. Approche différentielle

Pour le moment nous n'avons considéré que le point de vue "macro" ou global : on a supposé que le groupe G est donné avec une distance globale d_G possédant des propriétés d'invariance. Cependant, nous n'avons guère donné d'exemples de distances qui satisfassent les propriétés d'invariance qui nous sont utiles.

La raison est qu'il est assez rare de disposer de distances explicites avec lesquelles on puisse travailler directement. Il est plus fructueux d'adopter *un point de vue local* ou riemannien dans lequel les distances sont induites par des métriques locales et définies comme des distances de *plus court chemin* autrement dit des distances *géodésiques*. Pour cela, il nous faut mettre un peu de structure différentielle sur les groupes, ce qui nous conduit à rappeler brièvement la notion de variété et de sous-variété différentielle et celle de groupe de Lie. Nous n'en ferons qu'une présentation élémentaire (on pourra consulter pour plus de détails des ouvrages de références, par exemple [4, 6, 10, 2] ou se reporter au premier chapitre de [13]).

3.1. Variétés différentiables. Nous commençons par rappeler la notion de sous-variété qui est la plus intuitive et qui est en fait la plus utile pour nous.

3.1.1. *Sous-variétés différentiable de \mathbb{R}^q .*

DÉFINITION 10. On appelle sous-variété différentiable de classe C^p de \mathbb{R}^{n+k} de dimension n , un ensemble $M \subset \mathbb{R}^{n+k}$ tel que pour tout $m \in M$, il existe un ouvert $U \ni m$ de \mathbb{R}^{n+k} et une submersion $f : U \rightarrow \mathbb{R}^k$ de classe C^p telle que $M \cap U = f^{-1}(0)$ (on rappelle que f est une submersion si df_m est surjective en tout point).

EXERCICE 34. En utilisant la définition précédente ;

- (1) montrer que la sphère S^d est une sous-variété C^∞ de \mathbb{R}^{d+1} de dimension d ;
- (2) montrer le groupe spécial orthogonal $SO(d) \triangleq \{R \in \mathcal{M}_d(\mathbb{R}) \mid \det R = 1 \text{ et } R^T R = \text{Id}\}$ est une sous-variété de dimension $d(d-1)/2$
- (3) montrer que l'hyperboloïde $H_c^d = \{m = (x_0, \dots, x_d) \mid x_0^2 - x_1^2 - \dots - x_d^2 = c\}$ est une sous-variété C^∞ de dimension d de \mathbb{R}^{d+1} pour tout $c \neq 0$. Que se passe-t-il pour $c = 0$?

Cette définition est très efficace pour reconnaître des sous-variétés classiques en particulier parce qu'elle est *implicite* et ne fait pas apparaître de paramétrisation locale. Nous donnons sans démonstration (voir [2]) un résultat qui fait le lien avec la vision intuitive de sous-variété comme surface localement paramétrable (cf point (3) de la proposition).

PROPOSITION 4. *Les trois assertions suivantes sont équivalentes :*

- (1) M est une sous-variété de classe C^p et de dimension n de \mathbb{R}^{n+k} ,
- (2) Pour tout $m \in M$, il existe deux ouverts U et V de \mathbb{R}^{n+k} tels que $m \in U$ et $0 \in V$ et un C^p difféomorphisme $f : U \rightarrow V$ vérifiant $f(U \cap M) = V \cap (\mathbb{R}^n \times \{0\})$
- (3) Pour tout $m \in M$, il existe un voisinage U de m dans M , un ouvert $\Omega \ni 0$ dans \mathbb{R}^n et une application C^p $g : \Omega \rightarrow \mathbb{R}^{n+k}$ tel que (Ω, g) est une paramétrisation locale de $M \cap U$ au voisinage de m (ie g est un homéomorphisme de Ω dans $M \cap U$ et dg_0 est injective).

3.1.2. *Variétés abstraites.* Nous rappelons tout de même la définition d'une variété abstraite pour mémoire.

DÉFINITION 11. Une variété différentiable M de classe C^p est un espace topologique séparé muni d'un atlas C^p c'est à dire d'une famille $(U_i, \phi_i)_{i \in I}$ telle que

- (1) $(U_i)_{i \in I}$ est un recouvrement ouvert de M et pour tout $i \in I$, $\phi_i : U_i \rightarrow \Omega_i$ est un homéomorphisme sur un ouvert $\Omega_i \subset \mathbb{R}^{n_i}$ (carte locale)

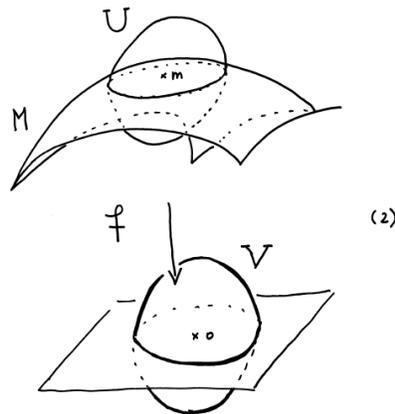


FIG. 3. Illustration du point 2) de la proposition

- (2) Pour tout $i, j \in I$ tel que $U_i \cap U_j \neq \emptyset$, l'application $\phi_j \circ \phi_i^{-1}$ est un C^p difféomorphisme de $\phi_i(U_i \cap U_j)$ sur $\phi_j(U_i \cap U_j)$ (compatibilité C^p des changements de cartes).

On dit que deux atlas sont équivalents, si la réunion des deux atlas constitue un atlas. On appelle ainsi structure différentiable sur M une classe d'équivalence d'atlas (qui contient un plus grand atlas, appelé atlas maximal).

Evidemment, une sous-variété est une variété ce qui se démontre à l'aide du théorème des fonctions implicites.

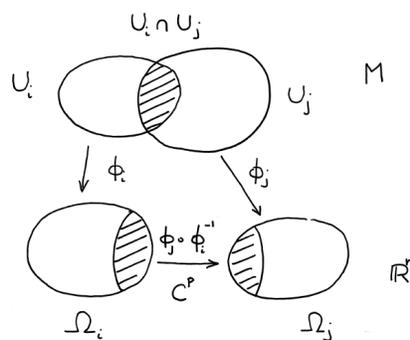


FIG. 4. Changement de carte pour une variété abstraite

3.1.3. *vecteurs tangents, espace tangent.* La définition des vecteurs tangents et espaces tangent est relativement directe pour les sous-variétés.

DÉFINITION 12. Soit M une sous-variété C^p de dimension n de \mathbb{R}^{n+k} . On dit que $v \in \mathbb{R}^{n+k}$ est un vecteur tangent en $m \in M$, si il existe une courbe $c : I \rightarrow \mathbb{R}^{n+k}$ où I est un intervalle ouvert contenant 0 tel que $c(I) \subset M$ (c est tracée sur M), $c(0) = m$ et c est dérivable en 0 de dérivée $c'(0) = v$.

REMARQUE 18. Cette caractérisation de l'espace tangent par les dérivées des courbes lisses est très similaire à celle que nous utiliserons pour parler de l'espace tangent sur des groupes de difféos.

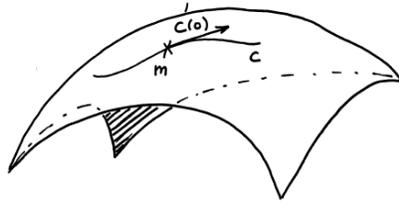


FIG. 5. Exemple de vecteur tangent

On montre alors facilement en utilisant la proposition 4.2) que l'ensemble des vecteurs tangent en un point m est un sèze de dimension n de \mathbb{R}^{n+k} noté $T_m M$ et appelé espace tangent de M en m . Le théorème suivant donne des caractérisations équivalentes de l'espace tangent qui peuvent être utiles pour déterminer explicitement les espaces tangents de variétés classiques.

THÉORÈME 3 (D'après [6]). Soient M une sous-variété C^p de dimension n de \mathbb{R}^{n+k} , $m \in M$ et U un voisinage de m dans \mathbb{R}^{n+k}

(1) Si $M \cap U = f^{-1}(0)$ où f est une C^p submersion de U dans \mathbb{R}^k alors

$$T_m M = \ker(df_m).$$

(2) Si f est un difféomorphisme de U dans un voisinage V de 0 dans \mathbb{R}^{n+k} avec $f(U \cap M) = V \cap (\mathbb{R}^n \times \{0\})$ alors

$$T_m M = (df_m)^{-1}(\mathbb{R}^n \times \{0\}).$$

(3) Si (Ω, g) est une paramétrisation de $M \cap U$ telle que $g(x) = m$, alors

$$T_m M = dg_x(\mathbb{R}^n).$$

EXERCICE 35. Identifier les espaces tangents de la sphère S^d et montrer que l'espace tangent en l'identité de $SO(d)$ est le sous-espace des matrices anti-symétriques.

Dans le cas d'une variété abstraite on procède de façon équivalente à ceci près que l'on ne peut pas définir directement la dérivée $c'(0)$ d'une courbe tracée sur M . On passe par les classe d'équivalence $c_0 \sim c_1$ si $(\Phi \circ c_0)'(0) = (\Phi \circ c_1)'(0)$ dans une carte local (U, ϕ) autour de m .

REMARQUE 19. Si M est une variété C^p de dimension n , la collection des espaces tangent $TM = (T_m M)_{m \in M}$ peut être muni d'une structure de variété C^{p-1} (fibré tangent) de dimension $2n$.

3.2. Groupes de Lie. Nous considérons le cas ici de groupes de Lie de dimension finie même si dans la suite nous aurons besoin de travailler en dimension infinie. On considère que tout les groupes de Lie sont des sous-variétés de \mathbb{R}^{n+k} de dimension n .

DÉFINITION 13. On dit qu'un groupe G est un groupe de Lie C^∞ , si G est muni d'une structure différentiable C^∞ pour laquelle les deux opérations $(g, g') \rightarrow gg'$ et $g \rightarrow g^{-1}$ (multiplication et passage à l'inverse) sont C^∞ .

- NOTATION 1. (1) On notera dans précédemment e l'élément neutre et $T_e G$ l'espace tangent sur G en e appelé algèbre de Lie de G et noté parfois \mathfrak{g} .
- (2) Pour tout $g \in G$, on note L_g (resp. R_g) la multiplication à gauche (resp. à droite) définie par $L_g(g') = gg'$ (resp. $R_g(g') = g'g$).

De nombreux exemples de groupes de Lie de dimension finie peuvent être donné par des groupes de matrices sur \mathbb{R} .

3.2.1. *Groupe $SO(d)$.* Le groupe des rotations directes est un exemple que nous avons déjà rencontré. Nous avons vu que dans ce cas, $e = \text{Id}$ et l'espace tangent en e s'identifie avec l'espace des matrices anti-symétriques :

$$\mathfrak{g} = \mathfrak{so}(d) = \{A \in \mathcal{M}_d(\mathbb{R}) \mid A = -A^T\}$$

3.2.2. *Groupe spécial euclidien $SE(d)$.* On considère cette fois le groupe spécial euclidien c'est à dire des applications $h : \mathbb{R}^d \rightarrow \mathbb{R}^d$ telle que $h(x) = Rx + t$ pour un couple $(R, t) \in SO(d) \times \mathbb{R}^d$. Comme la décomposition est unique, on identifie h et (R, t) . On vérifie alors que si $h = (R, t)$ et $h' = (R', t')$, alors $h(h'(x)) = RR'x + Rt' + t$ d'où $h \circ h' = (RR', Rt' + t)$. Le produit $(R, t)(R', t') = (RR', Rt' + t)$ est appelé le produit semi-direct. Pour rappeler que l'on utilise le produit semi-direct, on note généralement le produit semi-direct $SO(d) \ltimes \mathbb{R}^d$. Dans ce cas, l'espace tangent en $e = (\text{Id}, 0)$ est

$$\mathfrak{g} = \mathfrak{so}(d) \ltimes \mathbb{R}^d.$$

3.2.3. *Groupe $\text{Diff}_{\text{Id}}^\infty(\Omega)$* . Soit Ω un ouvert borné et notons $\text{Diff}_{\text{Id}}^\infty(\Omega)$ l'ensemble des C^∞ difféomorphismes sur \mathbb{R}^d dont la restriction au complémentaire de Ω vaut l'identité.

Il ne s'agit pas d'un groupe de Lie de dimension fini et la définition d'une structure différentiable C^∞ sur un tel espace est problématique (on pourra consulter [11] pour une exposition récente des résultats dans cette direction)

Nous nous bornerons ici à identifier quels peuvent être "raisonnablement" les espaces tangents en adaptant le point de vue de la dimension finie.

Appelons chemin C^∞ dans $\text{Diff}_{\text{Id}}^\infty(\Omega)$ une application $\Phi :]-\epsilon, \epsilon[\times \mathbb{R}^d \rightarrow \mathbb{R}^d$, C^∞ en les deux variables et telle que $x \rightarrow \Phi(t, x) \in \text{Diff}_{\text{Id}}^\infty(\Omega)$. Lorsqu'on suppose $\Phi(0, \cdot) = \text{Id}$ alors on voit que $x \rightarrow \frac{\partial}{\partial t} \Phi(0, x)$ est un champ de vecteur C^∞ sur \mathbb{R}^d nul en dehors de Ω . On peut donc identifier au moins formellement ici l'algèbre de Lie de $\text{Diff}_{\text{Id}}^\infty(\Omega)$ comme l'ensemble des champs de vecteurs C^∞ sur \mathbb{R}^d , nuls en dehors de Ω et noté $\chi_0^\infty(\Omega)$. On a donc

$$\mathfrak{g} = \chi_0^\infty(\Omega).$$

3.3. Applications adjointes. Nous revenons dans le cas d'un groupe de Lie classique. Par définition, la conjugaison par un élément $g \in G$ définit une application C^∞ $I_g : G \rightarrow G$ définie par $I_g(h) = ghg^{-1}$. En considérant la différentielle de I_g en $h = e$, on définit l'application adjointe $\text{Ad}_g : \mathfrak{g} \rightarrow \mathfrak{g}$.

EXERCICE 36. Vérifier en utilisant la relation $I_g \circ I_{g'} = I_{gg'}$ que $\text{Ad}_g \circ \text{Ad}_{g'} = \text{Ad}_{gg'}$. Vérifier que pour tout $g \in G$, $v \rightarrow \text{Ad}_g(v)$ définit un élément de $\text{GL}(\mathfrak{g})^2$.

Comme $g \rightarrow \text{Ad}_g$ est C^∞ , on peut différentier Ad en $g = e$. On note alors ad l'application $\text{ad} : \mathfrak{g} \rightarrow \text{GL}(\mathfrak{g})$ définie par $\text{ad}_u \triangleq (\text{dAd})_e(u) \in \text{GL}(\mathfrak{g})$ pour tout $u \in \mathfrak{g}$. On note souvent pour tout $u, v \in \mathfrak{g}$, $[u, v] \triangleq \text{ad}_u(v) = ((\text{dAd})_e(u))v$ appelé le crochet de deux éléments u et v .

EXERCICE 37. Vérifier que lorsque G est un groupe de Lie de matrices, alors $I_A(B) = ABA^{-1}$, $\text{Ad}_A(V) = AVA^{-1}$ et $\text{ad}_U(V) = UV - VU = [U, V]$ pour tous $A, B \in G$ et $U, V \in \mathfrak{g}$.

$(u, v) \rightarrow [u, v]$ est un produit sur \mathfrak{g} qui donne une structure d'algèbre de Lie à \mathfrak{g} et vérifie les propriétés classiques :

$$(1) [u, v] = -[v, u];$$

$$(2) [[u, v], w] + [[v, w], u] + [[w, u], v] = 0 \text{ (identité de Jacobi).}$$

²Un homomorphisme $R : G \rightarrow \text{GL}(V)$ où V est un espace vectoriel est appelé une représentation de G . Nous avons donc montré que Ad est une représentation de G appelée représentation adjointe. On voit d'ailleurs ici qu'un groupe de Lie de dimension finie peut-être vu comme un groupe de matrices.

4. Métriques riemanniennes

4.1. Définition d'une distance riemmanienne. La construction d'une distance riemmanienne sur une variété N (dans notre principale application on prendra $N = G$) passe par la donnée en tout point $\mathbf{n} \in N$ d'un produit scalaire $\langle \cdot, \cdot \rangle_{\mathbf{n}}$ sur l'espace tangent $T_{\mathbf{n}}N$ en \mathbf{m} qui varie de façon lisse avec \mathbf{n} . Pour toute courbe C^1 par morceaux $\mathbf{c} : [0, 1] \rightarrow N$ sur N , on définit alors la longueur de \mathbf{c} par

$$L(\mathbf{c}) = \sum_{i=0}^{n-1} \int_{\alpha_i}^{\alpha_{i+1}} |\mathbf{c}'(t)|_{\mathbf{c}(t)} dt$$

où $(\alpha_i)_{1 \leq i \leq n}$ est définie une partition de $[0, 1]$ en morceaux C^1 .

EXERCICE 38. Vérifier que si $\psi : [0, 1] \rightarrow [0, 1]$ est C^1 est croissante bijective, alors $L(\mathbf{c} \circ \psi) = L(\mathbf{c})$ (en particulier la longueur d'une courbe est bien une quantité invariante par changement de paramétrisation croissante; c'est une quantité géométrique au même titre que l'intégration d'une p -forme par un courant).

A partir des longueurs de courbes on définit une distance géodésique sur N par

$$d_N(\mathbf{n}, \mathbf{n}') = \inf\{L(\mathbf{c}) \mid \mathbf{c}(0) = \mathbf{n} \text{ et } \mathbf{c}(1) = \mathbf{n}', \mathbf{c} : [0, 1] \rightarrow N \text{ } C^1 \text{ par morceaux}\}$$

On vérifie que d_N est une *distance* sur N appelée distance géodésique.

4.2. Métriques invariantes. Dans le cas qui nous intéresse principalement, c'est-à-dire lorsque la variété est un groupe de Lie G , nous voulons *plus* qu'une simple distance mais nous souhaitons avoir des propriétés d'invariance sous l'action d'un sous-groupe.

Soit donc H un sous-groupe de G . Pour tout $\mathbf{h} \in H$, on note $R_{\mathbf{h}} : G \rightarrow G$ la multiplication à droite par \mathbf{h} définie par $R_{\mathbf{h}}(\mathbf{g}) \triangleq \mathbf{g}\mathbf{h}$. Si G est un groupe de Lie, alors $R_{\mathbf{h}}$ est C^∞ et pour tout $\mathbf{g} \in G$, $(dR_{\mathbf{h}})_{\mathbf{g}}$ est une bijection de $T_{\mathbf{g}}G$ dans $T_{\mathbf{g}\mathbf{h}}G$.

NOTATION 2. Pour simplifier les notations, pour tout $\mathbf{g} \in G$ on notera

$$\mathbf{u}\mathbf{h} \triangleq (dR_{\mathbf{h}})_{\mathbf{g}}(\mathbf{u}) \in T_{\mathbf{g}\mathbf{h}}G \text{ pour tout } \mathbf{u} \in T_{\mathbf{g}}G.$$

EXERCICE 39. En partant de l'égalité $R_{\mathbf{h}} \circ R_{\mathbf{h}'} = R_{\mathbf{h}\mathbf{h}'}$ montrer que $(\mathbf{u}\mathbf{h})\mathbf{h}' = \mathbf{u}(\mathbf{h}\mathbf{h}')$ pour tout $\mathbf{u} \in T_{\mathbf{g}}G$ et $\mathbf{h}, \mathbf{h}' \in H$.

L'invariance à droite par les éléments de H est équivalente à dire que $R_{\mathbf{h}}$ est une isométrie pour tout $\mathbf{h} \in H$.

PROPOSITION 5. Une condition suffisante au niveau des métriques locales pour que d_G soit H invariante à droite est que pour tout $\mathbf{h} \in H$:

$$(52) \quad \langle \mathbf{v}\mathbf{h}, \mathbf{v}'\mathbf{h} \rangle_{\mathbf{g}\mathbf{h}} = \langle \mathbf{v}, \mathbf{v}' \rangle_{\mathbf{g}} \text{ pour tous } \mathbf{v}, \mathbf{v}' \in T_{\mathbf{g}}G.$$

DÉMONSTRATION. En effet, si $c : [0, 1] \rightarrow G$ est un chemin C^1 par morceaux, on a alors $L(ch) = \sum_{i=0}^{n-1} \int_{\alpha_i}^{\alpha_{i+1}} |c'(t)h|_{c(t)h} dt = \sum_{i=0}^{n-1} \int_{\alpha_i}^{\alpha_{i+1}} |c'(t)|_{c(t)} dt = L(c)$ et donc $d_G(g, g') = d_G(gh, g'h)$ pour tout $h \in H$. \square

EXERCICE 40. Montrer que (52) est équivalent à

$$(53) \quad \langle v g, v' g \rangle_g = \langle v g', v' g' \rangle_{g'} \text{ pour tous } v, v' \in T_e G \text{ et } g' \in g H.$$

Montrer qu'il existe alors une famille $(Q(\cdot, \cdot)_{[g]})_{[g] \in G/H}$ de produits scalaires sur $T_e G$ tels que $\langle u g, v g \rangle_g = Q(u, v)_{[g]}$ pour tous $u, v \in T_e G$ et $g \in G$.

4.3. Exemples. Nous considérons ci-dessous deux cas emblématiques pour nous, celui de $SO(d)$ et celui de $\text{Diff}_{\text{Id}}^\infty(\Omega)$.

4.3.1. *Exemple de $SO(d)$.*

EXERCICE 41. Montrer que la métrique induite sur $SO(d)$ par la métrique euclidienne usuelle sur les matrices de $\mathcal{M}_d(\mathbb{R})$ dite métrique de Hilbert-Schmidt (ou de Frobenius) et définie par

$$\langle M, N \rangle_{\text{HS}} = \text{tr}(M^T N),$$

est invariante à droite.

Dans le cas de $SO(3)$ la distance géodésique associée à la métrique de Hilbert-Schmidt est explicite. On a (on verra plus tard la preuve de ce fait) que :

$$d_{SO(3)}(R, R') = \arccos\left(\frac{\text{tr}(R'R^{-1}) - 1}{2}\right).$$

4.3.2. *Exemple des déformations non-rigides.* Reprenons le cas $G = \text{Diff}_{\text{Id}}^\infty(\Omega)$. Dans cette partie, nous ne cherchons pas à donner un sens rigoureux à la notion de différentielle pour des applications sur G . Nous nous contenterons de raisonner pour le moment par analogie avec le cas de la dimension finie.

La multiplication à droite par un élément $R_{\varphi'}$ est donnée par $\varphi \rightarrow \varphi \circ \varphi'$. Si $\phi :]-\epsilon, \epsilon[\times \mathbb{R}^d \rightarrow \mathbb{R}^d$ est un chemin C^∞ dans G tel que $\phi(0, x) = \varphi(x)$, alors la dérivée en $t = 0$ de $R_{\varphi'} \circ \phi$ vaut $x \rightarrow \frac{\partial}{\partial t} \phi(0, \varphi'(x))$. En identifiant $x \rightarrow \frac{\partial}{\partial t} \phi(0, x)$ à un élément générique de $u \in T_\varphi G$, nous déduisons, au moins formellement, que si la différentielle de $R_{\varphi'}$ existe, on devrait avoir $(dR_{\varphi'})_\varphi(u) = u \circ \varphi'$. Ainsi, en utilisant la convention de la dimension finie, on peut noter $u \varphi' \triangleq u \circ \varphi'$ pour tout $u \in T_\varphi G$. Par conséquent, on a $T_\varphi G = \{u \circ \varphi \mid u \in \chi_0^\infty(\Omega)\}$. Supposons que l'on dispose en chaque $\varphi \in G$ d'un produit scalaire $\langle \cdot, \cdot \rangle_\varphi$. Alors, la condition d'invariance de la métrique par un sous-groupe $H \subset G$ s'écrit

$$\langle u \circ \varphi', v \circ \varphi' \rangle_{\varphi \circ \varphi'} = \langle u, v \rangle_\varphi, \text{ pour tous } u, v \in T_\varphi G.$$

4.4. Retour au cas général. L'invariance minimale qui nous intéresse est celle de la métrique par rapport au stabilisateur $G_{\mathfrak{m}_0}$ d'un élément $\mathfrak{m}_0 \in \mathfrak{M}$. On veut donc (cf (53))

$$\langle \mathbf{u}g, \mathbf{v}g \rangle_g = \langle \mathbf{u}g', \mathbf{v}g' \rangle_{g'}$$

pour tous $\mathbf{u}, \mathbf{v} \in T_e G$ et $g' \in gG_{\mathfrak{m}_0}$ ie $g\mathfrak{m}_0 = g'\mathfrak{m}_0$. Il suffit donc de définir une famille de métriques $(Q(\cdot, \cdot)_{[g]})_{[g] \in G/G_{\mathfrak{m}_0}}$ ou de façon équivalente $(Q(\cdot, \cdot)_m)_{m \in G\mathfrak{m}_0}$ sur $T_e G$ puis de définir

$$(54) \quad \langle \mathbf{u}g, \mathbf{v}g \rangle_g \triangleq Q(\mathbf{u}, \mathbf{v})_{g\mathfrak{m}_0}$$

REMARQUE 20. *On voit ici que la donnée d'une métrique $G_{\mathfrak{m}_0}$ invariante sur G est équivalente à la donnée en tout point \mathfrak{m} de l'orbite d'une métrique sur l'espace tangent en l'identité de G . Ceci peut-être comparé avec la donnée d'une métrique riemannienne sur M (lorsque M est une variété) qui est équivalente à la donnée d'une famille indexée par $\mathfrak{m} \in M$ de produits scalaires sur les espaces tangents $T_{\mathfrak{m}}M$ pour tout $\mathfrak{m} \in M$.*

En partant de la caractérisation donnée par l'équation (54), on traduit facilement la condition suffisante de H invariance donnée par (52) en

$$(55) \quad Q(\mathbf{u}, \mathbf{v})_{g\mathfrak{m}_0} = \langle \mathbf{u}g, \mathbf{v}g \rangle_g = \langle \mathbf{u}gh, \mathbf{v}gh \rangle_{gh} = Q(\mathbf{u}, \mathbf{v})_{gh\mathfrak{m}_0},$$

pour tous $\mathbf{u}, \mathbf{v} \in T_e G$ et $g \in G$. La condition de changement d'origine de \mathfrak{m}_0 en \mathfrak{m}'_0 devient alors

$$(56) \quad Q(\mathbf{u}, \mathbf{v})_{g\mathfrak{m}_0} = Q(\mathbf{u}, \mathbf{v})_{g\mathfrak{m}'_0},$$

pour tous $\mathbf{u}, \mathbf{v} \in T_e G$ et $g \in G$. Enfin, la condition d'invariance à droite de la métrique par le groupe G en entier est assurée dès que

$$\langle \mathbf{u}g, \mathbf{v}g \rangle_g = \langle \mathbf{u}, \mathbf{v} \rangle_e,$$

c'est-à-dire que toutes les métriques se déduisent d'une métrique unique définie en e et propagée partout par invariance à droite.

CHAPITRE 3

Construction de groupes de déformations

1. Flots

On considère Ω un ouvert de \mathbb{R}^d .

NOTATION 3. On note $C_0^1(\Omega, \mathbb{R}^d)$ l'ensemble des champs de vecteurs $\mathbf{u} : \overline{\Omega} \rightarrow \mathbb{R}^d$ qui soient continus sur $\overline{\Omega}$, C^1 sur Ω et tels que $|\mathbf{u}(\mathbf{x})| + \sum_{i=1}^d \left| \frac{\partial}{\partial x_i} \mathbf{u}(\mathbf{x}) \right| \rightarrow 0$ lorsque $\mathbf{x} \rightarrow \mathbf{y}$ pour tout $\mathbf{y} \in \partial\Omega$. Si Ω n'est pas borné, on impose de plus que $|\mathbf{u}(\mathbf{x})| + \sum_{i=1}^d \left| \frac{\partial}{\partial x_i} \mathbf{u}(\mathbf{x}) \right| \rightarrow 0$ lorsque $|\mathbf{x}| \rightarrow \infty$.

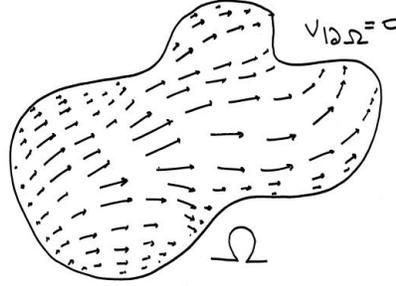


FIG. 1. Exemple de champs de vecteurs dans $C_0^1(\Omega, \mathbb{R}^d)$

NOTATION 4. Pour tout $\mathbf{u} \in C_0^1(\Omega, \mathbb{R}^d)$, on note

$$|\mathbf{u}|_{1,\infty} \triangleq \sup_{\mathbf{x} \in \Omega} \left(|\mathbf{u}(\mathbf{x})| + \sum_{i=1}^d \left| \frac{\partial}{\partial x_i} \mathbf{u}(\mathbf{x}) \right| \right).$$

DÉFINITION 14. On dit que \mathbf{V} est un espace admissible de champs de vecteurs si \mathbf{V} est espace de Hilbert qui s'injecte continûment dans $C_0^1(\Omega, \mathbb{R}^d)$ ie il existe $C > 0$ tel que pour tout $\mathbf{v} \in \mathbf{V}$ on a

$$|\mathbf{v}|_{1,\infty} \leq C|\mathbf{v}|_{\mathbf{V}}.$$

NOTATION 5. Soit \mathbf{V} un espace de champs de vecteurs admissible.

- (1) On note $L_{\mathbf{V}}^1 \triangleq \{\mathbf{v}(t, \mathbf{x}) \in \mathcal{F}([0, 1] \times \Omega, \mathbb{R}^d) \text{ mesurable} \mid \mathbf{v}(t, \cdot) \in \mathbf{V} \text{ et } \int_0^1 |\mathbf{v}_t|_{\mathbf{V}} dt < \infty\}$.

- (2) On note $C_{Id}(\Omega, \mathbb{R}^d) \triangleq \{ \varphi \in C(\overline{\Omega}, \mathbb{R}^d) \mid \sup_{\overline{\Omega}} |\varphi(x) - x| < \infty \}$. On munit $C_{Id}(\Omega, \mathbb{R}^d)$ de la distance $d_\infty(\varphi, \varphi') \triangleq \sup_{\Omega} |\varphi(x) - \varphi'(x)|$.
- (3) On note $C_{Id}([0, 1] \times \Omega, \mathbb{R}^d) \triangleq \{ \phi(t, x) \in C([0, 1] \times \overline{\Omega}, \mathbb{R}^d) \mid \sup_{[0, 1] \times \Omega} |\phi(t, x) - x| < \infty \}$. On munit $C_{Id}([0, 1] \times \Omega, \mathbb{R}^d)$ de la distance $d_{\infty, \infty}(\phi, \phi') \triangleq \sup_{[0, 1] \times \Omega} |\phi(t, x) - \phi'(t, x)|$.

EXERCICE 42. Vérifier que les espaces métriques $C_{Id}(\Omega, \mathbb{R}^d)$ et $C_{Id}([0, 1] \times \Omega, \mathbb{R}^d)$ sont complets.

THÉORÈME 4. Soit $v \in L^1_V$. Il existe une unique application $\phi^v \in C_{Id}([0, 1] \times \Omega, \mathbb{R}^d)$ solution de l'équation intégrale

$$\phi(t, x) = x + \int_0^t v(s, \phi(s, x)) ds,$$

pour tout $(x, t) \in \Omega \times [0, 1]$. De plus, $\phi^v(t, \cdot) \in \text{Hom}(\overline{\Omega})$ pour tout $t \in [0, 1]$ et vaut l'identité sur sa restriction à $\partial\Omega$.

NOTATION 6. On utilisera souvent dans la suite la notation $\phi_t(x)$ pour $\phi(t, x)$ et $v_t(x)$ pour $v(t, x)$.

DÉMONSTRATION. Montrons d'abord le lemme suivant :

LEMME 1. Considérons l'application $J : C_{Id}([0, 1] \times \Omega, \mathbb{R}^d) \rightarrow C_{Id}([0, 1] \times \Omega, \mathbb{R}^d)$ définie par :

$$J(\phi)(t, x) = x + \int_0^t v(s, \phi(s, x)) ds$$

pour tout $(t, x) \in [0, 1] \times \overline{\Omega}$. Alors $J^n = \underbrace{J \circ \dots \circ J}_{n \text{ fois}}$ est contractante pour n assez grand.

PREUVE DU LEMME. Vérifions tout d'abord que pour tout $\phi \in C_{Id}([0, 1] \times \Omega, \mathbb{R}^d)$, on a $\psi \triangleq J(\phi) \in C_{Id}([0, 1] \times \Omega, \mathbb{R}^d)$. On a

$$|\psi(t, x) - \psi(t, y)| \leq |x - y| + \int_0^t |\partial_2 v(s, \cdot)|_\infty |\phi(s, x) - \phi(s, y)| ds,$$

où $|\partial_2 v(s, \cdot)|_\infty = \sup_{x \in \Omega, u \in \mathbb{R}^d, |u| \leq 1} \left| \sum_{i=1}^d \frac{\partial}{\partial x_i} v(s, x) u_i \right|$. Par convergence dominée, on a donc $|\psi(t, x) - \psi(t, y)| \rightarrow 0$ lorsque $x \rightarrow y$ ce qui montre que $x \rightarrow \psi(t, x)$ est continue pour tout $t \in [0, 1]$. De plus

$$|\psi(t', \cdot) - \psi(t, \cdot)|_\infty \leq \left| \int_t^{t'} |v(s, \cdot)|_\infty ds \right| \rightarrow 0$$

lorsque $\mathbf{t}' \rightarrow \mathbf{t}$ par convergence dominée. Enfin,

$$\sup_{\mathbf{t}, \mathbf{x}} |\psi(\mathbf{t}, \mathbf{x}) - \mathbf{x}| \leq \int_0^1 |v(s, \cdot)|_\infty ds < \infty,$$

et donc nous avons $\psi \in C_{\text{Id}}([0, 1] \times \Omega, \mathbb{R}^d)$.

Vérifions maintenant que J^n est contractante pour n assez grand. En effet, nous établissons facilement par récurrence que

$$\begin{aligned} |J^n(\phi)(\mathbf{t}, \mathbf{x}) - J^n(\phi')(\mathbf{t}, \mathbf{x})| &\leq \int_0^{\mathbf{t}} |\partial_2 v(s, \cdot)|_\infty |J^{n-1}(\phi)(s, \mathbf{x}) - J^{n-1}(\phi')(s, \mathbf{x})| d\mathbf{r} \\ &\leq \frac{1}{n!} \left(\int_0^1 |\partial_2 v(s, \cdot)|_\infty ds \right)^n \sup_{s, \mathbf{x}} |\phi(s, \mathbf{x}) - \phi'(s, \mathbf{x})| \end{aligned}$$

ce qui donne le résultat puisque $\int_0^1 |\partial_2 v(s, \cdot)|_\infty ds \leq \int_0^1 |v_s|_{1, \infty} ds < \infty$. \square

Par suite, comme $C_{\text{Id}}([0, 1] \times \Omega, \mathbb{R}^d)$ est complet, nous avons par point fixe existence et unicité de $\phi \in C_{\text{Id}}([0, 1] \times \Omega, \mathbb{R}^d)$ telle que $J^n(\phi) = \phi$. En écrivant $J(J^n(\phi)) = J(\phi) = J^n(J(\phi))$, on déduit que $J(\phi) = \phi$. Comme on vérifie facilement que si $J(\phi) = \phi$ alors $J^n(\phi) = \phi$ pour tout n , on déduit que J a un unique point fixe.

De plus, on déduit que le point fixe vérifie $\phi(\mathbf{t}, \mathbf{x}) \in \Omega$ pour tout $\mathbf{t} \in [0, 1]$ et $\mathbf{x} \in \Omega$. En effet, si il existe (s, \mathbf{x}) tels que $\phi(s, \mathbf{x}) \notin \Omega$, alors en considérant $s_* = \inf\{s \in [0, 1] \mid \phi(s, \mathbf{x}) \notin \Omega\}$, on a par continuité de $s \rightarrow \phi(s, \mathbf{x})$ que $\mathbf{y} \triangleq \phi(s_*, \mathbf{x}) \in \partial\Omega$.

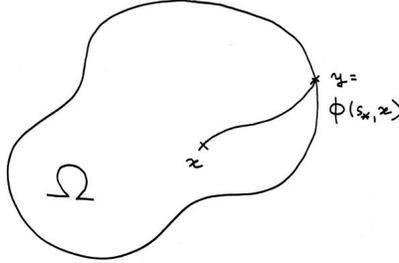


FIG. 2. Définition de s_*

Or comme $v(s, \mathbf{y}) = 0$ pour tout $s \in [0, 1]$, on a pour $0 \leq \mathbf{t} \leq s_*$ et on considérant le flot rétrograde $|\phi(\mathbf{t}, \mathbf{x}) - \mathbf{y}| \leq \int_0^{s_*} |v(s, \phi(s, \mathbf{x})) - v(s, \mathbf{y})| d\mathbf{r} \leq \int_0^{s_*} |\partial_2 v(s, \cdot)|_\infty |\phi(s, \mathbf{x}) - \mathbf{y}| d\mathbf{r} \leq \frac{1}{n!} \left(\int_0^1 |\partial_2 v(s, \cdot)|_\infty ds \right)^n \sup_{s \in [0, s_*]} |\phi(s, \mathbf{x}) - \mathbf{y}|$. En prenant la limite quand $n \rightarrow \infty$, on déduit que $\phi(s, \mathbf{x}) \equiv \mathbf{y}$ sur $[0, s_*]$ ce qui est absurde¹

¹Il faut noter ici au passage que le rôle important joué par le contrôle de la *dérivée* spatiale de v . Si nous enlevons ce contrôle, en particulier si les champs de vitesses ne sont pas Lipschitziens à \mathbf{t} fixé, alors on peut avoir deux trajectoires qui se rejoignent (chercher un exemple). Par contre, on peut être

Il nous reste à vérifier l'inversibilité de $\phi(t, \cdot) : \overline{\Omega} \rightarrow \overline{\Omega}$. Pour cela, il suffit de considérer le flot retourné en temps en posant $\tilde{v}(t, x) = -v(1-t, x)$. Si $\tilde{\phi}$ est le flot intégral pour \tilde{v} , on vérifie que $\tilde{\phi}_t(y) = \phi_{1-t}(x)$ pour $y = \phi_1(x)$ (cf fig 1). En effet, il suffit de remarquer par

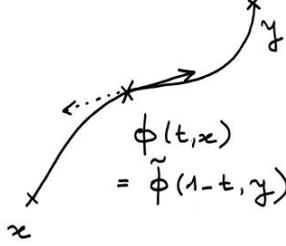


FIG. 3. Retournement du temps

changement de variable que $\phi_{1-t}(x) = y + \int_0^t \tilde{v}(s, \phi_{1-s}(x)) ds$ puis de conclure par unicité du point fixe de \tilde{J} . Ainsi $\tilde{\phi}_1 \circ \phi_1 = \text{Id}_{\overline{\Omega}}$. De même $\phi_1 \circ \tilde{\phi}_1 = \text{Id}_{\overline{\Omega}}$ d'où $\phi_1 \in \text{Hom}(\overline{\Omega})$ et vaut l'identité sur $\partial\Omega$. Comme le point 1 ne joue pas ici de rôle particulier, on déduit le même résultat pour ϕ_t . \square

On note $\text{Diff}^1(\Omega)$ l'ensemble des C^1 difféomorphismes sur Ω . On a alors le théorème suivant :

THÉORÈME 5. *Sous les mêmes hypothèses que celles du Théorème 4, la restriction de ϕ_t^v à Ω est dans $\text{Diff}^1(\Omega)$ et*

$$\partial_2 \phi^v(t, x) = \text{Id} + \int_0^t \partial_2 v(s, \phi(s, x)) \partial_2 \phi^v(s, x) ds,$$

pour tout $(t, x) \in [0, 1] \times \Omega$.

DÉMONSTRATION. Formellement, en partant de l'edo $\partial_1 \phi(t, x) = v(t, \phi(t, x))$ puis en dérivant par rapport à la deuxième variable, on a $\partial_1 \partial_2 \phi(t, x) = \partial_2 \partial_1 \phi(t, x) = \partial_2 v(t, \phi(t, x)) \partial_2 \phi(t, x)$ d'où en intégrant maintenant par rapport à la première variable, $\partial_2 \phi(t, x) = \partial_2 \phi(0, x) + \int_0^t \partial_2 v(s, \phi(s, x)) \partial_2 \phi(s, x) ds$.

On considère, pour tout $x \in \overline{\Omega}$, la solution $A(t, x) \in \mathcal{M}_d(\mathbb{R})$ solution de l'équation intégrale linéaire :

$$A(t, x) = \text{Id} + \int_0^t \partial_2 v(s, \phi(s, x)) A(s, x) ds.$$

beaucoup plus souple sur la régularité temporelle des champs de vitesses puisque ici la seule condition est une condition d'intégrabilité.

EXERCICE 43. Vérifier que la même technique de point fixe que pour le théorème précédent s'applique ici et que l'on a une solution $A \in C([0, 1] \times \overline{\Omega}, \mathcal{M}_d(\mathbb{R}))$

Il nous suffit maintenant de montrer que $\lim_{\epsilon \rightarrow 0} \frac{\phi(t, x + \epsilon h) - \phi(t, x)}{\epsilon} = A(t, x)h$.

On commence par un résultat plus faible. Nous allons montrer qu'il existe $K > 0$ tel que $|\phi(s, x + \epsilon h) - \phi(s, x)| \leq K|\epsilon h|$. En effet, on a

$$\phi(t, x + \epsilon h) - \phi(t, x) = \epsilon h + \int_0^t (v(s, \phi(s, x + \epsilon h)) - v(s, \phi(s, x))) ds$$

d'où

$$(57) \quad |\phi(t, x + \epsilon h) - \phi(t, x)| \leq |\epsilon h| + \int_0^t |\partial_2 v(s, \cdot)|_\infty |\phi(s, x + \epsilon h) - \phi(s, x)| ds.$$

On termine par un lemme de Grönwall dont nous donnons ici une version :

LEMME 2 (Lemme de Grönwall). *Soient $f, g : [0, 1] \rightarrow \mathbb{R}^+$ mesurables telles que f est bornée et il existe une constante $K > 0$ pour laquelle $f(t) \leq K + \int_0^t f(s)g(s)ds$ pour tout $t \in [0, 1]$. Alors,*

$$f(t) \leq K \exp\left(\int_0^t g(s)ds\right).$$

PREUVE DU LEMME DE GRÖNWALL. On montre par récurrence sur $n \geq 0$ que $f(t) \leq K \sum_{i=0}^n \frac{1}{n!} (\int_0^t g(s)ds)^n + \int \mathbb{1}_{0 < s_1 < \dots < s_{n+1}} f(s_1)g(s_1) \cdots g(s_{n+1})ds_1 \cdots ds_{n+1}$ puis on passe à la limite en n . \square

En appliquant Grönwall à (57), on obtient

$$(58) \quad |\phi(t, x + \epsilon h) - \phi(t, x)| \leq |\epsilon h| \exp\left(\int_0^t |\partial_2 v(s, \cdot)|_\infty ds\right) \leq |\epsilon h| \exp\left(\int_0^t |v_s|_{1, \infty} ds\right).$$

ce que nous avions annoncé.

Notons maintenant $\delta(t, \epsilon) \triangleq \phi(t, x + \epsilon h) - \phi(t, x) - \epsilon A(t, x)h$. On a

$$|\delta(t, \epsilon)| = \left| \int_0^t (v(s, \phi(s, x + \epsilon h)) - v(s, \phi(s, x)) - \epsilon \partial_2 v(s, \phi(s, x))A(s, x)h) ds \right|$$

Or $|\nu(s, \mathbf{b}) - \nu(s, \mathbf{a}) - \partial_2 \nu(s, \mathbf{a})(\mathbf{b} - \mathbf{a})| \leq \sqrt{d} |\mathbf{b} - \mathbf{a}| \sup_{c \in [\mathbf{a}, \mathbf{b}]} |\partial_2 \nu(s, c) - \partial_2 \nu(s, \mathbf{a})|$ d'où

$$\begin{aligned} |\delta(t, \epsilon)| &\leq \int_0^t |\partial_2 \nu(s, \phi(s, \mathbf{x})) (\phi(s, \mathbf{x} + \epsilon \mathbf{h}) - \phi(s, \mathbf{x})) - \epsilon \partial_2 \nu(s, \phi(s, \mathbf{x})) A(s, \mathbf{x}) \mathbf{h}| ds \\ &\quad + \underbrace{\sqrt{d} \int_0^t \sup_{c \in [\phi(s, \mathbf{x} + \epsilon \mathbf{h}), \phi(s, \mathbf{x})]} |\partial_2 \nu(s, c) - \partial_2 \nu(s, \phi(s, \mathbf{x}))|_\infty |\phi(s, \mathbf{x} + \epsilon \mathbf{h}) - \phi(s, \mathbf{x})| ds}_{\eta(s, \epsilon)} \\ &\leq \int_0^t |\partial_2 \nu(s, \phi(s, \mathbf{x}))| |\delta(s, \mathbf{x})| ds + \sqrt{d} \exp\left(\int_0^t |\nu_s|_{1, \infty} ds\right) |\epsilon \mathbf{h}| \int_0^t \eta(s, \epsilon) ds \end{aligned}$$

où la dernière inégalité s'obtient en utilisant (58). En appliquant maintenant le lemme de Grönwall, on obtient

$$\sup_{t \in [0, 1]} |\delta(t, \epsilon)| \leq |\epsilon \mathbf{h}| \underbrace{\sqrt{d} \exp\left(2 \int_0^1 |\nu_s|_{1, \infty} ds\right) \int_0^1 \eta(s, \epsilon) ds}_{=o(1)} = o(\epsilon),$$

ce qui termine la preuve du théorème. \square

2. Définition et premières propriétés du groupe G_V

DÉFINITION 15. *A partir de l'application flot $\nu \in L_V^1 \rightarrow \phi^\nu$, on définit*

$$(59) \quad G_V \triangleq \{ \phi_1^\nu \mid \nu \in L_V^1 \}.$$

L'ensemble G_V n'est rien d'autre que l'ensemble des C^1 diffeos obtenus par les flots au temps 1 des champs de vitesses dans L_V^1 . Le résultat fondamental est ici que G_V est un groupe.

THÉORÈME 6. *G_V est un sous-groupe de $\text{Diff}^1(\Omega)$.*

DÉMONSTRATION. On introduit le produit \star dans L_V^1 défini par

$$(\nu \star w)(t, \mathbf{x}) = 2w(2t, \mathbf{w}) \mathbb{1}_{t < 1/2} + 2\nu(2t - 1, \mathbf{x}) \mathbb{1}_{t \geq 1/2}.$$

qui n'est rien d'autre qu'une concaténation de w et ν sur $[0, 1]$. Par un changement de variable, on vérifie que

$$|\nu \star w|_{1, V} = |\nu|_{1, V} + |w|_{1, V}$$

puis que $\phi_1^\nu \circ \phi_1^w = \phi_1^{\nu \star w}$. \square

DÉFINITION 16. *On définit sur G_V ,*

$$d_V(\text{Id}, \varphi) \triangleq \inf\{ |\nu|_{1, V} \mid \nu \in L_V^1 \text{ et } \phi_1^\nu = \varphi \},$$

puis

$$d_V(\varphi, \varphi') \triangleq d_V(\text{Id}, \varphi \circ \varphi'),$$

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

THÉORÈME 7. *L'application d_V définit une distance invariante à droite sur G_V et G_V est complet pour d_V .*

DÉMONSTRATION. On commence par remarquer que l'on a $d_V(\varphi, \varphi') = \inf\{ |v|_{1,V} \mid \phi_1^v \circ \varphi = \varphi' \}$. Soient maintenant $\varphi, \varphi', \varphi'' \in G_V$ et $\epsilon > 0$. Il existe $v, v' \in L_V^1$ tels que $\phi_1^v \circ \varphi = \varphi'$, $\phi_1^{v'} \circ \varphi' = \varphi''$, $|v|_{1,V} \leq d_V(\varphi, \varphi') + \epsilon$ et $|v'|_{1,V} \leq d_V(\varphi', \varphi'') + \epsilon$. On a alors $\phi_1^{v \star v'} \circ \varphi = \varphi''$ d'où $d_V(\varphi, \varphi'') \leq |v \star v'|_{1,V} = |v|_{1,V} + |v'|_{1,V} \leq d_V(\varphi, \varphi') + d_V(\varphi', \varphi'') + 2\epsilon$. On a donc montré l'inégalité triangulaire.

EXERCICE 44. Vérifier l'invariance à droite et montrer que $d_V(\varphi, \varphi') = 0 \Rightarrow \varphi = \varphi'$.

Montrons que G_V est complet pour la distance d_V . Soit donc $(\varphi_n)_{n \in \mathbb{N}}$ une suite de Cauchy. Nous pouvons quitte à extraire une sous-suite supposer que $\sum_{n \geq 0} d_V(\varphi_n, \varphi_{n+1}) < \infty$. On choisit maintenant pour tout $n \geq 0$ un $v_n \in L_V^1$ tel que $\phi_1^{v_n} \circ \varphi_n = \varphi_{n+1}$ et $|v_n|_{1,V} \leq 2d_V(\varphi_n, \varphi_{n+1})$.

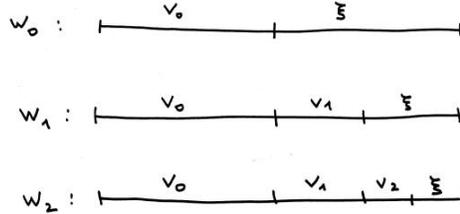


FIG. 4. Définition des w_n

On note $\xi \equiv 0$ le champ de L_V^1 constant 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 $\varphi_{n+1} = \phi_1^{w_n} \circ \varphi_0$ et $|w_n|_{1,V} = \sum_{k=0}^n |v_k|_{1,V} < 2 \sum_{k=0}^{\infty} d_V(\varphi_k, \varphi_{k+1}) < \infty$. Comme par ailleurs $|w_{n+1} - w_n|_{1,V} = |v_n|_{1,V}$ 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 $\phi_\infty = \phi_1^{w_\infty} \circ \varphi_0 \in G_V$ on vérifie que $\phi_\infty = \phi_1^{w_\infty - w_{n-1}} \circ \varphi_n$ d'où $d_G(\phi_\infty, \varphi_n) \leq |w_\infty - w_n|_{1,V} \rightarrow 0$ lorsque $n \rightarrow \infty$. \square

3. 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 17. 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 $|v|_{2,V} \triangleq (\int_0^1 |v_s|_V^2 ds)^{1/2}$ qui fait de L_V^2 un Hilbert.

REMARQUE 21. L_V^2 est séparable dès que V l'est. On se placera dorénavant dans ce cas.

PROPOSITION 6. On a pour tous $\varphi, \varphi' \in G_V$:

$$d_V(\varphi, \varphi') = \inf\{ |v|_{2,V} \mid \phi_1^v \circ \varphi = \varphi' \}.$$

REMARQUE 22. 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 $\varphi' = \phi_1^v \circ \varphi$. En particulier $G_V = \{ \phi_1^v \mid v \in L_V^2 \}$.

DÉMONSTRATION. On remarque d'abord que pour tout $v \in L_V^2$, on a $|v|_{1,V} \leq |v|_{2,V}$ et donc $d_V(\varphi, \varphi') \leq \inf\{ |v|_{2,V} \mid \phi_1^v \circ \varphi = \varphi' \}$.

Montrons l'autre inégalité. Soit $\epsilon > 0$ et $v \in L_V^1$ tel que $\phi_1^v \circ \varphi = \varphi'$ et $|v|_{1,V} \leq d_V(\varphi, \varphi') + \epsilon$. Soit $\eta > 0$ et

$$L(t) \triangleq \int_0^t (|v_s|_V + \eta) ds / \int_0^1 (|v_s|_V + \eta) ds.$$

L'application L est absolument continue strictement croissante bijective de $[0, 1]$ dans $[0, 1]$ d'inverse T absolument continue de dérivée presque partout : $\frac{dT}{dt} = \frac{\int_0^1 (|v_s|_V + \eta) ds}{|v_{T(t)}|_V + \eta}$. En notant $\tilde{v}_t = v_{T(t)} \frac{dT}{dt}$, on a $|\tilde{v}_t|_V = \frac{|v_{T(t)}|_V}{|v_{T(t)}|_V + \eta} \int_0^1 (|v_s|_V + \eta) ds \leq \int_0^1 (|v_s|_V + \eta) ds$ d'où $|\tilde{v}|_{2,V} \leq |v|_{1,V} + \eta$. Or $\phi_1^{\tilde{v}} = \phi_1^v$ (il suffit de faire un changement de variable) d'où $|v|_{2,V} + |\eta| \leq d_V(\varphi, \varphi') + \epsilon$. Comme η est arbitraire, on obtient $\inf\{ |v|_{2,V} \mid \phi_1^v \circ \varphi = \varphi' \} \leq d_V(\varphi, \varphi') + \epsilon$. \square

THÉORÈME 8. L'application flot $\phi : L_V^2 \rightarrow C_{Id}([0, 1] \times \Omega, \mathbb{R}^d)$ 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 $[0, 1] \times \overline{\Omega}$ sur $C_{Id}([0, 1] \times \Omega, \mathbb{R}^d)$.

DÉMONSTRATION. Soit $(v_n)_{n \geq 0}$ une suite de L_V^2 convergeant faiblement vers v . On a

$$(60) \quad \phi^{v_n}(t, x) - \phi^v(t, x) = \int_0^t (v_n - v)(s, \phi^v(s, x)) ds + \int_0^t v_n(s, \phi^{v_n}(s, x)) - v_n(s, \phi^v(s, x)) ds.$$

4. EXISTENCE DE GÉODÉSQUES MINIMISANTES ET SOLUTION DE PROBLÈMES VARIATIONNELS

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

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

d'où la suite ϵ_n est équicontinue en t . Or comme $w \rightarrow \int_0^t w(s, \phi^v(s, x)) ds$ 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 $|\phi^{v_n}(t, x) - \phi^v(t, x)| \leq |\epsilon_n|_\infty + \int_0^t |\partial_2 v_n(s, \cdot)|_\infty |\phi^{v_n}(s, x) - \phi^v(s, x)| ds$ tirée de (60), on obtient

$$|\phi^{v_n}(t, x) - \phi^v(t, x)| \leq |\epsilon_n|_\infty \exp\left(\int_0^1 |\partial_2 v_n(s, \cdot)|_\infty ds\right) \leq |\epsilon_n|_\infty \exp(K|v_n|_{2,V}) \rightarrow 0$$

car $\sup_{n \geq 0} |v_n|_{2,V} < \infty$ et $|\epsilon_n|_\infty \rightarrow 0$ lorsque $n \rightarrow \infty$ par convergence uniforme.

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

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

qui donne par Grönwall

$$|\phi^{v_n}(t, x) - \phi^{v_n}(t, y)| \leq |x - y| \exp(C \sup_{n \geq 0} |v_n|_{2,V}),$$

et

$$|\phi^{v_n}(t, x) - \phi^{v_n}(t + h, x)| \leq C\sqrt{h} \sup_{n \geq 0} |v_n|_{2,V}.$$

□

4. 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 9. *Pour tous $\varphi, \varphi' \in G_V$ il existe $u \in L_V^2$ tel que*

$$d_V(\varphi, \varphi') = |u|_{1,V} = |u|_{2,V}$$

et $\varphi_1 = \phi_1^u \circ \varphi$.

REMARQUE 23. La courbe $t \rightarrow \varphi_t \triangleq \phi_t^v \circ \varphi$ est une courbe joignant $\varphi = \varphi_0$ à $\varphi' = \varphi_1$ de longueur $d_V(\varphi, \varphi')$. L'égalité $|u|_{1,V} = |u|_{2,V}$ assure que $|u_t|_V \equiv Cste$ si bien que $t \rightarrow \varphi_t$ est une géodésique à vitesse constante entre φ et φ' .

DÉMONSTRATION. Soit $F \triangleq \{ u \in L_V^2 \mid \varphi' = \phi_1^u \circ \varphi \}$. Le résultat de continuité faible nous donne que F est fermé pour la topologie faible. De plus on déduit de la remarque 22 que F est non vide. Il faut montrer que l'infimum de l'énergie $|u|_{2,V}$ pour $u \in F$ est atteint. Soit donc $(u_n)_{n \geq 0}$ une suite minimisante. Comme $|u_n|_{2,V}$ est borné on peut supposer quitte à extraire une sous-suite, que $u_n \rightharpoonup u_\infty \in F$. Or $|u_\infty|_{2,V} \leq \liminf |u_n|_{2,V} = \inf_{u \in F} |u|_{2,V}$ (car la norme forte est s.c.i. pour la topologie faible) donc $d_V(\varphi, \varphi') \leq |u_\infty|_{1,V} \leq |u_\infty|_{2,V} = d_V(\varphi, \varphi')$ d'après la proposition 6. \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 10. Soit $E : G_V \rightarrow \mathbb{R}^+$ une application semi-continue inférieurement pour la convergence uniforme sur tous les compacts de $\overline{\Omega}$ et $R : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ une application croissante telle que $R(\rho) \rightarrow +\infty$ lorsque $\rho \rightarrow +\infty$. On pose

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

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

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

EXERCICE 45. Montrer le théorème 10.

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

Bibliographie

- [1] N. Aronszajn. La théorie des noyaux reproduisants et ses applications. *Proc. Cambridge Philos. Soc.*, pages 133–153, 1943.
- [2] M. Berger and B. Gostiaux. *Géométrie différentielle, variétés, courbes et surfaces*. Presses Universitaires de France, 1992.
- [3] L. Bookstein, F. Principal warps : Thin plate splines and the decomposition of deformations. *IEEE trans. PAMI*, 11(6) :567–585, 1989.
- [4] W. Boothby. *An Introduction to Differentiable Manifolds and Riemannian Geometry*. Academic Press, 2002. Edition originale 1986.
- [5] H. Brezis. *Analyse fonctionnelle*. Dunod, 1983.
- [6] S. Gallot, D. Hullin, and J. Lafontaine. *Riemannian Geometry*. Springer, 1993.
- [7] 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.
- [8] J. Glaunes, A. Trouvé, and L. Younes. Diffeomorphic matching of distributions : A new approach for unlabelled point-sets and sub-manifolds matching. In *Computer Vision and Pattern Recognition (Washington DC)*, 2004.
- [9] M. Hein and O. Bousquet. Kernels, associated structures and generalizations. Technical report, Max Plack Institute for Biological Cybernetics, July 2004.
- [10] S. Helgason. *Differential Geometry, Lie groups and Symmetric spaces*. Academic Press, 1978.
- [11] A. Kriegl and W. Michor, P. Regular infinite dimensional lie groups. *J. of Lie Theory*, 7(1) :61–99, 1997.
- [12] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. In *Proceedings of the Third Intl. Conf. on 3D Digital Imaging and Modeling*, pages 145–152, 2001.
- [13] L. Younes. *Invariance, déformations et reconnaissance de formes*. Springer, 2004.