

# Espaces de formes et applications à l'imagerie médicale

Nathan Kessler

Septembre 2023

## 1 Introduction et motivations

Le problème de l'anatomie computationnelle consiste à essayer d'évaluer quantitativement la variabilité entre des individus ou des structures biologiques. Peut-être l'un des pionniers de ce domaine est D'Arcy Thompson, qui montrait qu'on pouvait passer de la morphologie d'une espèce à l'autre via une simple transformation géométrique.

Aujourd'hui, cette idée a trouvé son application dans des problèmes d'imagerie médicale : étant donné un organe de deux individus différents, ils n'ont a priori pas exactement la même forme, mais il est parfois possible d'effectuer un "recalage" entre les deux formes, c'est-à-dire de trouver un système de coordonnées commun pour ensuite pouvoir les comparer. Cela peut-être obtenu en déformant habilement la première forme sur la deuxième.

Un tel recalage nous donne également des informations sur la différence anatomique entre ces deux organes : s'il a fallu grossir une zone de l'organe pour la recaler correctement sur le deuxième, cela signifie que cette zone est plus petite chez le premier individu. Si ensuite on se donne encore plus d'individus, on peut commencer à faire des statistiques et étudier précisément la variabilité de toutes ces zones.

Le point de vue qu'on prend dans ce papier est celui des LDDMM (large deformation diffeomorphic metric mapping) [2]. L'idée est de faire de l'espace des formes un espace riemannien, où les recalages entre deux formes sont obtenus via des trajectoires géodésiques.

Travailler dans un espace riemannien a beaucoup d'avantages : de nombreuses méthodes de statistiques sont encore disponibles dans les espaces riemanniens [7]. On peut par exemple généraliser : la notion de moyenne, le concept de régression linéaire, qui devient une régression selon une géodésique, l'analyse en composantes principales (PCA, méthode qui cherche à trouver les axes majeurs de variabilité dans un jeu de données), les algorithmes de clustering...

Pour obtenir une métrique riemannienne sur l'espace des formes, on va prendre un groupe de Lie (en l'occurrence le groupe des difféomorphismes de  $\mathbb{R}^d$ ) qui agit sur ces formes et munir ce groupe d'une métrique riemannienne invariante. Cette approche est un héritage d'une idée de Vladimir Arnold [1], qui

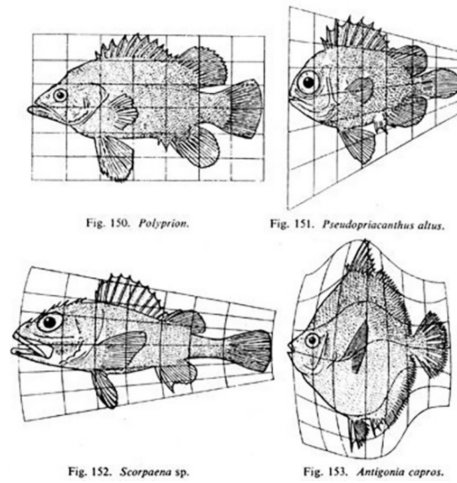


FIGURE 1 – Illustration de D’Arcy Thompson dans son livre ”On Growth and Form”.

décrit comme certaines équations de la physique, comme la rotation d’un objet ou l’évolution d’un fluide non visqueux, peuvent être décrites à travers l’action d’un groupe de Lie muni d’une métrique riemannienne invariante (respectivement  $SO_3$  et  $SDiff$ , l’espace des difféomorphismes qui préservent le volume).

## 2 Représentation des objets

Quelle représentation pour nos objets? Traditionnellement on peut voir notre objet comme l’image d’une immersion  $j : M \rightarrow \mathbb{R}^d$  pour  $d = 2$  ou  $3$  et où  $M$  est une variété. C’est notamment l’approche qui était employée dans les premiers articles sur le sujet (voir par exemple [10]), où on cherchait à trouver une métrique sur l’espace de telles immersions, quotienté par les difféomorphismes laissant l’image de l’immersion invariante.

L’action d’un difféomorphisme  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}^d$  sur une telle immersion  $j$  est alors :

$$\varphi.j = \varphi \circ j. \quad (1)$$

Il s’agit ici d’une action à gauche car la composition d’un difféomorphisme  $\varphi$  puis  $\psi$  sur l’objet donne naturellement

$$\psi.(\varphi.j) = (\psi \circ \varphi).j.$$

Le défaut d’une telle représentation est qu’elle est assez peu pratique pour manier des véritables données. En général, on a plutôt accès à une intensité en chaque point voire à un nuage de point. Il est alors dans ce cas plus judicieux de travailler avec une distribution positive  $\mu : \mathbb{R}^d \rightarrow \mathbb{R}_+$  qui va représenter la

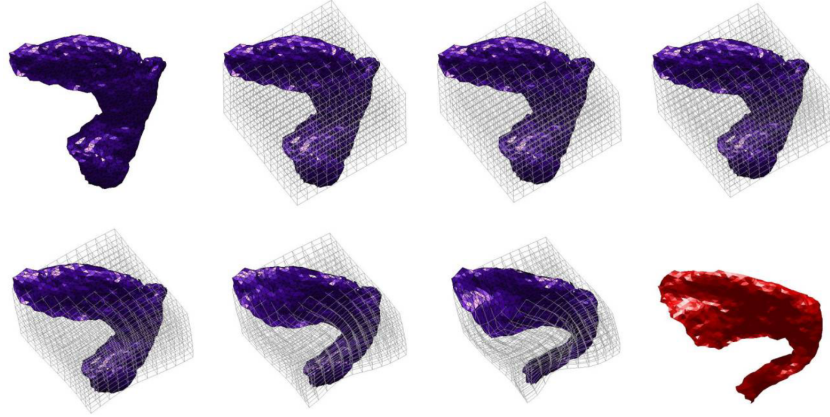


FIGURE 2 – Recalage de planum tempore entre deux individus provenant de [8]. Le recalage final est l'avant-dernière objet, tandis que l'objectif est l'objet en rouge.

densité de notre objet dans  $\mathbb{R}^d$  (il s'agit alors d'une somme de diracs dans le cas d'un nuage de points).

Cette représentation recouvre la précédente : étant donné une immersion  $j$  on peut définir la distribution  $\mu_j$  en définissant l'action contre une fonction  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  comme

$$(\mu_j | f) = \int_M f(j(m)) d\text{Vol}_M(m).$$

L'analogie pour l'action d'un difféomorphisme se trouve alors être :

$$\varphi.\mu = \mu \circ \varphi^{-1}.$$

Il s'agit toujours d'une action à gauche, qui récupère la même action que dans (1) pour le cas d'une immersion.

Variations : on peut demander à ce que  $\mu$  ait plusieurs canaux de sortie, ce qui revient à dire que  $\mu$  est la donnée de  $F$  distributions, où  $F \in \mathbb{N}^*$ , ou alors que  $\mu$  est une distribution positive de domaine  $\mathbb{R}^d \times \{0, 1, \dots, F - 1\}$ . Par exemple, un nuage de points avec des labels correspond aux telles distributions  $\mu$  qui sont des sommes de Dirac.

Aussi, si jamais l'objet qu'on veut représenter à une dimension intrinsèque  $d'$  strictement inférieure à  $d$  (le plus commun étant la surface d'un objet dans  $\mathbb{R}^3$ ) il peut être intéressant de garder la donnée d'espace tangent en chaque point. Cela peut se faire en voyant  $\mu$  non pas seulement comme une mesure de domaine  $\mathbb{R}^d$ , mais comme une mesure de domaine  $\mathbb{R}^d \times G_{d'}(d)$ , où  $G_{d'}(d)$  est l'espace des sous-espaces vectoriels de dimension  $d'$  dans  $\mathbb{R}^d$ . Cette approche est fortement inspirée du concept de varifold, notion d'abord introduite pour l'étude des surfaces minimales, et qui cherche partiellement à généraliser la notion de sous-variété (voir par exemple [9]). Dans ce cas, les difféomorphismes agissent

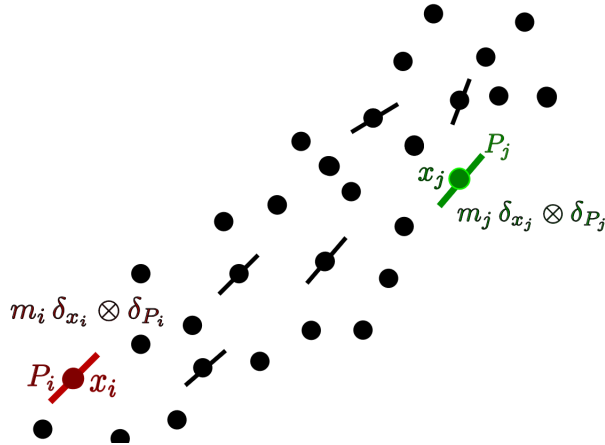


FIGURE 3 – Exemple de 'varifold' associé à un nuage de point avec la donnée d'un axe en chaque point, qu'on peut écrire  $\sum_i m_i \delta_{x_i} \otimes \delta_{P_i}$ , où  $x_i$  donne la position,  $P_i$  l'axe et  $m_i$  la masse accordée au point.

naturellement sur  $G_{d'}(d)$  : une base de  $E \in G_{d'}(d)$  est transportée par l'action de la différentielle du difféomorphisme.

On peut également trouver des variations plus exotiques, comme les IRM de diffusion où on peut voir  $\mu$  comme une distribution de domaine  $\mathbb{R}^d \times \mathbb{S}_d^{++}$ , où  $\mathbb{S}_d^{++}$  est l'ensemble des matrices définies positives. Ces matrices s'interprètent, en chaque point, comme la matrice de covariance du mouvement brownien de l'eau dans les tissus du cerveau, et idéalement le premier vecteur propre donne la direction des faisceaux de fibres nerveuses.

### 3 Métrique sur les difféomorphismes

On veut maintenant munir notre espace de formes d'une métrique, qui sera induite par une métrique sur le groupe de Lie des difféomorphismes. Plus précisément, on se donne une mesure  $\mu$  comme dans le paragraphe précédent, et on regarde son orbite sous les difféomorphismes, c'est-à-dire  $O_\mu = \{\varphi.\mu \mid \varphi \in \text{Diff}(\mathbb{R}^d)\}$ . On veut alors une distance sur le groupe des difféomorphismes de sorte que la distance sur  $O_\mu$  soit donnée par :

$$d(\varphi.\mu, \psi.\mu) = d(\varphi, \psi).$$

Cela suppose que la métrique qu'on met sur les difféomorphismes vérifie une condition de compatibilité, il faut qu'elle soit invariante à droite, c'est-à-dire, pour tous difféomorphismes  $\phi, \varphi, \psi$ ,

$$d(\varphi \circ \phi, \psi \circ \phi) = d(\varphi, \psi).$$

Cette condition s'exprime pour une métrique riemannienne localement de la manière suivante : on se donne  $\delta\varphi$  et  $\delta\psi$  deux éléments de l'espace tangent

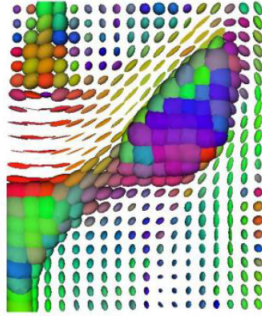


FIGURE 4 – Image obtenue d’une IRM de diffusion, où chaque ellipse représente une matrice symétrique définie positive. La couleur vert-rouge-bleu code l’axe préférentiel de l’ellipse. Vert : haut-bas, rouge : gauche-droite, bleu : devant-derrrière.

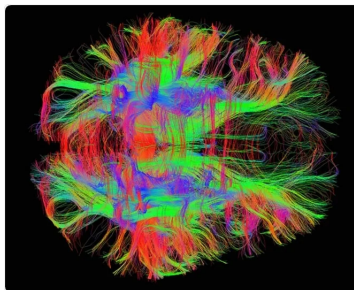


FIGURE 5 – Image plus haut niveau issue d’une IRM de diffusion, où on a estimé des faisceaux de fibres nerveuses.

en l'identité (ce qui correspond à deux champs de vecteurs). Alors pour tout difféomorphisme  $\phi$

$$\begin{aligned}\langle \delta\varphi, \delta\psi \rangle_{\text{Id}} &= \langle dR_\phi(\delta\varphi), dR_\phi(\delta\psi) \rangle_\phi \\ &= \langle (\delta\varphi) \circ \phi, (\delta\psi) \circ \phi \rangle,\end{aligned}$$

où  $R_\phi$  est la multiplication à droite par  $\phi$ , et  $dR_\phi$  la différentielle de cette application.

Ainsi, la condition d'invariance de la métrique à droite détermine automatiquement la métrique tant qu'on décide de comment se comporte le produit scalaire en l'espace tangent à l'identité. Cet espace n'est autre que l'espace vectoriel des champs de vecteurs  $C^\infty$ , il ne nous reste plus qu'à trouver un produit scalaire sur cet espace (cet espace n'étant pas métrisable, on va plutôt se restreindre à un sous-espace assez grand de champs de vecteurs, mais nous omettons cette distinction par la suite).

## 4 Métrique sur l'espace tangent en l'identité

Le choix du produit scalaire est une étape cruciale pour obtenir une métrique de déformation qui donne des géodésiques satisfaisantes. Notamment, si on se donne deux champs de vecteurs  $v$  et  $w$ , le premier choix de métrique qui nous vient à l'esprit est

$$\langle v, w \rangle = \int_{\mathbb{R}^d} \langle v(x), w(x) \rangle dx.$$

Cette métrique est cependant problématique. En effet, dès lors que le support de notre objet a une mesure nulle par rapport à la mesure de Lebesgue sur  $\mathbb{R}^d$  (ce qui est le cas pour un nuage de points) le déplacement de notre objet devient complètement gratuit, puisqu'il suffit de prendre un champ de vecteurs qui vaut la valeur voulue en le support de notre objet et nul ailleurs. De manière générale, cette métrique a le défaut de laisser deux vecteurs très proches avoir des directions très différentes, alors qu'on cherche plutôt à avoir des déformations assez lisses : il faut forcer les points proches à avoir des vitesses corrélées.

Une façon de régler ce problème, et qui est la solution que nous allons employer, est de pénaliser les dérivées de notre champ de vecteurs. Pour cela on peut prendre une métrique du type

$$\langle v, w \rangle = \sum_i a_i \int_{\mathbb{R}^d} \langle D^i v(x), D^i w(x) \rangle dx,$$

où  $D^i$  est la différentielle appliquée  $i$  fois et les  $a_i$  sont des coefficients qu'on décide de prendre tous strictement positifs.

Un tel produit scalaire serait assez difficile à évaluer en pratique, mais un calcul efficace est possible dans certains cas particuliers : c'est l'objet du paragraphe suivant.

## 5 Espaces à noyau reproduisant

En général, on appelle noyau défini positif sur un espace  $X$  (sans aucune structure a priori, même si on est plus intéressés au cas  $X = \mathbb{R}^d$ ) une fonction  $k(\cdot, \cdot)$  symétrique dont la sortie est une matrice carrée de dimensions  $d$  et vérifiant, pour tous  $x_i \in X$  et  $c_i \in \mathbb{R}^d$ ,

$$\sum_{i,j} c_i^T k(x_i, x_j) c_j \geq 0.$$

Un tel noyau défini positif définit un espace de Hilbert sur un sous-espace des applications  $X \rightarrow \mathbb{R}^d$ , plus exactement sur le complété de l'ensemble

$$\left\{ \sum_i k(x_i, \cdot) c_i \mid x_i \in X, c_i \in \mathbb{R}^d \right\}$$

muni du produit scalaire

$$\left\langle \sum_i k(x_i, \cdot) c_i, \sum_i k(y_i, \cdot) c'_i \right\rangle = \sum_{i,j} c_i^T k(x_i, y_j) c'_j.$$

Un tel espace est alors nommé espace à noyau reproduisant (reproducing kernel Hilbert space en anglais, ou RKHS). Il s'agit d'un concept abondamment utilisé en machine learning pour généraliser des algorithmes s'appuyant normalement sur un produit scalaire sur  $X$  (exemples : SVM et PCA).

Réciproquement, on se donne un espace de Hilbert  $V$  de champs de vecteurs de  $X$  dans  $\mathbb{R}^d$  (avec toujours aucune structure sur  $X$ ) ainsi que son dual  $V^*$ , et on demande que  $V^*$  contienne les distributions de Dirac  $c\delta_x$  où  $c \in \mathbb{R}^d$  (la sortie est vectorielle et non scalaire, d'où le vecteur  $c$ ). On a une application canonique  $K : V \rightarrow V^*$  qui envoie  $v$  sur  $Kv = \langle v, \cdot \rangle$ . Par le théorème de représentation de Riesz,  $K$  est un isomorphisme. On définit alors  $k$  par la formule

$$k(x, y)c = K^{-1}(c\delta_x)(y).$$

Cette fonction  $k$  est un noyau défini positif, et son espace à noyau reproduisant est  $V$ . Remarquons au passage que nous avons la formule

$$(c\delta_x | v) = \langle K^{-1}(c\delta_x), v \rangle_V \tag{2}$$

$$= \langle k(x, \cdot)c, v \rangle_V. \tag{3}$$

Ainsi l'aspect intéressant de cet objet pour nous est le suivant : étant donné deux distributions de champ de vecteurs  $v = \sum_i c_i \delta_{x_i}$  et  $w = \sum_i c'_i \delta_{y_i}$  dans  $V^*$ , le produit scalaire entre  $v$  et  $w$  a une forme simple :

$$\langle v, w \rangle = \langle v | K^{-1}w \rangle = \sum_{i,j} c_i^T k(x_i, y_j) c'_j.$$

Si  $k$  a une forme simple, le produit scalaire donc est facile à évaluer dans l'espace dual pour des sommes de Dirac. Par exemple, si  $V$  est muni du produit scalaire

$$\langle v, w \rangle = \sum_i \frac{1}{i!(2\sigma)^i} \int_{\mathbb{R}} \langle D^i v(x), D^i w(x) \rangle dx,$$

le noyau associé  $k$  a l'expression

$$k_{\sigma}(x, y) = e^{-(x-y)^2/2\sigma^2}.$$

Il s'agit du noyau gaussien, qui grossièrement tend à demander de la régularité jusqu'à une portée  $\sigma$ . On appelle  $V_{\sigma}$  l'espace de fonctions associé.

## 6 Shooting géodésique en pratique

Supposons que l'on ait deux objets  $\mu_0$  et  $\mu^*$  qui sont dans la même orbite via l'action des difféomorphismes, décrivons la géodésique les reliant. On se donne une application  $v : [0, 1] \rightarrow V_{\sigma}$  qui représente un champ de vecteurs dans le temps, et une application  $\phi : [0, 1] \rightarrow \text{Diff}(\mathbb{R}^d)$  vérifiant  $\phi(0) = \text{Id}$ ,  $\phi(1) \cdot \mu_0 = \mu^*$  et

$$\begin{aligned} \frac{d}{dt} \phi(t) &= dR_{\phi(t)}(v(t)) \\ &= v(t) \circ \phi(t). \end{aligned}$$

$\phi$  représente donc un chemin entre  $\mu_0$  et  $\mu^*$ . Etant donné la métrique invariante à droite mise sur le groupe des difféomorphismes, pour que  $\phi$  soit une géodésique il faut minimiser la fonctionnelle (ou le lagrangien)

$$\mathcal{L}(v) = \frac{1}{2} \int_0^1 |v(t)|_{V_{\sigma}}^2 dt.$$

On fait maintenant l'hypothèse que  $\mu_0$  est un nuage de points,  $\mu_0 = \sum_i m_i \delta_{x_i}$ . On a alors  $\mu_t = \phi(t) \cdot \mu_0 = \sum_i m_i \delta_{x_i(t)}$ . On minimise le lagrangien en passant par la formulation hamiltonienne : on attache à chaque point  $x_i$  un vecteur moment  $p_i(t) \in \mathbb{R}^d$  et on introduit le hamiltonien

$$\begin{aligned} \mathcal{H}(x(t), p(t), v(t)) &= \sum_i \langle p_i(t), \dot{x}_i(t) \rangle_{\mathbb{R}^d} - \frac{1}{2} |v(t)|_{V_{\sigma}}^2 \\ &= \sum_i (p_i \delta_{x_i} |v) - \frac{1}{2} |v|_{V_{\sigma}}^2 \\ &= \sum_i \langle k_{\sigma}(x_i, \cdot) p_i, v \rangle_{V_{\sigma}} - \frac{1}{2} |v|_{V_{\sigma}}^2, \end{aligned}$$

où on a retiré la dépendance en temps à partir de la deuxième ligne pour la lisibilité. Le passage de la deuxième à la troisième a été obtenu avec (2).

D'après le principe du maximum de Pontryagin, une trajectoire qui minimise le lagrangien vérifie  $\frac{\partial \mathcal{H}}{\partial v} = 0$ , on en déduit donc  $v = \sum_i k_\sigma(x_i, \cdot) p_i$ , et la forme réduite du hamiltonien :

$$\mathcal{H}(x, p) = \frac{1}{2} \sum_{i,j} p_i^T k_\sigma(x_i, x_j) p_j.$$

La dynamique est alors donnée par les équations  $\dot{x} = \frac{\partial \mathcal{H}}{\partial p}$  et  $\dot{p} = -\frac{\partial \mathcal{H}}{\partial x}$  qu'on peut aisément exprimer.

Même dans le cas où on a une centaine de milliers de points, ces équations peuvent s'intégrer numériquement en utilisant habilement les capacités d'un processeur graphique (GPU) : en effet, la plupart des termes durs à calculer sont des sommes double qui parcourent les points, mais ces calculs peuvent être massivement parallélisés sur un GPU. Il existe désormais des interfaces de programmation haut niveau qui permettent d'effectuer facilement ce genre de routines, comme la librairie python KeOps [3, 5] qui a été créée en ce but.

Il reste cependant quelques problèmes pratiques : premièrement, intégrer les équations hamiltoniennes donnera bien une géodésique entre  $\mu_0$  et  $\phi(1) \cdot \mu_0 = \mu_1$ , mais on n'a a priori pas que  $\mu_1 \neq \mu^*$ . Il faut donc trouver quel  $p(0)$  donne bien la géodésique entre  $\mu_0$  et  $\mu^*$ .

Deuxièmement, en pratique  $\mu^*$  n'est pas dans la même orbite pour les difféomorphismes que  $\mu_0$ , par exemple si  $\mu_0$  et  $\mu^*$  sont des nuages de points, simplement car ils n'ont pas le même nombre de points. Pourtant cette différence n'est qu'une forme de bruit pour nous, on va donc prendre un objectif  $\mu_1$  dans l'orbite de  $\mu_0$  et évaluer la différence entre  $\mu_1$  et  $\mu^*$  pour une certaine métrique  $d(\cdot, \cdot)$ . On arrive ainsi au problème de minimisation de la fonctionnelle suivante :

$$\begin{aligned} L(p(0)) &= \frac{1}{2} \int_0^1 |v(t)|_{V_\sigma}^2 dt + \lambda d(\mu_1, \mu^*) \\ &= \frac{1}{2} \sum_{i,j} p_i^T(0) k_\sigma(x_i(0), x_j(0)) p_j(0) + \lambda d(\mu_1, \mu^*), \end{aligned}$$

où  $\lambda$  est un paramètre à régler. La deuxième ligne a été obtenue en sachant que le hamiltonien (i.e. l'énergie) est constant au cours du temps.

En pratique, la métrique qu'on choisit entre  $\mu_1$  et  $\mu^*$  est, comme pour les vitesses, une norme à noyaux  $|\mu_1 - \mu^*|_{V_{\sigma'}}$  sur les champs scalaires. La raison intuitive pour cela est la suivante : un nuage de points correspond à un genre de "densité" mais où la précision est assez faible, au contraire d'une densité représentée par une fonction lisse. On voit ici une densité comme une distribution qui doit être évaluée contre des fonctions. Comme ici notre précision est "faible", il faut évaluer notre distribution contre des fonctions particulièrement lisses. C'est pourquoi on décide de comparer deux distributions de points (i.e. deux formes) via une norme duale d'un espace de fonctions lisses  $V_{\sigma'}$ . Autre avantage d'une norme à noyaux : pour les raisons citées ci-dessus on peut aisément les calculer numériquement.

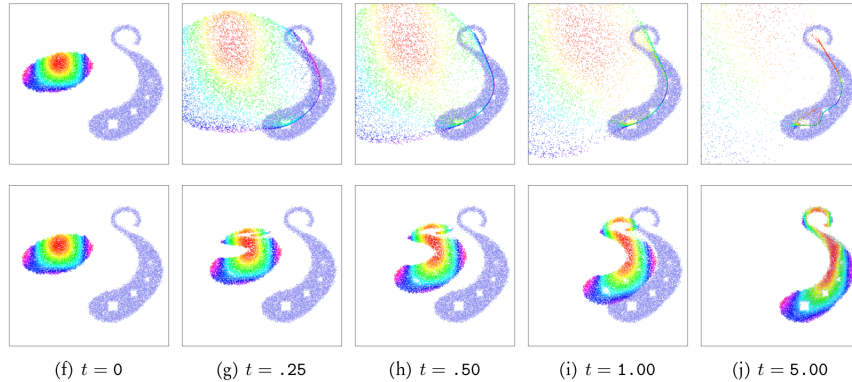


FIGURE 6 – Comparaison de correspondances en employant dans le terme d’erreur respectivement un noyau gaussien et un terme de transport optimal, images tirées de [4].

On a ainsi un algorithme permettant de trouver une géodésique entre  $\mu_0$  et  $\mu_1$  une forme proche de  $\mu^*$  : il suffit d’effectuer une descente de gradient sur la fonctionnelle  $L$  ci-dessus, en intégrant numériquement  $\mu_1$  à chaque étape. Nul besoin de calculer des milliers de dérivées pour écrire l’algorithme, l’expression littérale de celles-ci est calculée directement via des bibliothèques comme PyTorch [11]. Cela économise de nombreux calculs laborieux.

Néanmoins, faire une descente de gradient directement sur ce problème n’aboutit pas toujours au résultat attendu. Une réponse à ce problème est donnée dans [6], qui propose de remplacer l’erreur de norme à noyaux en une erreur donnée par un terme de transport optimal, ce qui permet au gradient lors de la descente d’indiquer une direction plus plausible pour faire correspondre les deux formes. Le terme de transport optimal peut se calculer très rapidement grâce à la bibliothèque GeomLoss.

## 7 Conclusion et ouvertures

On a présenté dans ce court papier comment il est possible de faire d’un espace de formes un espace riemannien, en pouvant déformer une forme en une autre. L’approche qu’on a présenté permet une implémentation numérique, et les bibliothèques actuelles permettent de le coder assez aisément et sur des volumes de données assez grands. Cela nous permet d’expliquer la source de variabilité dans certains jeux de données médicaux et de faire des recalages propres.

Il est à noter que la déformation entre certaines données reste cependant intraitable pour des raisons computationnelles. Donnons l’exemple des transcritomiques spatiales : il s’agit d’un grand nuage de points en 2D ou 3D, dans lequel chaque point représente un gène qui a été aperçu, et où on garde la donnée de quel gène a été observé. Sur un seul individu, un tel nuage peut cor-

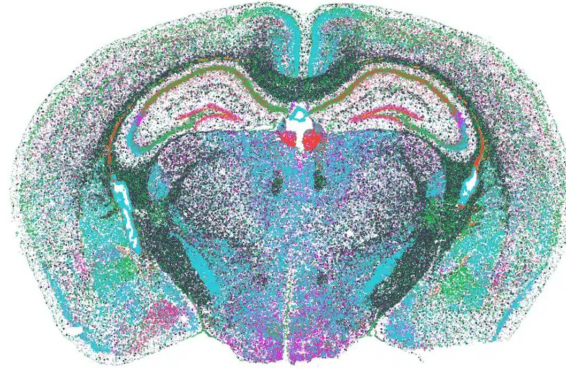


FIGURE 7 – Exemple de transcriptomiques spatiales issues d’une tranche de cerveau de souris, avec 483 types de gènes distincts, chacun représenté par une couleur distincte.

respondre à plusieurs milliards de points, avec jusqu’à environ 1000 types de gènes différents observés.

De plus il n’est pas toujours possible d’effectuer un recalage ainsi, quand la variabilité entre deux formes dépasse une simple déformation. La question se pose alors de s’il est toujours possible de définir un espace riemannien pertinent pour ces formes, en autorisant des transformations plus violentes. Peut-être qu’il devient nécessaire de mieux comprendre les formes que nous avons sous la main, et de faire de l’apprentissage sur les données pour comprendre quelles transformations sur les caractéristiques de l’objet sont les plus pertinentes. Précisons qu’à l’heure actuelle, il n’y a pas de moyen automatique standardisé de choisir le produit scalaire sur l’espace de champs de vecteurs  $V$ , alors qu’il y aurait possiblement de nombreux autres hyperparamètres à fixer pour des transformations plus permissives.

Précisons aussi qu’avec l’afflux grandissant de données d’imagerie issues du secteur médical ou biologique, les méthodes de déformation peuvent connaître de plus en plus d’applications. On se permet de citer comme application éventuelle l’étude de la variabilité conformationnelle dans les données de cryomicroscopie électronique : on observe de manière très bruitée un complexe macromoléculaire avec une rotation et une conformation aléatoire, parmi celles qui sont possibles. On répète l’expérience de multiples fois, et le but est d’obtenir la “variété” des conformations possibles. La métrique riemannienne développée dans ce papier permet de se déplacer d’une conformation à l’autre.

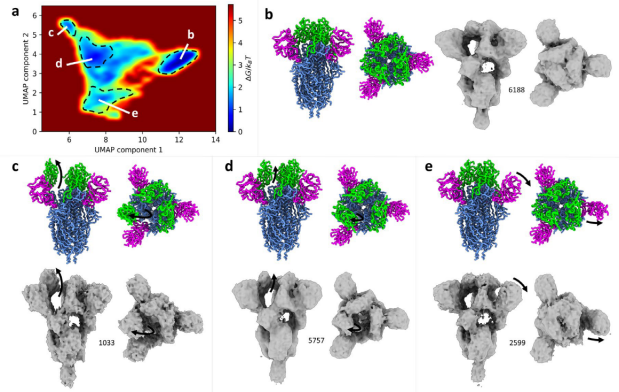


FIGURE 8 – Volumes issus de cryomicroscopie-électronique sur la protéine spike de SARS-CoV-2 [12]. En haut à gauche, un résumé du paysage conformationnel est représenté, et dans les autres cases on a reconstruit les conformations de certaines zones.

## Références

- [1] Vladimir ARNOLD. « Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits ». fr. In : *Annales de l'Institut Fourier* 16.1 (1966), p. 319-361. DOI : 10.5802/aif.233. URL : <http://www.numdam.org/articles/10.5802/aif.233/>.
- [2] Mirza Faisal BEG et al. « Computing Large Deformation Metric Mappings via Geodesic Flows of Diffeomorphisms ». In : *International Journal of Computer Vision* 61 (fév. 2005), p. 139-157. DOI : 10.1023/B:VISI.0000043755.93987.aa.
- [3] Benjamin CHARLIER et al. « Kernel Operations on the GPU, with Autodiff, without Memory Overflows ». In : *Journal of Machine Learning Research* 22.74 (2021), p. 1-6. URL : <http://jmlr.org/papers/v22/20-275.html>.
- [4] Jean FEYDY. « Analyse de données géométriques, au delà des convolutions ». Theses. Université Paris-Saclay, juill. 2020. URL : <https://theses.hal.science/tel-02945979>.
- [5] Jean FEYDY et al. « Fast geometric learning with symbolic matrices ». In : *Advances in Neural Information Processing Systems* 33 (2020).
- [6] Jean FEYDY et al. *Optimal Transport for Diffeomorphic Registration*. 2017. arXiv : 1706.05218 [math.NA].
- [7] Tom FLETCHER. « 2 - Statistics on manifolds ». In : *Riemannian Geometric Statistics in Medical Image Analysis*. Sous la dir. de Xavier PENNEC, Stefan SOMMER et Tom FLETCHER. Academic Press, 2020, p. 39-74. ISBN :

978-0-12-814725-2. DOI : <https://doi.org/10.1016/B978-0-12-814725-2.00009-1>. URL : <https://www.sciencedirect.com/science/article/pii/B9780128147252000091>.

- [8] Joan Alexis GLAUNÈS. « Transport par difféomorphismes de points, de mesures et de courants pour la comparaison de formes et l'anatomie numérique ». 2005PA132033. Thèse de doct. 2005, xiv-156 p. URL : <http://www.theses.fr/2005PA132033>.
- [9] Ulrich MENNE. « The Concept of Varifold ». In : *Notices of the American Mathematical Society* 64.10 (oct. 2017), p. 1148-1152. DOI : 10.1090/noti1589. URL : <https://doi.org/10.1090%2Fnoti1589>.
- [10] Peter W. MICHOR et David MUMFORD. *Riemannian geometries on spaces of plane curves*. 2006. arXiv : math/0312384 [math.DG].
- [11] Adam PASZKE et al. *PyTorch: An Imperative Style, High-Performance Deep Learning Library*. 2019. arXiv : 1912.01703 [cs.LG].
- [12] Rémi VUILLEMOT, Isabelle ROULLER et Slavica JONIC. « MDTOMO: Continuous conformational variability analysis in cryo electron subtomogram data using flexible fitting based on Molecular Dynamics simulations ». In : (fév. 2023). DOI : 10.1101/2023.02.25.529934.