

# Introduction à la Géométrie Riemannienne par l'étude des Espaces de Formes

Jean Feydy

Polycopié co-écrit avec les élèves du groupe de lecture éponyme,  
en première année au département de mathématiques de l'ENS Ulm :

Basile Coron	Arnaud Eteve
Paul Fraux	Vadim Lebovici
Robin Lemaire	Yohan Mandin-Huble
Nicolas Masson	Pascal Millet
Manolis Perrot	Eva Philippe
Arthur Touati	

2016–2017

Lino Benedetto	Raphaël Barboni
Valentin Carlier	Louis-Pierre Chaintron
Léo Hahn Lecler	Massil Hihat
Emmanuel Kammerer	Antoine Le Calvez
Grégoire Macqueron	Luc Pointurier
Lucas Rey	Maxime Sylvestre
Ivan Yakovlev	

2018–2019





# Résumé

**Qu'est-ce qu'une forme ?** Comment écrire mathématiquement qu'une baleine ressemble plus à un merlu qu'à un labrador (quoique...)? Pour répondre à ces questions, qui sont aujourd'hui d'un intérêt crucial en vision par ordinateur et en imagerie médicale, une idée féconde consiste à munir des espaces quotients de courbes ou de surfaces d'une structure métrique convenable, où deux objets "semblables" seraient proches. Comment y parvenir? Et quelles seront alors les propriétés des espaces construits?

Dans un premier temps, on présentera les concepts élémentaires de géométrie riemannienne dans un cadre simple : variétés, métriques, connections, géodésiques, distances induites et courbures. On pourra alors appliquer ces notions à des exemples concrets : les possibilités sont nombreuses! Formules explicites et étude poussée dans les espaces de courbes du plan, problématiques de contrôlabilité dans les espaces de formes "abstraites", calculs effectifs sur les nuages de points pour des applications en biologie et en anatomie humaine... Il s'agira de découvrir un véritable écosystème de travaux, allant des considérations sur les actions par transport de groupes de difféomorphismes aux outils diagnostiques de demain.

**What is a shape ?** How can we mathematically write that a dolphin looks much more similar to a fish than to a dog (although...)? To answer those questions, now of crucial interest in the fields of computer vision and medical imaging, a seminal idea is to equip curves or surfaces quotient spaces of a suitable metric structure : "similar" objects shall be close to each other. How can we do so? What would then be the properties of the constructed spaces?

First, we will introduce the elementary tools of Riemannian geometry in a simple way : manifolds, metrics, connections, geodesics, induced distances and curvatures. Then, we shall apply these notions to concrete situations : possibilities are countless! Explicit formulas and extensive studies in the spaces of plane curves, controlability problems in "abstract" shapes spaces, effective computations on points clouds with applications to biology and human anatomy... Walking through a colourful collection of papers, we shall discover an active field of research, which links diffeomorphisms groups actions to the future of diagnosis.

# Informations pratiques

## Contact :

- Adresses mail : [jean.feydy@ens.fr](mailto:jean.feydy@ens.fr), [jean.feydy@gmail.com](mailto:jean.feydy@gmail.com) (utilisez les deux).
- Bureau : sous les toits du DMA, les lundis et mercredis à coup sûr.
- Page web du cours : [www.math.ens.fr/~feidy/Teaching/index.html](http://www.math.ens.fr/~feidy/Teaching/index.html).  
Vous y trouverez la version la plus à jour du poly, mais aussi de petits carnets interactifs en Matlab/Python qui vous permettront d'expérimenter et de voir nos algorithmes en action.

## Planning hebdomadaire :

- Samedi-Dimanche : comptez un bon week-end de travail...
- Lundi ou Mardi 16h30–17h30 : réunion pour faire le point dans mon bureau, avec un plan détaillé, une ébauche de chapitre et des questions.
- Mardi soir : envoi d'une version finale du chapitre du polycopié à l'adresse [jean.feydy@gmail.com](mailto:jean.feydy@gmail.com). Je me charge de l'impression.
- Mercredi après-midi, de 15h à 16h30 en salle R : exposé.

## Échéances :

- 26 septembre : Séance d'introduction, Jean Feydy.  
. Répartition des rôles pour la première session.
- 3 octobre : Séance 2, Emmanuel et Léo.
- 10 octobre : Séance 3, Massil et Ivan.
- 17 octobre : Séance 4, Luc et Lucas.
- 24 octobre : Séance 5, Raphaël et Antoine.
- 7 novembre : Séance 6, Lino et Valentin.
- 14 novembre : Séance 7, Maxime et Grégoire.  
. Répartition des rôles pour la deuxième session.
- 21 novembre : Séance banalisée pour les partiels ?
- 28 novembre : Séance 8, . et .
- 5 décembre : Séance 9, . et .
- 12 décembre : Séance 10, . et .
- 19 décembre : Séance 11, Louis-Pierre et .
- 9 janvier : Séance 12, . et .
- 16 janvier : Séance 13, . et .
- 17 janvier : Séance 14, . et .

# Table des matières

<b>Bibliographie</b>	<b>9</b>	
<b>1 Introduction</b>	<b>11</b>	<i>Séance 1</i>
1.1 Géométries non-euclidiennes . . . . .	11	<i>Jean Feydy</i>
1.2 Métriques locales sur un ouvert de $\mathbb{R}^n$ . . . . .	13	
1.3 La géométrie Riemannienne : l'exemple du tore . . . . .	17	
1.4 Applications en vue : l'anatomie computationnelle . . . . .	18	
1.5 Variétés : un petit guide de survie . . . . .	22	
<b>I Éléments de Géométrie Riemannienne</b>	<b>25</b>	
<b>2 Intuitions sur la courbure, Vocabulaire</b>	<b>27</b>	<i>Séance 2</i>
2.1 Introduction . . . . .	27	<i>[Lee97], Chap. 1-2</i>
2.2 Intuitions sur la courbure . . . . .	27	
2.3 Variétés . . . . .	28	<i>Arnaud Eteve,</i>
2.4 Fibrés Vectoriels . . . . .	29	<i>Léo Hahn-Lecler,</i>
2.5 Tenseurs . . . . .	33	<i>Emmanuel</i>
2.6 Fibré Tensoriels, Champs de Tenseur . . . . .	34	<i>Kammerer</i>
2.7 Conclusion . . . . .	34	<i>Un peu énumératif</i>
<b>3 Métriques Riemanniennes</b>	<b>35</b>	<i>Séance 3</i>
3.1 Métriques Riemanniennes . . . . .	35	<i>[Lee97], Chap. 3</i>
3.2 Constructions élémentaires associées à une métrique Riemannienne . . . . .	40	
3.3 Généralisation des métriques Riemanniennes . . . . .	41	<i>Paul Fraux,</i>
3.4 Les espaces canoniques de la géométrie Riemannienne . . . . .	42	<i>Manolis Perrot,</i>
		<i>Massil Hihat,</i>
		<i>Ivan Yakovlev</i>
 		<i>Séance 4</i>
<b>4 Connexions et courbes géodésiques</b>	<b>47</b>	<i>Sans souci</i>
4.1 Dériver des champs de vecteurs? . . . . .	47	<i>[Lee97], Chap. 4</i>
4.2 Connexions affines . . . . .	47	
4.3 Dérivation d'un champ de vecteur sur une courbe paramétrée . . . . .	50	<i>Basile Coron,</i>
4.4 Géodésiques . . . . .	51	<i>Yohan Mandin</i>
4.5 Transport parallèle . . . . .	53	<i>-Huble, Lucas</i>
		<i>Rey, Louis-</i>
		<i>Pierre Chaintron</i>
 		<i>Séance 5</i>
<b>5 Géodésiques Riemanniennes</b>	<b>55</b>	<i>Technique</i>
5.1 Connexion associée à une métrique Riemannienne . . . . .	55	<i>[Lee97], Chap. 5</i>
5.2 Géodésiques et isométries . . . . .	59	
5.3 Carte exponentielle . . . . .	60	<i>Nicolas Masson,</i>
		<i>Pascal Millet,</i>
		<i>Raphaël Barboni,</i>
		<i>Antoine Le Calvez</i>
		<i>Équilibré</i>

	5.4	Voisinages normaux et coordonnées normales . . . . .	62
	5.5	Géodésiques des espaces classiques . . . . .	64
<i>Séance 6</i>	<b>6</b>	<b>Le plan hyperbolique</b>	<b>67</b>
<i>[CFK<sup>+</sup> 97]</i>	6.1	Cas unidimensionnel . . . . .	67
	6.2	Cinq modèles en dimension 2 . . . . .	69
<i>Vadim Lebovici,</i>	6.3	Géodésiques . . . . .	72
<i>Robin Lemaire,</i>	6.4	Propriétés diverses . . . . .	76
<i>Valentin Carlier,</i>	6.5	Le sixième modèle . . . . .	80
<i>Lino Benedetto</i>	6.6	Conclusion . . . . .	81
<i>Très joli</i>			
<i>Séance 7</i>	<b>7</b>	<b>Distances et géodésiques</b>	<b>83</b>
<i>[Lee97], Chap. 6</i>	7.1	Longueur d'une courbe, distance et topologie induite . . . . .	83
	7.2	Courbes minimisantes et géodésiques . . . . .	85
<i>Eva Philippe,</i>	7.3	Complétude : Le théorème de Hopf-Rinow . . . . .	92
<i>Arthur Touati,</i>			
<i>Grégoire Macqueron,</i>			
<i>Maxime Sylvestre</i>			
<i>Équilibré</i>			
	<b>II</b>	<b>Application à l'étude des Espaces de Formes</b>	<b>93</b>
<i>Séance 8</i>	<b>8</b>	<b>Un article fondateur : les triangles de Kendall</b>	<b>95</b>
<i>[Ken84]</i>	8.1	Introduction et motivation à l'étude des triangles . . . . .	95
<i>Basile Coron,</i>	8.2	Un espace de formes : la sphère des triangles . . . . .	96
<i>Paul Fraux,</i>	8.3	Généralisation . . . . .	107
<i>Massil Hihat,</i>	8.4	Etude des formes du plan, $\Sigma_2^k$ . . . . .	108
<i>Lino Benedetto</i>	8.5	Statistiques sur la sphère . . . . .	113
<i>Élémentaire</i>			
<i>et élégant</i>			
<i>Séance 9</i>	<b>9</b>	<b>Étude statistique dans un espace de formes simple</b>	<b>117</b>
<i>[Ken84]</i>	9.1	Généralisation de l'analyse de Kendall à des polygones quelconques . . . . .	117
<i>Grégoire Macqueron,</i>	9.2	Statistiques sur la sphère algorithmique d'ACP . . . . .	120
<i>Maxime Sylvestre</i>	9.3	Application sur les petits vers . . . . .	123
<i>Très visuel</i>			
<i>Séance 10</i>	<b>10</b>	<b>Travailler avec des données non-étiquetées</b>	<b>127</b>
<i>[CT13]</i>	10.1	Introduction . . . . .	127
<i>Valentin Carlier,</i>	10.2	Distances sur un espace de formes . . . . .	127
<i>Louis-Pierre</i>	10.3	Normes à noyau . . . . .	131
<i>Chaintron</i>	10.4	Application : recalcul des formes . . . . .	134
<i>Retour aux</i>	10.5	Conclusion . . . . .	136
<i>fondamentaux</i>			
<i>Séance 11</i>	<b>11</b>	<b>Géodésiques Riemanniennes, équations Hamiltoniennes</b>	<b>137</b>
<i>Antoine Le Calvez</i>	11.1	Métrique riemannienne, distance de chemins . . . . .	137
<i>Raphaël Barboni</i>	11.2	Application numérique : un code de tir géodésique . . . . .	140
<i>Deux résultats</i>			
<i>remarquables</i>			
<i>Séance 12</i>	<b>12</b>	<b>Cométries à noyaux, méthode LDDMM</b>	<b>141</b>
<i>Emmanuel Kammerer</i>	12.1	Introduction . . . . .	141
<i>Léo Hahn-Lecler</i>	12.2	Cométries à noyaux : l'algorithme . . . . .	141
	12.3	Comprendre les cométries à noyaux . . . . .	143
	12.4	conclusion . . . . .	147

<b>13 Méthodes à noyaux pour l'appariement de nuages de points</b>	<b>149</b>	<i>Bonus</i>
13.1 Objectifs . . . . .	149	<i>[Gla05]</i>
13.2 Difféomorphismes . . . . .	150	
13.3 Retour sur le problème d'appariement . . . . .	153	<i>Vadim Lebovici</i>
13.4 Élargissement du modèle aux déformations rigides . . . . .	154	<i>et Arthur Touati</i>
13.5 Méthodes à noyaux . . . . .	154	<i>Fondamental</i>
13.6 Équations Hamiltoniennes . . . . .	158	
<b>14 Appariement de courbes et surfaces polygonales</b>	<b>165</b>	<i>Bonus</i>
14.1 Introduction . . . . .	165	<i>[CT13]</i>
14.2 Transport et appariement de mesures . . . . .	165	
14.3 Courants . . . . .	175	<i>Arnaud Eteve et</i>
14.4 Varifolds . . . . .	176	<i>Yohan Mandin</i>
14.5 Retour au problème d'appariement . . . . .	181	<i>-Huble</i>
14.6 Conclusion . . . . .	182	<i>Aspects</i>
		<i>pratiques</i>
<b>15 Construction d'Atlas, étude statistique et aide au diagnostic</b>	<b>185</b>	<i>Bonus</i>
15.1 Introduction . . . . .	185	<i>[CCT14]</i>
15.2 Ajout d'un terme fonctionnel . . . . .	185	
15.3 Etude de population avec moyenne à estimer . . . . .	190	<i>Eva Philippe</i>
15.4 Limitations et contraintes d'implémentation . . . . .	194	<i>et Robin Lemaire</i>
15.5 Conclusion . . . . .	196	<i>Applications</i>
		<i>réelles</i>
<b>III Critiques et perspectives</b>	<b>199</b>	
<b>16 Conséquences pratiques du choix des métriques à noyaux</b>	<b>201</b>	<i>Séance 12</i>
16.1 Introduction. . . . .	201	<i>[MMM12]</i>
16.2 Le modèle de déformation et son exploitation. . . . .	201	
16.3 Un cas typique pour comprendre la dynamique des landmarks. . . . .	204	<i>Manolis Perrot</i>
16.4 Intuitions sur les géodésiques et comportements problématiques. . . . .	206	
16.5 Courbure positive et points conjugués. . . . .	207	
16.6 Courbure négative et comportements chaotiques. . . . .	208	
<b>17 Remise en question du cadre d'étude riemannien</b>	<b>211</b>	<i>Séance 13</i>
17.1 Introduction . . . . .	211	<i>[LP13], [Ash07]</i>
17.2 Éléments de théorie des groupes de Lie . . . . .	211	
17.3 La méthode SVF . . . . .	212	<i>Nicolas Masson</i>
17.4 Récapitulatif des notions vues dans les théories Riemanniennes et Lie . . . . .	218	
17.5 Comparaison expérimentale . . . . .	219	
17.6 Conclusion . . . . .	219	
<b>18 Autres utilisations de la géométrie Riemannienne en imagerie médicale</b>	<b>221</b>	<i>Séance 14</i>
18.1 Synthèse sur la géométrie Riemannienne . . . . .	221	<i>[Pen09]</i>
18.2 Quelques méthodes vectorielles utilisées en analyse d'image et leurs analogues Riemanniennes . . . . .	225	<i>Pascal Millet</i>
18.3 Application de la géométrie Riemannienne à l'analyse d'images de tenseurs . . . . .	228	
18.4 Pour aller plus loin : conclusion et perspectives . . . . .	234	



# Bibliographie

- [Arg14] Sylvain Arguillere. *Géométrie sous-riemannienne en dimension infinie et applications à l'analyse mathématique des formes*. PhD thesis, Université Pierre et Marie Curie-Paris VI, 2014.
- [Arn66] Vladimir Arnold. Sur la géométrie différentielle des groupes de lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. In *Annales de l'institut Fourier*, volume 16, pages 319–361, 1966.
- [Ash07] John Ashburner. A fast diffeomorphic image registration algorithm. *Neuroimage*, 38(1) :95–113, 2007.
- [BBM14] Martin Bauer, Martins Bruveris, and Peter W Michor. Overview of the geometries of shape spaces and diffeomorphism groups. *Journal of Mathematical Imaging and Vision*, 50(1-2) :60–97, 2014.
- [CCT14] Benjamin Charlier, Nicolas Charon, and Alain Trouvé. The fshape framework for the variability analysis of functional shapes. *Foundations of Computational Mathematics*, pages 1–71, 2014.
- [CFK<sup>+</sup>97] James W Cannon, William J Floyd, Richard Kenyon, Walter R Parry, et al. Hyperbolic geometry. *Flavors of geometry*, 31 :59–115, 1997.
- [CT13] Nicolas Charon and Alain Trouvé. The varifold representation of nonoriented shapes for diffeomorphic registration. *SIAM Journal on Imaging Sciences*, 6(4) :2547–2580, 2013.
- [Gla05] Joan Glaunes. Transport par difféomorphismes de points, de mesures et de courants pour la comparaison de formes et l'anatomie numérique. *These de sciences, Université Paris*, 13, 2005.
- [Jan12] Robert T Jantzen. Geodesics on the torus and other surfaces of revolution clarified using undergraduate physics tricks with bonus : Nonrelativistic and relativistic kepler problems. *arXiv preprint arXiv :1212.6206*, 2012.
- [Ken84] David G Kendall. Shape manifolds, procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society*, 16(2) :81–121, 1984.
- [Lee97] John M Lee. *Riemannian Manifolds : An Introduction to Curvature*, volume 176. Springer Science & Business Media, 1997.
- [LGP<sup>+</sup>14] Herve Lombaert, Leo Grady, Xavier Pennec, Nicholas Ayache, and Farida Cheriet. Spectral log-demons : diffeomorphic image registration with very large deformations. *International journal of computer vision*, 107(3) :254–271, 2014.
- [LP13] Marco Lorenzi and Xavier Pennec. Geodesics, parallel transport & one-parameter subgroups for diffeomorphic image registration. *International journal of computer vision*, 105(2) :111–127, 2013.

- [MM07] Peter W Michor and David Mumford. An overview of the riemannian metrics on spaces of curves using the hamiltonian approach. *Applied and Computational Harmonic Analysis*, 23(1) :74–113, 2007.
- [MMM12] Mario Micheli, Peter W Michor, and David Mumford. Sectional curvature in terms of the cometric, with applications to the riemannian manifolds of landmarks. *SIAM Journal on Imaging Sciences*, 5(1) :394–433, 2012.
- [MPS<sup>+</sup>11] Tommaso Mansi, Xavier Pennec, Maxime Sermesant, Hervé Delingette, and Nicholas Ayache. ilogdemons : A demons-based registration algorithm for tracking incompressible elastic biological tissues. *International journal of computer vision*, 92(1) :92–111, 2011.
- [Pen09] Xavier Pennec. Statistical computing on manifolds : from riemannian geometry to computational anatomy. In *Emerging Trends in Visual Computing*, pages 347–386. Springer, 2009.
- [T<sup>+</sup>42] Darcy Wentworth Thompson et al. On growth and form. *On growth and form.*, 1942.



# Chapitre 1

## Introduction

*Séance 1*

*Jean Feydy*

### 1.1 Géométries non-euclidiennes

*Géométrie riemannienne...* En sortant de prépa, on ne connaît bien souvent que “La” géométrie, *euclidienne*, classique. Du cadre restreint des anciens, on s’est tout de même un petit peu émancipé : étude de  $\mathbb{R}^n$  muni du produit scalaire canonique – ou pas –, espaces  $L^2$ , espaces de Hilbert. Sans, jamais, remettre en question les concepts géométriques sous-jacents : droite engendrée copie de la droite réelle et structure additive globale. Eh bien, plutôt qu’un long discours, attaquons ensemble l’entame du chapitre III de *La Science et l’Hypothèse*, d’Henri Poincaré.

Les géométries non euclidiennes — Toute conclusion suppose des prémisses ; ces prémisses elles-mêmes ou bien sont évidentes par elles-mêmes et n’ont pas besoin de démonstration, ou bien ne peuvent être établies qu’en s’appuyant sur d’autres propositions, et comme on ne saurait remonter ainsi à l’infini, toute science déductive, et en particulier la géométrie, doit reposer sur un certain nombre d’axiomes indémontrables. Tous les traités de géométrie débutent donc par l’énoncé de ces axiomes. Mais il y a entre eux une distinction à faire : quelques-uns, comme celui-ci par exemple : « deux quantités égales à une même troisième sont égales entre elles », ne sont pas des propositions de géométrie, mais des propositions d’analyse. Je les regarde comme des jugements analytiques a priori, je ne m’en occuperai pas.

Mais je dois insister sur d’autres axiomes qui sont spéciaux à la géométrie. La plupart des traités en énoncent trois explicitement :

- 1° Par deux points ne peut passer qu’une droite ;
- 2° La ligne droite est le plus court chemin d’un point à un autre.
- 3° Par un point on ne peut faire passer qu’une parallèle à une droite donnée.

Bien que l’on se dispense généralement de démontrer le second de ces axiomes, il serait possible de le déduire des deux autres et de ceux, beaucoup plus nombreux, que l’on admet implicitement sans les énoncer, ainsi que je l’expliquerai plus loin.

On a longtemps cherché en vain à démontrer également le troisième axiome, connu sous le nom de postulatum d'Euclide. Ce qu'on a dépensé d'efforts dans cet espoir chimérique est vraiment inimaginable. Enfin au commencement du siècle et à peu près en même temps, deux savants, un Russe et un Hongrois, Lobatchevsky et Bolyai établirent d'une façon irréfutable que cette démonstration est impossible ; ils nous ont à peu près débarrassés des inventeurs de géométries sans postulatum ; depuis lors l'Académie des Sciences ne reçoit plus guère qu'une ou deux démonstrations nouvelles par an.

La question n'était pas épuisée ; elle ne tarda pas à faire un grand pas par la publication du célèbre mémoire de Riemann intitulé : *Ueber die Hypothesen welche der Geometrie zum Grunde liegen*. Cet opuscule a inspiré la plupart des travaux récents dont je parlerai plus loin et parmi lesquels il convient de citer ceux de Beltrami et de Helmholtz.

La Géométrie de Lobatchevsky. — S'il était possible de déduire le postulatum d'Euclide des autres axiomes, il arriverait évidemment qu'en niant le postulatum, et en admettant les autres axiomes, on serait conduit à des conséquences contradictoires ; il serait donc impossible d'appuyer sur de telles prémisses une géométrie cohérente.

Or c'est précisément ce qu'a fait Lobatchevsky. Il suppose au début que :

L'on peut par un point mener plusieurs parallèles à une droite donnée ;

Et il conserve d'ailleurs tous les autres axiomes d'Euclide. De ces hypothèses, il déduit une suite de théorèmes entre lesquels il est impossible de relever aucune contradiction et il construit une géométrie dont l'impeccable logique ne le cède en rien à celle de la géométrie euclidienne.

Les théorèmes sont, bien entendu, très différents de ceux auxquels nous sommes accoutumés et ils ne laissent pas de déconcerter un peu d'abord.

Ainsi la somme des angles d'un triangle est toujours plus petite que deux droits et la différence entre cette somme et deux droits est proportionnelle à la surface du triangle.

Il est impossible de construire une figure semblable à une figure donnée mais de dimensions différentes.

Si l'on divise une circonférence en  $n$  parties égales, et qu'on mène des tangentes aux points de division, ces  $n$  tangentes formeront un polygone si le rayon de la circonférence est assez petit ; mais si ce rayon est assez grand, elles ne se rencontreront pas.

Il est inutile de multiplier ces exemples ; les propositions de Lobatchevsky n'ont plus aucun rapport avec celles d'Euclide, mais elles ne sont pas moins logiquement reliées les unes aux autres.

## 1.2 Métriques locales sur un ouvert de $\mathbb{R}^n$

On l'a vu : d'autres *géométries*, d'autres jeux d'axiomes que celui retenu par Euclide sont donc cohérents. Mais comment nous représenter ces mondes – on préférera dire : ces espaces – aux règles inattendues ? Il suffit de pousser la lecture de *La Science et l'Hypothèse* jusqu'au chapitre suivant :

Le monde non euclidien. — Si l'espace géométrique était un cadre imposé à chacune de nos représentations, considérée individuellement, il serait impossible de se représenter une image dépouillée de ce cadre, et nous ne pourrions rien changer à notre géométrie.

Mais il n'en est pas ainsi, la géométrie n'est que le résumé des lois suivant lesquelles se succèdent ces images. Rien n'empêche alors d'imaginer une série de représentations, de tout point semblables à nos représentations ordinaires, mais se succédant d'après des lois différentes de celles auxquelles nous sommes accoutumés.

On conçoit alors que des êtres dont l'éducation se ferait dans un milieu où ces lois seraient ainsi bouleversées pourraient avoir une géométrie très différente de la nôtre.

Supposons, par exemple, un monde renfermé dans une grande sphère et soumis aux lois suivantes :

La température n'y est pas uniforme ; elle est maxima au centre, et elle diminue à mesure qu'on s'en éloigne, pour se réduire au zéro absolu quand on atteint la sphère où ce monde est renfermé.

Je précise davantage la loi suivant laquelle varie cette température. Soit  $R$  le rayon de la sphère limite ; soit  $r$  la distance du point considéré au centre de cette sphère. La température absolue sera proportionnelle à  $R^2 - r^2$ .

Je supposerai de plus que, dans ce monde, tous les corps aient même coefficient de dilatation, de telle façon que la longueur d'une règle quelconque soit proportionnelle à sa température absolue.

Je supposerai enfin qu'un objet transporté d'un point à un autre, dont la température est différente, se met immédiatement en équilibre calorifique avec son nouveau milieu.

Rien dans ces hypothèses n'est contradictoire ou inimaginable.

Un objet mobile deviendra alors de plus en plus petit à mesure qu'on se rapprochera de la sphère limite.

Observons d'abord que, si ce monde est limité au point de vue de notre géométrie habituelle, il paraîtra infini à ses habitants.

Quand ceux-ci, en effet, veulent se rapprocher de la sphère limite, ils se refroidissent et deviennent de plus en plus petits. Les pas qu'ils font sont donc aussi de plus en plus petits, de sorte qu'ils ne peuvent jamais atteindre la sphère limite.

Si, pour nous, la géométrie n'est que l'étude des lois suivant lesquelles se meuvent les solides invariables ; pour ces êtres ima-

ginaires, ce sera l'étude des lois suivant lesquelles se meuvent les solides déformés par ces différences de température dont je viens de parler.

Sans doute, dans notre monde, les solides naturels éprouvent également des variations de forme et de volume dues à l'échauffement ou au refroidissement. Mais nous négligeons ces variations en jetant les fondements de la géométrie ; car, outre qu'elles sont très faibles, elles sont irrégulières et nous paraissent par conséquent accidentelles.

Dans ce monde hypothétique, il n'en serait plus de même, et ces variations suivraient des lois régulières et très simples.

D'autre part, les diverses pièces solides dont se composerait le corps de ses habitants, subirait les mêmes variations de forme et volume.

Je ferai encore une autre hypothèse ; je supposerai que la lumière traverse des milieux diversement réfringents et de telle sorte que l'indice de réfraction soit inversement proportionnel à  $R^2 - r^2$ . Il est aisé de voir que, dans ces conditions, les rayons lumineux ne seraient pas rectilignes, mais circulaires.

Pour justifier ce qui précède, il me reste à montrer que certains changements survenus dans la position des objets extérieurs peuvent être corrigés par des mouvements corrélatifs des êtres sentants qui habitent ce monde imaginaire ; et cela de façon à restaurer l'ensemble primitif des impressions subies par ces êtres sentants.

Supposons en effet qu'un objet se déplace, en se déformant, non comme un solide invariable, mais comme un solide éprouvant des dilatations inégales exactement conformes à la loi de température que j'ai supposée plus haut. Qu'on me permette pour abrégé le langage, d'appeler un pareil mouvement déplacement non euclidien.

Si un être sentant se trouve dans le voisinage, ses impressions seront modifiées par le déplacement de l'objet, mais il pourra les rétablir en se mouvant lui-même d'une manière convenable. Il suffit que finalement l'ensemble de l'objet et de l'être sentant, considéré comme formant un seul corps, ait éprouvé un de ces déplacements particuliers que je viens d'appeler non euclidiens. Cela est possible si l'on suppose que les membres de ces êtres se dilatent d'après la même loi que les autres corps du monde qu'ils habitent.

Bien qu'au point de vue de notre géométrie habituelle les corps se soient déformés dans ce déplacement et que leurs diverses parties ne se retrouvent plus dans la même situation relative, cependant nous allons voir que les impressions de l'être sentant sont redevenues les mêmes.

En effet, si les distances mutuelles des diverses parties ont pu varier, néanmoins les parties primitivement en contact sont revenues en contact. Les impressions tactiles n'ont donc pas changé.

D'autre part, en tenant compte de l'hypothèse faite plus haut

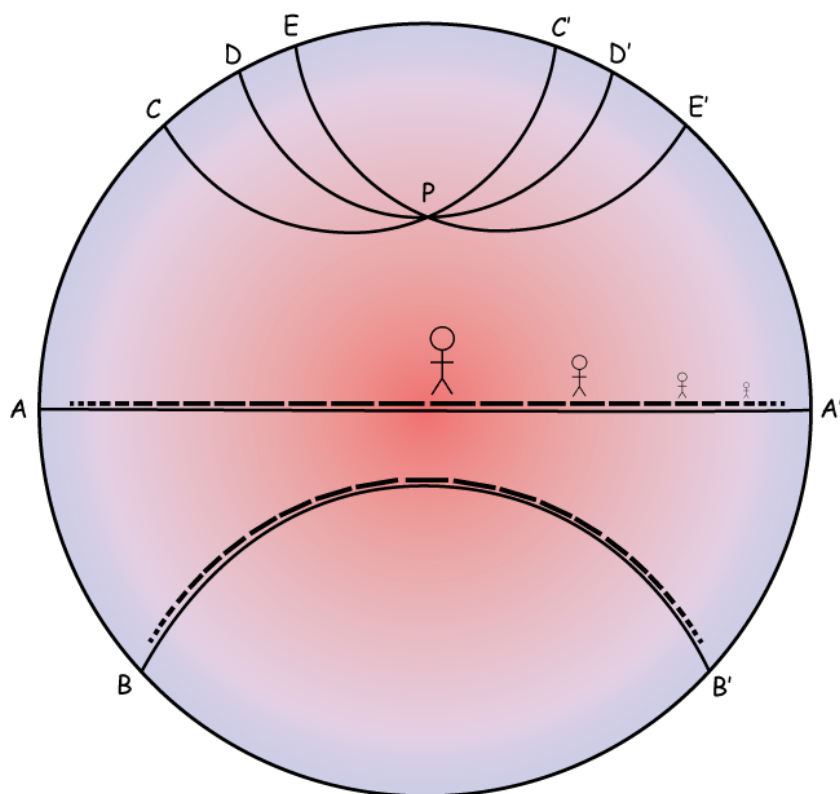


FIGURE 1.1 – Le disque de Poincaré. Pour emprunter le plus court chemin, un habitant de ce monde réduit aura toujours tendance à passer par le centre du disque, où la température élevée lui permet de faire des pas de géant. La simplicité de la métrique permet d’expliciter les “droites” géodésiques : ce sont les arcs de cercle orthogonaux au bord du disque. Par le point  $P$  passent donc une infinité de “droites” n’intersectant pas la droite  $(AA')$ .  
Image modifiée du site de John D. Norton, University of Pittsburgh.

au sujet de la réfraction et de la courbure des rayons lumineux, les impressions visuelles seront aussi restées les mêmes.

Ces êtres imaginaires seront donc comme nous conduits à classer les phénomènes dont ils seront témoins et à distinguer parmi eux, les « changements de position » susceptibles d’être corrigés par un mouvement volontaire corrélatif.

S’ils fondent une géométrie, ce ne sera pas comme la nôtre, l’étude des mouvements de nos solides invariables ; ce sera celle des changements de position qu’ils auront ainsi distingués, et qui ne sont autres que les « déplacements non euclidiens », ce sera la géométrie non euclidienne.

Ainsi des êtres comme nous, dont l’éducation se ferait dans un pareil monde, n’auraient pas la même géométrie que nous.

Pour *construire* des géométries nouvelles, M. Poincaré propose donc simplement de partir d’un ouvert de l’espace euclidien – ici, la boule  $\Omega = B(0, R)$  de  $\mathbb{R}^2$  –, et de le *déformer* localement :

son champ de températures,

$$K(x) = R^2 - \|x\|^2, \quad (1.1)$$

peut être compris comme un champ de dilatation/compression de l'espace.

Afin de s'adresser au sens physique de chacun, M. Poincaré s'encombre de plusieurs notions a priori indépendantes : température, coefficient de dilatation, indice de réfraction... Puisque nous avons ici à cœur de faire de la géométrie, seules les courbes, les distances, les volumes nous intéressent. On se contentera donc, pour caractériser notre espace, de fournir une formule donnant la longueur infinitésimale ressentie d'un petit segment  $[x, x + dx]$  en tout point du disque :

$$d(x \rightarrow x + dx) = \frac{1}{R^2 - \|x\|^2} \|dx\| \quad (1.2)$$

ou, ce qui est plus pratique, sous la forme quadratique

$$d^2(x \rightarrow x + dx) = \frac{1}{(R^2 - \|x\|^2)^2} dx^2. \quad (1.3)$$

La forme quadratique  $d_x^2 : dx \mapsto d^2(x \rightarrow x + dx)$ , que l'on appellera *métrique Riemannienne locale* au point  $x$ , permet alors de définir la longueur d'un chemin  $\gamma : [0, 1] \rightarrow \Omega$  :

$$\ell(\gamma) = \int_0^1 \sqrt{d_{\gamma(t)}^2(\gamma'(t))} dt. \quad (1.4)$$

On peut par exemple calculer le "rayon" du disque de Poincaré associé au chemin radial  $\gamma(t) = (t, 0)$ . Pour la métrique euclidienne, habituelle, on a tout simplement

$$\ell_{\text{eucl}}(\gamma) = \int_0^1 \|\gamma'(t)\|_{\gamma(t), \text{eucl}} dt \quad (1.5)$$

$$= \int_0^1 \|(1, 0)\|_{\gamma(t), \text{eucl}} dt \quad (1.6)$$

$$= \int_0^1 \sqrt{1^2 + 0^2} dt = 1. \quad (1.7)$$

Pour la métrique intrinsèque  $d^2$ , on a par contre

$$\ell_{\text{poinc}}(\gamma) = \int_0^1 \|\gamma'(t)\|_{\gamma(t), \text{poinc}} dt \quad (1.8)$$

$$= \int_0^1 \|(1, 0)\|_{\gamma(t), \text{poinc}} dt \quad (1.9)$$

$$= \int_0^1 \frac{\|(1, 0)\|_{\gamma(t), \text{eucl}}}{1 - \|\gamma(t)\|^2} dt \quad (1.10)$$

$$= \int_0^1 \frac{\sqrt{1^2 + 0^2}}{1 - t^2} dt \quad (1.11)$$

$$= \int_0^1 \frac{1}{1+t} \cdot \frac{1}{1-t} dt = +\infty \quad (1.12)$$

Le disque de Poincaré est donc un espace métrique *non borné* : un univers dans une coquille de noix. Son étude explicite sera l'objet de notre sixième séance, et nous verrons que le Postulum d'Euclide n'y est pas vérifié – la figure 1.1 permet de s'en faire une bonne intuition.

## 1.3 La géométrie Riemannienne : l'exemple du tore

On l'a vu, déformer judicieusement l'espace euclidien permet de construire des *modèles* aux théories axiomatiques non-euclidiennes développées par Lobatchevski et Bolyai. L'idée clé de la géométrie Riemannienne est de généraliser cette procédure à *tous* les étirements réguliers de la métrique euclidienne sur  $\mathbb{R}^n$  : on s'autorisera à considérer tout champ de matrices  $n \times n$  symétriques définies positives,  $g : x \in \Omega \mapsto g(x) \in S_n^{++}$ , avec  $g$  de classe  $C^\infty$  et  $\Omega$  un ouvert de  $\mathbb{R}^n$  – ou mieux, une variété de dimension  $n$  comme défini section 1.5.

L'idée est féconde : prenons l'exemple du tore. Il existe a priori deux manières de l'étudier : en tant que surface immergée dans l'espace ambiant  $\mathbb{R}^3$ , ou comme quotient du plan  $\mathbb{R}^2$  par l'action des translations entières – le fameux carré aux bords recollés. Cette dernière représentation est de loin la plus pratique : plutôt que de travailler avec des points  $(x, y, z)$  dont il faut toujours s'assurer qu'ils ne quittent pas la surface, on utilise des couples  $(r, \theta)$  sans se poser d'autre question que celle du “modulo  $2\pi$ ”.

Malheureusement, la structure métrique du carré recollé n'est pas représentative de l'idée sous-jacente de “bouée” que nous avons tous en tête : avec un parallélépipède, impossible de rendre le “petit équateur” plus court que le grand – voir figure 1.2. Pour réunir le meilleur des deux mondes, il faudra en fait déformer le carré par une métrique appropriée, qui accorde d'autant plus de poids à  $d\theta$  que le point de base  $(r, \theta)$  est proche de l'axe de révolution du donut. Vous pourrez trouver des précisions à ce sujet sur la page web du cours et dans les excellentes notes de Robert T. Jantzen [Jan12].

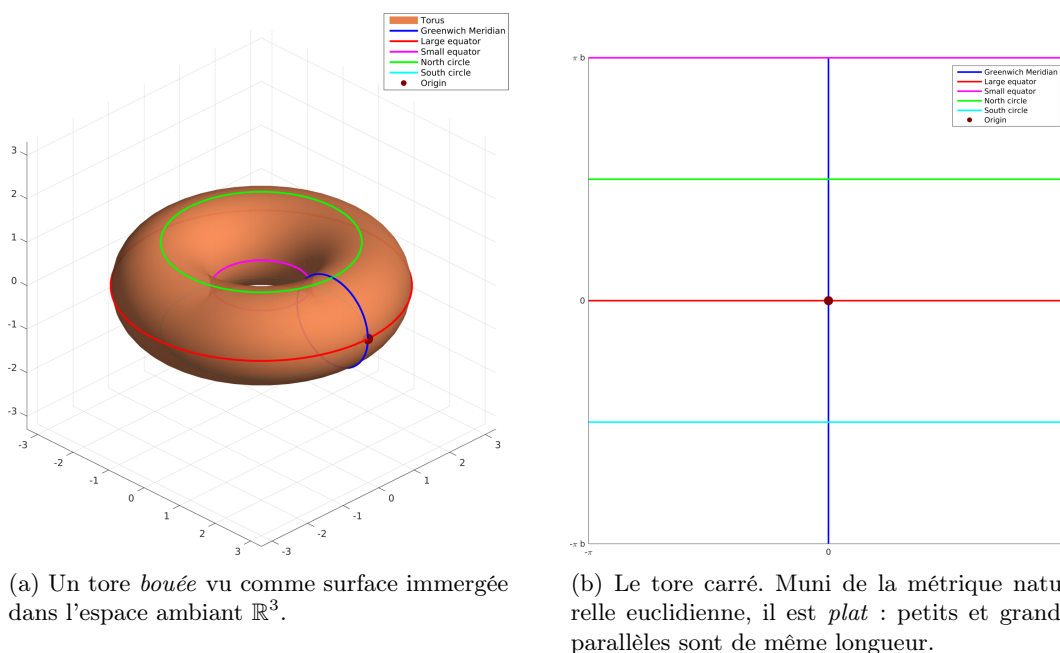


FIGURE 1.2 – Deux manières de représenter un tore. La géométrie Riemannienne permet d'étudier la structure métrique de la bouée avec le confort pratique offert par une paramétrisation bidimensionnelle.

## 1.4 Applications en vue : l'anatomie computationnelle

Plutôt que d'étudier de manière fine les polygones, cercles et autres structures algébriques sur l'espace euclidien, Riemann nous incite donc à nous concentrer sur les seules propriétés *métriques* de nos espaces, qui peuvent maintenant être courbes et définis localement. Ce point de vue a de nombreuses applications, dont la plus célèbre est sans aucun doute la théorie de la relativité d'Einstein. Au fil de ce cours, nous nous concentrerons plutôt sur un débouché tout à fait inattendu : l'étude des espaces de formes.

**Régression et suivi longitudinal** On connaît tous les courbes de croissances présentes sur nos carnets de santé : dûment remplies, elles permettent de suivre notre évolution, de la comparer à une population "normale" et de détecter un éventuel retard de croissance – voir Figure 1.3. Maintenant, serait-il possible de faire la même chose avec, disons, la forme du cœur? Attention : il ne s'agit pas ici de remplacer une grandeur scalaire – la taille – par un simple vecteur "largeur/longueur/volume/que-sais-je". Ce traitement simpliste éclaterait les cœurs similaires – mêmes rapports de taille – en une ribambelle de points... Si l'on s'intéresse véritablement à la *forme* du cœur, aux atrophies/hypertrophies éventuelles de ses ventricules, il faudra nécessairement travailler dans un espace quotient – sur lequel la notion de régression n'est a priori pas définie.

**Création d'atlas et analyse statistique** Autre axe de recherche, dans la même veine : l'étude statistique d'une population de formes. On le sait, l'étude approfondie d'un jeu d'indicateurs scalaires – poids, âge, tension artérielle... – peut nous apprendre beaucoup sur l'état de santé d'un patient. En comparant celui-ci à un humain moyen, en le replaçant au sein d'une population connue, on peut le classer dans tel ou tel groupe, détecter l'apparition de certaines pathologies. Alors, au XXI<sup>e</sup> siècle, sera-t-il possible de faire la même chose avec des données plus complexes? On pense par exemple à des images de colonne vertébrale (scoliose...), de fonds d'œils rétiniens (glaucome...) ou à la forme de l'hippocampe, corrélée à la présence de démences dégénératives comme la maladie d'Alzheimer – voir Figures 1.4-1.5.

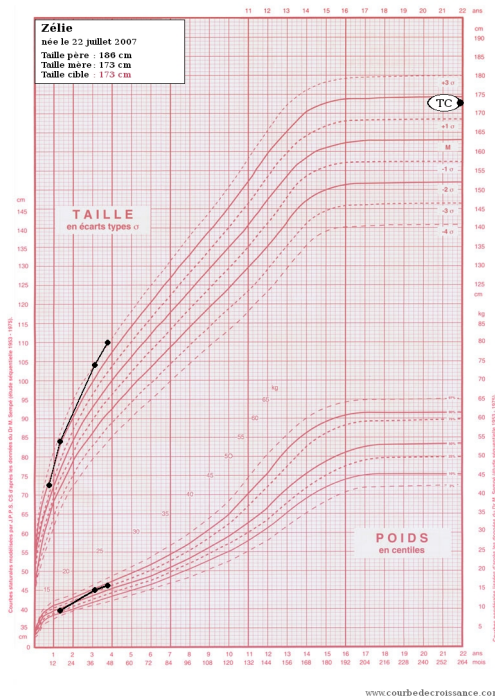
**Matching et transfert de données** Enfin, un dernier type de problèmes issu de la neuro-imagerie : la création de cartes sémantiques sur les surfaces corticales des sujets observés. Pensons par exemple à une collection de visages. Il est aisé pour un être humain de "découper", *segmenter* celui-ci en régions anatomiques bien définies : nez, yeux, bouches, oreilles... Toutes sont d'apparences, de topologies bien différenciées : il est donc possible d'apprendre à un ordinateur à reconnaître ces différents point saillants sur une image.

Mais comment procéder lorsque "toutes les régions se ressemblent"? Sur une simple image anatomique, difficile en effet de distinguer les circonvolutions du cortex visuel de celles du lobe frontal... On s'efforcera donc de transférer une *carte* – dessinée par des experts – d'un cerveau *modèle* typique à un cerveau donné, de manière anatomiquement crédible – voir Figure 1.6.

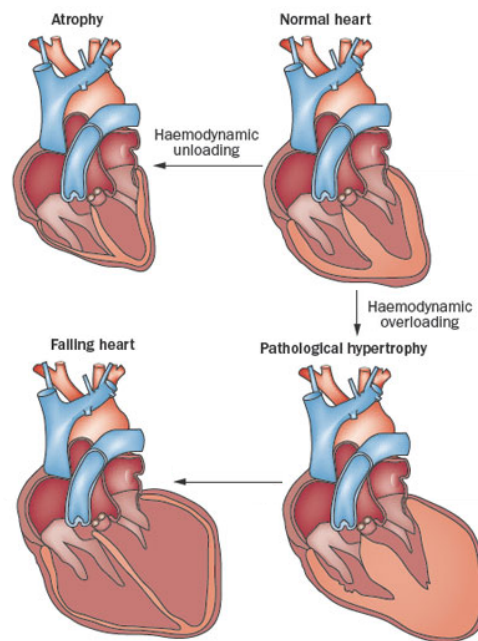
Dans les trois situation précédentes, tout serait simple si l'on disposait d'une structure algébrique/vectorielle raisonnable sur nos données – qui sont ici des "formes" – : moyenne, régression linéaire et analyse de la variance de l'échantillon sont toujours données par des formule pleines de « + » et de « × »... Malheureusement, une théorie additive des formes a peu de chance de voir le jour : quel sens pourrait-on donner à la somme de deux cœurs?

Qu'à cela ne tienne : une notion de distance, une métrique Riemannienne suffira. La *moyenne* ne sera plus qu'un point minimisant la somme des distances aux observations et la droite de *régression*, une courbe géodésique approchant au mieux les données aux instants d'observations.



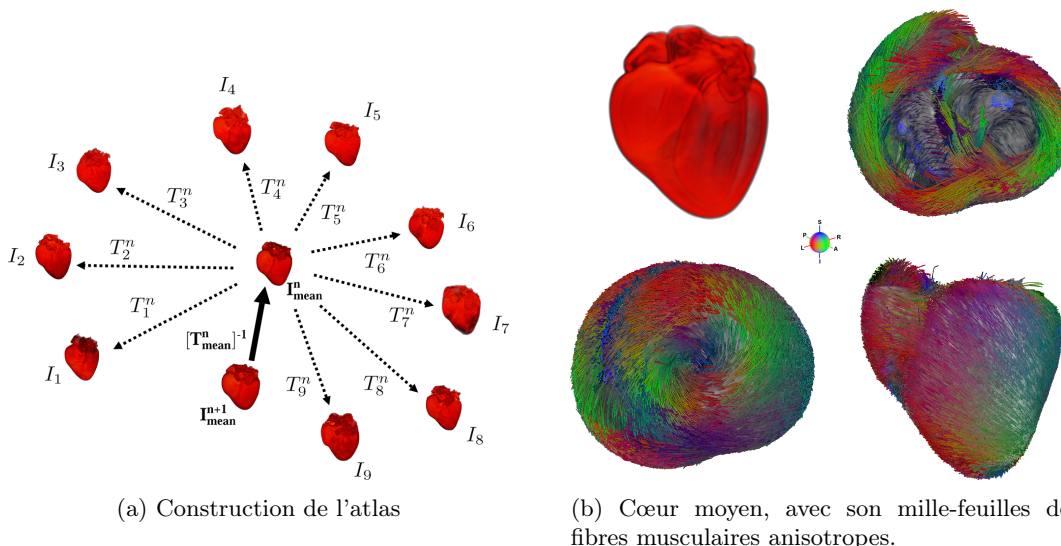


(a) Courbe de croissance.



(b) Exemples de malformations cardiaques. Comment les détecter avant qu'il ne soit trop tard ?

FIGURE 1.3 – Le suivi longitudinal et les techniques de régression méritent d'être généralisés aux formes. (b) est tiré de *Vascular endothelial growth factor in heart failure*, Taimieh et Al., *Nature Reviews Cardiology* 10 (Septembre 2013).



(a) Construction de l'atlas

(b) Cœur moyen, avec son mille-feuilles de fibres musculaires anisotropes.

FIGURE 1.4 – Estimation d'un cœur moyen à partir de sept cœurs de chiens. Images tirées de *A computational framework for the statistical analysis of cardiac diffusion tensors : application to a small database of canine hearts*, Peyrat et Al., *IEEE transactions on medical imaging*, 2007.

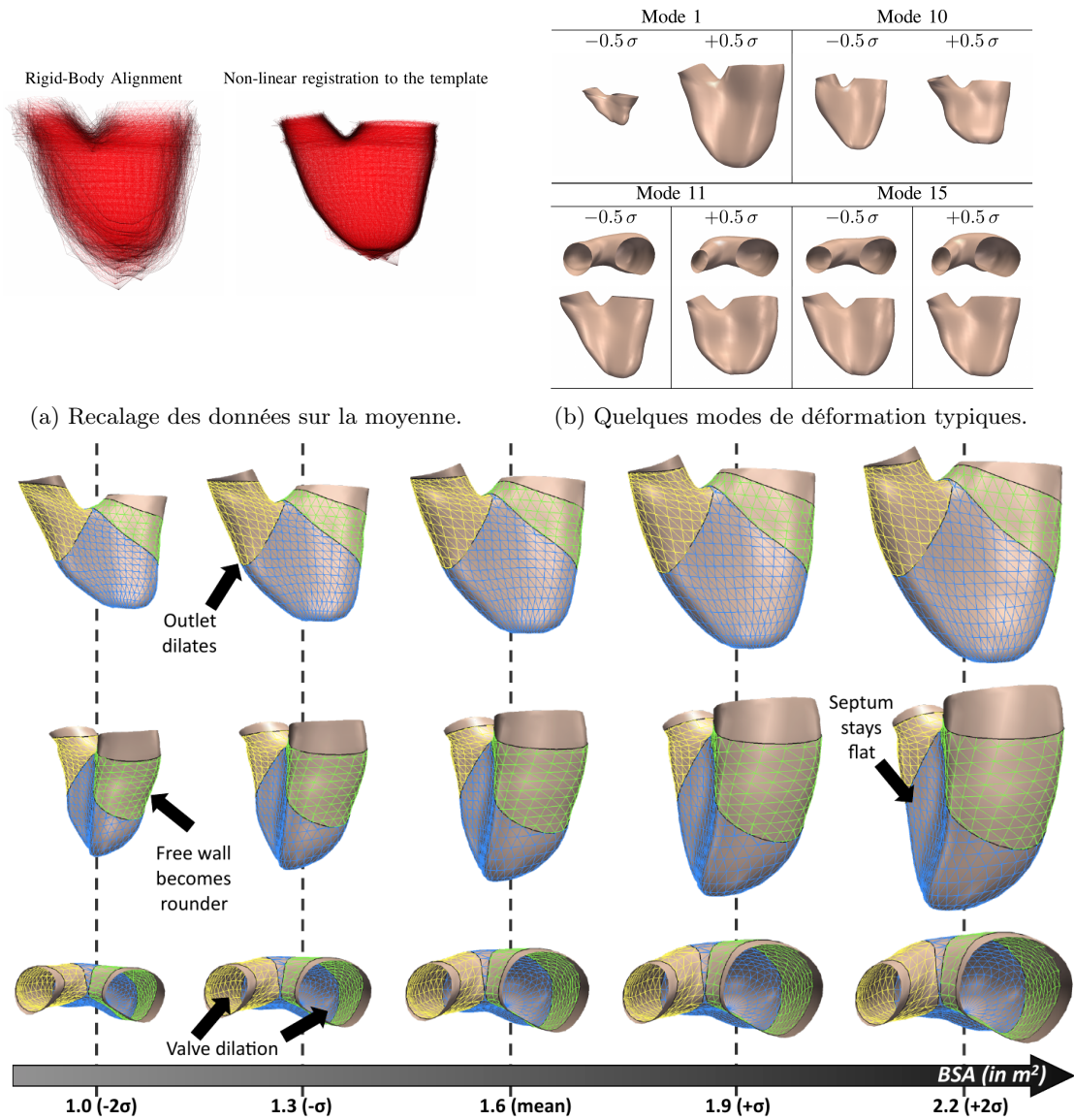
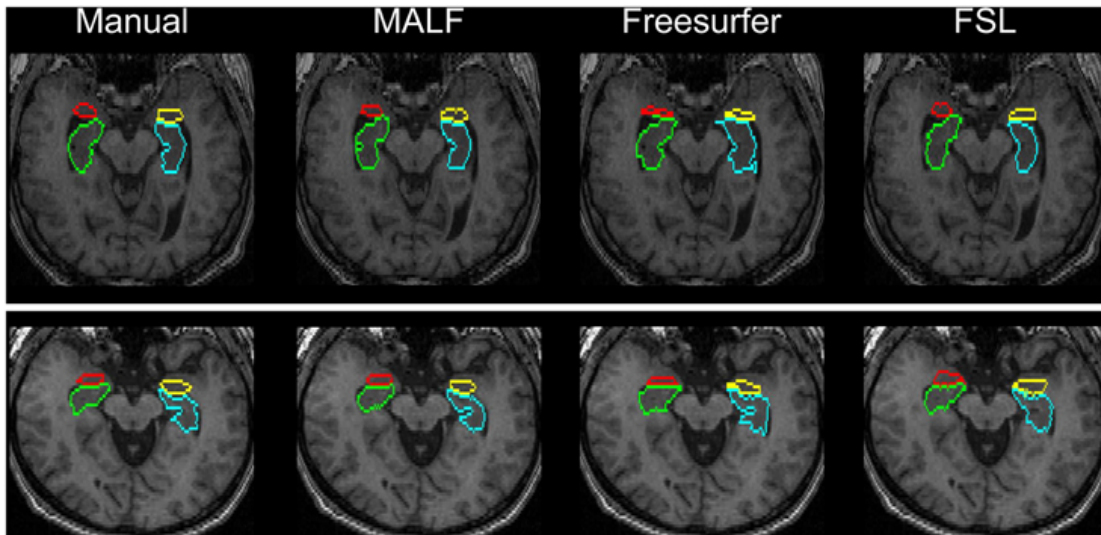
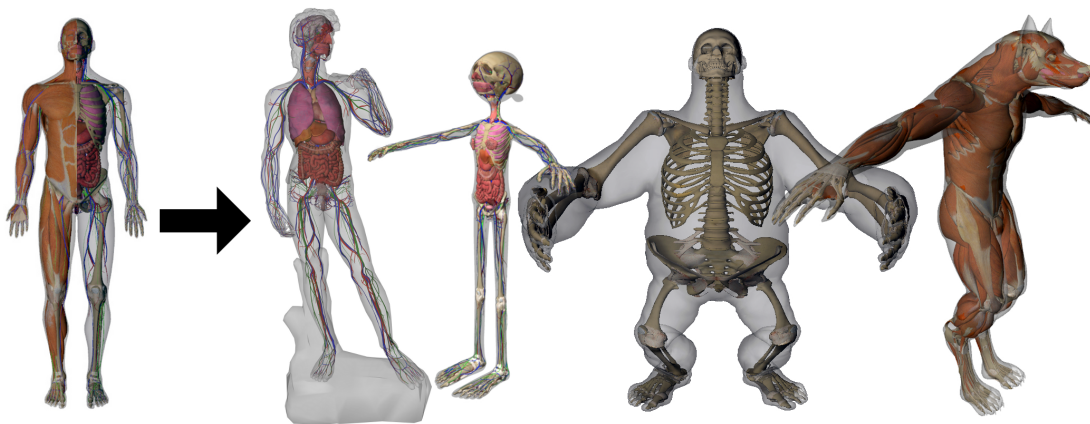


FIGURE 1.5 – Autre estimation d’atlas de données cardiaques – ventricules droits.  
Images tirées de *A statistical model for quantification and prediction of cardiac remodelling : Application to tetralogy of fallot*, Mansi et Al., IEEE transactions on medical imaging, 2011.



(a) Exemples de segmentation de régions du cerveau. Deux sujets sont observées, quatre méthodes sont utilisées pour transporter une carte pré-établie.

Image tirée de *Segmentation of brain magnetic resonance images based on multi-atlas likelihood fusion : testing using data with a broad range of anatomical and photometric profiles*, Tang et Al., *Frontiers in Neurosciences*, 03 March 2015.



(b) À partir d'un modèle de référence, on peut inférer les mécanismes internes de silhouettes arbitraires : Le David, Olive, Brutus, un loup-garou... Des débouchés dans le domaine de l'imagerie médicale et de l'animation sont envisagés.

Image tirée de *Anatomy Transfer*, Dicko et Al., *ACM Transactions on Graphics*, 2013.

FIGURE 1.6 – Deux exemples de transfert de modèle anatomique.

On modélisera ici nos formes par des nuages de points, éléments d'un espace de *landmarks*

$$\mathcal{L}_p^n = \{(x_1, \dots, x_n) \in \mathbb{R}^p \mid \forall i \neq j, x_i \neq x_j\}, \quad (1.13)$$

que l'on peut voir comme un ouvert de  $\mathbb{R}^{np}$ . (On pourrait bien sûr en considérer le quotient à numérotation  $\sigma \in S_n$  près, à similitude près, ou bien considérer des espaces de courbes, de surfaces voire de mesures, mais cela nous encombrerait plus qu'autre chose dans ce cours à l'agenda serré.)

Viendront alors trois questions : Comment construire une métrique locale anatomiquement pertinente sur  $\mathcal{L}_p^n$  ? Quelles seront alors les propriétés de l'espace métrique induit ? Et surtout, sera-t-on capable d'implémenter de manière efficace les algorithmes de régression, matching, création d'atlas que nous demandent les médecins/neurologues/biologistes ? C'est ce que nous verrons dans la deuxième moitié du semestre.

## 1.5 Variétés : un petit guide de survie

Dans le cadre de ce groupe de lecture – de niveau licence, rappelons-le –, on se bornera à travailler sur des ouverts de  $\mathbb{R}^d$  : nos espaces de landmarks définis Eq. (1.13), qui sont déjà bien assez riches tout en restant conceptuellement très simples. Malheureusement, ce n'est pas le point de vue adopté par notre manuel – de niveau master – et les articles de recherche rencontrés dans la littérature. Le bon cadre n'est en effet pas celui – extrinsèque – des ouverts de  $\mathbb{R}^n$ , mais plutôt celui des *variétés lisses*, qui englobe les surfaces, les espaces courbes, l'espace-temps relativiste... En fait, tout ce qui ressemble *localement* à un ouvert de  $\mathbb{R}^n$  – voire d'un espace de Banach de dimension infinie.

Mettons donc à plat ce qu'il est nécessaire de savoir sur les variétés pour ne pas être complètement perdu.

**Sous-variétés, variétés ?** Nous souhaitons travailler avec des espaces qui sont “localement semblables” à  $\mathbb{R}^d$ ... Comme souvent en mathématiques, il y a en fait deux manières de définir et caractériser ces objets : en parler de manière *extrinsèque*, avec le vocabulaire des sous-variétés plongées dans  $\mathbb{R}^n$  ; ou bien le faire de manière *intrinsèque* – on parlera alors de variété abstraite, ou, tout simplement, de *variété*. Le théorème de plongement de Whitney permettra alors d'identifier les deux notions, en montrant que toute variété abstraite de dimension  $d$  peut être plongée dans l'espace euclidien  $\mathbb{R}^{2d}$ .

**Définition 1.1.** Soit  $M \subset \mathbb{R}^n$ . On dit que  $M$  est une sous variété de dimension  $d$  si et seulement si en chaque point  $x_0 \in M$ , il existe un voisinage  $U$  de  $x_0$ , un difféomorphisme  $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n$  tel que

$$\varphi(U \cap M) = \mathbb{R}^d \times \{0\} \cap \varphi(U). \quad (1.14)$$

Suivant le polycopié de géométrie différentielle de M. Viterbo, dont on s'inspirera ici fortement, on supposera ici que tous les difféomorphismes sont de classe  $C^\infty$ .

Le difféomorphisme local  $\varphi$ , qui permet d'identifier  $M$  avec  $\mathbb{R}^d$  au voisinage de  $x_0$ , s'appelle une *carte locale*. L'application réciproque  $\varphi|_{\mathbb{R}^d}^{-1} : \mathbb{R}^d \rightarrow M$  s'appelle quand à elle *un système de coordonnées locales*.

Bien entendu, les exemples classiques de sous-variété de l'espace sont la sphère, le tore... Le résultat suivant permet de jongler entre les différentes représentations que l'on pourrait en donner :

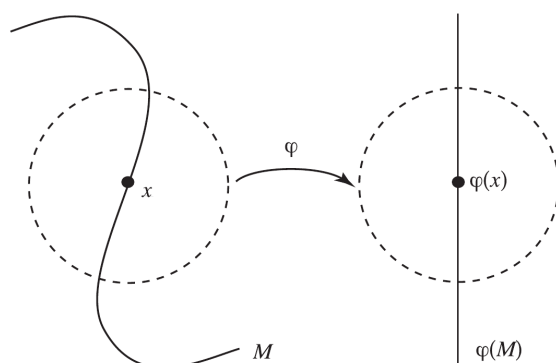


FIGURE 1.7 – Carte pour une sous-variété de dimension 1 du plan.

**Proposition 1.1.** Soit  $M \subset \mathbb{R}^n$ . Alors  $M$  est une sous-variété de dimension  $d$  si et seulement si pour tout point  $x_0 \in M$ , il existe un voisinage  $U$  de  $x_0$  dans  $\mathbb{R}^n$  tel que l'une de ces propriétés soit vérifiée :

**Graphe** il existe un changement linéaire de coordonnées  $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , une application  $f : \mathbb{R}^d \rightarrow \mathbb{R}^{n-d}$  qui soit  $C^\infty$  telle que

$$U \cap M = U \cap \{A(z, f(z)) \mid z \in \mathbb{R}^d\}. \quad (1.15)$$

**Équation** Il existe une application  $F : U \rightarrow \mathbb{R}^{n-d}$  telle que  $dF_0$  est surjective, et  $U \cap M = F^{-1}(0)$ .

**Nappe paramétrée** Il existe  $j : \mathbb{R}^d \rightarrow \mathbb{R}^n$  de classe  $C^\infty$ , définie sur un voisinage  $V$  de 0, telle que  $j(0) = x_0$ ,  $dj(0)$  est injective, et que la restriction

$$j : V \rightarrow M \cap U \quad (1.16)$$

soit bijective, continue, d'inverse continue.

*Exercice 1.1.* Montrer que le cercle unité, la sphère unité, les ouverts de  $\mathbb{R}^n$  sont des sous-variétés.

Nous terminons cette première leçon avec la définition formelle d'une variété abstraite :

**Définition 1.2.** Soit  $M$  un espace métrisable, réunion dénombrable de compacts. On appelle structure de variété lisse de dimension  $d$  sur  $M$  la donnée d'une famille  $(U_\lambda, \varphi_\lambda)_{\lambda \in \Lambda}$  telle que :

- les  $U_\lambda$  sont des ouverts, et forment un recouvrement de  $M$ ,
- les  $\varphi_\lambda : U_\lambda \rightarrow \mathbb{R}^d$  sont des homéomorphismes sur l'ouvert  $\varphi_\lambda(U_\lambda)$  tels que pour tous  $\alpha, \beta$

$$\varphi_{\alpha\beta} = \varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \rightarrow \varphi_\beta(U_\alpha \cap U_\beta) \quad (1.17)$$

est une application  $C^\infty$ .

Au lieu de s'encombrer d'un espace ambiant qui passe souvent très mal au quotient, on se concentre donc sur le seul ressenti "interne" des habitants de l'espace : pensons au tore, souvent représenté comme une bouée, mais qu'il est plus simple de voir comme un carré aux bords recollés – Cf. les RPG japonais, Final Fantasy en tête.

Rendez-vous la semaine prochaine pour la fin de cette introduction au vocabulaire de la géométrie différentielle !



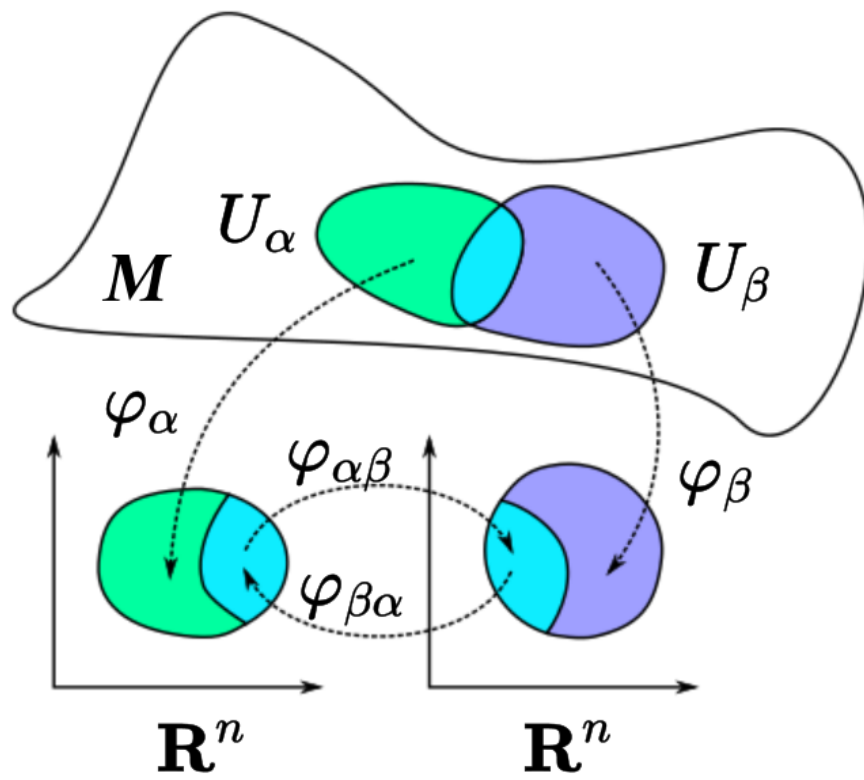


FIGURE 1.8 – Atlas de cartes sur une variété abstraite.  
Image tirée de Wikipédia.

Première partie

Éléments de Géométrie  
Riemannienne





# Chapitre 2

## Intuitions sur la courbure, Vocabulaire

*Séance 2*

*[Lee97], Chap. 1-2*

*Arnaud Eteve,  
Léo Hahn-Lecler,  
Emmanuel  
Kammerer*

*Un peu énumératif*

### 2.1 Introduction

La géométrie Riemannienne permet de réintroduire, dans la géométrie différentielle, les concepts géométriques d'angles, de distances, sur des objets courbes. Nous introduirons donc la notion de courbure et son lien avec la topologie. Puis nous nous équiperons de l'attirail nécessaire dans la perspective d'introduire les distances et les angles : tenseurs, variétés, fibrés...

### 2.2 Intuitions sur la courbure

#### 2.2.1 Dans le plan

Pour étudier d'autres courbes que les cercles et les droites, il peut être utile de mesurer à quel point une courbe est courbée. On se donne donc  $\gamma$  une courbe  $C^\infty$  définie par son paramétrage de  $\mathbb{R} \rightarrow \mathbb{R}^2$ . On suppose de plus que  $\|\dot{\gamma}(t)\| = 1$ .

**Définition 2.1** (Courbure). Soit  $p$  un point de la courbe. La courbure de  $\gamma$  en  $p = \gamma(t)$  est  $\kappa(t) = \|\ddot{\gamma}(t)\|$ .

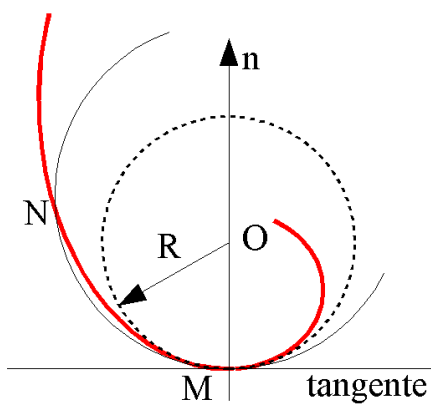
**Définition 2.2** (Cercle osculateur). Mieux qu'une droite tangente, il existe un unique cercle  $C$  approchant  $\gamma$  à l'ordre 2 en  $p$ , qu'on appelle le cercle *osculateur* à  $\gamma$  en  $p$ . C'est le cercle (paramétré) dont les vecteurs vitesse et accélération coïncident en  $p$  avec ceux de la courbe. On peut facilement montrer que son centre est sur la droite orthogonale à  $\dot{\gamma}$  en  $p$ , et son rayon est égal à l'inverse de la courbure,  $R = \frac{1}{\kappa}$ .

La courbure ainsi définie ne permet pas de savoir "dans quel sens ça tourne". On peut pour ce faire affecter un signe à la courbure en choisissant un champ de vecteurs normaux à la courbe, noté  $N$ , que l'on voit comme une direction canonique. On note alors la courbure signée  $\kappa_N$ ,  $\kappa_N$  est positive si  $N$  et  $\ddot{\gamma}$  ont le même sens, négative sinon.

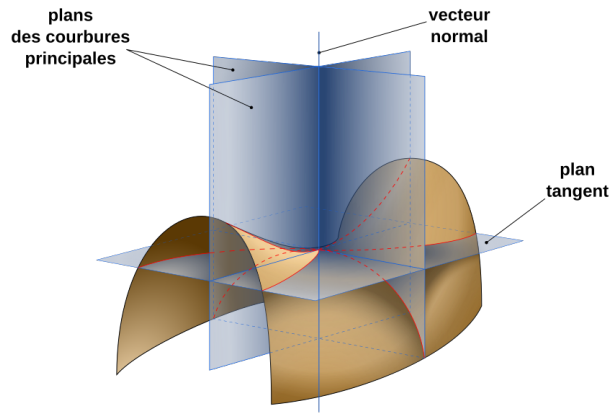
#### 2.2.2 Dans l'espace

Soit  $S$  une surface lisse – i.e. de classe  $C^\infty$  – dans  $\mathbb{R}^3$ .

Arnaud Eteve,  
 Léo Hahn-Lecler,  
 Emmanuel  
 Kammerer  
 Un peu énumératif



(a) Le cercle osculateur en M à la courbe rouge. De manière générique, la courbe *traverse* le cercle osculateur : le terme d'ordre 3 n'a aucune raison de s'annuler.



(b) Courbures principales sur une selle de cheval. Quelle que soit l'orientation choisie, les deux courbures sont de signe opposé : la selle est une surface de courbure gaussienne *négative*.

FIGURE 2.1 – Illustrations sur la courbure d'une trajectoire et d'une surface, tirées de Wikipédia.

**Définition 2.3** (Courbure Principale). Soit  $p \in S$ , on note  $n$  la normale à  $S$  en  $p$ . Soit  $\Pi$  un plan contenant  $n$  et  $p$  : l'intersection de  $\Pi$  avec  $S$  est (localement) une courbe planaire, dont on note  $\kappa_\Pi$  la courbure en  $p$ . Lorsqu'on parcourt tous les plans  $\Pi$  possibles,  $\kappa_\Pi$  atteint un min et un max, notés  $\kappa_1$  et  $\kappa_2$ , elles sont appelées courbures principales de  $S$  en  $p$ .

Les courbures principales ne sont pas invariantes par isométries (ici, les isométries sont des applications préservant la longueur des chemins tracés dans la surface). Cependant, le théorème suivant nous donne un invariant :

**Théorème 2.1** (Theorema Egregium). *Le produit  $\kappa_1\kappa_2$  est intrinsèque (i.e. invariant par isométrie : il ne dépend que de la structure métrique "dans" la surface), on l'appelle la courbure Gaussienne.*

Exemples : le plan et le demi-cylindre ont une courbure gaussienne égale à 0 alors qu'une sphère de rayon  $R$  a une courbure gaussienne égale à  $\frac{1}{R^2}$ .

## 2.3 Variétés

Introduisons quelques définitions pour généraliser les courbes et les surfaces. Nous renvoyons le lecteur au guide de survie pour la définition des variétés lisses.

Soit  $M$  une variété,  $x_0 \in M$  et  $U \subset M$  un voisinage de  $x_0$  : on identifiera  $U$  avec son image dans  $\mathbb{R}^n$ , ce qui nous donne des coordonnées locales lisses autour de  $x_0$ .

Le plan (vectoriel) tangent à une surface en un point est le plan ayant pour vecteur normal la normale  $N$  à la surface en  $p$ . Mais c'est aussi l'ensemble des vecteurs vitesse en  $p$  des courbes dans la surface passant par  $p$ . De même, pour une sous-variété  $M$  de  $\mathbb{R}^n$ , l'espace tangent en  $p$  est l'ensemble des vecteurs vitesse en  $p$  des courbes à valeurs dans  $M$  passant par  $p$ .

De manière encore plus générale, on définit :

**Définition 2.4** (Espace Tangent, Tangent space). Soit  $p \in M$ , on considère l'ensemble des courbes de  $M$  qui passent par  $p$  quotienté par la relation d'équivalence suivante :  $\gamma_1 \sim \gamma_2$  ssi ils

ont le même vecteur vitesse en  $p$ . Cet espace est un espace vectoriel de même dimension que  $M$  noté  $T_pM$ .

**Définition 2.5** (Sous Variété, Submanifold). Soit  $M_1$  et  $M_2$  des variétés,  $M_1$  est une sous variété (régulière) de  $M_2$  si on dispose d'une injection  $i : M_1 \rightarrow M_2$  et  $i$  réalise un homéomorphisme de  $M_1 \rightarrow i(M_1)$ .

De cela on obtient des coordonnées locales sur  $M_1$  à partir de celles de  $M_2$  et on peut voir  $T_pM_1$  comme un sous espace de  $T_pM_2$

## 2.4 Fibrés Vectoriels

La réunion disjointe des espaces tangents peut être considérée en même temps comme une union d'espaces vectoriels et comme une variété. C'est un exemple de fibré vectoriel.

**Définition 2.6** (Fibré Vectoriel, Vector Bundle). Un fibré vectoriel de dimension  $k$  est la donnée d'un couple de variétés lisses,  $E$  (l'espace total) et  $M$  (la base), ainsi que d'une application  $\pi : E \rightarrow M$  (la projection) vérifiant :

- Chaque  $E_p = \pi^{-1}(p)$  (appelé la fibre de  $p$ ) a une structure d'espace vectoriel
- Pour tout  $p \in M$ , il existe un voisinage  $U$  de  $p$  et un difféomorphisme  $\varphi : \pi^{-1}(U) \rightarrow U \times \mathbb{R}^k$  appelé "trivialisatation locale" telle que  $\pi = \pi_1 \circ \varphi$  où  $\pi_1$  est la projection sur  $U$ .
- Les restriction de  $\pi$  à chaque fibre sont des isomorphismes.

Exemples :  $TM$  (fibré tangent) dont on a parlé plus tôt,  $T^*M = \coprod_{p \in M} (T_pM)^*$ ,  $NM =$

$$\coprod_{p \in M} (T_pM)^\perp.$$

D'ailleurs, l'espace des phases  $(x, \dot{x})$  en physique correspond au fibré tangent  $TR$ .

*Remarque 2.1.* La notion de fibré vectoriel est une généralisation de celle de produit cartésien : un produit cartésien est un fibré vectoriel, et tout fibré vectoriel ressemble localement à un produit cartésien. Il est par contre possible d'exhiber des fibrés vectoriels qui sont des produits cartésiens "ensemblistes", mais dont la structure topologique n'est pas "redressable" : on pense par exemple aux twists sur le produit  $S^1 \times \mathbb{R}$ , comme le ruban de Moebius.

*Remarque 2.2.* En introduisant une notion de recollement lisse entre plans tangents voisins, les fibrés tangents/cotangents permettent d'exprimer simplement les équations différentielles portant sur le couple (position, vitesse)  $(q, v) \in TM$  ou mieux, le couple (position, moment)  $(q, p) \in T^*M$ . C'est une notion bien connue en physique : celle de l'espace des phases d'un problème, qui permet d'exprimer très simplement les équations de la mécanique.

Prenons l'exemple d'une masse ponctuelle  $m$  en chute libre dans un champ de gravité uniforme  $g$ . En repérant la position de la masse par son altitude  $z$ , et en appliquant les lois de Newton, on obtient (avec les bonnes conventions de signe) l'équation du second ordre :

$$m\ddot{z} = -mg. \quad (2.1)$$

Pour obtenir existence et unicité des solutions au problème de Cauchy, il est de bon goût de réécrire notre dynamique sous la forme d'une équation du *premier ordre* sur l'espace des phases :

$$\begin{pmatrix} \dot{z} \\ \dot{v} \end{pmatrix} = X(z, v) = \begin{pmatrix} v \\ -g \end{pmatrix}. \quad (2.2)$$

[Lee97], Chap. 1-2

Arnaud Eteve,  
Léo Hahn-Lecler,  
Emmanuel  
Kammerer  
Un peu énumératif

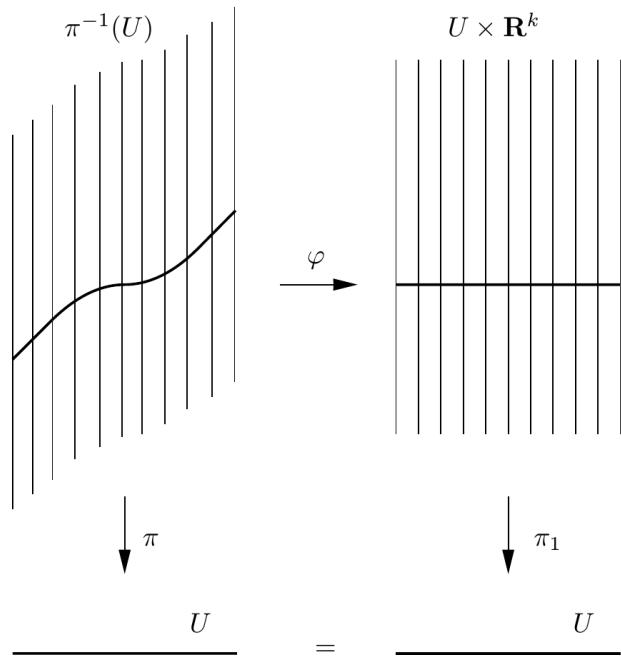
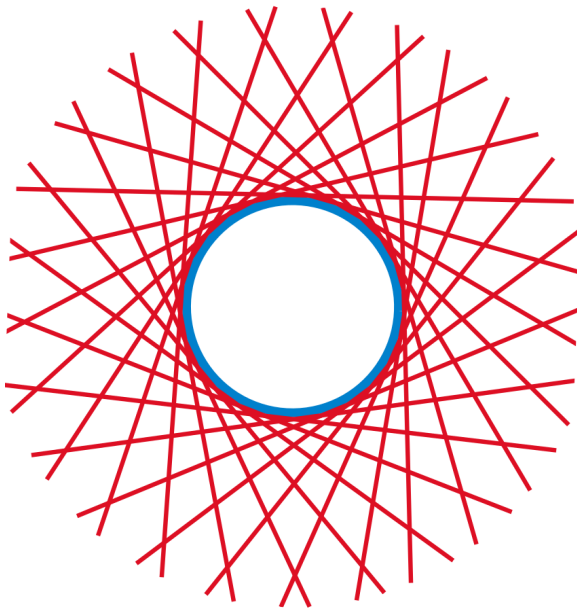
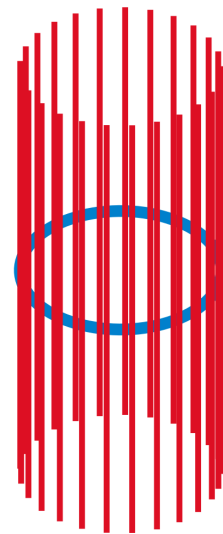


FIGURE 2.2 – Définition intrinsèque et locale de la notion de *fibré vectoriel*. Image tirée de [Lee97].

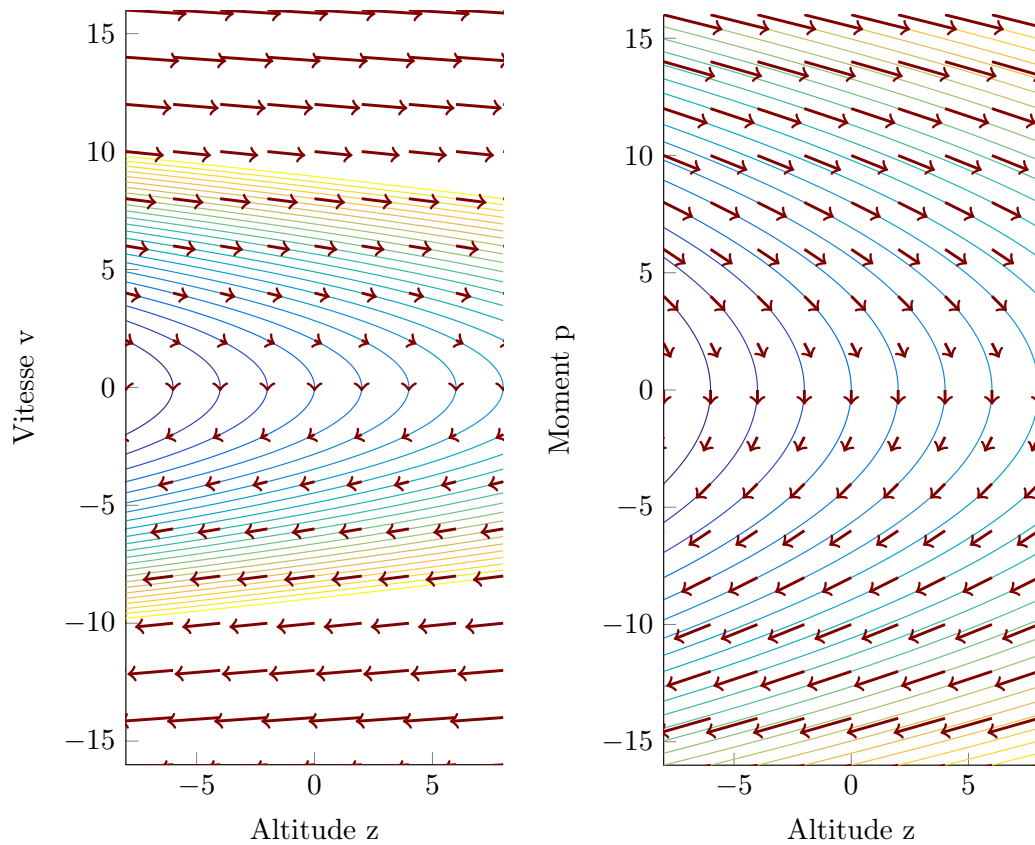


(a) Fibres dessinées tangentes au cercle.



(b) L'espace de base  $E = S^1 \times \mathbb{R}$ .

FIGURE 2.3 – Deux manières de voir le fibré tangent au cercle  $S^1$ . Twister la représentation (b) pour obtenir un ruban de Moebius permet d'obtenir un fibré vectoriel sur le cercle qui n'est *pas* équivalent au produit cartésien – exercice !



(a) Flot des équations de Newton dans l'espace des phases  $(z, v)$ . Le champ de vecteur  $X(z, v)$  est tangent aux lignes de niveau de l'énergie mécanique : cette dernière est donc une quantité conservée au cours du mouvement.

(b) Flot Hamiltonien dans l'espace des phases  $(z, p)$ . Les trajectoires restent à énergie constante, avec un renseignement supplémentaire : plus les lignes de niveau du Hamiltonien sont serrées, plus le gradient est fort, et plus le point courant bouge vite dans l'espace des phases.

FIGURE 2.4 – Flot des équations de la mécanique classique dans l'espace des phases (en rouge) pour la chute libre d'une masse  $m = 2$  dans un champ de gravité uniforme  $g = 1$ . Les lignes de contours correspondent aux iso-énergies.

[Lee97], Chap. 1-2

Arnaud Eteve,  
Léo Hahn-Lecler,  
Emmanuel  
Kammerer  
Un peu énumératif

L'étude de notre système est donc ramenée à celle du flot du champ de vecteur  $X$  sur le produit cartésien  $\mathbb{R} \times \mathbb{R}$  : c'est, à peu de choses près, ce que les classes prépa nous ont appris sur la mécanique classique... Mais depuis Newton, les physiciens ont trouvé des manières bien plus élégantes – et donc efficaces – de formuler ce problème. Le formalisme Lagrangien exprime ainsi qu'entre deux états  $(q_0, v_0)$  et  $(q_1, v_1)$ , un système physique emprunte une trajectoire qui minimise l'intégrale au cours du temps de l'action – une grandeur physique ad hoc. Plus pratique encore le formalisme Hamiltonien propose de substituer au vecteur tangent  $v$  un vecteur cotangent  $p$ , le moment, défini de manière appropriée. Dans notre cas, il s'agit simplement du moment inertiel  $p = mv \in (\mathbb{R})^* = \mathbb{R}$ . En les variables  $(z, p)$ , on réécrit alors le système 2.2 :

$$\begin{cases} \dot{z} &= +p/m \\ \dot{p} &= -mg \end{cases} \quad (2.3)$$

Le coup de génie est alors de remarquer que, pour le Hamiltonien

$$H(z, p) = E_c + Epp = \frac{1}{2} \frac{p^2}{m} + mgz, \quad (2.4)$$

on a

$$\begin{cases} \dot{z} &= +\frac{\partial H}{\partial p} \\ \dot{p} &= -\frac{\partial H}{\partial z} \end{cases} \quad \text{i.e.} \quad \begin{pmatrix} \dot{z} \\ \dot{p} \end{pmatrix} = R_{-90^\circ} \nabla H \quad (2.5)$$

Le flot "Hamiltonien" des systèmes physiques n'est donc autre qu'un flot de gradient "tourné d'un quart de tour". De cette information, on peut immédiatement déduire que le Hamiltonien – ici, l'énergie mécanique – est une quantité conservée au cours du mouvement... Et de nombreuses autres propriétés qui forment l'objet d'étude de la *mécanique géométrique* et la motivation principale de la *géométrie symplectique*.

L'espace des phases d'une variable  $(x, y, z) \in \mathbb{R}^3$  n'est autre que le produit  $\mathbb{R}^3 \times \mathbb{R}^3$  des couples  $((x, y, z), (\dot{x}, \dot{y}, \dot{z}))$  – i.e. le fibré tangent  $T\mathbb{R}^3$ . La généralisation présentée ici permet donc simplement de parler d'espaces des phases d'une variable  $q \in M$  quelconque.

C'est un point crucial dans l'élaboration de notre cours : comme le problème des géodésiques est moralement un problème "Lagrangien", nous verrons que les courbes *géodésiques* sont en fait solutions d'une équation différentielle d'ordre 2 (analogue aux équations de Newton), que l'on peut exprimer dans l'espace des phases  $T^*M$  comme le flot du Hamiltonien  $H(q, p) = \frac{1}{2}(p, K_q p)$  où  $K_q$  est la cométrique au point  $q$  – tout ceci sera bien sûr précisé en temps voulu. La puissance du point de vue Hamiltonien – qui capture l'essence de la propriété "être une géodésique" sous une forme compacte et pratique – sera au cœur des algorithmes que nous développerons dans la deuxième partie du semestre.

Pour montrer que les exemples précédents sont des fibrés vectoriels sans avoir à construire une structure de variété lisse sur l'ensemble total, on peut s'appuyer sur le lemme suivant :

**Lemme 2.1.** *Soit  $M$  une variété lisse,  $E$  un ensemble et  $\pi : E \rightarrow M$  une application surjective. Soit  $\{U_\alpha\}$  un recouvrement ouvert de  $M$ , et des applications bijectives  $\varphi_\alpha : \pi^{-1}(U_\alpha) \rightarrow U_\alpha \times \mathbb{R}^k$  satisfaisant  $\pi_1 \circ \varphi_\alpha = \pi$  et dès que  $U_\alpha \cap U_\beta \neq \emptyset$ , l'application  $\varphi_\alpha \circ \varphi_\beta^{-1} : U_\alpha \cap U_\beta \times \mathbb{R}^k \rightarrow U_\alpha \cap U_\beta \times \mathbb{R}^k$  est de la forme  $\varphi_\alpha \circ \varphi_\beta^{-1}(p, V) = (p, \tau(p)V)$ , où  $\tau$  est une application lisse  $\tau : U_\alpha \cap U_\beta \rightarrow GL(k, \mathbb{R})$ .*

*Alors  $E$  a une unique structure de  $k$ -fibré vectoriel lisse sur  $M$  pour laquelle les applications  $\varphi_\alpha$  sont les trivialisations locales. La fonction  $\tau$  est appelée fonction de transition.*

En d'autres termes, il suffit d'exhiber des trivialisations locales qui se superposent "correctement".

**Définition 2.7** (Section). Soit  $\pi : E \rightarrow M$  un fibré vectoriel, une section de  $E$  est une application lisse  $F : M \rightarrow E$  telle que  $\pi \circ F = Id_M$ .

C'est donc la donnée lisse en tout point  $p$  d'un "vecteur"  $F(p)$  vivant dans la fibre  $\pi^{-1}(p)$ .

## 2.5 Tenseurs

La géométrie Riemannienne est construite à l'aide de tenseurs. Ces objets sont en quelque sorte des généralisations des matrices. Ils peuvent représenter en particulier des formes multilinéaires, des matrices ou des produits scalaires.

Pour alléger les notations, il est d'usage d'utiliser la sommation d'Einstein : dès lors qu'un indice est hors contexte c'est qu'on le somme sur toutes les valeurs possibles. C'est par exemple utile pour le calcul de la différentielle d'une composée, ou pour écrire des produits matriciels de manière compacte.

**Définition 2.8** (Tenseur, Cotenseur). Soit  $V$  un espace vectoriel de dimension finie, un  $k$ -tenseur est une application  $k$  linéaire de  $V^k \rightarrow \mathbb{R}$  un  $l$ -cotenseur est une application  $l$  linéaire de  $V^{*l} \rightarrow \mathbb{R}$

un  $\binom{k}{l}$ -tenseur est une application multilinéaire de  $V^{*l} \times V^k \rightarrow \mathbb{R}$

On note respectivement  $T^k(V)$ ,  $T_l(V)$  et  $T_l^k(V)$  ces ensembles de tenseurs.

On a en particulier :  $T^1(V) = V^*$ ,  $T_1(V) = V$  et  $T_1^1(V) = End(V)$ .

*Remarque 2.3.* On peut voir un tenseur d'ordre 2 comme la décomposition d'une matrice sur l'espace de base  $V$  et son dual  $V^*$  : en effet, une matrice  $M$  de taille  $n \times n$  est entièrement déterminée par la donnée de l'application bilinéaire  $(Y, X) \in (\mathbb{R}^n)^* \times \mathbb{R}^n \mapsto (Y, MX) = Y^t M X$ .

*Remarque 2.4.* La terminologie "tensorielle" est plus lourde que celle des matrices, qui nous suffirait pourtant ici – aucun tenseur d'ordre supérieur à 2 n'est nécessaire pour parler de géodésiques. Elle a néanmoins l'avantage de désambiguïser la confusion qui est souvent faite entre un espace  $V$  et son dual  $V^*$ .

On le voit en cours d'analyse fonctionnelle, le dual topologique d'un espace de dimension infinie est rarement identifiable à l'espace de départ. Mais même dans le cas des espaces de landmarks, celui des ouverts de  $\mathbb{R}^n$ , il est important de faire la distinction : c'est que si on a topologiquement  $(\mathbb{R}^n)^* \simeq \mathbb{R}^n$ , la structure *métrique* sur ces deux espaces peut elle être complètement chamboulée.

Supposons par exemple  $\mathbb{R}^n$  muni d'un produit scalaire  $\langle \cdot, \cdot \rangle_G$  donné par une matrice  $n \times n$  symétrique définie positive  $G$  :

$$\forall v \in \mathbb{R}^n, \|v\|_G^2 = v^T G v. \quad (2.6)$$

De même qu'un vecteur est identifié à la matrice colonne de ses coordonnées dans la base canonique, on représente toujours une forme linéaire  $p \in (\mathbb{R}^n)^*$  par la matrice ligne de ses coordonnées dans la base duale canonique  $(e_1^*, \dots, e_n^*)$  :

$$p(x) = p_1 x_1 + p_2 x_2 + \dots + p_n x_n = p^T x. \quad (2.7)$$

Arnaud Eteve,  
Léo Hahn-Lecler,  
Emmanuel  
Kammerer  
Un peu énumératif

La norme d'opérateur d'une forme linéaire  $p$  est alors donnée par :

$$\|p\|_{G^*} = \max_{\langle v, v \rangle_{G=1}} p(v) \quad (2.8)$$

$$= \max_{\langle v, v \rangle_{G=1}} p^\top v \quad (2.9)$$

$$= \max_{u^\top u=1} p^\top G^{-1/2} u \quad (2.10)$$

$$= p^\top G^{-1/2} \cdot \frac{G^{-1/2} p}{\|G^{-1/2} p\|_2}, \quad (2.11)$$

avec le changement de variable  $u = G^{1/2} v$  – racine carrée d'une matrice de  $S_n^+$ . Autrement dit,

$$\|p\|_{G^*}^2 = p^\top G^{-1} p. \quad (2.12)$$

À une métrique  $G$  sur l'espace primal, il faut donc associer la cométrique  $K = G^{-1}$  sur l'espace dual. Si l'identification est tentante dans  $\mathbb{R}^n$  muni du produit scalaire canonique, c'est uniquement parce que  $I_n^{-1} = I_n$  : dans notre cours, il faudra aller plus loin.

## 2.6 Fibré Tensoriels, Champs de Tenseur

Étant donné un fibré tangent  $TM$ , on peut lui associer le fibré vectoriel pour lequel on a remplacé les  $T_p M$  par les  $T_l^k(T_p M)$ .

**Définition 2.9** (Fibré Tensoriel, Tensor Bundle). Soit  $M$  une variété, on construit un fibré vectoriel de la manière suivante : la fibre associée à  $p$  est  $T_l^k(T_p M)$ . On définit ainsi  $T_l^k M = \coprod_{p \in M} T_l^k(T_p M)$ , le  $\binom{k}{l}$ -fibré tensoriel associé à  $M$ .

**Définition 2.10** (Champ de tenseur, Tensor Field). Un champ de tenseurs est une section lisse de  $T_l^k M$ .

Exemple : un champ qui associe à chaque point  $p$  un produit scalaire  $\langle \cdot, \cdot \rangle_p$ .

Dans le cas des espaces de landmarks, qui sont des ouverts de  $\mathbb{R}^{nd}$ , l'inclusion donne une carte globale pour toute la variété : ces notions a priori locales de champs *lisses* se confondent avec la notion habituelle d'application à coordonnées lisses. Autrement dit, un champ de tenseur  $T_1^1$  lisse  $g$  sur un ouvert  $\Omega$  de  $\mathbb{R}^n$  n'est rien d'autre qu'une application  $g : \Omega \rightarrow M_{n,n}$  de classe  $C^\infty$  :

$$\forall x \in \Omega, g(x) = \begin{pmatrix} g_{1,1}(x) & g_{1,2}(x) & \cdots & g_{1,n}(x) \\ g_{2,1}(x) & g_{2,2}(x) & \cdots & g_{2,n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ g_{n,1}(x) & g_{n,2}(x) & \cdots & g_{n,n}(x) \end{pmatrix} \quad (2.13)$$

où les coordonnées  $g_{i,j} : \mathbb{R} \rightarrow \mathbb{R}$  sont infiniment différentiables.

## 2.7 Conclusion

Si, au début, on a pu appréhender la notion de courbure d'une surface, nous avons ensuite bien équipé nos variétés de fibrés qui nous permettront de construire ainsi une géométrie.



# Chapitre 3

## Métriques Riemanniennes

Séance 3  
[Lee97], Chap. 3

Paul Fraux,  
Manolis Perrot,  
Massil Hihat,  
Ivan Yakovlev  
Sans souci

### 3.1 Métriques Riemanniennes

#### 3.1.1 Définitions

**Définition 3.1** (Métrique). Une métrique Riemannienne sur une variété lisse (smooth manifold)  $M$  est une application continue  $g$  de  $M$  dans l'ensemble des formes bilinéaires symétriques et définies positives sur  $T_pM$ . Autrement dit,  $g$  associe à un  $p \in M$  un produit scalaire de  $T_pM$ .

#### Notations :

- Pour  $(X, Y) \in T_pM^2$ , on note  $\langle X, Y \rangle := g(p)(X, Y) = g(X, Y)$ .
- Un couple variété et métrique Riemannienne est appelé variété Riemannienne.
- S'il n'y a pas de confusion possible, on parlera de métrique et pas de métrique Riemannienne.
- $(X, Y) \in T_pM^2$  sont dit orthonormaux si  $\langle X, Y \rangle = 0$ .

**Définition 3.2** (Normes, angles). On peut alors définir une norme et des angles sur  $T_p$ , avec

$$|X| = \langle X, X \rangle^{1/2} \quad (3.1)$$

et l'angle entre  $X$  et  $Y$  pour  $(X, Y) \in T_p$  comme l'unique réel  $\theta$  de  $[0, \pi]$  tel que

$$\cos(\theta) = \frac{\langle X, Y \rangle}{|X| \cdot |Y|}. \quad (3.2)$$

**Définition 3.3** (tiré en arrière, pull-back). Si  $f$  est différentiable de  $M$  dans  $N$  (deux variétés), et si  $N$  est munie d'une métrique  $g^N$ , alors on peut définir une métrique sur  $M$  par :

$$v, w \in T_pM \quad (f^*g^N)(v, w) = g^N(df_p(v), df_p(w)). \quad (3.3)$$

On voit aisément que cela définit bien une forme bilinéaire symétrique et positive, par contre elle est définie positive si et seulement si la différentielle  $df_p$  est injective.

C'est à partir de cette métrique que d'autres notions clés vont être définies comme celle d'isométrie entre variétés riemanniennes. Avant de voir ces notions, manipulons cette définition à travers un exemple.

[Lee97], Chap. 3

Paul Fraux, Exemple 3.1. Regardons l'exemple du tore dans  $\mathbb{R}^3$ .

Manolis Perrot, On se place dans un repère  $(Oxyz)$ , avec le  $\theta$  des coordonnées polaires. On considère alors  
 Massil Hihat, l'application entre le rectangle  $[-b\pi, b\pi] \times [-\pi, \pi]$  et le tore dans  $\mathbb{R}^3$  :

Ivan Yakovlev

Sans souci

$$f(r, \theta) = R(r)(\cos \theta, \sin \theta, 0) + Z(r)(0, 0, 1) \quad (3.4)$$

$$\text{avec } R(r) = a + b \cos(r/b) \quad (3.5)$$

$$Z(r) = b \sin(r/b). \quad (3.6)$$

C'est une paramétrisation du tore, où à  $\theta$  fixé  $r$  décrit le "petit" cercle (section du tube) représenté ci-dessous, et pour  $r$  fixé  $\theta$  décrit le "grand" cercle (de hauteur  $z$  constante).

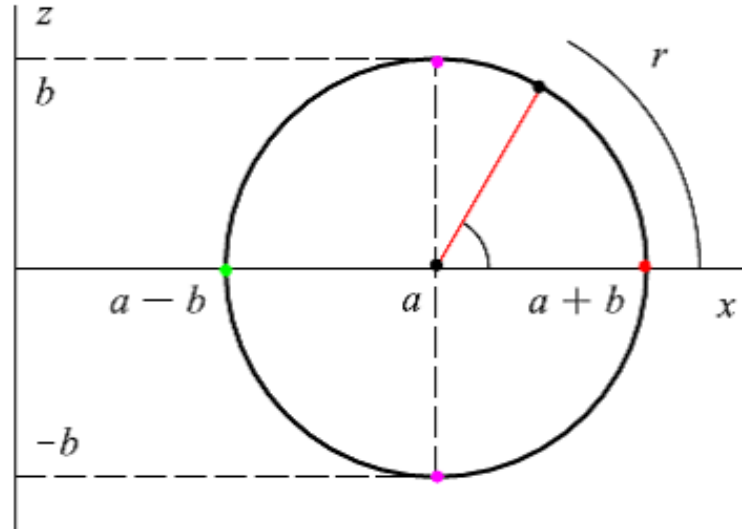


FIGURE 3.1 – Le tore vu en coupe : on fait tourner ce cercle autour de  $Oz$  pour obtenir le tore.

Notre but est de munir le carré d'une métrique qui est cohérente avec la vision qu'on a du tore de  $\mathbb{R}^3$  muni de la métrique euclidienne. Pour cela on peut calculer la métrique pull-back, qui en revenant à la définition nous demande de calculer la différentielle de  $f$  :

$$Df(r, \theta) = \begin{pmatrix} \cos(\theta)R'(r) & -\sin(\theta)R(r) \\ \sin(\theta)R'(r) & \cos(\theta)R(r) \\ Z'(r) & 0 \end{pmatrix} = \begin{pmatrix} -\cos(\theta)\sin(r/b) & -\sin(\theta)(a + b\cos(r/b)) \\ -\sin(\theta)\sin(r/b) & \cos(\theta)(a + b\cos(r/b)) \\ \cos(r/b) & 0 \end{pmatrix}$$

ainsi  $g_{(r,\theta)} = \begin{pmatrix} 1 & 0 \\ 0 & R(r)^2 \end{pmatrix}$  est la matrice de la métrique dans la base canonique.

et  $(e_{\hat{r}}, e_{\hat{\theta}}) = \left( \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1/R(r) \end{pmatrix} \right)$  est une base orthonormée du plan tangent au point  $(r, \theta)$ .

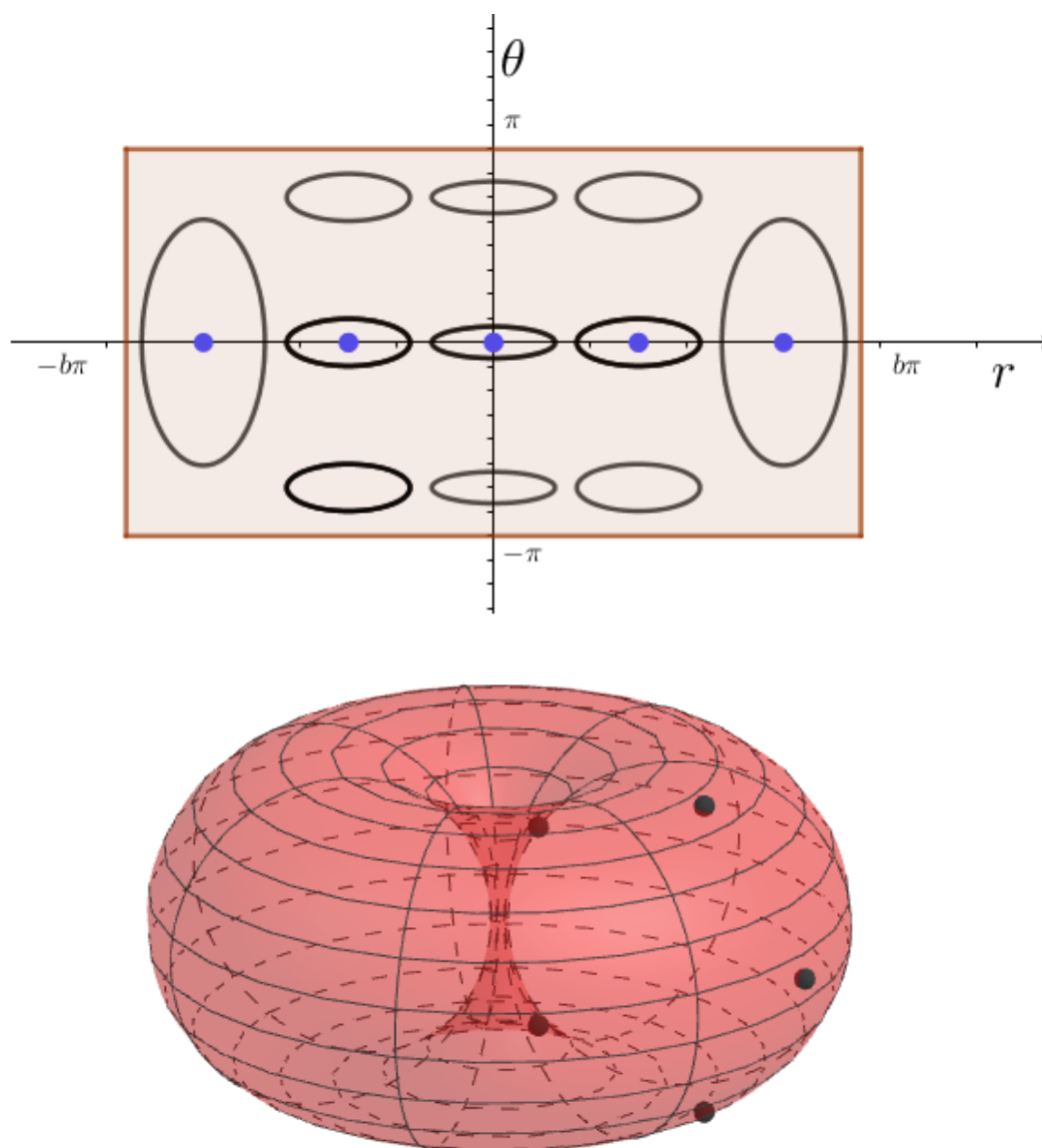


FIGURE 3.2 – Représentation des ellipses de la métrique pull-back sur le rectangle

**Définition 3.4** (Isométries). Soit  $(M, g)$  et  $(\tilde{M}, \tilde{g})$  deux variétés Riemanniennes. Un difféomorphisme  $\varphi : M \rightarrow \tilde{M}$  est appelé une isométrie si

$$\varphi^* \tilde{g} = g. \quad (3.7)$$

On dit que  $(M, g)$  et  $(\tilde{M}, \tilde{g})$  sont isométriques s'il existe une isométrie entre les deux.

*Remarque 3.1.* Si on se place dans l'ensemble des isométries entre une même variété. Il est facile de vérifier que la composition de deux isométries et que l'inverse d'une isométrie restent des isométries, celles-ci forment donc un groupe pour la composition.

[Lee97], Chap. 3

Paul Fraux, Manolis Perrot, Massil Hihat, Ivan Yakovlev  
 Sans souci

*Remarque 3.2.* La relation être isométrique étant une relation d'équivalence sur la classe des variétés Riemanniennes, la géométrie Riemannienne s'intéresse principalement aux propriétés conservés par isométries.

*Exemple 3.2.* Si on considère une application linéaire  $A$  de  $\mathbb{R}^n$  dans  $\mathbb{R}^n$  qui est une isométrie dans le sens ci-dessus : on a  $A^*g_{eucl} = g_{eucl}$ , ce qui est bien équivalent au fait que  $AA^T = I_2$ .

**Coordonnées locales :** Soit  $(E_1, \dots, E_n)$  une base locale<sup>1</sup> de TM, et  $(\varphi^1, \dots, \varphi^n)$  sa base duale. Alors une métrique Riemannienne peut s'écrire localement comme :

$$g = g_{ij}\varphi^i \otimes \varphi^j = \sum_{i,j} g_{ij}\varphi^i \otimes \varphi^j$$

où l'on a que  $\varphi^i \otimes \varphi^j(X, Y) = \varphi^i(X)\varphi^j(Y)$  (produit tensoriel).

Les coefficients de la matrice  $(g_{i,j})_{i,j} = \langle E_i, E_j \rangle_{i,j}$  sont symétriques en  $i$  et  $j$  et dépendent de manière lisse de  $p$ .

### 3.1.2 Exemples

Un premier exemple est  $\mathbb{R}^n$  muni de la métrique euclidienne  $g$ , qui est le simple produit scalaire  $(x, y) \mapsto x^T y$  avec l'identification  $T_x \mathbb{R}^n = \mathbb{R}^n$ . La matrice de  $g$  dans la base canonique est alors  $I_n = (\delta_{i,j})$ .

D'autres exemples de métrique Riemannienne arrivent naturellement quand on considère des sous-variétés, des produits ou des quotients de variétés.

**Définition 3.5** (Métrique sur une sous-variété). Soit  $(\tilde{M}, \tilde{g})$  une variété Riemannienne. Soit  $M$  une sous variété de  $\tilde{M}$ . Alors la restriction de  $\tilde{g}$  aux vecteurs tangents de  $M$  est une métrique Riemannienne car la restriction d'un produit scalaire à un sous espace vectoriel est encore un produit scalaire.

C'est ainsi qu'on a précédemment défini une métrique sur le tore, et c'est aussi de cette manière qu'on définit la métrique de  $S^n \subset \mathbb{R}^{n+1}$ .

Dans la pratique, on supposera souvent  $\tilde{M} \subset \mathbb{R}^m$  et la donnée d'un paramétrage local, i.e. un plongement<sup>2</sup> d'un ouvert  $U \subset \mathbb{R}^n$  dans  $\tilde{M}$  dont l'image est un ouvert de  $M$ ). En se donnant un tel paramétrage  $X : U \rightarrow \mathbb{R}^m$  d'une sous variété  $M \subset \mathbb{R}^m$  avec la métrique induite, on dispose de la formule ci-dessous pour calculer cette métrique dans la base canonique de  $U$  :

$$g = \sum_{i=1}^m \left( \sum_{j=1}^n \frac{\partial X^i}{\partial u^j} du^j \right)^2.$$

où  $du^i \cdot du^j$  est le produit tensoriel symétrisé, c.-à-d. :

$$du^i \cdot du^j = \frac{1}{2} (du^i \otimes du^j + du^j \otimes du^i).$$

*Exercice 3.1.* Retrouver directement à partir de cette formule la métrique induite sur le tore définie dans l'exemple 3.1.

1. i.e.  $((E_1(p), \dots, E_n(p)))$  est une base de  $T_p M$  pour  $p \in U$

2. en anglais embedding : application injective dont les différentielles sont injectives

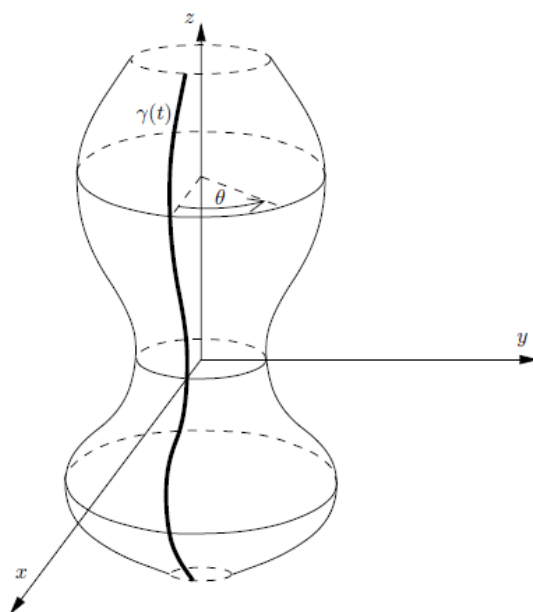


FIGURE 3.3 – Une surface de révolution

*Exercice 3.2.* Soit  $\gamma(t) = (a(t), b(t))$ ,  $t \in I$  (un intervalle ouvert) une courbe lisse injective dans le plan  $xz$ . On suppose que  $\forall t \in I, a(t) > 0$  et  $\dot{\gamma}(t) \neq 0$

Soit  $M \subset \mathbb{R}^3$  la surface obtenue par révolution de  $\gamma$  autour de l'axe des  $z$

- Montrer que  $M$  est une sous variété de  $\mathbb{R}^3$ , et que  $M$  est un plongement si  $\gamma$  en est un.
- Montrer que  $\varphi(\theta, t) = (a(t)\cos(\theta), a(t)\sin(\theta), b(t))$  est un paramétrage local de  $M$  au voisinage de tout point.
- Calculer l'expression de la métrique induite sur  $M$  sur le point de coordonnée  $(\theta, t)$ .
- appliquer ce calcul au cas du tore donné par  $(a(t), b(t)) = (2 + \cos(t), \sin(t))$ .

**Produit cartésien de variétés :** Soit  $(M_1, g_1)$  et  $(M_2, g_2)$  deux variétés Riemanniennes. Le produit  $M_1 \times M_2$  a une structure naturelle d'espace métrique  $g = g_1 \oplus g_2$  appelée métrique produit et définie par :

$$g(X_1 + X_2, Y_1 + Y_2) = g_1(X_1, Y_1) + g_2(X_2, Y_2), \quad (X_i, Y_i) \in T_{p_i} M_i$$

avec l'identification  $T_{p_1, p_2} M_1 \times M_2 = T_{p_1} M_1 \oplus T_{p_2} M_2$

Alors, toutes coordonnées locales  $(x^1, \dots, x^n)$  pour  $M_1$  et  $(x^{n+1}, \dots, x^{n+m})$  pour  $M_2$  donne les coordonnées  $(x^1, \dots, x^{n+m})$  pour  $M_1 \times M_2$ . Donc la matrice de  $(g_{ij})$  vérifie

$$(g_{ij}) = \begin{pmatrix} (g_1)_{ij} & 0 \\ 0 & (g_2)_{i,j} \end{pmatrix}.$$

Paul Froux,  
Manolis Perrot,  
Massil Hihat,  
Ivan Yakovlev  
Sans souci

**Quotient de variétés :** Soit  $\pi : \tilde{M} \rightarrow M$  un revêtement<sup>3</sup> (covering map) lisse.

On appelle transformation couvrante (covering transformation) une application lisse  $\varphi : \tilde{M} \rightarrow \tilde{M}$  telle que  $\pi = \pi \circ \varphi$ .

Si  $g$  est une métrique Riemannienne sur  $M$ , alors  $\tilde{g} := g \circ \pi$  est une métrique Riemannienne sur  $\tilde{M}$  invariante par transformation couvrante.

Dans ce cas, on appelle  $\tilde{g}$  la métrique couvrante (covering metric) et  $\pi$  un recouvrement Riemannien (a Riemannian covering).

## 3.2 Constructions élémentaires associées à une métrique Riemannienne

### 3.2.1 Monter ou baisser des indices

Une propriété élémentaire mais très importante d'une métrique Riemannienne est qu'elle permet de convertir des vecteurs en covecteurs et vice versa.

Soit  $g$  une métrique sur  $M$ , on définit l'application (dit "flat" ou "bémol") de  $TM$  dans  $T^*M$  comme suit : pour le vecteur  $X \in T_pM$  on définit le covecteur  $X^\flat \in T_p^*M$  par

$$X^\flat(Y) := g(X, Y)$$

On dit que  $X^\flat$  est obtenu de  $X$  en baissant un indice. En coordonnées :

$$X^\flat = g(X^i \partial_i, \cdot) = g_{ij} X^i dx^j$$

Donc la matrice de flat dans une base et sa base duale est celle de  $g$ . Cette matrice est inversible, ainsi cette opération est aussi inversible. Donc on peut définir l'opérateur inverse (dit "dièse" ou "sharp")  $\omega \rightarrow \omega^\sharp$ . En coordonnées,  $\omega^\sharp$  a les composantes :

$$\omega^\sharp := g^{ij} \omega_j,$$

où, par définition,  $g^{ij}$  sont les composantes de la matrice inverse  $(g_{ij})^{-1}$ .

L'application la plus importante de l'opérateur dièse est de permettre d'étendre la notion de gradient aux variétés Riemanniennes. Sans la métrique riemannienne (i.e. le produit scalaire sur chaque espace tangent), on n'a pas de façon naturelle (indépendant de la base) d'identifier  $T_pM$  avec  $T_p^*M$ .

Soit  $f$  une fonction lisse à valeurs réelles sur la variété Riemannienne  $(M, g)$ , on définit le champ de vecteur  $grad(f) := df^\sharp$ . On a alors en remontant les définitions que :

$$\forall Y \in TM, df(Y) = \langle grad(f), Y \rangle.$$

### 3.2.2 Élément de volume et integration

**Proposition :** Sur toute variété Riemannienne  $(M, g)$ , il existe une unique  $n$ -forme (c'est à dire un  $(n, 0)$ -tenseur alterné)  $dV$  satisfaisant la propriété que  $dV(E_1, \dots, E_N) = 1$  chaque fois que  $(E_1, \dots, E_N)$  est une base orthonormée pour quelque espace tangent  $T_pM$ .

<sup>3</sup> application continue et surjective telle que  $\forall p \in M, \exists U$  ouvert de  $M : \pi^{-1}(U) = \bigsqcup V_i$  et  $V_i$  est un ouvert de  $\tilde{M}$  homéomorphe à  $U$

Cette  $n$ -forme (noté aussi  $dV_g$ ) est appelé *l'élément de volume*. L'importance de l'élément de volume est qu'il permet d'intégrer des applications. Si  $f$  est une fonction lisse avec le support compact, alors  $f dV$  est une  $n$ -form avec le support compact. Alors l'intégral  $\int_M f dV$  est bien défini et est appelé l'intégral de  $f$  sur  $M$ . De même, le *volume* de  $M$  est défini comme  $\int_M dV = \int_M 1 dV$ .

### 3.3 Généralisation des métriques Riemanniennes

Il y a autres méthodes de mesurer les "longueurs" des vecteurs tangents. On obtient d'autres métriques en retirant des hypothèses dans la définition de la métrique Riemannienne.

#### 3.3.1 Métriques pseudo-Riemanniennes

C'est un champ  $g$  de 2-tenseur symétrique et non-dégénéré, c'est-à-dire que le seul vecteur orthogonal à tout l'espace est le vecteur nul :  $g(X, Y) = 0$  pour tout  $Y \in T_p M$  si et seulement si  $X = 0$ . Mais elle n'est pas nécessairement positive.

Un exemple notable d'une telle métrique est la métrique de Minkowski dans  $\mathbb{R}^4$  : c'est un invariant fondamental dans la théorie de la relativité restreinte d'Einstein. On peut l'exprimer succinctement en disant qu'en l'absence de gravité, les lois de la physique sont les mêmes dans tout système de coordonnées, dans lequel la métrique a une expression  $dx_1^2 + dx_2^2 + dx_3^2 - dt^2$ . La théorie de la relativité générale inclut les effets gravitationnels en permettant la métrique de varier point à point.

#### 3.3.2 Métrique de Finsler

Dans ce cas, sur chaque espace tangent on a une norme pas nécessairement engendrée par le produit scalaire. L'intérêt pour ces métriques est motivé par des applications dans la théorie de plusieurs variables complexes.

#### 3.3.3 Métrique sous-Riemannienne

Une métrique sous-riemannienne (également appelée métrique riemannienne singulière) est une métrique sur  $S$  — un champ de  $k$ -plan ou un sous-fibre de  $TM$ ). Cela signifie que les longueurs ne sont définies que pour les vecteurs de  $S$ , et que les courbes dont on peut mesurer les longueurs sont celles dont les vecteurs tangents sont tous en  $S$ .

L'une des motivations pour étudier les métriques sous-riemanniennes vient de la théorie du contrôle. Dans ce sujet, on a une variété avec un champ de vecteur dépendant de paramètres appelés les contrôles. L'objectif est de faire varier les contrôles pour obtenir une courbe de solution avec les propriétés souhaitées. Un bon exemple de restrictions infinitésimales est un problème de parking : la voiture a trois degrés de liberté, mais ne peut en contrôler que deux (roues et déplacement avant-arrière).

Paul Fraux,  
Manolis Perrot,  
Massil Hihat,  
Ivan Yakovlev  
Sans souci

## 3.4 Les espaces canoniques de la géométrie Riemannienne

### 3.4.1 L'espace Euclidien

L'espace Euclidien  $\mathbb{R}^n$  est l'exemple le plus simple de variété Riemannienne. On prend alors comme métrique  $\tilde{g}$  le produit scalaire usuel sur  $T_p\mathbb{R}^n$  (que l'on identifie à  $\mathbb{R}^n$  pour tout  $p$  de l'espace). Plus généralement, si  $V$  est un espace vectoriel de dimension  $n$  doté d'un produit scalaire, on peut y définir la métrique Riemannienne  $g(X, Y) = \langle X, Y \rangle$  sur  $T_pV = V$ . En choisissant une base orthonormée  $(E_1, \dots, E_n)$  de  $V$ , l'application qui à  $(x^1, \dots, x^n) \in \mathbb{R}^n$  associe  $X = x^i E_i \in V$  définit une isométrie de  $(\mathbb{R}^n, \tilde{g})$  dans  $(V, g)$ .

### 3.4.2 La sphère

La sphère  $S_R^n$  de  $\mathbb{R}^{n+1}$ , munie de la métrique  $g'$  induite par celle sur  $\mathbb{R}^{n+1}$ , est une autre variété Riemannienne canonique. La première propriété qui nous vient à l'esprit est la grande symétrie de la sphère.

On dit que la variété Riemannienne  $M$  est *homogène*, s'il existe un groupe de Lie opérant de manière lisse et transitivement par isométries sur  $M$ . Donc, pour chaque couple de points de  $M$  on peut trouver une isométrie qui envoie l'un sur l'autre.

Si  $p \in M$ , on dit que  $M$  est *isotrope* dans  $p$  s'il existe un groupe de Lie opérant par isométries transitivement sur l'ensemble des vecteurs unitaires de  $T_pM$ .

Plus simplement, une variété homogène a les mêmes propriétés géométriques à chaque point, et une variété isotrope a les mêmes propriétés à chaque direction.

La sphère est homogène et isotrope à chaque point. On peut le voir facilement : l'action du groupe orthogonal  $O(n+1)$  sur  $\mathbb{R}^{n+1}$  préserve la sphère, la métrique euclidienne, et opère transitivement sur les bases orthonormées sur la sphère.

Outre sa grande symétrie, la sphère est localement équivalente de manière conforme à l'espace euclidien.

**Définition :** deux métriques  $g$  et  $g'$  sur une variété  $M$  sont dites *conformes* l'une avec l'autre s'il existe une fonction positive  $f \in C^\infty(M)$  telle que  $g = fg'$ . Deux variétés Riemanniennes  $(M, g)$  et  $(M', g')$  sont dites *équivalentes de manière conforme* s'il existe un difféomorphisme  $\varphi : M \rightarrow M'$  tel que  $g' \circ \varphi$  soit conforme à  $g$ .

En fait, deux métriques sont conformes si et seulement si elles définissent les mêmes angles mais pas nécessairement les mêmes longueurs.

L'équivalence de manière conforme entre  $\mathbb{R}^n$  et la sphère  $S_R^n \subset \mathbb{R}^{n+1}$  privée d'un point est donnée par la projection stéréographique depuis le pôle Nord, ou par la projection Mercator.

#### La projection stéréographique :

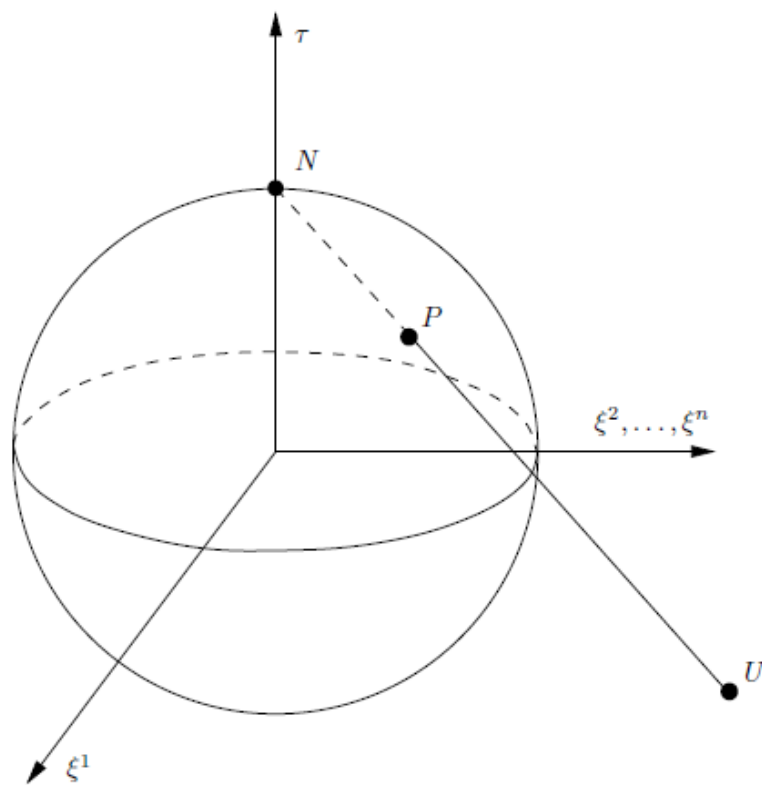
C'est l'application  $\sigma : S_R^n - \{N\} \rightarrow \mathbb{R}^n$  qui au point  $P = (\xi^1, \dots, \xi^n, \tau) \in S_R^n - \{N\}$  associe  $u$  dans  $\mathbb{R}^n$  où  $U = (u^1, \dots, u^n, 0) \in \mathbb{R}^{n+1}$  est le point d'intersection entre la droite (NP) et l'hyperplan  $\{\tau = 0\}$ .

Donc  $U$  est caractérisé par la relation  $\vec{N}U = \lambda \vec{N}P$  pour un  $\lambda$  non nul. On note  $N = (0, R)$ ,  $U = (u, 0)$  et  $P = (\xi, \tau) \in \mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$ .

On obtient que :

$$\sigma(\xi, \tau) = \frac{R\xi}{R - \tau}.$$



FIGURE 3.4 – Projection stéréographique depuis le pôle Nord :  $B = \sigma(A)$

Paul Fraux, On vérifie que  $\sigma$  est inversible, d'inverse :  
Manolis Perrot,  
Massil Hihat,  
Ivan Yakovlev  
Sans souci

$$\sigma^{-1}(u) = \left( \frac{2R^2 u}{|u|^2 + R^2}, R \frac{|u|^2 - R^2}{|u|^2 + R^2} \right)$$

On peut montrer que cette projection rendent bien équivalents la sphère et le plan. En fait, on peut interpréter cette propriété par le fait que la projection stéréographique conserve les angles entre des vecteurs tangents à la sphère et leurs images par  $\sigma$ .

De plus, les méridiens sont des droites passant par l'image du pôle Sud, et les parallèles sont des cercles.

**Quelques belles propriétés de la projection stéréographique** En fait, la projection stéréographique est la restriction de l'inversion de pôle  $N$  et de rayon approprié. Il est connu que l'inversion préserve les cercles (sphères, hypersphères). Donc les cercles sur la sphère projettent vers cercles. Notez que les droites sont les cercles de rayon infini. Alors, quelques exercices à ce sujet.

*Exercice.* Soit  $P$  un polyèdre qui est équivalent combinatoirement à un cube (c'est à dire il a 8 sommets, 6 faces quadrangulaire, à chaque sommet il y a trois faces). Montrer que si 7 de sommets se trouvent sur une sphère, le dernier sommet aussi y se trouve.

*Exercice* Montrer qu'il est impossible de construire le centre du cercle en utilisant uniquement la règle. (Indication : il existe une projection de le plan sur un autre plan, qui projete le cercle vers le cercle, mais ne projete pas le centre vers le centre).

**La projection de Mercator** La projection de Mercator consiste à enrouler un cylindre autour de la sphère, projeter les points de la sphère depuis le centre, puis dérouler le cylindre.

Comme pour le tore, on peut expliciter le mapping en coordonnées polaires :

$$M : \mathbb{R} \times [-\pi, \pi] \rightarrow \mathbb{R}^3, \\ (Z, \theta) \mapsto R(Z)(\cos \theta, \sin \theta, 0) + h(Z)(0, 0, 1) \quad (3.8)$$

$$\text{avec } R(Z) = \arccos(\arctan(Z/a)) \quad (3.9)$$

$$h(Z) = \arcsin(\arctan(Z/a)) \quad (3.10)$$

puis calculer la métrique pull-back associée à  $M^{-1}$  sur la carte Mercator. Cette projection est souvent employée par les marins, et pour les planisphères usuels. Elle ne conserve pas les distances (comparer la taille du Groenland et de l'Afrique), mais on peut se faire une idée de la déformation avec la figure suivante. Faites attention aux cercles à chaque point - ils ne sont pas déformés.

### 3.4.3 Plan hyperbolique

Un autre cadre canonique est le plan hyperbolique et ses généralisations. Son étude fera l'objet de la séance 6.

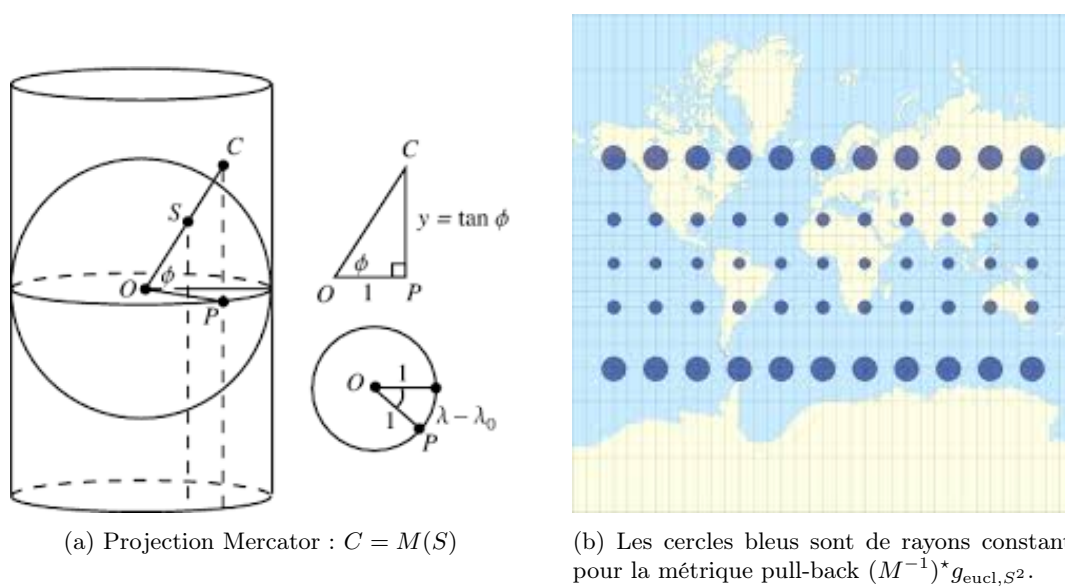


FIGURE 3.5 – Projection de Mercator et déformations induites

*Séance 3 46*  
*[Lee97], Chap. 3*

*Chapitre 3. Métriques Riemanniennes*

*Paul Faux,*  
*Manolis Perrot,*  
*Massil Hihat,*  
*Ivan Yakovlev*  
*Sans souci*

# Chapitre 4

## Connexions et courbes géodésiques

Séance 4

[Lee97], Chap. 4

Basile Coron,  
Yohan Mandin  
-Huble, Lucas  
Rey, Louis-  
Pierre Chaintron  
Technique

### 4.1 Dériver des champs de vecteurs ?

Intuitivement, une géodésique correspond à une courbe de  $M$  aussi droite que possible. Dans l'espace euclidien usuel  $\mathbb{R}^n$ , les géodésiques sont précisément les droites, qui se caractérisent comme les courbes d'accélération nulle. Considérons une variété plongée dans un espace euclidien : on pourrait de même envisager de calculer l'accélération d'une courbe  $\gamma(t)$  par dérivation double, puis projeter orthogonalement ce vecteur sur l'espace tangent à  $M$  en  $\gamma(t)$ . Dans une variété quelconque, cela ne marche pas : dériver dans un système de coordonnées quelconque donne un vecteur vitesse indépendant de ce système, mais il n'en va pas de même pour l'accélération.

*Exemple 4.1.* Considérons le cercle dans le plan. Il peut être paramétré par  $\gamma : t \mapsto (\cos(t), \sin(t))$ . Le champ des vecteurs vitesse obtenu est alors  $\dot{\gamma} : t \mapsto (-\sin(t), \cos(t))$  et en dérivant une fois de plus, on obtient  $\ddot{\gamma} : t \mapsto (-\cos(t), -\sin(t))$ .

En coordonnées polaires, on peut utiliser le paramétrage  $t \mapsto (1, t)$ , qui donne le champ de vecteurs vitesse  $t \mapsto (0, 1)$  : on retrouve bien ce qui précède. Néanmoins cela ne fonctionne plus pour l'accélération : dériver deux fois donne le vecteur nul.

Il manque donc une définition intrinsèque de l'accélération. Les vecteurs vitesses évoluant dans des fibres tangentes a priori très différentes, on ne peut définir l'accélération comme on l'a fait pour la vitesse (les positions appartenaient toutes au même espace). La première étape consiste à introduire un nouvel opérateur, la connexion, qui va permettre de relier les espaces tangents. Elle permet de définir la notion de dérivée covariante d'un champ de vecteurs selon une courbe, pour enfin arriver aux géodésiques et au transport parallèle.

### 4.2 Connexions affines

**Définition 4.1.** Soit  $M$  une variété différentielle. On note  $T(M)$  l'ensemble des champs de vecteurs tangents à  $M$ , i.e. des applications lisses  $X : M \rightarrow TM$  telles que

$$\forall q \in M, X(q) \in T_q M. \quad (4.1)$$

Pour  $q$  dans  $M$ ,  $X \in T(M)$  et  $f \in C^\infty(M, \mathbb{R})$ , la dérivée directionnelle de  $f$  en  $q$  selon  $X$  est donnée par :

$$(Xf)(q) = \left. \frac{d}{dt} \right|_{t=0} f(q + tX(q)) = \left. \frac{d}{dt} \right|_{t=0} f(\gamma(t)), \text{ avec } \gamma(0) = q, \dot{\gamma}(0) = X(q) \quad (4.2)$$

[Lee97], Chap. 4

Basile Coron,  
Yohan Mandin  
-Huble, Lucas  
Rey, Louis-  
Pierre Chaintron  
Technique

qui est une expression bien définie si  $X(q)$  est un vecteur de l'espace tangent. Une connexion affine sur  $M$  est une application  $\nabla : T(M) \times T(M) \rightarrow T(M)$  qui vérifie les propriétés suivantes :

<b>Linéarité à gauche</b>	Pour $f, g \in C^\infty(M, \mathbb{R})$ ,	$\nabla_{fX_1+gX_2}Y = f\nabla_{X_1}Y + g\nabla_{X_2}Y$
<b><math>\mathbb{R}</math>-linéarité à droite</b>	Pour $a, b \in \mathbb{R}$ ,	$\nabla_X(aY_1 + bY_2) = a\nabla_XY_1 + b\nabla_XY_2$
<b>Leibnitz-linéarité à droite</b>	Pour $f \in C^\infty(M, \mathbb{R})$ ,	$\nabla_X(fY) = f\nabla_XY + (Xf)Y$

*Remarque 4.1.* A priori, le comportement de  $\nabla_X(Y)$  en  $p$  dépend de celui de  $X$  et  $Y$  sur toute la variété, et pas seulement en  $p$ . C'est pourquoi la propriété de linéarité à gauche est globale et distincte d'une simple  $\mathbb{R}$ -linéarité.

*Exemple 4.2.* Sur  $\mathbb{R}^n$  (ou plus généralement sur tout ouvert  $U$  de  $\mathbb{R}^n$ , la connexion naturelle, appelée connexion euclidienne, est :

$$\nabla_X Y = \sum_{i=1}^n (XY^i) \partial_i = (XY^i) \partial_i \tag{4.3}$$

en utilisant la notation sommatoire d'Einstein, ce qu'on fera toujours par la suite. Définie à partir d'un système de coordonnées locales arbitraire,  $\partial_i$  est l'application qui à  $p \in M$  associe le  $i$ -ème vecteur d'une base de  $T_pM$  convenue. Dans  $\mathbb{R}^n$ , c'est l'application constante donnant le  $i$ -ième vecteur de la base canonique. Le cas euclidien permet de voir un champ de vecteurs comme une application  $f \in C^\infty(U, \mathbb{R}^n)$ . On a alors plus simplement

$$\nabla_X Y(p) = dY(p)(X(p)) \tag{4.4}$$

*Remarque 4.2.* On a vu au chapitre précédent qu'une structure métrique était *en gros* la donnée d'une base orthonormée  $(e_1(q), e_2(q), \dots, e_n(q))$  en chaque point  $q$  d'une carte. Il est donc raisonnable de penser que, pour comprendre les courbes géodésiques, il faille décrire la manière dont évolue un champs de bases (ou, ce qui est équivalent,  $n$  champs de vecteurs  $e_1(\cdot), \dots, e_n(\cdot)$ ) avec le point courant  $q$ .

De même qu'une *métrique riemannienne* est la notion infinitésimale qui correspond à la *longueur* d'une courbe, une *connexion* doit être vue comme la *différentielle* par rapport au point de base  $q$  du transport d'un vecteur  $Y$  le long d'un déplacement  $Xdt$  :

$$\ll (\nabla_X Y)dt \text{ quantifie la manière dont } Y \text{ varie le long de } Xdt \gg \tag{4.5}$$

Cette interprétation qualitative correspond intuitivement à l'expression 4.4 pour la connexion euclidienne : on s'intéresse à la façon dont varie un champ de vecteurs relativement à un autre. La *linéarité* en la première variable est donc tout naturelle : aller deux fois plus loin ( $X \rightarrow 2X$ ), c'est donner à  $Y$  deux fois plus d'espace pour varier. La "*Leibniz-linéarité*" en la seconde variable est alors une simple conséquence de la formule de dérivation d'un produit.

Dans notre manuel, l'étude des variations d'une base orthonormale avec la position est découpée en deux temps : les exemples explicites, les liens avec les propriétés géométriques ne seront abordées qu'à partir du chapitre suivant. Pour aujourd'hui, nous nous contenterons d'une approche *axiomatique*, proche de celle adoptée par un cours d'analyse fonctionnelle ou d'EDP mathématique : en nous contentant des axiomes de linéarité/Leibniz-linéarité ci-dessus, valables au premier ordre, nous montrerons des résultats forts sur la résolution d'équations différentielles liées aux géodésiques et aux problèmes de transport.

Les lemmes suivants montrent que l'opérateur  $\nabla$  est local : sa valeur en un point ne dépend que du comportement de  $X$  et  $Y$  dans un voisinage arbitrairement petit de ce point. Cela s'aperçoit dans l'exemple de la connexion euclidienne.

**Lemme 4.1.** Soit  $\nabla$  une connexion affine sur  $M$ ,  $p \in M$  ainsi que  $X_1, X_2, Y_1, Y_2 \in T(M)$ . Si  $X_1 = X_2$  et  $Y_1 = Y_2$  au voisinage de  $p$ , alors  $\nabla_{X_1} Y_1(p) = \nabla_{X_2} Y_2(p)$

*Démonstration.* On va prouver le résultat pour  $Y$ , le cas de  $X$  étant analogue. Par linéarité de  $\nabla$  en  $Y$ , on peut se ramener à  $Y$  nulle sur un voisinage  $U$  de  $p$ . Montrons alors que  $\nabla_X Y(p) = 0$ . On prend  $\varphi \in C^\infty(M, \mathbb{R})$  à support dans  $U$  et valant 1 en  $p$ .  $\varphi Y$  est alors identiquement nulle sur  $M$ , ce qui permet d'utiliser la linéarité en la seconde variable, car on s'est ramené à une propriété globale :

$$\nabla_X(\varphi Y) = \nabla_X(0 \cdot \varphi Y) = 0 \cdot \nabla_X(\varphi Y) = 0$$

Par Leibnitz-linéarité,

$$0 = \nabla_X(\varphi Y) = \varphi \nabla_X(Y)(p) + (X\varphi)Y$$

Évaluant chaque membre en  $p$ , on a bien  $\nabla_X(Y)(p) = 0$ . □

**Lemme 4.2.**  $\nabla_X Y(p)$  ne dépend que de  $X(p)$ , et du comportement de  $Y$  au voisinage de  $p$ .

*Démonstration.* Par linéarité, il suffit de montrer que  $X(p) = 0$  entraîne  $\nabla_X Y(p) = 0$ . Considérons un voisinage  $U$  de  $p$  sur lequel on dispose d'un système de coordonnées permettant la décomposition  $X = X^i \partial_i$ . Le lemme précédent assure que la valeur de  $\nabla_X Y$  en  $p$  ne dépend que du comportement de  $X$  et  $Y$ , ce qui permet d'utiliser localement les propriétés de la connexion :

$$\nabla_X Y(p) = \nabla_{X^i \partial_i} Y(p) = X^i(p) \nabla_{\partial_i} Y(p) = 0 \quad (4.6)$$

□

Soit maintenant  $U$  un ouvert de  $M$  muni d'un système de coordonnées locales. Chaque  $\nabla_{\partial_i} \partial_j$  se décompose dans la base  $(\partial_k)$ , d'où  $n^3$  fonctions  $\Gamma_{ij}^k \in C^\infty(M, \mathbb{R})$  telles que :

$$\nabla_{\partial_i} \partial_j = \Gamma_{ij}^k \partial_k \quad (4.7)$$

Ces fonctions sont appelés symboles de Christoffel de la connexion  $\nabla$ .

*Remarque 4.3.* L'espace tangent en tout point de  $\mathbb{R}^n$  est  $\mathbb{R}^n$  lui-même : on utilise systématiquement la base canonique. Par conséquent les applications  $\partial_i$  sont constantes, donc de différentielle nulle : la connexion euclidienne a tous ses symboles de Christoffel nuls. Le lemme suivant montre que l'action de  $\nabla$  sur  $U$  est entièrement déterminée par ses symboles de Christoffel.

**Lemme 4.3.** Soit  $\nabla$  une connexion affine sur  $M$ ,  $U$  un ouvert de  $M$ ,  $X = X^i \partial_i$ ,  $Y = Y^j \partial_j$  dans le système de coordonnées locales associé à  $U$ . On a :

$$\nabla_X Y = (XY^k + X^i Y^j \Gamma_{ij}^k) \partial_k \quad (4.8)$$

*Démonstration.* C'est un simple développement :

$$\nabla_X Y = \nabla_X(Y^j \partial_j) \quad \text{par décomposition dans la base canonique,} \quad (4.9)$$

$$= (XY^j) \partial_j + Y^j \nabla_{X^i \partial_i} \partial_j \quad \text{par Leibniz-linéarité en } Y, \quad (4.10)$$

$$= (XY^j) \partial_j + X^i Y^j \nabla_{\partial_i} \partial_j \quad \text{par linéarité en } X, \quad (4.11)$$

$$= (XY^j) \partial_j + X^i Y^j \Gamma_{ij}^k \partial_k \quad \text{par définition des } \Gamma_{ij}^k. \quad (4.12)$$

□

Basile Coron,  
Yohan Mandin  
-Huble, Lucas  
Rey, Louis-  
Pierre Chaintron  
Technique

Ce résultat est du type « opérateur algébrique déterminé par des paramètres scalaires ». C'est l'analogie pour les connexions (i.e. les applications linéaires à gauche, Leibniz-linéaires à droite) du célèbre résultat de prépa : « une application linéaire est entièrement déterminée par l'image de la base canonique ». Les symboles de Christoffel sont donc moralement les « coefficients » de la connexion : nous verrons au chapitre suivant comment on peut les relier aux coefficients de la métrique  $g$ . Nous allons maintenant voir comment construire des connexions sur une variété. Sans surprise, il suffira de donner explicitement les « coordonnées dans la base canonique », les fameux symboles de Christoffel : le lemme suivant montre comment construire toutes les connexions affines sur une variété couverte par un seul système de coordonnées (un ouvert de  $\mathbb{R}^n$  par exemple).

**Lemme 4.4.** Soit  $M$  une telle variété différentielle. La donnée d'une connexion affine sur  $M$  revient exactement à la donnée de  $n^3$  fonctions  $\Gamma_{ij}^k \in C^\infty(M, \mathbb{R})$  par la formule :

$$\nabla_X Y = (XY^k + X^i Y^j \Gamma_{ij}^k) \partial_k \quad (4.13)$$

*Démonstration.* À une connexion sur  $M$  on associe  $n^3$  telles fonctions, qui sont entièrement déterminées par la connexion. Réciproquement, avec  $n^3$  fonctions  $\Gamma_{ij}^k \in C^\infty(M, \mathbb{R})$  un calcul montre que la formule ci-dessus définit bien une connexion ayant pour coefficients de Christoffel les  $\Gamma_{ij}^k$ .  $\square$

En recollant les connexions affines obtenues sur les ouverts d'un revêtement de  $M$ , on obtient alors le résultat suivant :

*Propriété 4.1.* Sur toute variété il existe une connexion affine.

*Démonstration.* On recouvre  $M$  par des ouverts  $(U_\alpha)$  : sur chacun d'eux on a une connexion  $\nabla^\alpha$ . On se donne une partition de l'unité  $(\varphi_\alpha)$ , i.e. une famille de fonctions lisses telle que  $\varphi_\alpha$  soit à support dans  $U_\alpha$  pour tout  $\alpha$ , et  $\sum_\alpha \varphi_\alpha = 1$  sur  $M$ . Alors  $\nabla : (X, Y) \rightarrow \sum_\alpha \varphi_\alpha \nabla_X^\alpha Y$  est une connexion sur  $M$ .  $\square$

*Remarque 4.4.* Une combinaison linéaire de connexions n'en est pas nécessairement une, à cause de la Leibniz-dérivation.  $\nabla$  et  $\nabla'$  étant deux connexions, ni  $\frac{1}{2}\nabla$  ni  $\nabla + \nabla'$  n'en sont en général. Ce qui précède fonctionne grâce à la condition de normalisation  $\sum \varphi_\alpha = 1$ .

### 4.3 Dérivation d'un champ de vecteur sur une courbe paramétrée

Soit  $\gamma$  une courbe dans  $M$  paramétrée par un intervalle  $I$  de  $\mathbb{R}$ .

**Définition 4.2.** Un champ de vecteurs le long de la courbe  $\gamma$  est une application  $V$  de classe  $C^\infty$  de  $I$  dans  $TM$ , telle que  $V(t) \in T_{\gamma(t)}M$  pour tout  $t$  dans  $I$ . On note  $\mathfrak{X}(\gamma)$  l'espace des champs de vecteurs le long de  $\gamma$ .

Usuellement, un champ de vecteurs tangents est simplement une application  $V : M \rightarrow TM$ . Un tel champ peut être « restreint » en un champ de vecteurs le long de  $\gamma$  par  $t \mapsto V_{\gamma(t)}$ . On dit qu'un champ de vecteur le long d'une courbe est extensible s'il est la restriction d'un champ de vecteurs tangents.

*Exemple 4.3.* Les cas suivants seront utiles par la suite :

- $t \mapsto \dot{\gamma}(t)$  est un champ de vecteurs le long de  $\gamma$ .



- Les applications  $\partial_i : I \mapsto TM$  qui associent à  $t$  le  $i$ -ème vecteur d'une base fixée de  $T_{\gamma(t)}M$  sont des champs de vecteurs le long de  $\gamma$  extensibles, car restriction des champs  $\partial_i$  habituels.
- S'il existe  $t_0$  et  $t_1$  distincts tels que  $\gamma(t_0) = \gamma(t_1)$ , un champ de vecteurs  $f$  le long de  $\gamma$  tel que  $f(t_0) \neq f(t_1)$  n'est pas extensible, car un champ  $V : M \rightarrow TM$  ne peut envoyer un même point sur deux vecteurs tangents distincts.

Une connexion affine va alors permettre de définir une notion voisine de la « dérivation » le long d'une courbe.

**Lemme 4.5.** *Pour toute courbe  $\gamma$ , une connexion affine  $\nabla$  détermine un unique opérateur :*

$$D_t : \mathfrak{T}(\gamma) \rightarrow \mathfrak{T}(\gamma) \quad (4.14)$$

*vérifiant*

**$\mathbb{R}$ -linéarité**  $D_t$  est linéaire.

**Leibnitz-linéarité** Pour toute  $f \in C^\infty(I)$ ,  $D_t(fV) = \dot{f}V + fD_tV$

**Compatibilité avec  $\nabla$**  Pour  $V$  extensible d'extension  $\tilde{V}$ ,  $\forall t_0 \in I$ ,  $D_tV(t_0) = \nabla_{\dot{\gamma}(t_0)}(\tilde{V} \circ \gamma)(t_0)$ .

Un tel opérateur dépend seulement localement des valeurs de  $\gamma$ .  $\dot{\gamma}$  n'étant pas un élément de  $T(M)$ , i.e. un champ de vecteurs tangents défini sur  $M$ , la notation  $\nabla_{\dot{\gamma}(t_0)}$  peut paraître abusive. Néanmoins  $\nabla_X(Y)(p)$  ne dépend que de  $X(p)$ , et du comportement de  $Y$  au voisinage de  $p$  : la décomposition de  $\nabla$  par les symboles de Christoffel assure que tout ceci à un sens. Si l'espace tangent  $T_{\gamma(t_0)}M$  est identifié à  $M$ , on peut aussi voir  $\dot{\gamma}(t_0)$  comme un champ constant de  $M$  dans  $TM$ .

*Démonstration.* Fixons  $t_0$  dans  $I$  ainsi qu'un voisinage de  $\gamma(t_0)$  pour lequel on dispose d'une base locale de l'espace tangent en chaque point :  $\{E_i\}$  de sections associées  $\{\partial_i\}$ . Dans cette base un champ de vecteurs  $V$  étendu en  $\tilde{V}$  s'écrit :

$$\tilde{V} = \sum_j \tilde{V}^j \partial_j = \tilde{V}^j \partial_j \quad (4.15)$$

En utilisant (4.8), on montre l'unicité en calculant que

$$D_tV = \nabla_{\dot{\gamma}}(\tilde{V} \circ \gamma) = \nabla_{\dot{\gamma}}(\tilde{V} \circ \gamma)^i \partial_i \quad (4.16)$$

$$= (\dot{\gamma}(\tilde{V} \circ \gamma)^k + \dot{\gamma}^i(\tilde{V} \circ \gamma)^k \Gamma_{ij}^k) \partial_k \quad (4.17)$$

$$= \left( \dot{V}^k + V^j \dot{\gamma}^i (\Gamma_{ij}^k \circ \gamma) \right) \partial_k \quad (4.18)$$

Pour l'existence, on vérifie que l'opérateur défini par cette formule a les propriétés souhaitées.  $\square$

## 4.4 Géodésiques

Ayant construit une dérivation formelle sur les courbes, on peut revenir à notre objet initial : l'étude des géodésiques.

**Définition 4.3.** On appelle géodésique toute courbe  $\gamma$  telle que  $D_t\dot{\gamma}$  est nul, ce que l'on note souvent <sup>1</sup> :

$$\nabla_{\dot{\gamma}}\dot{\gamma} = 0 \quad (4.19)$$

1. Il s'agit d'un abus, dans la mesure où  $\dot{\gamma}$  n'est pas un champ  $M \rightarrow TM$

[Lee97], Chap. 4

Basile Coron,  
Yohan Mandin  
-Huble, Lucas  
Rey, Louis-  
Pierre Chaintron  
Technique

**Théorème 4.1** (Existence et Unicité des géodésiques). Soient  $p \in M$ ,  $X \in T_p M$ ,  $t_0 \in \mathbb{R}$ , il existe un intervalle ouvert  $I$  de  $\mathbb{R}$  contenant  $t_0$  et une géodésique  $\gamma : I \rightarrow M$  telle que  $\dot{\gamma}(t_0) = X$ . De plus deux telles géodésiques définies respectivement sur des intervalles  $I$  et  $I'$  contenant  $t_0$  sont égales sur leur domaine de définition commun  $(I \cap I')$ . On peut alors considérer une géodésique définie sur le plus grand intervalle possible, et parler de la géodésique passant par  $p$  et de vitesse initiale  $X$

*Démonstration.* Localement on dispose de la formule 4.18 pour un champ de vecteurs quelconque  $V$ . En remplaçant  $V$  par  $\dot{\gamma}$  la condition  $D_t \dot{\gamma} = 0$  donne l'équation différentielle d'ordre 2 en  $\gamma(t) = (\gamma^1(t), \dots, \gamma^n(t))$  :

$$\forall k, \ddot{\gamma}^k + \sum_i \sum_j \dot{\gamma}^i \dot{\gamma}^j \Gamma_{ij}^k(\gamma(t)) = 0. \quad (4.20)$$

Le théorème de Cauchy-Lipschitz donne immédiatement l'existence. Pour l'unicité, on prend deux géodésiques comme dans l'énoncé du théorème, définies sur  $I$  et  $I'$  contenant tous deux  $t_0$ . Par la partie unicité du théorème de Cauchy-Lipschitz elles sont égales sur un certain voisinage  $I''$  inclus dans  $(I \cap I')$ . On peut donc prendre  $\beta$  le sup des  $b$  tels que les deux solutions soient égales sur  $[t_0, \beta]$ . Si  $\beta$  n'est pas au bord de  $I$  ou  $I'$ , l'application de la partie unicité du théorème de Cauchy-Lipschitz en  $\beta$  contredit la maximalité de  $\beta$ .  $\square$

*Remarque 4.5.* Dans le cas de la connexion euclidienne, tous les symboles de Christoffel sont nuls et l'équation des géodésiques devient  $\ddot{\gamma} = 0$  : les géodésiques sont des droites, parcourues à vitesse constante (unitaire si on prend un paramétrage normal)

*Remarque 4.6.* L'équation (4.20) peut être vue comme une équation de *contrôle* d'ordre 2 sur le point courant  $\gamma(t)$  : la commande de vitesse  $\dot{\gamma}$  est ajustée en temps réel pour rester "constante" au sens de  $\nabla$ ,

$$\dot{\gamma}(t + dt) = \dot{\gamma}(t) + dt \cdot B_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t)) \quad (4.21)$$

où  $B_{\gamma(t)}$  est une forme bilinéaire. Le "problème" est ici que ce contrôle *quadratique* peut *exploser en temps fini* : j'en veux pour preuve le cas classique de l'équation différentielle réelle

$$\dot{y} = 1 + y^2, \quad (4.22)$$

dont les solutions  $t \mapsto \tan(t - t_0)$  ne peuvent être définies sur des intervalles de longueurs supérieures à  $\pi$ , ou de l'équation

$$\dot{y} = y^2, \quad (4.23)$$

dont les solutions  $t \mapsto -\frac{1}{t-t_0}$  divergent en temps fini dans une direction et s'écrasent dans l'autre.

*Remarque 4.7.* Le théorème de Cauchy-Lipschitz ne permet donc pas a priori de disposer de géodésiques définies pour tout temps. En anticipant sur les chapitres suivants et le lien entre géodésiques "au sens des connexions" et géodésiques "qui minimisent la longueur", il n'est pas difficile de produire des contre-exemples explicites.

Ainsi, sur la bande  $] -1, +1[ \times \mathbb{R}$  munie de la métrique euclidienne canonique  $g_{\text{eucl}}$ , on pourra montrer que la courbe  $\gamma : t \in ] -\sqrt{2}, +\sqrt{2}[ \mapsto (t/\sqrt{2}, t/\sqrt{2})$  est une géodésique (au sens de la connexion métrique associée) de vitesse unitaire qui ne peut être prolongée : les bords du domaine sont atteints!

Si l'on veut des géodésiques qui vont à l'infini "pour de vrai" (ce qui n'est pas un notion intrinsèque), il suffit de considérer le changement de coordonnées

$$\varphi : \begin{array}{l} \mathbb{R} \times \mathbb{R} \rightarrow ] -1, +1[ \times \mathbb{R} \\ (x, y) \mapsto (2 \arctan(x)/\pi, y) \end{array} \quad (4.24)$$

et de munir  $\mathbb{R} \times \mathbb{R}$  de la métrique pull-back  $\varphi^* g_{\text{eucl}}$ . La courbe  $\varphi^{-1} \circ \gamma : ]-\sqrt{2}, +\sqrt{2}[ \rightarrow \mathbb{R} \times \mathbb{R}$  est alors une géodésique pour la connexion associée, qui explose à l'infini aux deux bornes de l'intervalle de définition – on sait en effet que  $\varphi$  est une isométrie entre  $(\mathbb{R} \times \mathbb{R}, \varphi^* g_{\text{eucl}})$  et  $(]-1, +1[ \times \mathbb{R}, g_{\text{eucl}})$ .

## 4.5 Transport parallèle

On a vu que les connexions sont des opérateurs locaux au sens où l'on ne considère le transport d'un champ de vecteurs que selon des déplacements infinitésimaux (représentés par les vecteurs tangents “ $X$ ”). On peut tout à fait voir les choses de manière globale en considérant des déplacements selon non plus un vecteur mais une courbe tout entière : c'est la notion de *transport parallèle*. Soit  $M$  une variété munie d'une connexion  $\nabla$ .

**Définition 4.4** (Champ de vecteurs parallèle). On dit d'un champ de vecteurs le long d'une courbe  $\gamma$  qu'il est parallèle le long de  $\gamma$  si  $D_t V = 0$ .

On a alors le résultat suivant :

**Théorème 4.2** (Transport parallèle). Soit  $\gamma : I \rightarrow M$ ,  $t_0 \in I$ ,  $X \in T_{\gamma(t_0)}M$ . Il existe un unique champ de vecteurs  $V$  le long de  $\gamma$  parallèle le long de  $\gamma$  tel que  $V(t_0) = X$ .

*Démonstration.* En reprenant la formule 4.18, la condition de parallélisme de  $V$  s'écrit :

$$\forall k, \dot{V}^k(t) + \sum_i \sum_j V^j(t) \dot{\gamma}^i(t) \Gamma_{ij}^k(\gamma(t)) = 0. \quad (4.25)$$

Il s'agit maintenant d'une équation différentielle d'ordre 1 linéaire : le théorème de Cauchy-Lipschitz permet de conclure.  $\square$

*Remarque 4.8.* La linéarité de l'équation est cruciale pour obtenir l'existence du champ parallèle sur toute la courbe  $\gamma$ . On a en effet remplacé l'équation de contrôle quadratique (4.21) par le contrôle linéaire

$$V(t + dt) = V(t) + dt \cdot B_{\gamma(t)}(\dot{\gamma}(t), V(t)) \quad (4.26)$$

où  $q \mapsto B_q(\cdot, \cdot)$  est une application continue de  $M$  dans l'espace des formes bilinéaires muni de la norme d'opérateur – nous travaillons en dimension finie. Si  $\|B_q\|$  était uniformément borné, le très utile *lemme de Grönwall* permettrait alors d'assurer un résultat intuitivement crédible :

$$\|V(t)\| \leq \|V(0)\| e^{\|B_q\|_{\infty} t}, \quad (4.27)$$

et donc, par complétude de  $\mathbb{R}$ , la bonne définition de solutions sur  $\mathbb{R}$  tout entier – Cf. un grand classique de spé, le “théorème des bouts”.

Notez qu'en toute généralité,  $\|B_q\|$  n'est pas uniformément bornée... Mais comme elle est continue et que l'image par  $\gamma$  de tout intervalle  $[0, T]$  est compacte, on peut toujours prolonger “raisonnablement” l'intervalle de définition d'une solution de l'équation (4.25) et donc, en utilisant les arguments habituels, obtenir des prolongement à  $\mathbb{R}$  tout entier des solutions.

Soient alors  $\gamma$  une courbe dans  $M$  définie sur un intervalle  $I$  et  $t_0, t_1$  deux réels de  $I$ . On peut définir une fonction  $P_{t_0 t_1} : T_{\gamma(t_0)}M \rightarrow T_{\gamma(t_1)}M$  qui à un vecteur  $V_0$  associe le vecteur  $V_1$  défini par  $V_1 = V(t_1)$  avec  $V$  l'unique champ de vecteurs parallèle le long de  $\gamma$  tel que  $V(t_0) = V_0$ . On peut donc “transporter” un vecteur tangent le long de la courbe  $\gamma$ . Réciproquement, on peut revenir du global au local par un taux d'accroissement classique.

Séance 4 54  
[Lee97], Chap. 4

Chapitre 4. Connexions et courbes géodésiques

Basile Coron,  
Yohan Mandin  
-Huble, Lucas  
Rey, Louis-  
Pierre Chaintron  
Technique

# Chapitre 5

## Géodésiques Riemanniennes

Séance 5

[Lee97], Chap. 5

Nicolas Masson,  
Pascal Millet,  
Raphaël Barboni,  
Antoine Le Calvez  
Équilibré

### 5.1 Connexion associée à une métrique Riemannienne

On cherche à associer de manière naturelle une connexion à une variété  $M$  sur laquelle est définie une métrique Riemannienne. On demande pour cela deux propriétés (compatibilité avec la métrique et symétrie) qui sont naturelles dans le sens où elles déterminent une unique connexion qui coïncide avec celle induite par la connexion euclidienne si on plonge la variété  $M$  dans  $\mathbb{R}^n$  (ce qui est toujours possible d'après un théorème de Nash).

#### 5.1.1 Cas d'une sous-variété de $\mathbb{R}^n$

Dans ce cas, on peut naturellement induire une connexion  $\nabla^\perp$  sur  $M$  à partir de la connexion  $\nabla$  euclidienne. En effet, considérons deux champs de vecteurs  $X$  et  $Y$  sur  $M$  que l'on prolonge de manière  $C^\infty$  à  $\mathbb{R}^n$  en  $\tilde{X}$  et  $\tilde{Y}$ . On définit alors :

$$\nabla_X^\perp Y := \pi^\perp(\nabla_{\tilde{X}} \tilde{Y}) \quad (5.1)$$

Où  $\pi^\perp$  est la projection orthogonale sur l'espace tangent à  $M$ .

*Démonstration.* L'application est bien définie. En effet,  $\nabla_{\tilde{X}} \tilde{Y}(P)$  ne dépend que de  $\tilde{X}(P)$  et de  $\tilde{Y}$  le long d'une courbe passant par  $P$  et de vecteur tangent  $\tilde{X}(P)$  en  $P$ . Cela résulte de l'expression de  $\nabla_X Y$  sur un voisinage  $U$  de  $P$ , démontrée au chapitre précédent :

$$\nabla_{\tilde{X}} \tilde{Y} = (\tilde{X} \tilde{Y}^k + \tilde{X}^i \tilde{Y}^j \Gamma_{ij}^k) \partial_k \quad (5.2)$$

Grâce à cette remarque, on ne notera plus les  $\tilde{\phantom{x}}$  dans la suite. Le caractère  $C^\infty$  provient du fait que la projection  $M \rightarrow L(\mathbb{R}^n), P \mapsto \pi_P^\perp$  est  $C^\infty$ . Les deux premiers axiomes de linéarité sont évidents avec la linéarité de  $\pi^\perp$ . Soit  $f \in C^\infty(M)$ , on a  $f\tilde{Y}$  qui prolonge  $fY$ , donc on peut écrire :

$$\nabla_X^\perp fY = \pi^\perp((Xf)Y + f\nabla_X Y) \quad (5.3)$$

$$= (Xf)\pi^\perp Y + f\pi^\perp \nabla_X Y \quad \text{linéarité de } \pi^\perp \quad (5.4)$$

$$= (Xf)Y + f\nabla_X^\perp Y \quad \text{car } Y \text{ est déjà dans l'espace tangent} \quad (5.5)$$

□

Nicolas Masson,  
Pascal Millet,  
Raphaël Barboni,  
Antoine Le Calvez  
Équilibré

### 5.1.2 Application au tore

On va calculer les coefficients de Christoffel dans le système de coordonnées global sur le tore vu dans un chapitre précédent (le paramètre  $\theta$  désigne l'angle autour de l'axe de révolution et le paramètre  $\varphi$  permet de reperer le point sur le cercle qui engendre le tore par révolution autour de l'axe). De sorte qu'on a :

$$P(\theta, \varphi) = \begin{pmatrix} (b \cos \varphi + a) \cos \theta \\ (b \cos \varphi + a) \sin \theta \\ b \sin \varphi \end{pmatrix} \quad (5.6)$$

On utilise la définition des  $\partial_i$  pour calculer une base de l'espace tangent au point  $P(\theta, \varphi)$ . Comme le tore est plongé dans  $\mathbb{R}^3$ , ils peuvent être exprimés dans la base canonique :

$$\partial_\theta(\theta, \varphi) = \begin{pmatrix} -(b \cos \varphi + a) \sin \theta \\ +(b \cos \varphi + a) \cos \theta \\ 0 \end{pmatrix} \quad (5.7)$$

$$\partial_\varphi(\theta, \varphi) = b \begin{pmatrix} -\sin \varphi \cos \theta \\ -\sin \varphi \sin \theta \\ \cos \varphi \end{pmatrix} \quad (5.8)$$

Il est également utile de calculer les coordonnées du vecteur normal à la surface du tore en  $P(\theta, \varphi)$

$$n(\theta, \varphi) = \begin{pmatrix} \cos \varphi \cos \theta \\ \cos \varphi \sin \theta \\ \sin \varphi \end{pmatrix} \quad (5.9)$$

Pour avoir les symboles de Christoffel, il faut exprimer  $\nabla_{\partial_i}^\perp \partial_j$ . Pour cela on utilise la définition de la connexion tangentielle à partir de la connexion euclidienne  $\nabla$  (on notera également  $\partial_i$  les champs prolongés à l'espace tout entier). Dans le cas de la connexion euclidienne  $\nabla$  on a :  $\nabla_X Y(P) = dY_P(X(P))$ .

$$\nabla_{\partial_\theta} \partial_\varphi(P) = \frac{-b \sin \varphi}{a + b \cos \varphi} \partial_\theta \quad (5.10)$$

$$\nabla_{\partial_\varphi} \partial_\theta(P) = \frac{-b \sin \varphi}{a + b \cos \varphi} \partial_\theta \quad (5.11)$$

$$\nabla_{\partial_\theta} \partial_\theta(P) = (b \cos \varphi + a) \left( \frac{\sin \varphi}{b} \partial_\varphi - \cos \varphi n \right) \quad (5.12)$$

$$\nabla_{\partial_\varphi} \partial_\varphi(P) = -bn \quad (5.13)$$

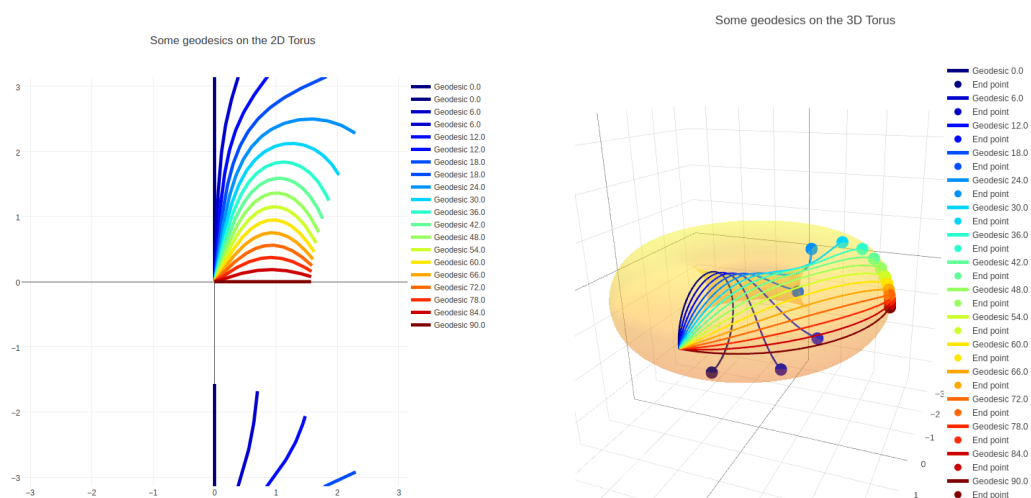
D'où on déduit avec la définition :

$$\Gamma_{\theta\varphi}^\theta(P) = \frac{-b \sin \varphi}{a + b \cos \varphi} \quad \text{et} \quad \Gamma_{\theta\varphi}^\varphi(P) = 0 \quad (5.14)$$

$$\Gamma_{\varphi\theta}^\theta(P) = \frac{-b \sin \varphi}{a + b \cos \varphi} \quad \text{et} \quad \Gamma_{\varphi\theta}^\varphi(P) = 0 \quad (5.15)$$

$$\Gamma_{\theta\theta}^\theta(P) = 0 \quad \text{et} \quad \Gamma_{\theta\theta}^\varphi(P) = (b \cos \varphi + a) \left( \frac{\sin \varphi}{b} \right) \quad (5.16)$$

$$\Gamma_{\varphi\varphi}^\theta(P) = 0 \quad \text{et} \quad \Gamma_{\varphi\varphi}^\varphi(P) = 0 \quad (5.17)$$



(a) Sur la carte 2D, on peut faire les calculs simplement – modulo les Christoffels. (b) En 3D, le résultat est bien conforme à nos attentes !

FIGURE 5.1 – Géodésiques de longueur  $3\pi/2$  tirées à partir du grand équateur, sur le tore de rayon médian 2 et de section 1. Les trajectoires sont intégrées par une simple méthode d'Euler, sur un domaine rectangulaire périodique  $(\theta, \varphi) \in S^1 \times S^1$ .

Équation géodésique :

$$\ddot{\gamma}^\theta + 2\dot{\gamma}^\theta \dot{\gamma}^\varphi \frac{-b \sin \varphi}{a + b \cos \varphi} = 0 \quad (5.18)$$

$$\ddot{\gamma}^\varphi + (\dot{\gamma}^\theta)^2 \frac{\sin \varphi}{b} (b \cos \varphi + a) = 0 \quad (5.19)$$

### 5.1.3 Compatibilité avec $g$

On aimerait se rapprocher du cas de la connexion euclidienne  $\nabla$  de  $\mathbb{R}^n$  qui vérifie :

$$\nabla_X \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle \quad (5.20)$$

Du fait de l'expression de la connexion euclidienne et de la bilinéarité du produit scalaire.

**Définition 5.1.** On dit qu'une connexion  $\nabla$  est compatible avec  $g$  si, pour tous champs de vecteurs  $X, Y, Z$  :

$$\nabla_X \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle \quad (5.21)$$

*Remarque 5.1.* Dans le cas d'un champ scalaire  $f$  on a :  $\nabla_X f = (Xf)$ .

En particulier  $\nabla_X \langle Y, Z \rangle = (X \langle Y, Z \rangle)$ , ce que l'on notera simplement  $X \langle Y, Z \rangle$ .

**Lemme 5.1.** Étant données une connexion  $\nabla$  et une métrique  $g$ , les propositions suivantes sont équivalentes :

1.  $\nabla$  est compatible avec  $g$
2. Pour toute courbe  $\gamma$  sur  $M$  extensible et pour tous champs de vecteurs  $U$  et  $V$  le long de  $\gamma$ ,  $D_t \langle X, Y \rangle = \langle D_t X, Y \rangle + \langle X, D_t Y \rangle$

[Lee97], Chap. 5

Nicolas Masson,  
 Pascal Millet,  
 Raphaël Barboni,  
 Antoine Le Calvez  
 Équilibré

3. Pour toute courbe  $\gamma$  sur  $M$ , si  $U$  et  $V$  sont deux champs de vecteurs parallèles le long de  $\gamma$ ,  $\langle U, V \rangle$  est constant.
4. Pour toute courbe  $\gamma$  et pour tout réels  $t_0$  et  $t_1$  de l'intervalle de définition de  $\gamma$ , le transport parallèle le long de  $\gamma$  de  $\gamma(t_0)$  à  $\gamma(t_1)$  est une isométrie. Autrement dit le transport parallèle conserve la norme des vecteurs.

*Démonstration.* Traitons les implications suivantes :

1)  $\Rightarrow$  2) Par définition de  $D_t$  et car  $\nabla$  est compatible avec  $g$ ,

$$D_t \langle X, Y \rangle = \nabla_{\dot{\gamma}} \langle X, Y \rangle \quad (5.22)$$

$$= \langle \nabla_{\dot{\gamma}} X, Y \rangle + \langle X, \nabla_{\dot{\gamma}} Y \rangle \quad (5.23)$$

$$= \langle D_t X, Y \rangle + \langle X, D_t Y \rangle \quad (5.24)$$

2)  $\Rightarrow$  3) Évident car si  $U$  et  $V$  sont parallèles le long de  $\gamma$ , par définition  $D_t U = D_t V = 0$ . Donc  $D_t \langle U, V \rangle = 0$  et  $\langle U, V \rangle$  est constant le long de  $\gamma$ .

3)  $\Rightarrow$  4) Cas particulier de 3) avec  $U = V$ .

4)  $\Rightarrow$  3) On utilise les identités de polarisation.

3)  $\Rightarrow$  1) On repasse par la définition de la dérivation directionnelle suivant le champ  $X$  au point  $P$  :

$$X \langle Y, Z \rangle (P) = \lim_{t \rightarrow 0} \frac{1}{t} [\langle Y, Z \rangle (\gamma(t)) - \langle Y, Z \rangle (\gamma(0))] \quad (5.25)$$

où  $\gamma$  est un chemin tel que  $\gamma(0) = P$  et  $\gamma'(0) = X(P)$ . On a supposé que transporter parallèlement les vecteurs le long de  $\gamma$  ne change pas le produit scalaire. On notera  $X(P)'_P$  le vecteur du champ  $X$  au point  $P'$  transporté parallèlement au point  $P$  le long de  $\gamma$ .

On a donc

$$X(P) \langle Y, Z \rangle = \lim_{t \rightarrow 0} \frac{1}{t} [\langle Y(\gamma(t))_P, Z(\gamma(t))_P \rangle - \langle Y(P), Z(P) \rangle] \quad (5.26)$$

On fait apparaître le produit intermédiaire  $\langle Y(\gamma(t))_P, Z(P) \rangle$ .

On utilise alors les limites démontrées au chapitre précédent (avec le fait que  $\gamma'(0) = X(P)$ ) :

$$\lim_{t \rightarrow 0} \frac{1}{t} [Y(\gamma(t))_P - Y_P] = \nabla_X Y(P) \quad (5.27)$$

$$\lim_{t \rightarrow 0} \frac{1}{t} [Z(\gamma(t))_P - Z_P] = \nabla_X Z(P) \quad (5.28)$$

On obtient l'égalité voulue avec la continuité du produit scalaire sur  $T_P M^2$  (vraie pour une application bilinéaire en dimension finie).  $\square$

*Remarque 5.2.* A priori l'unicité d'une connexion  $\nabla$  compatible avec la métrique  $g$  n'est pas assurée. On cherche donc dans la suite à renforcer les hypothèses pour avoir une connexion canonique associée à la métrique  $g$ .

### 5.1.4 Symétrie

**Définition 5.2.** On dit qu'une connexion est symétrique si elle vérifie :

$$\nabla_X Y - \nabla_Y X = [X, Y] \quad (5.29)$$

Où  $[X, Y]$  désigne un champ de vecteur appelé crochet de Lie de  $X$  et  $Y$ , qui quantifie le défaut de commutativité des flots de  $X$  et  $Y$  – c'est une notion abordée en cours de géométrie différentielle,



qui ne dépend que des champs de vecteurs et de la structure de variété. Nous ne l'utiliserons pas par la suite. Pour la démonstration de l'unicité de la connexion de Levi-Civita, il suffit de savoir que  $[X, Y]$  ne dépend pas de  $\nabla$ .

*Exemple 5.1.* La connexion tangentielle sur  $M$  sous-variété de  $\mathbb{R}^n$  est à la fois symétrique et compatible avec la métrique.

**Théorème 5.1** (Théorème fondamental de la géométrie riemannienne). *Soit  $(M, g)$  une variété riemannienne. Il existe une unique connexion linéaire compatible avec  $g$  et symétrique. Elle est appelée connexion de Levi-Civita.*

*Démonstration.* On montre ici l'unicité. Puisque  $\nabla$  respecte  $g$ , on a :

$$X\langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle \quad (5.30)$$

$$Y\langle Z, X \rangle = \langle \nabla_Y Z, X \rangle + \langle Z, \nabla_Y X \rangle \quad (5.31)$$

$$Z\langle X, Y \rangle = \langle \nabla_Z X, Y \rangle + \langle X, \nabla_Z Y \rangle \quad (5.32)$$

Par symétrie, on obtient :

$$X\langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_Z X \rangle + \langle Y, [X, Z] \rangle \quad (5.33)$$

$$Y\langle Z, X \rangle = \langle \nabla_Y Z, X \rangle + \langle Z, \nabla_X Y \rangle + \langle Z, [Y, X] \rangle \quad (5.34)$$

$$Z\langle X, Y \rangle = \langle \nabla_Z X, Y \rangle + \langle X, \nabla_Y Z \rangle + \langle X, [Z, Y] \rangle \quad (5.35)$$

En sommant les deux premières lignes, en soustrayant la troisième puis en réarrangeant les termes, il vient :

$$\langle \nabla_X Y, Z \rangle = \frac{1}{2} \left( X\langle Y, Z \rangle + Y\langle Z, X \rangle - Z\langle X, Y \rangle - \langle Y, [X, Z] \rangle + \langle X, [Z, Y] \rangle \right) \quad (5.36)$$

Cette expression ne dépend pas de  $\nabla$ . Donc, puisqu'un vecteur  $V$  est uniquement déterminé par la forme linéaire  $\langle V, \cdot \rangle$ , le champ de vecteur  $\nabla_X Y$  est uniquement déterminé par cette égalité.

Pour l'existence, on pose pour tout système de coordonnées locales les symboles de Christoffel (qui déterminent entièrement la connexion) :

$$\Gamma_{ij}^k = \frac{1}{2} g^{kl} (\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij}), \quad (5.37)$$

où les  $g^{kl}$  désignent les coefficients de la cométrique  $g^{-1}$ . Il faudrait alors démontrer que la connexion est bien définie, qu'elle vérifie la propriété de symétrie et la compatibilité avec  $g$  mais nous ne le ferons pas ici. L'important pour la suite est de retenir que l'on peut calculer les symboles de Christoffel de la connexion de Levi-Civita à partir des coefficients de la métrique.  $\square$

*Remarque 5.3.* La connexion tangentielle sur une sous-variété de  $\mathbb{R}^n$  est donc la connexion de Levi-Civita sur cette sous-variété.

## 5.2 Géodésiques et isométries

Dans ce cours, nous avons opté pour une définition locale des géodésiques. Au lieu de les voir comme des courbes minimisant la longueur entre deux points, on a préféré adopter un point de vue différentiel en les définissant comme les courbes "d'accélération nulle". Le lien entre point de vue métrique (intuitif), variationnel (utile) et différentiel (pratique) sera fait au chapitre 7. Pour l'instant, il faut donc vérifier explicitement que les géodésiques Riemanniennes "différentielles" se comportent bien vis à vis d'objets métriques définis précédemment : les isométries.

Nicolas Masson,  
 Pascal Millet,  
 Raphaël Barboni,  
 Antoine Le Calvez  
 Équilibré

**Proposition 5.1.** Soit  $\varphi : (M, g) \rightarrow (\tilde{M}, \tilde{g})$  une isométrie.

1.  $\varphi$  envoie la connexion Riemannienne  $\nabla$  de  $(M, g)$  sur la connexion Riemannienne  $\tilde{\nabla}$  de  $\tilde{g}$ , au sens où :

$$\varphi^*(\nabla_X Y) = \tilde{\nabla}_{d\varphi X}(d\varphi Y) \quad (5.38)$$

2. Si  $\gamma$  est une courbe et  $V$  un champ de vecteurs le long de  $\gamma$ , alors

$$\varphi^* D_t V = \tilde{D}_t(d\varphi V) \quad (5.39)$$

3.  $\varphi$  envoie une géodésique sur une géodésique : si  $\gamma$  est la géodésique de  $M$  partant de  $p$  avec la vitesse  $V$ , alors  $\varphi \circ \gamma$  est la géodésique de  $\tilde{M}$  partant de  $\varphi(p)$  avec la vitesse  $d\varphi V$

*Démonstration.* Pour le premier point, on définit  $\varphi^* \tilde{\nabla}$  par  $\varphi^* \tilde{\nabla}_X Y = d\varphi^{-1}(\tilde{\nabla}_{d\varphi X}(d\varphi Y))$ . On montre que  $\varphi^* \tilde{\nabla}$  est une connexion (appelée *connexion pullback*) symétrique et compatible avec  $g$  (par linéarité de  $d\varphi$ ), ce qui implique par unicité de la connexion Riemannienne que  $\varphi^* \tilde{\nabla} = \nabla$ . Pour le deuxième point, on définit un opérateur  $\varphi^* \tilde{D}_t : \mathfrak{T}(\gamma) \rightarrow \mathfrak{T}(\gamma)$  par une formule similaire, et on conclut par unicité. Enfin, le troisième point se déduit immédiatement du deuxième.  $\square$

### 5.3 Carte exponentielle

Comment travailler efficacement sur une variété? Étant donné un point de référence  $p$ , une idée forte sera de décrire un voisinage (aussi grand que possible!) de celui-ci en associant à “tout” vecteur  $V$  de l’espace tangent  $T_p M$  le point “d’arrivée” obtenu en avançant tout droit en direction de  $V$  pendant un temps 1. On a vu qu’à un vecteur tangent à  $M$  au point  $p$ ,  $V \in T_p M$ , on peut associer une unique géodésique maximale solution du problème de Cauchy :

$$(\text{Shoot}_{p,V}) : \begin{cases} \gamma(0) & = p \\ \dot{\gamma}(0) & = V \end{cases} \quad (5.40)$$

Reste à voir si cette géodésique n’explose pas aux bords de la variété avant le temps 1.

**Définition 5.3.** On définit le domaine de la carte exponentielle par :

$$\epsilon_p := \{V \in T_p M \mid \gamma_V \text{ est définie sur un intervalle contenant } [0, 1]\} \quad (5.41)$$

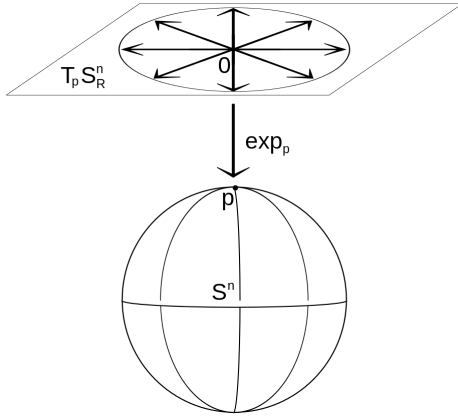
et la carte exponentielle par :

$$\exp_p : \begin{cases} \epsilon_p & \rightarrow M \\ V & \mapsto \gamma_V(1) \end{cases}. \quad (5.42)$$

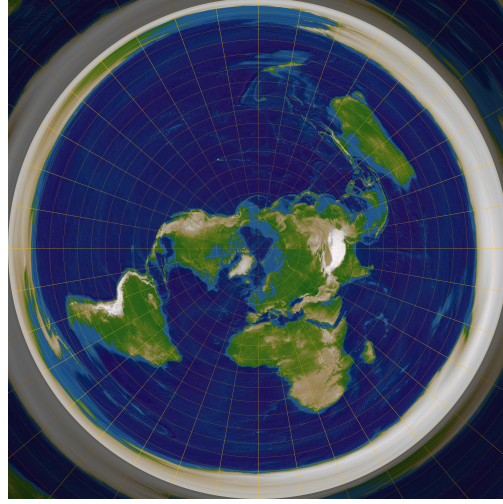
La carte exponentielle fait ainsi le lien entre l’espace vectoriel  $T_p M$  et la variété courbe  $M$  via les géodésiques.

*Exemple 5.2* (Exponentielle complexe). Si  $M = S^1$  muni de la métrique canonique, et qu’on prend pour point de base  $p = 1$ , on peut identifier  $T_p M$  avec la droite imaginaire  $i\mathbb{R}$ . Sans surprise, la carte exponentielle riemannienne agissant par “enroulements” est donnée par... l’exponentielle complexe!

*Exemple 5.3* (Projection azimutale équidistante). Nous démontrerons que les géodésiques sur la sphère sont bien celles que l’on attend à la fin du chapitre. En cartographie, la carte exponentielle a alors un nom : c’est la projection azimutale équidistante, qui est par exemple utilisée sur l’emblème des Nations Unies – voir Figure 5.2.



(a) Construction de la carte.



(b) Système de coordonnées radiales au pôle Nord. Notons qu'ici, la carte est définie sur le plan tout entier – même si la bijectivité n'est assurée qu'au voisinage du point de repère.

FIGURE 5.2 – Projection azimutale équidistante au pôle Nord.

**Lemme 5.2** (Lemme de changement d'échelle). *Pour tout  $V \in T_p M$  et  $c, t \in \mathbb{R}$ , lorsque cela a un sens*

$$\gamma_{cV}(t) = \gamma_V(ct) \quad (5.43)$$

*De plus lorsqu'un membre est défini, l'autre aussi (et on a égalité).*

*Autrement dit, on a équivalence entre "aller  $c$  fois plus vite" et "attendre  $c$  fois plus longtemps".*

*Démonstration.* On peut supposer que  $\gamma_V(ct)$  est défini (sinon, on s'y ramène en changeant  $c$  et  $V$ ). On a donc  $\gamma_V : I \rightarrow M$  avec  $ct \in I$ . On pose

$$\tilde{\gamma} : \begin{cases} J \rightarrow M \\ t \mapsto \gamma_V(ct) \end{cases} \quad (5.44)$$

où  $J = \{t \in I : ct \in I\}$ . Il suffit de démontrer que  $\tilde{\gamma}$  est une géodésique de vecteur dérivé en 0 égal à  $cV$  pour avoir l'existence du second membre et égalité. On repasse par les coordonnées locales. On note  $D_t$  et  $\tilde{D}_t$  les dérivations suivant  $\gamma$  et  $\tilde{\gamma}$ .

$$\tilde{D}_t \dot{\tilde{\gamma}}(t) = \left( \frac{d}{dt} \dot{\tilde{\gamma}}^k(t) + \Gamma_{ij}^k(\tilde{\gamma}(t)) \dot{\tilde{\gamma}}^i(t) \dot{\tilde{\gamma}}^j(t) \right) \partial_k \quad (5.45)$$

$$= (c^2 \dot{\gamma}^k(ct) + c^2 \Gamma_{ij}^k(\gamma(ct)) \dot{\gamma}^i(ct) \dot{\gamma}^j(ct)) \partial_k \quad (5.46)$$

$$= c^2 D_t \dot{\gamma}(ct) = 0 \quad (5.47)$$

Le point clé est donc que, tant pour l'accélération  $\ddot{\gamma}$  que pour le contrôle quadratique en  $\dot{\gamma}$ , dilater l'axe du temps par  $c$  fait sortir un facteur  $c^2$ .  $\square$

*Remarque 5.4.* La carte exponentielle est donc la carte des géodésiques partant d'un point donné de la variété. Si  $V \in TM$ ,  $\gamma_V$  est donné par :

$$\gamma_V(t) = \exp_p(tV) \quad (5.48)$$

Nicolas Masson,  
Pascal Millet,  
Raphaël Barboni,  
Antoine Le Calvez  
Équilibré

**Proposition 5.2.**  $\epsilon_p \subset T_p M$  est étoilé par rapport à 0.

*Démonstration.* C'est une conséquence du lemme précédent avec  $t = 1$  et  $c \in [0, 1]$  □

**Proposition 5.3.** La carte exponentielle est de classe  $C^\infty$ .

*Démonstration.* L'équation géodésique se traduit par un flot lisse dans l'espace des phases  $TM$  (les symboles de Christoffel sont lisses d'après l'expression en fonction des coefficients de la métrique). Le théorème d'intégration des flots (Cauchy-Lipschitz) garantit alors que l'évaluation au temps 1 dépend de manière  $C^\infty$  de la condition initiale, et donc en particulier de la vitesse initiale – le point de départ étant fixé à  $p$ . □

**Proposition 5.4.** Si  $\varphi : (M, g) \rightarrow (\tilde{M}, \tilde{g})$  est une isométrie, alors pour tout  $p \in M$  le diagramme suivant commute :

$$\begin{array}{ccc} T_p M & \xrightarrow{\text{d}\varphi} & T_{\varphi(p)} \tilde{M} \\ \exp_p \downarrow & & \downarrow \exp_{\varphi(p)} \\ M & \xrightarrow{\varphi} & \tilde{M} \end{array} \quad (5.49)$$

*Exercice 5.1.* Démontrer la propriété précédente.

## 5.4 Voisinages normaux et coordonnées normales

La carte exponentielle joue un rôle essentiel en géométrie Riemannienne dans la mesure où elle fournit un système de coordonnées locales particulièrement adapté : les coordonnées normales. Mais avant de pouvoir définir ces dernières, il nous faut élucider certains points à propos de la carte :

**Lemme 5.3.** Pour tout  $p \in M$ , il existe un voisinage  $\mathcal{V}$  de l'origine de  $T_p M$  et un voisinage  $\mathcal{U}$  de  $p$  dans  $M$  tels que  $\exp_p : \mathcal{V} \rightarrow \mathcal{U}$  est un difféomorphisme

*Démonstration.* Découle du théorème d'inversion locale. Il suffit de montrer que la différentielle de  $\exp_p$  est inversible en 0. On fait même mieux : on montre que c'est l'identité. □

Introduisons un peu de vocabulaire pour la suite : tout voisinage ouvert  $\mathcal{U}$  de  $p$  difféomorphe par  $\exp_p$  à un voisinage étoilé de 0 dans  $T_p M$  est appelé *voisinage normal de  $p$* . Si  $\epsilon > 0$  est tel que  $B_\epsilon(0) \subset T_p M$  est un *voisinage normal de  $p$*  (où le rayon est défini par rapport à la métrique  $g$ ), alors  $\exp_p(B_\epsilon(0))$  est appelée *boule géodésique*.

Désormais, grâce au lemme précédent, il nous est possible de définir un nouveau système de coordonnées locales. Si  $\{E_i\}$  est une base orthonormée de  $T_p M$ , on a un isomorphisme naturel  $E$  de  $\mathbb{R}^n$  dans  $T_p M$ . Ainsi, si  $\mathcal{U}$  est un voisinage normal de  $p$ , on dispose du paramétrage local :

$$\varphi := E^{-1} \circ \exp_p^{-1} : \mathcal{U} \rightarrow \mathbb{R}^n. \quad (5.50)$$

Tout paramétrage de ce type est appelé *système de coordonnées Riemanniennes normales* centré en  $p$ . Dans un tel système, il est naturel de définir la *distance radiale* :

$$r(x) := \left( \sum_i (x^i)^2 \right)^{1/2} \quad (5.51)$$

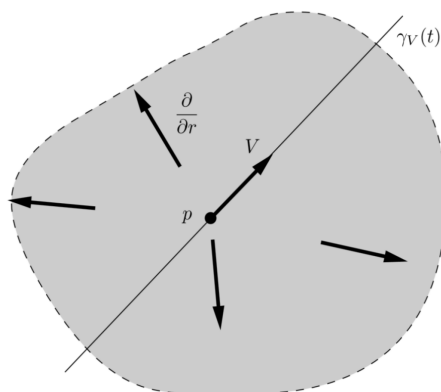


FIGURE 5.3 – Coordonnées Riemanniennes normales

ainsi que le *champ de vecteurs unitaires radial*  $\frac{\partial}{\partial r}$  :

$$\frac{\partial}{\partial r} := \frac{x^i}{r(x)} \frac{\partial}{\partial x^i} \quad (5.52)$$

Ces dernières définitions sont tout à fait intuitives dans un espace euclidien, et la proposition suivante permet de transposer cette intuition dans une variété Riemannienne quelconque :

**Proposition 5.5.** *Soit  $(\mathcal{U}, (x^i))$  un système de coordonnées Riemanniennes normales centré en  $p \in M$ . On a alors :*

1. Pour tout  $V = V^i \partial_i$ , la géodésique  $\gamma_V$  s'écrit en coordonnées normales :

$$\gamma_V(t) = (tV^1, \dots, tV^n) \quad (5.53)$$

tant que  $\gamma_V$  est dans  $\mathcal{U}$ .

2. Les coordonnées de  $p$  sont  $(0, \dots, 0)$ .
3. Les coefficients de la métrique en  $p$  pour ces coordonnées sont  $g_{ij} = \delta_{ij}$ .
4. Toute boule euclidienne  $\{x : r(x) < \varepsilon\}$  contenue dans  $\mathcal{U}$  est une boule géodésique.
5. En tout point  $q$  de  $\mathcal{U} - \{p\}$ ,  $\frac{\partial}{\partial r}$  est le vecteur vitesse de la géodésique de  $p$  vers  $q$ . Par conséquent,  $\frac{\partial}{\partial r}$  est unitaire pour  $g$ .
6. Les dérivées partielles premières des  $g_{ij}$  sont nulles en  $p$ , ainsi que les coefficients de Christoffel.

Les coordonnées normales étant un outil essentiel pour calculer en géométrie Riemannienne, ces propriétés doivent être comprises. De plus, comme évoqué précédemment, elles transposent l'intuition que l'on peut avoir dans un espace euclidien : une géodésique est une ligne droite dans un système de coordonnées normales. En revanche il est compliqué d'estimer précisément la distance entre deux points de la variété qui sont différents du point d'origine à l'aide d'un tel système.

Nicolas Masson,  
Pascal Millet,  
Raphaël Barboni,  
Antoine Le Calvez  
Équilibré

## 5.5 Géodésiques des espaces classiques

### 5.5.1 L'espace euclidien

Si on se place dans le système de coordonnées canonique, on voit par la formule donnant les coefficients de Christoffel en fonction de la métrique que ces coefficients sont tous nuls. Ainsi, l'équation différentielle des géodésiques est très simple, et montre que les géodésiques sont les droites affines auxquelles nous sommes habitués.

### 5.5.2 Les sphères

Sur la  $n$ -sphère de rayon  $R$ ,  $S_R^n \subset \mathbb{R}^{n+1}$  munie de la métrique induite, les calculs cessent d'être triviaux. En dimension 3, ils restent tout de même faisables : on peut paramétrer la sphère  $S_R^2$  en coordonnées sphériques :

$$(x, y, z) = (R \sin \varphi \cos \theta, R \sin \varphi \sin \theta, R \cos \varphi), \quad (5.54)$$

Reste alors à dérouler :

1. On montre que la métrique s'écrit en ces coordonnées :  $g_R = R^2 d\varphi^2 + R^2 \sin^2 \varphi d\theta^2$ .
2. On calcule les coefficients de Christoffel dans ces coordonnées.
3. En utilisant les équations des géodésiques, on vérifie que les méridiens  $(\theta(t), \varphi(t)) = (\theta_0, t)$  sont des géodésiques
4. On généralise ce constat à tous les grands cercles par homogénéité : les isométries de l'espace ambiant  $\mathbb{R}^{n+1}$  permettent d'envoyer un méridien sur n'importe quel grand cercle, et envoient une géodésique sur une géodésique d'après la proposition 5.1.
5. On conclut alors par unicité des solutions au problème de Cauchy que toutes les géodésiques sont des grands cercles.

Si les points 4 et 5 sont satisfaisants à tout point de vue, les trois premières étapes, calculatoires, laissent un petit goût d'inachevé... Il est en fait possible de s'en passer par un argument de symétrie : Soit une géodésique

$$\gamma(t) = (x^1(t), \dots, x^{n+1}(t)) \quad (5.55)$$

partant du pôle nord  $N$  dont la vitesse initiale  $V$  est un multiple de  $\partial_1$ . Il paraît intuitivement évident par symétrie que cette géodésique doit rester sur le méridien  $x^2 = \dots x^n = 0$  – voir Figure 5.4. Pour le montrer rigoureusement, supposons l'existence d'un temps  $t_0$  et d'un indice  $2 \leq i \leq n$  tels que  $x^i(t_0) \neq 0$ . On peut considérer la réflexion  $\varphi : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^{n+1}$  qui envoie  $x^i$  sur  $-x^i$  en conservant les autres directions : elle induit une isométrie de la sphère dans elle-même qui préserve  $N$  ainsi que  $V$ . Par conséquent, la proposition 5.1 associée à l'unicité des solutions au problème de Cauchy

$$(\text{Shoot}_{N,V}) : \begin{cases} \gamma(0) &= N \\ \dot{\gamma}(0) &= V \end{cases} \quad (5.56)$$

permet d'assurer que  $\varphi \circ \gamma = \gamma$ , ce qui est faux car par hypothèse  $\varphi(\gamma(t_0)) \neq \gamma(t_0)$ .

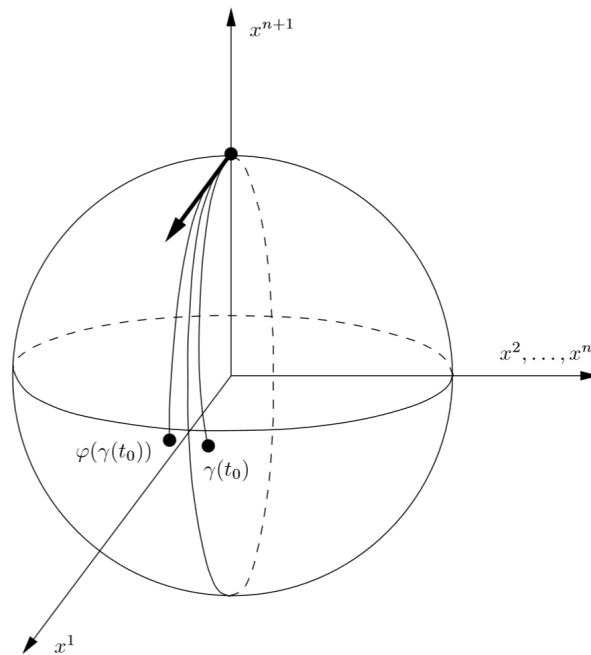


FIGURE 5.4 – Géodésiques sur la sphère : preuve par symétrie et homogénéité.

**Conclusion** Les géodésiques des espaces homogènes que sont le plan et la sphère sont donc particulièrement simples à décrire. Pour le tore, les surfaces arbitraires ou les espaces de formes et de landmarks, nous ne pourrons pas en dire autant... Avant d'attaquer ces problèmes difficiles (et riches!), nous consacrerons le chapitre qui vient à l'étude explicite des géodésiques du dernier espace homogène canonique : le plan hyperbolique, ou disque de Poincaré, introduit lors de la première séance. Ce sera l'occasion de réviser les notions clés introduites depuis le début du cours, et de nous forger une intuition des espaces à courbures négatives.

*Séance 5 66*  
*[Lee97], Chap. 5*

*Nicolas Masson,*  
*Pascal Millet,*  
*Raphaël Barboni,*  
*Antoine Le Calvez*  
*Équilibré*

*Chapitre 5. Géodésiques Riemanniennes*



# Chapitre 6

## Le plan hyperbolique

Séance 6

[CFK<sup>+</sup> 97]

Vadim Lebovici,  
Robin Lemaire,  
Valentin Carlier,  
Lino Benedetto  
Très joli

### Introduction

Depuis Euclide et ses 5 postulats de la géométrie, les mathématiciens ont longtemps cherché à démontrer le cinquième postulat, ou axiome des parallèles, à partir des quatre autres. Bien leur en prit, car en cherchant à démontrer l'absurdité d'une géométrie dans laquelle cet axiome n'est pas respecté, ils ont ouvert la voie aux géométries non-euclidiennes, dont la géométrie hyperbolique, qui nous intéresse ici. Elle peut être définie notamment en gardant les quatre premiers postulats d'Euclide, et en substituant au cinquième celui-ci : "Soient un point et une droite, alors il existe au moins une droite parallèle à ce point." On notera également que le mot "droite" est ici synonyme de "géodésique", soit une courbe à accélération nulle, ou encore un chemin minimisant la distance entre deux points.

On utilisera dans la suite cinq modèles analytiques de géométrie hyperbolique parmi les plus importants. Trois d'entre eux sont des modèles conformes dus à Henri Poincaré, qu'il imagina lors de son étude des automorphismes complexes.

### 6.1 Cas unidimensionnel

#### 6.1.1 Un premier modèle d'espace hyperbolique

Dans une vision classique, l'espace et le temps sont indépendants et, dans un espace  $n$ -dimensionnel, avec la donnée d'une  $n + 1$  coordonnée représentant le temps, la métrique Euclidienne suffit à décrire les phénomènes physiques. Mais en théorie de la relativité, le temps et l'espace sont liés, et la métrique Euclidienne ne rend plus compte des phénomènes physiques. On définit alors le modèle de Minkowski :  $\mathbb{R}^{n+1}$  munit de la "norme"  $x_1^2 + \dots + x_n^2 - x_{n+1}^2$ .

On appelle alors *cône de lumière* l'ensemble des points de norme 0. Pour un point  $(x_1, \dots, x_n, t)$  sur ce cône, la distance à l'origine "en espace" est égale au temps parcouru, ce qui exprime que la célérité de la lumière dans le vide est une constante.

Le produit scalaire associé à la norme définie précédemment est :  $x * y = x_1 y_1 + \dots + x_n y_n - x_{n+1} y_{n+1}$  ce qui permet de définir :

**Définition 6.1.** On définit l'espace hyperbolique de dimension  $n$  comme l'ensemble :

$$H^n = \{x \in \mathbb{R}^{n+1} : x * x = -1\}, \quad (6.1)$$

Vadim Lebovici,  
Robin Lemaire,  
Valentin Carlier,  
Lino Benedetto  
Très joli

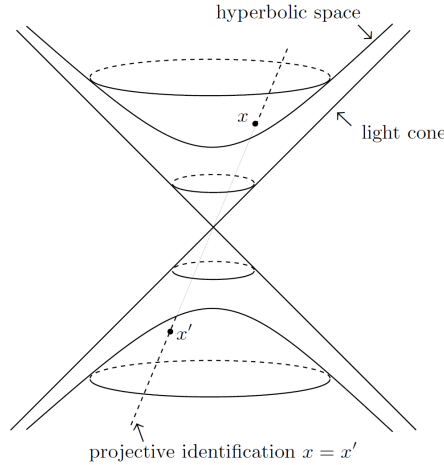


FIGURE 6.1 – Espace de Minkowski, cône de lumière et espace hyperbolique.

où l'on identifie les deux feuillet de l'hyperboloïde par quotient " $x = -x$ ". On peut le voir comme une sphère de  $\mathbb{R}^{n+1}$  de rayon au carré -1.

### 6.1.2 Cas $n = 1$

Comparons l'étude du cercle euclidien avec celle de l'espace  $H^1$ .

Si on paramètre le cercle unité dans l'espace euclidien par un chemin lisse  $p : ]-\infty, +\infty[ \rightarrow S^1$  avec  $p(0) = (0, 1)$ , on peut écrire  $p(t) = (x(t), y(t))$  et  $x^2 + y^2 = 1$ . Dans le plan hyperbolique si  $p_H : ]-\infty, +\infty[ \rightarrow H^1$  est une chemin lisse avec  $p_H = (1, 0)$ , on peut écrire  $p_H(t) = (x(t), y(t))$  et  $x^2 - y^2 = -1$ .

En différenciant ces équations on obtient :

$$\text{Cas euclidien : } p(t) \cdot p'(t) = 2x(t)x'(t) + 2y(t)y'(t) = 0, \quad (6.2)$$

$$\text{Cas hyperbolique : } p(t) * p'(t) = 2x(t)x'(t) - 2y(t)y'(t) = 0. \quad (6.3)$$

Ce qui exprime que le vecteur vitesse est orthogonal au vecteur position sur le cercle dans chacun des deux cas. On peut en particulier écrire :

$$\text{Cas euclidien : } p'(t) = k(t)(-y(t), x(t)), \quad (6.4)$$

$$\text{Cas hyperbolique : } p'(t) = k(t)(+y(t), x(t)), \quad (6.5)$$

et si on suppose de plus que  $|p'(t)| = 1$  alors  $k = \pm 1$ .

Dans un cas on parcourt alors le cercle unité dans le sens positif et on peut identifier  $t$  à la longueur d'arc parcourue depuis l'origine. On peut alors identifier :  $x(t) = \cos(t)$  et  $y(t) = \sin(t)$ .

Dans l'autre on parcourt la branche droite de l'hyperbole unité dans le plan de Minkowski a vitesse hyperbolique constante égale à 1. On peut alors identifier  $t$  à la longueur d'arc hyperbolique et  $x(t) = \sinh(t)$  et  $y(t) = \cosh(t)$ .

En conséquence directe des équations (6.4-6.5), on retrouve les formules classiques :

$$\sin' = \cos, \quad \cos' = -\sin, \quad \sinh' = \cosh \quad \cosh' = \sinh. \quad (6.6)$$

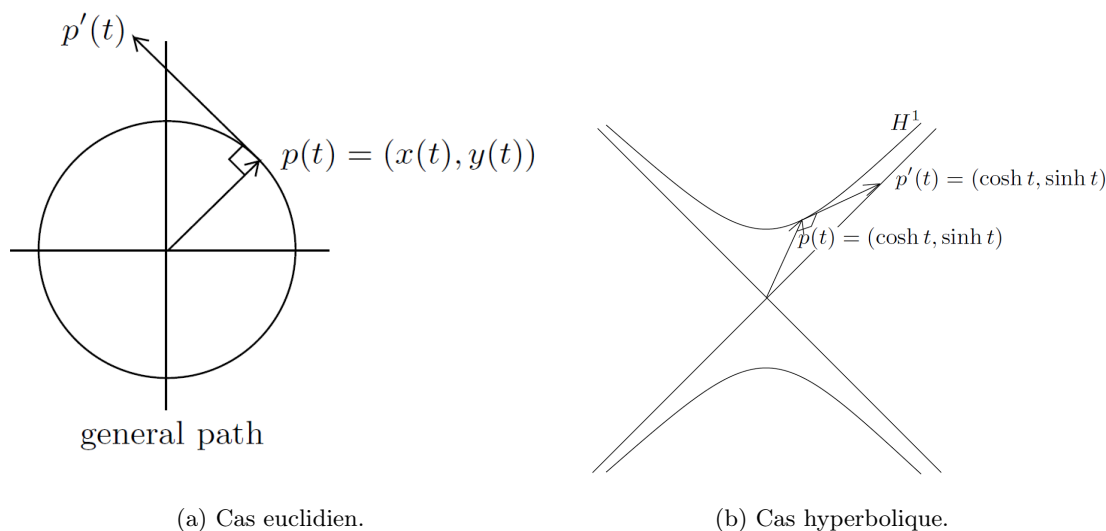


FIGURE 6.2 – Comparaison entre cercle et hyperbole, entre trigonométrie classique et trigonométrie hyperbolique.

## 6.2 Cinq modèles en dimension 2

Comme évoqué précédemment, il existe différents modèles de géométrie hyperboliques qui sont équivalents : Cette diversité sera un atout pour résoudre de nombreux problèmes.

### 6.2.1 Définitions

**Définition 6.2.** On définit les modèles suivants du plan hyperbolique :

$$\text{Le modèle du demi-espace : } H = \{(1, x_2, x_3) \mid x_3 > 0\} \quad (6.7)$$

$$\text{Muni de la métrique : } ds_H^2 = \frac{dx_2^2 + dx_3^2}{x_3^2} \quad (6.8)$$

$$\text{Le modèle du disque : } I = \{(x_1, x_2, 0) \mid x_1^2 + x_2^2 < 1\} \quad (6.9)$$

$$\text{Muni de la métrique : } ds_I^2 = 4 \frac{dx_1^2 + dx_2^2}{(1 - x_1^2 - x_2^2)^2} \quad (6.10)$$

$$\text{Le modèle de l'hémisphère : } J = \{(x_1, x_2, x_3) \mid x_1^2 + x_2^2 + x_3^2 = 1 \text{ et } x_3 > 0\} \quad (6.11)$$

$$\text{Muni de la métrique : } ds_J^2 = \frac{dx_1^2 + dx_2^2 + dx_3^2}{x_3^2} \quad (6.12)$$

$$\text{Le modèle de Klein : } K = \{(x_1, x_2, 1) \mid x_1^2 + x_2^2 < 1\} \quad (6.13)$$

$$\text{Muni de la métrique : } ds_K^2 = \frac{dx_1^2 + dx_2^2}{(1 - x_1^2 - x_2^2)^2} + \frac{(x_1 dx_1^2 + x_2 dx_2^2)}{(1 - x_1^2 - x_2^2)^2} \quad (6.14)$$

$$\text{Le modèle de l'hyperboloïde : } L = \{(x_1, x_2, x_3) \mid x_1^2 + x_2^2 - x_3^2 = 1 \text{ et } x_3 > 0\} \quad (6.15)$$

$$\text{Muni de la métrique : } ds_L^2 = dx_1^2 + dx_2^2 - dx_3^2 \quad (6.16)$$

Vadim Lebovici,  
Robin Lemaire,  
Valentin Carlier,  
Lino Benedetto  
Très joli

**Proposition 6.1.** *On constate immédiatement les propriétés d'invariance suivantes :*

- $ds_H^2$  est invariante par translation sur  $x_2, x_3$  ( mais pas  $x_1$  ),
- $ds_H^2$  est invariante par dilatation de l'espace  $((x_2, x_3) \mapsto (hx_2, hx_3)$  pour  $h > 0$ ).
- $ds_I^2$  et  $ds_J^2$  sont invariantes par rotations d'axe  $x_3$ .

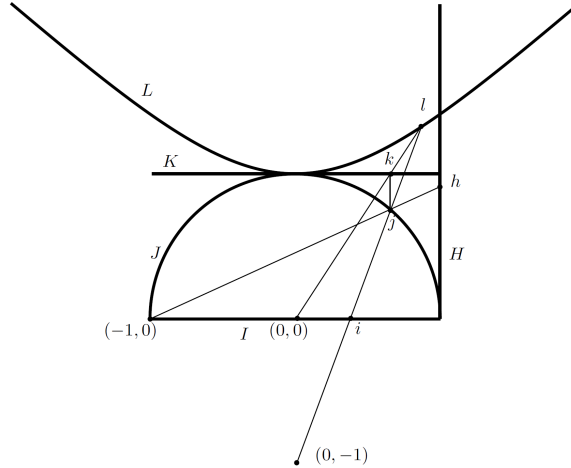


FIGURE 6.3 – Les cinq modèles du plan hyperbolique vus dans le plan  $(x_1, x_3)$ , et les projections associées qui sont des isométries.

**Théorème 6.1.** *Ces cinq espaces sont isométriques.*

*Démonstration.* Nous allons décrire des isométries entre ces espaces. Pour cela on utilisera  $J$  comme espace référent, et on trouvera des isométries de  $J$  vers les autres modèles ou inversement. On rédige la preuve pour  $H$  et  $I$ , qui sont nos espaces d'études privilégiés – les deux autres cas étant analogues.

Nos cinq modèles étant définis dans  $\mathbb{R}^3$ , on définira les fonctions comme suit :

$\alpha : J \rightarrow H$  est la projection stéréographique à partir du point  $(-1, 0, 0)$  sur  $H$  :

$$\alpha : J \rightarrow H, (x_1, x_2, x_3) \mapsto \left(1, \frac{2x_2}{x_1+1}, \frac{2x_3}{x_1+1}\right). \quad (6.17)$$

Comme projection,  $\alpha$  est bien à valeurs dans  $H$  et est définie sur tout  $J$ . De plus, on montre que la métrique  $ds_H^2$  est la métrique *pull back* de  $ds_J^2$  par  $\alpha$  :

Soient  $j = (x_1, x_2, x_3)$  et  $h = \alpha(x) = (1, y_2, y_3)$  (avec  $y_i = \frac{2x_i}{x_1+1}$ ). Alors en “notations physiciennes”,  $dy = d\alpha_x \cdot dx$ , i.e.

$$dy_i = \frac{2}{x_i+1} \left( dx_i - \frac{x_i}{x_1+1} dx_1 \right) \quad (6.18)$$

On sait de plus que  $x_1^2 + x_2^2 + x_3^2 = 1$ . En notation physiciennes, l'équation du plan tangent s'écrit donc

$$x_1 dx_1 = -(x_2 dx_2 + x_3 dx_3) \quad (6.19)$$

Ainsi, on a :

$$\alpha^*(ds_H^2) = \frac{dy_2^2 + dy_3^2}{y_3^2} \quad (6.20)$$

$$= \frac{(x_1 + 1)^2}{4x_3^2} \frac{4}{(x_1 + 1)^2} \left( \sum_{i=2}^3 dx_i^2 + \frac{2}{x_1 + 1} x_1 dx_1^2 + \frac{dx_1^2}{(x_1 + 1)^2} (1 - x_1^2) \right) \quad (6.21)$$

$$= \frac{1}{x_3^2} \sum_{i=1}^3 dx_i^2 \quad (6.22)$$

$$= ds_J^2 \quad (6.23)$$

$$(6.24)$$

Ce qui montre que  $H$  et  $J$  sont isométriquement équivalents.

De même, si  $\beta : J \rightarrow I$  ( $x_1, x_2, x_3 \mapsto (\frac{x_1}{x_3+1}, \frac{x_2}{x_3+1}, 0)$ ) est la projection centrale depuis le point  $(0, 0, -1)$  sur  $I$ ,  $\beta$  est bien définie de  $J$  dans  $I$ , et on montre que  $ds_I^2$  est la métrique *pull back* de  $ds_J^2$  par  $\beta$  :

$$\beta^*(ds_I^2) = ds_J^2. \quad (6.25)$$

Avec  $j = (x_1, x_2, x_3)$  et  $i = \beta(j) = (y_1, y_2, 0)$  respectivement dans  $J$  et  $I$ , on a

$$\forall i \in \llbracket 1, n \rrbracket, dy_i = \frac{1}{x_3 + 1} \left( dx_i - \frac{x_i}{x_3 + 1} dx_3 \right) \quad (6.26)$$

De plus, comme  $x_1^2 + x_2^2 + x_3^2 = 1$  on a  $\sum x_i dx_i = 0$ . On a alors :

$$\sum_{i=1}^2 dy_i^2 = \frac{1}{x_3 + 1} \left( \sum_{i=1}^2 dx_i^2 - 2 \sum_{i=1}^2 \frac{x_i}{x_3 + 1} dx_3 dx_i + \sum_{i=1}^2 \frac{x_i^2}{(x_3 + 1)^2} dx_3^2 \right) \quad (6.27)$$

$$= \frac{1}{x_3 + 1} \left( \sum_{i=1}^2 dx_i^2 + 2 \frac{x_3}{x_3 + 1} dx_3^2 + \frac{1 - x_3^2}{(x_3 + 1)^2} dx_3^2 \right) \quad (6.28)$$

$$= \frac{1}{(x_3 + 1)^2} \sum_{i=1}^3 dx_i^2 \quad (6.29)$$

De plus,  $1 - y_1^2 - y_2^2 = \frac{2x_3}{x_3+1}$ . D'où  $\beta^*(ds_I^2) = ds_J^2$ .

On utilise enfin  $\gamma : K \rightarrow J$  la projection verticale (par rapport à la 3-ème composante) et  $\delta : L \rightarrow J$  la projection stéréographique de centre  $(0, 0, -1)$  : des calculs similaires à celui pour  $\alpha$  montrent que ce sont des isométries pour les modèles munis de leur métrique riemannienne respective.  $\square$

Nous avons ainsi nos cinq modèles de géométrie hyperboliques, et les manières de passer des uns aux autres. On peut remarquer que deux des quatre isométries présentées sont des projections stéréographiques d'une sphère sur un plan.

## 6.2.2 Projections stéréographiques

**Définition 6.3** (Projection stéréographique). Soit  $\Sigma$  une sphère de dimension 2 dans l'espace euclidien  $\mathbb{R}^3$ , et  $P$  un plan. Si  $P$  est tangent à  $\Sigma$  en  $S$ , on note  $N$  le point opposé à  $S$  dans  $\Sigma$  (On verra ici  $S$  comme le pôle sud et  $N$  comme le pôle nord). Sinon, on note  $S$  un point d'intersection

Vadim Lebovici,  
Robin Lemaire,  
Valentin Carlier,  
Lino Benedetto  
Très joli

de la sphère et de la projection orthogonale de son centre sur  $P$ . Si  $x \in S \setminus N$ , alors on note  $\pi(x)$  l'intersection entre l'hyperplan  $P$  et la droite (euclidienne) reliant  $N$  et  $x$ . C'est la *projection stéréographique* de  $x$  sur  $P$ .

Par exemple, la fonction  $\alpha$  de  $J$  dans  $H$  est la restriction sur  $J$  de la projection stéréographique sur le plan  $x_1 = 1$ , dont l'image est  $H$ .

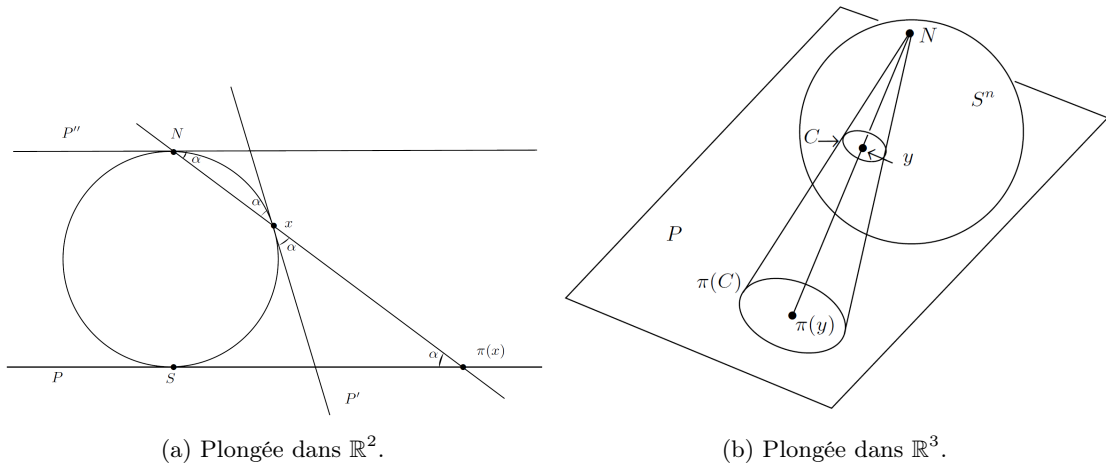


FIGURE 6.4 – Projection stéréographique.

**Théorème 6.2** (Conformité). *Si  $\Sigma$  est une sphère de  $\mathbb{R}^3$ ,  $P$  un plan et  $\pi$  la projection stéréographique de  $\Sigma$  sur  $P$ , alors  $\pi$  est une application conforme, c'est-à-dire qu'elle conserve les angles.*

**Théorème 6.3.** *Étant donnés  $\Sigma$ ,  $P$ ,  $\pi$  définis comme précédemment, et si  $C$  est un cercle inclus dans  $\Sigma$ , alors  $\pi(C)$  est un cercle généralisé de dimension  $c$  inclus dans  $P$  (c'est-à-dire que  $\pi(C)$  est un cercle ou une droite).*

*Démonstration.* Plutôt que d'écrire les preuves de ces résultats, nous regarderons le chapitre 9 du film *Dimensions* d'Etienne Ghys visionnable gratuitement sur Internet.  $\square$

### 6.3 Géodésiques

Pour déterminer la forme des géodésiques dans chacun des cinq modèles exposés précédemment, nous allons tout d'abord le faire dans le demi-espace  $H$  puis en déduire les résultats dans les autres modèles en utilisant les propriétés des projections. Le principal outil pour la détermination de la nature des géodésiques dans  $H$  est le *principe de rétraction* qui permet d'établir l'existence de géodésiques sous certaines conditions.

**Théorème 6.4** (Principe de rétraction). *Soient  $X$  une variété Riemannienne et  $C : ]a, b[ \rightarrow X$  un plongement d'intervalle de  $\mathbb{R}$ . Supposons de plus qu'il existe une application  $r : X \rightarrow \text{Im}(C)$  appelée rétraction qui réduit les distances au sens où : si on restreint la métrique sur  $X$  à  $\text{Im}(C)$  alors la métrique pullback par  $r$  définie sur tout  $X$  est inférieure ou égale à la métrique originale sur  $X$ .*

*Alors  $\text{Im}(C)$  contient une géodésique entre chaque paire de ses points. Autrement dit, on n'a jamais intérêt à sortir de  $\mathfrak{S}(C)$  pour rester "droit".*

*Démonstration.* Pour démontrer ce fait, il suffit de considérer un chemin arbitraire entre deux points de  $Im(C)$  et de constater que la rétraction de ce chemin par  $r$  est de longueur inférieure ou égale à la longueur du chemin initial.  $\square$

**Théorème 6.5** (Existence de géodésiques). *Dans le modèle  $H$  de l'espace hyperbolique, les droites verticales sont des géodésiques.*

*Démonstration.* Soit une droite verticale  $C : ]0, \infty[ \rightarrow H$  telle que  $C(t) = (1, x_2, t) \in H$  où  $x_2$  est fixé. On définit la rétraction  $r$  sur  $H$  par la formule :  $r(1, z_2, t) = (1, x_2, t)$  : c'est simplement la projection horizontale sur notre droite.

La métrique pullback est alors :  $\frac{dx_3^2}{x_3^2}$  et on a bien :  $\frac{dx_3^2}{x_3^2} \leq \frac{dx_2^2 + dx_3^2}{x_3^2}$ . D'où on déduit, par principe de rétraction, que l'image de  $C$  contient une géodésique entre chaque paire de ses points.

De plus si l'on considère deux points de l'image de  $C$  (ie sur la droite verticale) et n'importe quel chemin entre les deux, si ce chemin sort de l'image de  $C$  alors on a un des  $dx_i^2 > 0$  autour de ce point et le chemin est strictement plus long que le chemin restant dans  $C$ . De plus,  $Im(C)$  étant une droite verticale il existe une unique géodésiques reliant deux points de  $C$  restant dans l'image de  $C$ . Ce qui permet de conclure qu'il existe un unique plus court chemin entre deux points d'une droite verticale et que ce chemin est tracé le long de la droite.  $\square$

**Théorème 6.6** (Classification des géodésiques de  $H$ ). *Les géodésiques de  $H$  sont exactement les droites verticales et les demi-cercles d'extrémités appartenant à la limite de  $H$  :  $\partial H = \{(1, x_2, 0), x_2 \in \mathbb{R}\}$ .*

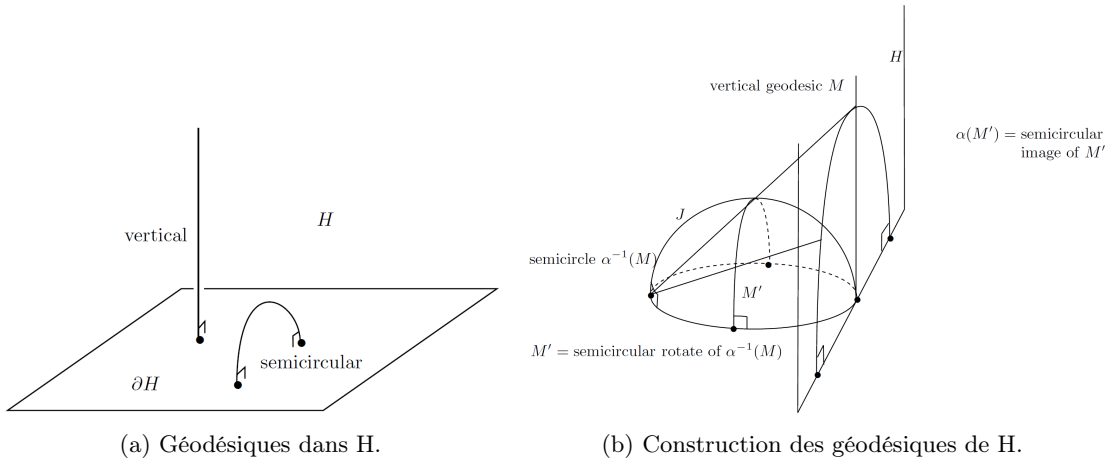


FIGURE 6.5 – Géodésiques de  $H$ .

**Lemme 6.1** (Isométries et demi-cercles).

1. *Les isométries euclidiennes de  $H$  qui envoient  $\partial H$  sur lui-même sont des isométries hyperboliques de  $H$ .*
2. *Les similitudes de  $H$  de la forme :  $(1, x_2, t) \mapsto (1, rx_1, rt)$  avec  $r > 0$  sont des isométries hyperboliques.*
3. *Les isométries euclidiennes de  $J$  sont des isométries hyperboliques de  $J$ .*
4. *Si  $p$  et  $q$  sont deux points de  $H$  qui ne sont pas sur la même droite verticale alors il existe un unique demi-cercle limite passant par  $p$  et  $q$ .*

Vadim Lebovici,  
Robin Lemaire,  
Valentin Carlier,  
Lino Benedetto  
Très joli

5. Deux demi-cercles limites sont les mêmes à une isométrie hyperbolique près.

*Démonstration.*

- **Les droites verticales sont les géodésiques de  $H$**  : Le théorème 6.5 et le (1) du lemme 6.1 nous assure que les droites verticales sont des géodésiques de  $H$ , toutes identiques à une isométrie hyperbolique près, et contenant chacun un plus court chemin entre chaque paire de leurs points.
- **Les demi-cercles orthogonaux-limites sont des géodésiques.** Considérons la droite verticale  $M$  de  $H$  passant par  $(1, 0, 0)$ .
  - **$M$  devient un grand cercle sur  $J$**  : son image réciproque par la projection stéréographique  $\alpha$  de  $J$  sur  $H$  par rapport au point de coordonnées  $(1, 0, 0)$  (voir Figure) est bien un grand cercle de  $J$ .
  - **La métrique sur  $J$  est invariante par rotation autour de l'axe passant par le pôle Nord** : cf proposition 6.1
  - **Le grand cercle de  $J$  parallèle à  $H$  est une géodésique** : En effet, c'est l'image d'une géodésique de  $H$  par deux isométries.
  - **Le projeté sur  $H$  de ce cercle est un demi-cercle orthogonal-limite** : En effet, la projection stéréographique de la sphère entière  $J'$  sur le plan entier  $H'$  envoie les grands cercles de  $J'$  sur les cercles de  $H'$  par théorèmes 6.2 et 6.3, ce qui assure que la corestriction de cette projection  $\alpha : J \rightarrow H$  envoie les demi-grands cercles sur les demi-cercles limites.
  - **Tout demi-cercle orthogonal-limite est une géodésique** : La métrique sur  $H$  étant invariante par translation et homothétie (cf (2) du lemme 6.1), tout demi-cercle orthogonal-limite est isométrique au demi-cercle construit dans cette preuve, qui est une géodésique, d'où l'on conclut.
- **On conclut par théorème de Cauchy.** A une couple  $(\vec{x}, \vec{v})$  où  $\vec{x}$  indique la position et  $\vec{v}$  la vitesse en un point dans  $H$ , on a établi l'existence d'une géodésique droite ou d'une géodésique semi-circulaire passant pas  $\vec{x}$  et tangent à  $\vec{v}$  en  $\vec{x}$ . Par théorème de Cauchy, il n'y en a pas d'autres.

□

De ce résultat fondamental on déduit, grâce aux propriétés sur la projection stéréographique, les géodésiques des quatre autres modèles, grâce aux projections exhibées précédemment entre ces modèles.

**Théorème 6.7** (Caractérisation des géodésiques des modèles standards du plan hyperbolique).

- Les géodésiques de  $J$  sont les demi-cercles limites.
- Les géodésiques de  $I$  sont les diamètres et les segments circulaires limites.
- Les géodésiques de  $K$  sont les segments droits.
- Les géodésiques de  $L$  sont les intersections de  $L$  avec des sous-espaces vectoriels de  $\mathbb{R}^{n+1}$  de dimension 2.

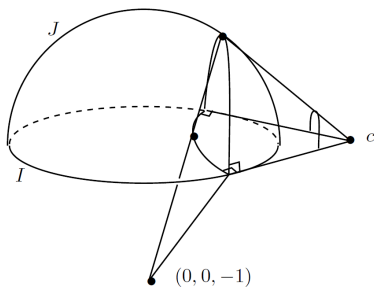
*Démonstration.* On procède comme suit :

- Par théorèmes 6.1 et 6.2 les demi-cercles limites de  $J$  sont isométriques aux demi-cercles orthogonaux et aux droites verticales de  $H$ . Les géodésiques de  $J$  sont donc les demi-cercles limites.
- En outre, les demi-cercles limites de  $J$  sont isométriques aux diamètres et aux segments circulaires limites de  $I$ , qui sont donc les géodésiques de  $I$ .

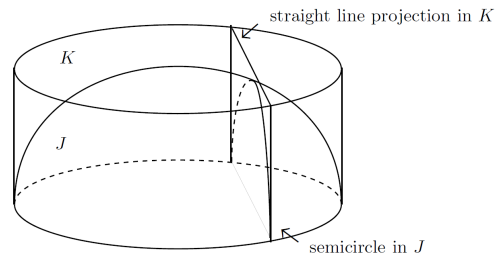


- Les demi-cercles de  $J$  sont isométriques aux segments droits de  $K$  via la projection verticale, qui sont donc les géodésiques de  $K$ .
- Les segments droits de  $K$  sont isométriques via la projection centrale aux intersections de  $L$  avec des sous-espaces vectoriels de  $\mathbb{R}^3$  de dimension 2. Ces intersections sont donc les géodésiques de  $L$ .

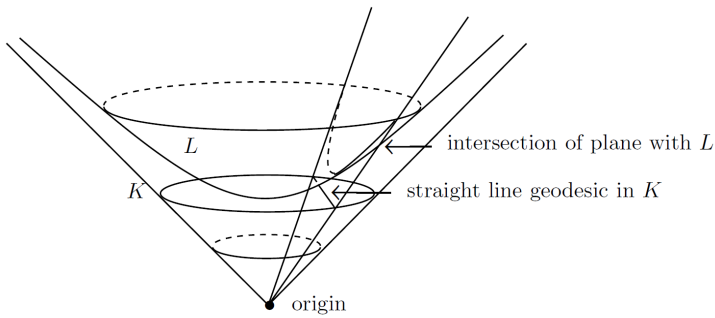
□



(a) Projection des géodésiques de  $J$  sur celles de  $I$ .



(b) Projection des géodésiques de  $J$  sur celles de  $K$ .



(c) Projection des géodésiques de  $K$  sur celles de  $L$ .

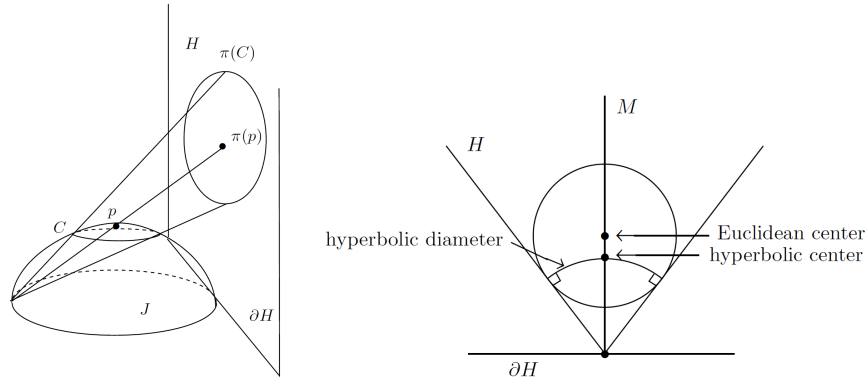
FIGURE 6.6 – Liens entre les géodésiques des différents modèles.

Vadim Lebovici,  
Robin Lemaire,  
Valentin Carlier,  
Lino Benedetto  
Très joli

## 6.4 Propriétés diverses

On se place dans les trois modèles conformes  $H$ ,  $I$ ,  $J$  définis plus haut.

*Propriété 6.1.* Dans les modèles étudiés, les sphères hyperboliques sont des sphères euclidiennes, mais le centre hyperbolique et le centre euclidien ne coïncident pas.



(a) La projection d'un cercle de  $J$  sur  $H$  est un cercle euclidien et hyperbolique.

(b) Le centre hyperbolique d'un cercle.

FIGURE 6.7 – Lien entre cercles euclidiens et cercles hyperboliques.

Un cercle hyperbolique centré en  $p = (0, 0, 1) \in J$  est clairement un cercle euclidien, dont la projection sur  $H$  reste un cercle euclidien et hyperbolique. On peut ensuite translater ce cercle dans  $H$  et le projeter de nouveau dans  $J$ , puis lui faire subir une éventuelle rotation. Toutes ces opérations laissant invariants à la fois les cercles euclidiens et les cercles hyperboliques, la propriété est vérifiée.

*Propriété 6.2.* Sur le plan hyperbolique, si  $\mathcal{D}$  est une droite (une géodésique), alors les courbes à distance  $r$  de chaque côté de  $\mathcal{D}$  ne sont pas des droites (des géodésiques).

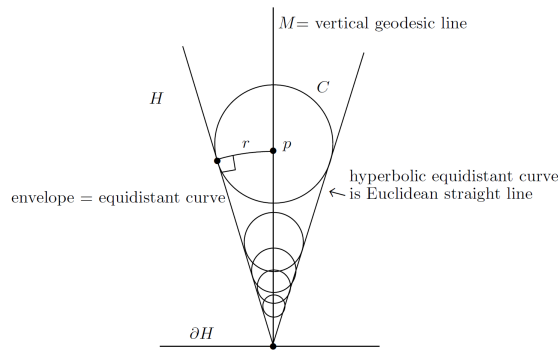


FIGURE 6.8 – Courbes équidistantes dans  $H$ .

Comme on peut voir sur la Figure 6.8, les courbes en question sont des droites euclidiennes non verticales, et non des demi-cercles “géodésiques”.

**Propriété 6.3.** La somme des angles d'un triangle du plan hyperbolique est inférieure à  $\pi$ . L'aire d'un triangle d'angles  $\alpha, \beta, \gamma$  est  $\pi - \alpha - \beta - \gamma$  (théorème de Gauss-Bonnet). Si  $\alpha, \beta, \gamma$  sont des angles de somme inférieure à  $\pi$ , alors il existe un unique triangle possédant ces angles à congruence près.

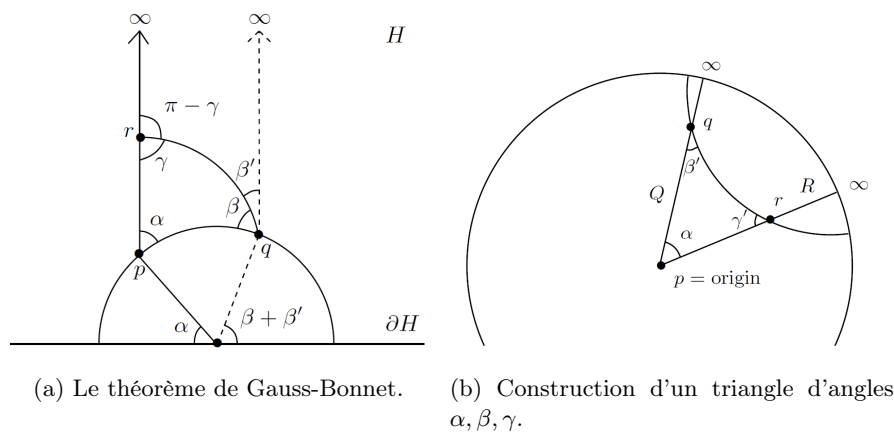


FIGURE 6.9 – La somme des angles d'un triangle hyperbolique est inférieure à  $\pi$ .

Si  $\delta = pqr$  est un triangle, on peut par isométries se ramener au cas où  $p, q$  appartiennent au cercle unité et le côté  $pr$  est vertical. Alors on utilise le fait que  $pqr = pq\infty - rq\infty$ , et on obtient facilement l'aire des triangles idéaux (ayant au moins un point à l'infini).

**Propriété 6.4.** Si  $\delta = pqr$  est un triangle hyperbolique, et  $x \in pq$ , alors il existe  $y \in pr \cup qr$  tel que pour la distance hyperbolique,  $d(x, y) \leq \ln(1 + \sqrt{2})$ . On dit que les triangles de l'espace hyperboliques sont uniformément fins.

On se place dans le pire cas. En effet comme nous avons vu, on peut supposer que  $\delta$  a un côté sur le cercle unité et un côté vertical, puis on peut déplacer  $r$  à l'infini, et  $p$  et  $q$  sur la bordure  $\partial H$ . Toutes ces opérations ne font qu'augmenter le triangle. On se place ensuite dans le cas où  $x$  est à égale distance de  $pr$  et  $qr$ . Alors  $d(x, pr) = \int_{\pi/4}^{\pi/2} \frac{d\psi}{\sin \psi} = \ln(1 + \sqrt{2})$

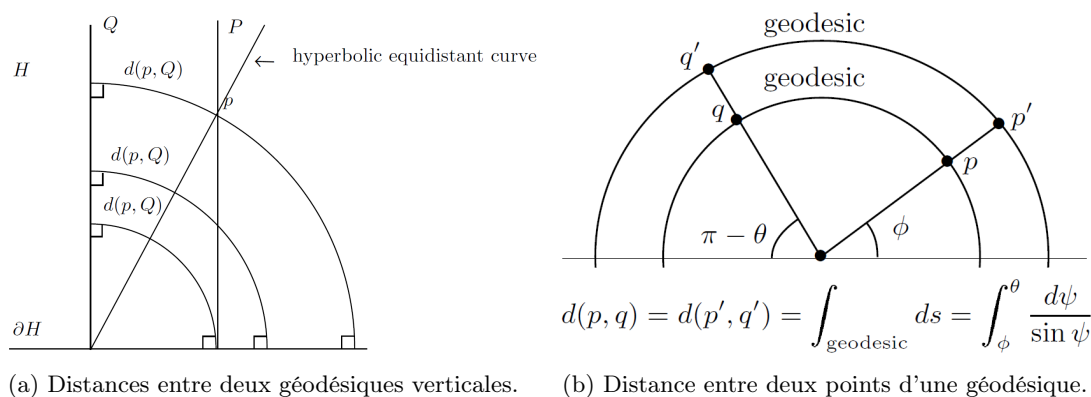


FIGURE 6.10 – Envoyer un sommet à l'infini ne peut qu'épaissir le triangle.

Vadim Lebovici,  
Robin Lemaire,  
Valentin Carlier,  
Lino Benedetto  
Très joli

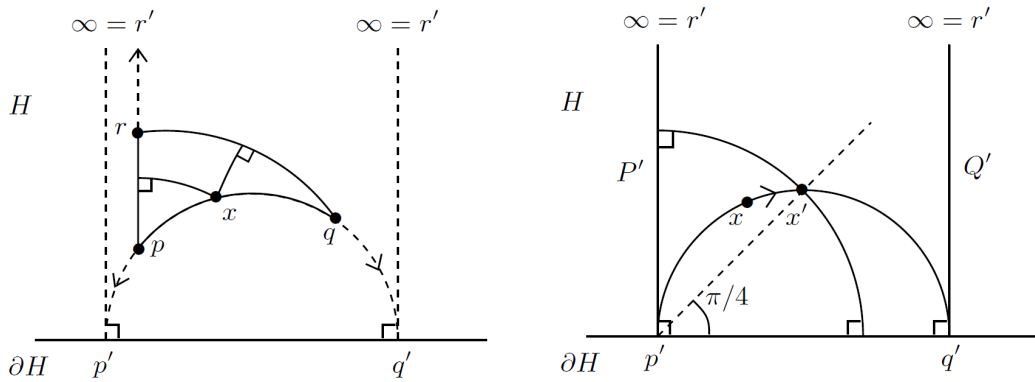


FIGURE 6.11 – Le triangle  $pqr$  et le triangle idéal  $p'q'\infty$ , uniformément fin.

*Propriété 6.5.* Dans le plan hyperbolique, le ratio de l'aire d'un disque à sa circonférence est inférieure à 1, et tend vers 1 quand le rayon tend vers l'infini. L'aire et la circonférence sont toutes deux des fonctions exponentielles du rayon hyperbolique.

On se place dans l'espace du disque  $I$  et on calcule un élément de distance le long d'un arc de cercle  $ds = 2\frac{dr}{1-r^2}$  et un élément d'aire  $dA = \frac{4r}{(1-r^2)^2} drd\theta$ . Si alors  $\rho$  est le rayon hyperbolique et  $R$  le rayon euclidien  $R$ , on a

$$A = 2\pi(\cosh \rho - 1)C = 2\pi \sinh \rho R = \frac{\cosh \rho - 1}{\sinh \rho} \quad (6.30)$$

D'où la propriété.

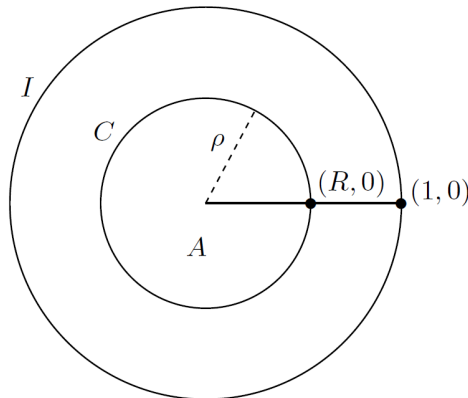


FIGURE 6.12 – Le rayon hyperbolique  $\rho$ , l'aire  $A$  et la circonférence  $C$  : l'essentiel de l'aire d'une boule hyperbolique se trouve sur les bords.

*Propriété 6.6.* Dans l'espace  $H$ , si  $\Sigma$  est un cercle centré en  $x \in \partial H$ , alors l'inversion par  $\Sigma$  induit une isométrie hyperbolique de  $H$  qui échange l'intérieur et l'extérieur de  $\Sigma$  dans  $H$ .

La preuve de cette propriété repose sur l'équivalence entre  $H$  et  $J$ , en effet on peut supposer, quitte à passer  $\Sigma$  par une isométrie, que c'est le cercle image du cercle  $x_1 = 0 |x_2^2 + x_3^2 = 1 \in J$ , et comme la métrique sur  $J$  est invariante par symétrie par rapport au plan  $x_1 = 0$ , la propriété est vérifiée.

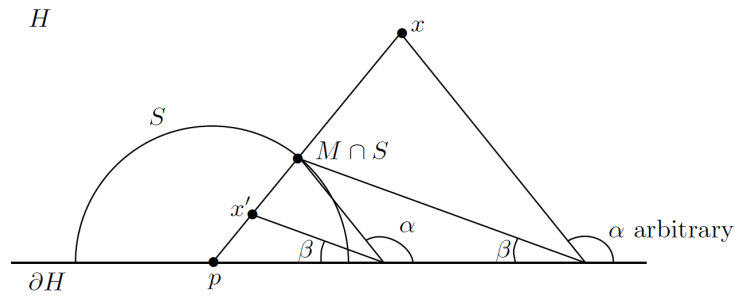


FIGURE 6.13 – Inversion par rapport à un cercle.

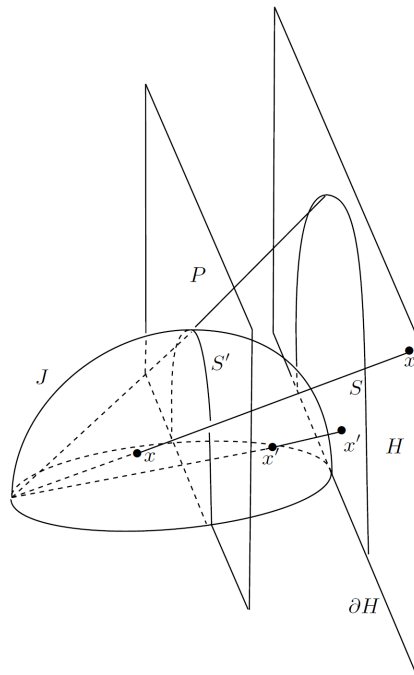


FIGURE 6.14 – L'inversion est une isométrie hyperbolique, qui correspond à une symétrie sur le modèle hémisphérique.

Vadim Lebovici,  
Robin Lemaire,  
Valentin Carlier,  
Lino Benedetto  
Très joli

## 6.5 Le sixième modèle

Le sixième modèle d'espace hyperbolique présenté ici est une approximation combinatoire du demi-espace hyperbolique  $H$ . L'idée est de paver  $H$  à l'aide de carrés de taille diminuant près de la limite de  $H$ . On construit donc ce pavage, en considérant l'image du carré unité de côté bas-gauche placé en  $(0,1)$  par l'application  $x \mapsto 2^j(x + (k, 0))$  où  $(j, k) \in \mathbb{Z}^2$ , et on s'intéresse à son *graphe dual* dont les sommets sont les carrés, reliés entre eux par une arête s'ils sont adjacents.

**Déplacements** Dans le graphe dual, la longueur d'un chemin est définie comme le nombre de côtés rencontrés par le chemin tracé.

*Propriété 6.7.* On a les faits suivants :

- Un chemin aléatoire tracé sur le graphe dual tend presque sûrement vers l'infini.
- Si  $\gamma$  est une géodésique reliant  $p$  et  $q$  dans le graphe dual alors il existe une géodésique  $\delta$  reliant  $p$  et  $q$  qui monte tout droit initialement, puis est horizontale sur une longueur de 5 au maximum puis redescend vers  $q$  de telle sorte que pour tout point de  $\delta$  il existe un point de  $\gamma$  à distance au plus 1, et réciproquement.

*Remarque 6.1.* Dans le graphe dual, il existe plusieurs géodésiques reliant deux points quelconques a priori.

Dans la propriété, ce qu'on appelle infini est la limite de  $H$ .

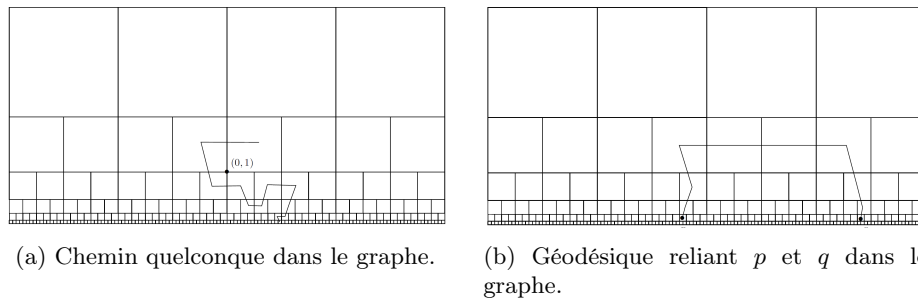


FIGURE 6.15 – Déplacements dans le graphe dual.

*Propriété 6.8.* Les triangles du graphe dual sont uniformément fins, de distance minimale 8.

On constate que la longueur d'un déplacement horizontal est au plus 5 par la propriété 6.7 et que la longueur du déplacement horizontal de gauche est au plus 3 compte-tenu du fait que le nombre de carrés à traverser sur un niveau horizontal est divisé par deux lorsqu'on monte d'un niveau.

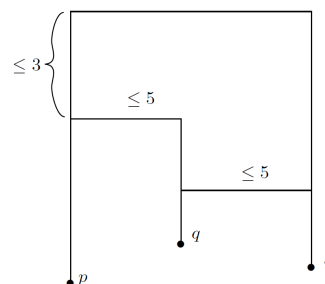


FIGURE 6.16 – Un triangle de côtés des géodésiques dans le graphe dual.

## 6.6 Conclusion

Ce dernier point de vue, combinatoire, fait ressortir de manière forte la propriété 6.4 sur la finesse des triangles : c'est une notion qui "passe au discret", et qui ne dépend donc pas de la structure analytique – très confortable – du plan hyperbolique.

Une idée extrêmement forte de Mikhaïl Gromov a été, dans le courant des années 80, de prendre cette propriété pour *définition* de l'hyperbolicité :

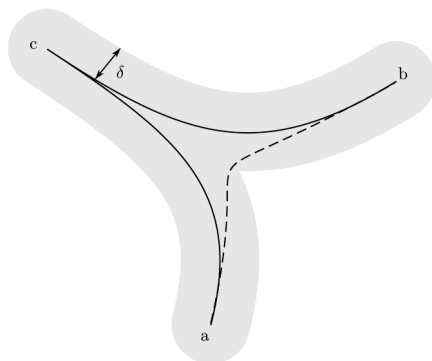


FIGURE 6.17 – Un espace métrique est  $\delta$ -hyperbolique au sens de Gromov si tous ses triangles sont  $\delta$ -fins : pour tous points  $a, b, c$ , le segment géodésique  $[a, b]$  est inclus dans un  $\delta$ -voisinage de  $[a, c] \cup [b, c]$ .

Un exemple tout naturel nous vient de la théorie des graphes : les arbres, ou graphes acycliques, sont des espaces métriques 0-hyperboliques. Ainsi, du point de vue de la structure métrique, un espace hyperbolique est "presque" un arbre !

Cela n'a l'air de rien, mais c'est exactement le genre d'idées qui fait la différence entre les mathématiques du XIX<sup>e</sup> siècle et la géométrie contemporaine. Des belles preuves analytiques, du travail sur les projections stéréographiques accessibles au calcul direct, on est passé à l'étude d'espaces a priori beaucoup moins "sympathiques" et pourtant si naturels : pensez aux pavages d'Escher !

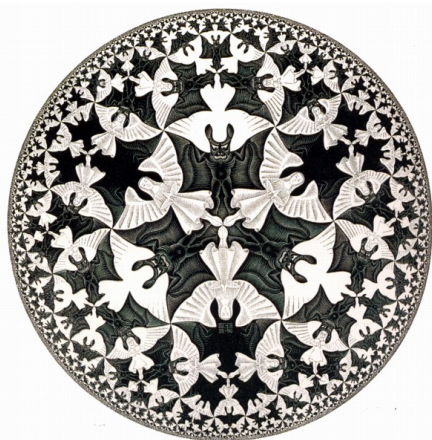


FIGURE 6.18 – Circle Limit IV (Heaven and Hell) de M.C. Escher : que se passerait-il si anges et démons jouaient au téléphone arabe ?

Vadim Lebovici,  
Robin Lemaire,  
Valentin Carlier,  
Lino Benedetto  
Très joli

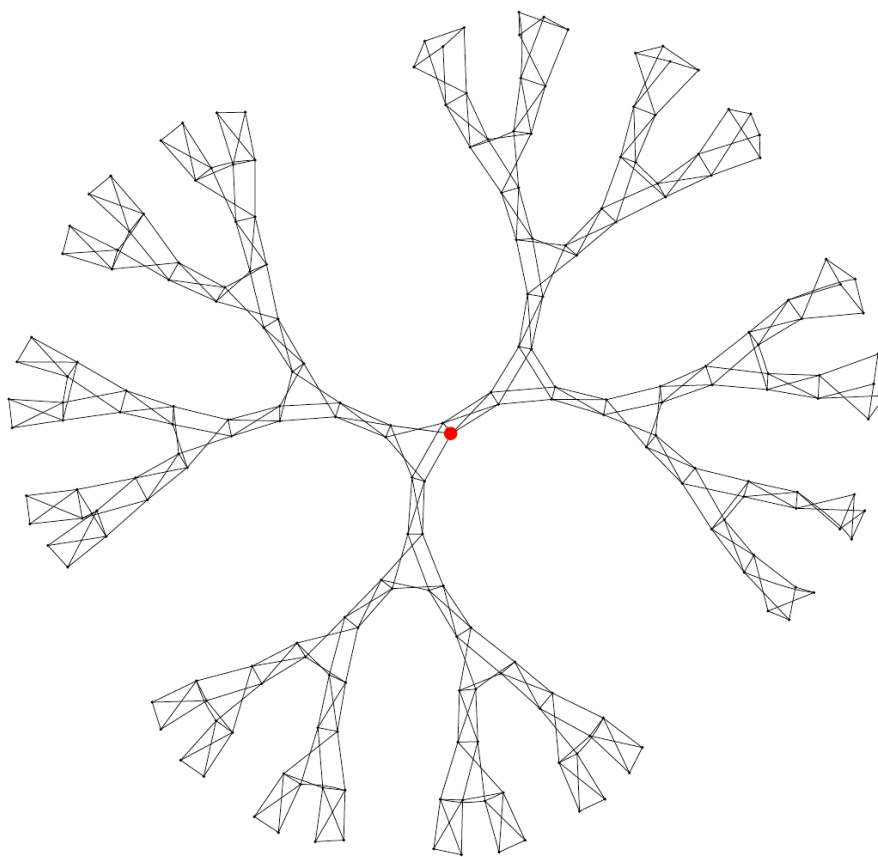


FIGURE 6.19 – Un exemple d’espace métrique hyperbolique non-trivial : le graphe de Cayley (ici tracé au voisinage de l’unité pour un bon système de générateurs) du groupe  $SL_2(\mathbb{Z})$ , qui agit naturellement sur  $H$  par isométries. C’est l’analogie hyperbolique de la grille infinie classique, graphe de Cayley du groupe  $\mathbb{Z} \times \mathbb{Z}$  qui agit par isométries sur le plan euclidien.

**Et les espaces de formes dans tout ça ?** La géométrie hyperbolique se trouve donc au carrefour de la topologie, de la théorie des groupes et de la théorie des applications conformes – la motivation initiale de Poincaré. Dans le cadre de notre cours, nous n’aborderons les questions d’hyperbolicité qu’au travers des géodésiques : l’essentiel sera de retenir que, si la courbure est négative (une notion que l’on abordera à la toute fin du semestre), alors deux géodésiques initialement proches divergent à vitesse exponentielle l’une de l’autre lorsqu’on les prolonge.

Ainsi, une petite incertitude sur la carte exponentielle pourra se traduire par une très grande incertitude sur la variété : en courbure négative, la procédure de *shooting* géodésique est extrêmement sensible à la direction de tir. Pour d’autres illustrations et intuitions sur “la vie en courbure négative”, je ne saurais que trop vous recommander la lecture d’un petit texte de vulgarisation dû à Étienne Ghys : *Poincaré et son disque* disponible à l’adresse suivante : <http://perso.ens-lyon.fr/ghys/articles/disque-poincare.pdf>.



# Chapitre 7

## Distances et géodésiques

Séance 7

[Lee97], Chap. 6

Eva Philippe,  
Arthur Touati,  
Grégoire Macqueron,  
Maxime Sylvestre  
Équilibré

Dans ce chapitre, nous allons étudier la relation entre les géodésiques et la notion de distance sur une variété riemannienne. Nous allons formuler rigoureusement et démontrer l'intuition que l'on a sur la nature des géodésiques : pour minimiser la distance parcourue entre deux points si on se balade sur une variété, on doit emprunter une géodésique.

### 7.1 Longueur d'une courbe, distance et topologie induite

#### 7.1.1 Longueur

Nous commençons par définir la classe de courbes avec lesquelles nous allons travailler.

**Définition 7.1** (Courbe régulière par morceaux). Une courbe lisse  $\gamma : I \rightarrow M$  est dite *régulière* si pour tout  $t \in I$ ,  $\dot{\gamma}(t) \neq 0$ . Elle est dite *régulière par morceaux* si  $I = [a, b]$  et qu'il existe une subdivision finie  $a = a_0 < \dots < a_k = b$  telle que  $\gamma|_{]a_{i-1}, a_i[}$  est régulière pour  $i = 1, \dots, k$ . Les vecteurs vitesses sont bien définis sur les  $]a_{i-1}, a_i[$  et on pose

$$\dot{\gamma}(a_i^-) = \lim_{t \rightarrow a_i^+} \dot{\gamma}(t) \quad (7.1)$$

$$\dot{\gamma}(a_i^+) = \lim_{t \rightarrow a_i^-} \dot{\gamma}(t). \quad (7.2)$$

Remarquons tout de suite qu'une géodésique est régulière, car la norme de son vecteur vitesse est constante. Dans la suite, on appellera des courbes régulières par morceaux des *courbes admissibles*. C'est avec cette classe de courbes qu'on mesurera la distance entre deux points d'une variété riemannienne. Nous commençons par définir la notion de longueur d'une courbe :

**Définition 7.2** (Longueur d'une courbe). Si  $M$  est une variété riemannienne et  $\gamma : [a, b] \rightarrow M$  une courbe, on définit sa *longueur*  $L(\gamma)$  par :

$$L(\gamma) := \int_a^b |\dot{\gamma}(t)| dt \quad (7.3)$$

Si  $\gamma : [a, b] \rightarrow M$  est admissible et si  $a = a_0 < \dots < a_k = b$  est une subdivision adaptée, la longueur de  $\gamma$  est la somme des longueurs des  $\gamma|_{]a_{i-1}, a_i[}$ .

Eva Philippe,  
Arthur Touati,  
Grégoire Macqueron,  
Maxime Sylvestre  
Équilibré

On écrit parfois  $L_g(\gamma)$  pour indiquer la dépendance de cette valeur à la métrique (la métrique intervient dans le calcul de la norme de  $\dot{\gamma}(t)$ ). La longueur d'une courbe ne dépend pas de la paramétrisation, comme le formalisent les énoncés suivants :

**Définition 7.3** (Reparamétrisation d'une courbe). Si  $\gamma : [a, b] \rightarrow M$  est une courbe, on appelle *reparamétrisation* de  $\gamma$  toute courbe de la forme  $\tilde{\gamma} = \gamma \circ \varphi$  où  $\varphi : [c, d] \rightarrow [a, b]$  est  $\mathcal{C}^\infty$ , bijective d'inverse  $\mathcal{C}^\infty$ .

**Lemme 7.1.** Si  $\gamma$  est une courbe et  $\tilde{\gamma}$  une de ses reparamétrisations, alors  $L(\gamma) = L(\tilde{\gamma})$ .

*Démonstration.* Si  $\gamma : [a, b] \rightarrow M$  et  $\tilde{\gamma} = \gamma \circ \varphi$  où  $\varphi : [c, d] \rightarrow [a, b]$ , alors :

$$L(\tilde{\gamma}) = \int_c^d |\dot{\tilde{\gamma}}(t)| dt = \int_c^d |\dot{\varphi}(t) \dot{\gamma}(\varphi(t))| dt = \int_a^b |\dot{\gamma}(t)| dt = L(\gamma) \quad (7.4)$$

□

On démontre maintenant un lemme utile pour la suite :

**Lemme 7.2.** Si  $\gamma : [a, b] \rightarrow M$  est une courbe admissible, alors il existe une reparamétrisation  $\tilde{\gamma}$  de  $\gamma$  de vecteur vitesse unitaire.

*Démonstration.* On note  $l = L(\gamma)$ , pour  $t \in [a, b]$  on définit  $s(t) = \int_a^t |\dot{\gamma}(t)| dt$ . On vérifie que  $s : [a, b] \rightarrow [0, l]$  est  $\mathcal{C}^\infty$  bijective d'inverse  $\mathcal{C}^\infty$ . On pose alors  $\tilde{\gamma} = \gamma \circ s^{-1}$  et on vérifie que  $\tilde{\gamma}$  convient. □

### 7.1.2 Distance riemannienne

Soit  $M$  une variété riemannienne connexe. On définit la *distance riemannienne* entre deux points  $p$  et  $q$  de  $M$  :

$$d(p, q) = \inf \{ L(\gamma) : \gamma \text{ courbe admissible reliant } p \text{ et } q \} \quad (7.5)$$

$M$  étant localement connexe par arcs lisses – comme toute variété –, elle est *globalement* connexe par arcs admissibles : la distance ci-dessus est donc bien définie.

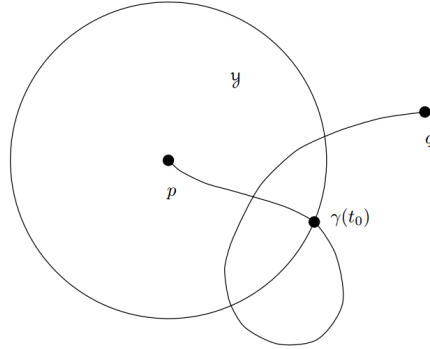
**Lemme 7.3.** Avec la distance  $d$  ainsi définie, toute variété riemannienne est un espace métrique dont la topologie induite coïncide avec la topologie donnée pour la variété.

*Démonstration.* Pour montrer que  $d$  est bien une distance, le seul point difficile est  $d(p, q) \neq 0$  pour  $p \neq q$ . Montrons d'abord que l'on peut comparer la distance riemannienne et la distance euclidienne : il existe une boule géodésique fermée  $\bar{\mathcal{Y}}$  de rayon  $\epsilon$  autour de  $p$  et  $c, C$  deux constantes positives telles que pour tout  $x \in \bar{\mathcal{Y}}$  et  $V \in T_x M$  :

$$c|V|_{\bar{g}} \leq |V|_g \leq C|V|_{\bar{g}} \quad (7.6)$$

( $|\cdot|_{\bar{g}}$  étant la norme euclidienne).

Le champ de métrique  $g$  étant donné par une matrice symétrique définie positive dans  $T_p M$ , d'après le théorème spectral on dispose d'une application continue  $x \mapsto (l_1(p) \geq l_2(p) \geq \dots \geq l_n(p) > 0)$  où les  $l_i(p)$  sont les  $n$  valeurs propres de  $g(p)$ . Pour tout  $V \in T_p M$ , on a donc  $l_n(p) \langle V, V \rangle \leq \langle V, g(p)V \rangle \leq l_1(p) \langle V, V \rangle$  (produit scalaire euclidien). Comme  $l_n$  et  $l_1$  sont des fonctions continues qui ne s'annulent pas, on dispose bien d'une boule géodésique fermée de rayon strictement positif autour de  $p$  sur laquelle elles admettent respectivement un minimum et un maximum.

FIGURE 7.1 – Démonstration de  $d(p, q) > 0$ 

Alors, pour toute courbe admissible  $\gamma$  dont l'image est dans  $\bar{\mathcal{Y}}$  :

$$cL_{\bar{g}}(\gamma) \leq L_g(\gamma) \leq CL_{\bar{g}}(\gamma). \quad (7.7)$$

Si  $p \neq q$ , on peut diminuer  $\epsilon$  jusqu'à avoir  $q \notin \bar{\mathcal{Y}}$ . Alors toute courbe admissible  $\gamma : [a, b] \rightarrow M$  reliant  $p$  à  $q$  doit traverser la sphère géodésique (frontière de  $\mathcal{Y}$ ), d'après le lemme de passage des douanes. On note  $t_0$  le premier instant où  $\gamma(t)$  est sur cette sphère.

On a alors :

$$L_g(\gamma) \geq L_g(\gamma|_{[a, t_0]}) \geq cL_{\bar{g}}(\gamma|_{[a, t_0]}) \geq cd_{\bar{g}}(p, \gamma(t_0)) = c\epsilon \quad (7.8)$$

Ces inégalités étant vraies pour toute courbe admissible  $\gamma$  reliant  $p$  à  $q$ , on a en passant à l'inf :  $d(p, q) \geq c\epsilon > 0$ . D'où  $d$  est une distance.

Enfin, pour comparer les deux topologies il suffit de voir qu'on peut construire une base pour la topologie de la variété à partir de petites boules euclidiennes dans des ouverts de la forme  $\mathcal{Y}$ . Or sur ces ensembles la distance euclidienne et la distance riemannienne sont équivalentes. Donc les ouverts de cette base sont aussi des ouverts pour la distance riemannienne et elle engendre la topologie riemannienne car une topologie métrique est engendrée par de petites boules métriques. Ainsi, ces deux topologies sont identiques.  $\square$

## 7.2 Courbes minimisantes et géodésiques

### 7.2.1 Courbe minimisante

**Définition 7.4.** (courbe minimisante)

Une courbe admissible  $\gamma$  sur une variété riemannienne est dite *minimisante* si  $L(\gamma) \leq L(\tilde{\gamma})$  pour toute courbe  $\tilde{\gamma}$  de mêmes extrémités, autrement dit,  $L(\gamma) = d(p, q)$  où  $p$  et  $q$  sont les extrémités de  $\gamma$ .

Le but de cette section est de montrer que toute courbe minimisante est une géodésique. Pour cela nous allons considérer  $L$  comme une fonctionnelle, c'est-à-dire une fonction qui prend pour argument des fonctions. La recherche d'une courbe minimisante revient à chercher le minimum de  $L$ . On a donc envie de dire qu'une courbe  $\gamma$  est minimisante si la « dérivée » de  $L$  s'annule en  $\gamma$ . C'est l'objet du calcul des variations, très utile dans la formulation moderne de la mécanique classique. On commence cette section en donnant un sens à cette dérivation sur un espace de fonction, dans le cadre qui nous intéresse, c'est-à-dire celui des courbes sur une variété riemannienne.

Eva Philippe,  
 Arthur Touati,  
 Grégoire Macqueron,  
 Maxime Sylvestre  
 Équilibré

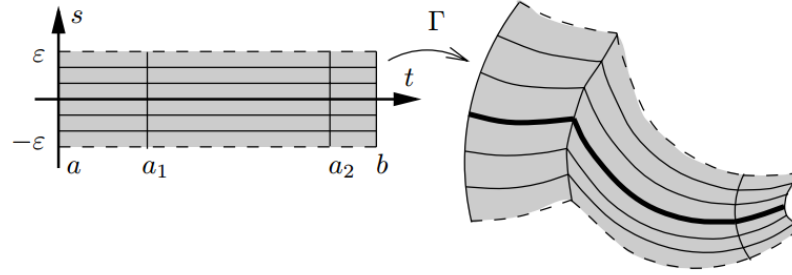


FIGURE 7.2 – Une famille admissible

## 7.2.2 Familles admissibles

**Définition 7.5.** Une *famille admissible* de courbes est une application continue

$$\Lambda : ]-\epsilon, \epsilon[ \times [a, b] \rightarrow M \quad (7.9)$$

lisse sur tout rectangle de la forme  $]-\epsilon, \epsilon[ \times [a_{i-1}, a_i]$  pour une certaine subdivision finie  $a = a_0 < a_1 < \dots < a_k = b$  et telle que  $\Lambda_s : t \mapsto \Lambda(s, t)$  soit une courbe admissible pour tout  $s \in ]-\epsilon, \epsilon[$ .

On notera

$$\partial_t \Lambda(s, t) := \frac{d}{dt} \Lambda_s(t) \quad \partial_s \Lambda(s, t) := \frac{d}{ds} \Lambda^{(t)} \quad (7.10)$$

où  $\Lambda^{(t)} : s \mapsto \Lambda(s, t)$  pour tout  $t \in [a, b]$ .

## 7.2.3 Lemme de symétrie

**Lemme 7.4.** Soit  $\Lambda : ]-\epsilon, \epsilon[ \times [a, b] \rightarrow M$  une famille admissible de courbes sur une variété riemannienne. Sur tous les rectangles  $]-\epsilon, \epsilon[ \times [a_{i-1}, a_i]$  où  $\Lambda$  est lisse, on a :

$$D_s \partial_t \Lambda = D_t \partial_s \Lambda \quad (7.11)$$

*Démonstration.* C'est un résultat local : on se place donc en un point  $\Lambda(s_0, t_0)$  et on se donne un système de coordonnées  $(x^i)$  autour de ce point où l'on écrit  $\Lambda(s, t) = (x^1(s, t), \dots, x^n(s, t))$ . On a alors :

$$\partial_t \Lambda = \frac{\partial x^k}{\partial t} \partial_k \quad (7.12)$$

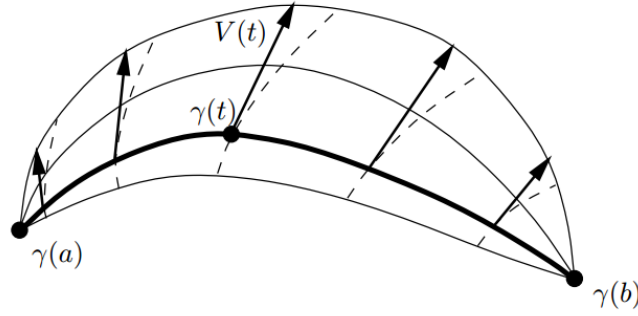
$$\partial_s \Lambda = \frac{\partial x^k}{\partial s} \partial_k \quad (7.13)$$

On utilise maintenant l'expression de la dérivée covariante le long d'une courbe dans la base  $(\partial_k)$  :

$$D_s \partial_t \Lambda = \left( \frac{\partial^2 x^k}{\partial s \partial t} + \frac{\partial x^i}{\partial t} \frac{\partial x^j}{\partial s} \Gamma_{ji}^k \right) \partial_k \quad (7.14)$$

$$D_t \partial_s \Lambda = \left( \frac{\partial^2 x^k}{\partial t \partial s} + \frac{\partial x^i}{\partial s} \frac{\partial x^j}{\partial t} \Gamma_{ji}^k \right) \partial_k \quad (7.15)$$

En permutant les indices  $i$  et  $j$  dans la deuxième expression et en utilisant  $\Gamma_{ji}^k = \Gamma_{ij}^k$ , on obtient la formule voulue.  $\square$

FIGURE 7.3 – Tout champ de vecteurs le long de  $\gamma$  est le champ de variation d’une variation de  $\gamma$ 

### 7.2.4 Formule de la première variation

On montre maintenant comment on peut associer à une courbe admissible  $\gamma$  une famille admissible.

**Définition 7.6** (Variation d’une courbe admissible). Si  $\gamma : [a, b] \rightarrow M$  est une courbe admissible, une *variation* de  $\gamma$  est une famille admissible  $\Lambda$  telle que  $\Lambda(0, t) = \gamma(t)$  pour  $t \in [a, b]$ .

On dit qu’une variation est *propre* si  $\Lambda(s, a) = \gamma(a)$  et  $\Lambda(s, b) = \gamma(b)$  pour tout  $s$ .

Si  $\Lambda$  est une variation de  $\gamma$ , on définit le *champ de variation* de  $\Lambda$  par  $V(t) = \partial_s \Lambda(0, t)$  (c’est un champ de vecteurs le long de  $\gamma$ ).

On dit qu’un champ de vecteurs le long d’une courbe  $\gamma$  est propre si  $V(a) = V(b) = 0$ . Le champ de variation d’une variation propre est lui-même propre.

Il y a équivalence entre variation d’une courbe et champ de vecteurs le long d’une courbe comme le montre le lemme suivant.

**Lemme 7.5.** Si  $\gamma$  est une courbe admissible et  $V$  un champ de vecteurs le long de  $\gamma$ , alors il existe une variation  $\Lambda$  de  $\gamma$  telle que  $V$  soit son champ de variation. Si  $V$  est propre,  $\Lambda$  peut être choisi propre.

**Proposition 7.1** (Formule de la première variation). Soit  $\gamma : [a, b] \rightarrow M$  une courbe admissible de vitesse unitaire,  $\Lambda$  une variation propre de  $\gamma$ ,  $V$  son champ de variation et  $a = a_0 < \dots < a_k = b$  une subdivision adaptée à  $\gamma$ . Alors :

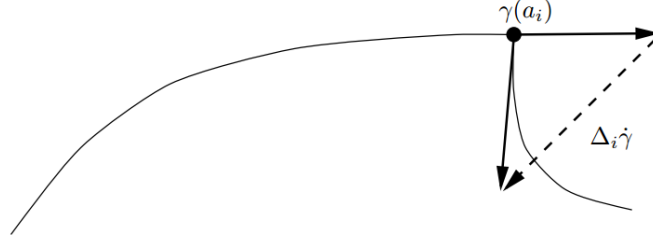
$$\left. \frac{d}{ds} \right|_{s=0} L(\Lambda_s) = - \int_a^b \langle V, D_t \dot{\gamma} \rangle dt - \sum_{i=1}^{k-1} \langle V(a_i), \Delta_i \dot{\gamma} \rangle \quad (7.16)$$

où  $\Delta_i \dot{\gamma} = \dot{\gamma}(a_i^+) - \dot{\gamma}(a_i^-)$  est le “saut” du champ de vecteur tangent  $\dot{\gamma}$  au point  $a_i$ .

*Démonstration.* On peut déjà faire remarquer que

$$\left. \frac{d}{ds} \right|_{s=0} L(\Lambda_s) = \lim_{ds \rightarrow 0} \frac{L(t \mapsto \gamma(t) + ds \cdot V(t)) - L(t \mapsto \gamma(t))}{ds} \quad (7.17)$$

est une forme linéaire en  $V$ . Ce que nous donne la formule, c’est le gradient associé à cette forme linéaire pour le produit scalaire  $(f, g) \mapsto \int_a^b \langle f, g \rangle dt$ , qui est  $-D_t \dot{\gamma}$ .

FIGURE 7.4 –  $\Delta_i \dot{\gamma}$ , le “saut” de  $\dot{\gamma}$  en  $a_i$ 

On notera  $T(s, t) = \partial_t \Lambda(s, t)$  et  $S(s, t) = \partial_s \Lambda(s, t)$ .

Sur tout intervalle  $[a_{i-1}, a_i]$ ,  $\Lambda$  est lisse,  $L(\Lambda_s)$  est lisse et le domaine d’intégration est compact donc on peut dériver sous l’intégrale :

$$\frac{d}{ds} L(\Lambda_s|_{[a_{i-1}, a_i]}) = \frac{d}{ds} \int_{a_{i-1}}^{a_i} \langle T, T \rangle^{\frac{1}{2}} dt \quad (7.18)$$

$$= \int_{a_{i-1}}^{a_i} \frac{\partial}{\partial s} \langle T, T \rangle^{\frac{1}{2}} dt \quad (7.19)$$

Or, comme la connexion  $\nabla$  choisie est compatible avec la métrique  $g$ , on a (se référer au chapitre 5, lemme 5.1)

$$\frac{\partial}{\partial s} \langle T, T \rangle dt = 2 \langle D_s T, T \rangle. \quad (7.20)$$

De plus, le fait d’avoir choisi une connexion symétrique permet d’écrire grâce au lemme de symétrie :

$$D_s T = D_t S \quad (7.21)$$

Ainsi, on comprend a posteriori le choix particulier de travailler avec la connexion de Levi-Civita, l’unique connexion compatible avec la métrique qui soit symétrique. Le calcul donne :

$$\int_{a_{i-1}}^{a_i} \frac{\partial}{\partial s} \langle T, T \rangle^{\frac{1}{2}} dt = \int_{a_{i-1}}^{a_i} \frac{1}{2} \langle T, T \rangle^{-\frac{1}{2}} 2 \langle D_s T, T \rangle dt \quad (7.22)$$

$$= \int_{a_{i-1}}^{a_i} \frac{1}{|T|} \langle D_t S, T \rangle dt. \quad (7.23)$$

Avec  $s = 0$  on a  $S(0, t) = V(t)$  et  $T(0, t) = \dot{\gamma}(t)$ , d’où :

$$\frac{d}{ds} \Big|_{s=0} L(\Lambda_s|_{[a_{i-1}, a_i]}) = \int_{a_{i-1}}^{a_i} \langle D_t V, \dot{\gamma} \rangle dt \quad (7.24)$$

$$= \int_{a_{i-1}}^{a_i} \left( \frac{d}{dt} \langle V, \dot{\gamma} \rangle - \langle V, D_t \dot{\gamma} \rangle \right) dt \quad (7.25)$$

$$= \langle V(a_i), \dot{\gamma}(a_i^-) \rangle - \langle V(a_{i-1}), \dot{\gamma}(a_{i-1}^+) \rangle - \int_{a_{i-1}}^{a_i} \langle V, D_t \dot{\gamma} \rangle dt. \quad (7.26)$$

En sommant sur  $i$  et en se souvenant que  $V(a_0) = V(a_k) = 0$  car  $V$  est une variation propre, on obtient bien l’expression voulue.  $\square$

Comme le montre le lemme 7.1, la contrainte d'avoir un vecteur vitesse unitaire pour tout  $t$  de la proposition précédente n'en est pas vraiment une, mais plutôt une exigence de calcul. On peut maintenant préciser le lien entre les courbes minimisantes, les points critiques de  $L$  et les géodésiques sous la forme d'un lemme et d'un théorème.

### 7.2.5 Les courbes minimisantes sont des géodésiques

**Lemme 7.6.** *Si  $\gamma : [a, b] \rightarrow M$  est une courbe admissible minimisante, alors  $\gamma$  est un point critique pour  $L$ , c'est-à-dire que pour toute variation  $\Lambda$  de  $\gamma$  on a  $\frac{d}{ds}L(\Lambda_s) = 0$  en  $s = 0$ .*

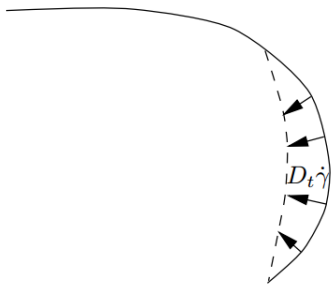
**Théorème 7.1.** *Une courbe admissible de vecteur vitesse unitaire est un point critique de  $L$  si et seulement si c'est une géodésique.*

*Démonstration.* Si  $\gamma$  est une courbe admissible de vecteur vitesse unitaire, on peut utiliser la formule de la première variation. Si  $\gamma$  est une géodésique :  $D_t\dot{\gamma} = 0$  et  $\Delta_t\dot{\gamma} = 0$  donc le membre de droite de la formule de la première variation est nul et  $\frac{d}{ds}L(\Lambda_s) = 0$  en  $s = 0$ .

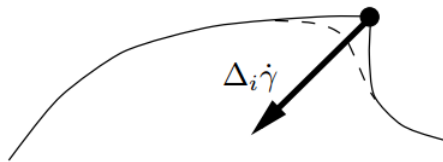
Si  $\gamma : [a, b] \rightarrow M$  est un point critique pour  $L$ , soit  $a = a_0 < \dots < a_k = b$  une subdivision adaptée à  $\gamma$ . Comme tout champ de vecteur propre le long de  $\gamma$  est le champ de variation d'une certaine variation propre de  $\gamma$ , le membre de droite doit être nul pour tout champ de vecteurs propre le long de  $\gamma$ . La première étape est de montrer que  $D_t\dot{\gamma} = 0$  sur tous les intervalles  $[a_{i-1}, a_i]$ . Sur un tel intervalle, on choisit comme champ de vecteur propre  $\varphi D_t\dot{\gamma}$  où  $\varphi$  est  $\mathcal{C}^\infty$ , strictement positive sur  $[a_{i-1}, a_i]$  et nulle ailleurs. L'équation devient alors :

$$0 = - \int_{a_{i-1}}^{a_i} \varphi |D_t\dot{\gamma}|^2 dt \quad (7.27)$$

La fonction intégrée est strictement positive donc  $D_t\dot{\gamma} = 0$  sur tous les intervalles  $[a_{i-1}, a_i]$ . La deuxième étape est de montrer que  $\Delta_i\dot{\gamma} = 0$  pour tout  $i$ . Pour un certain  $i$ , en utilisant une fonction de la forme de  $\varphi$  et en se donnant un système de coordonnées on peut construire un champ de vecteur  $V$  le long de  $\gamma$  tel que  $V(a_i) = \Delta_i\dot{\gamma}$  et  $V(a_j) = 0$  pour  $j \neq i$ . L'équation devient alors  $-\Delta_i\dot{\gamma} = 0$ . Finalement, comme les vecteurs vitesse sont égaux de part et d'autres de chaque  $a_i$ , par unicité des géodésiques,  $\gamma|_{[a_i, a_{i+1}]}$  est le prolongement de  $\gamma|_{[a_{i-1}, a_i]}$  et ainsi  $\gamma$  est lisse sur  $[a, b]$ .  $\square$



(a) Si  $D_t\dot{\gamma} \neq 0$ , on peut déformer  $\gamma$  le long de son vecteur accélération.



(b) Si  $\Delta_i\dot{\gamma} \neq 0$ , on peut "arrondir le coin".

Eva Philippe,  
 Arthur Touati,  
 Grégoire Macqueron,  
 Maxime Sylvestre  
 Équilibré

### 7.2.6 Les géodésiques sont localement minimisantes

Nous allons maintenant démontrer une forme de réciproque au théorème précédent : les géodésiques sont localement minimisantes.

**Définition 7.7.** Une courbe  $\gamma : I \rightarrow M$  est dite *localement minimisante* si pour tout  $t_0 \in I$ , il existe un voisinage  $\mathcal{U}$  de  $t_0$  dans  $I$  tel que  $\gamma|_{\mathcal{U}}$  est minimisante entre toute paire de points de  $\mathcal{U}$ .

Une courbe minimisante est bien sûre toujours localement minimisante, mais la réciproque est fautive : on se place sur  $S^2$  et on prend deux points  $p$  et  $q$  tels que  $d(p, q) < \pi$ . La géodésique qui les relie (c'est-à-dire un grand cercle) peut être scindée en deux géodésiques : la plus longue est localement minimisante mais pas minimisante.

La preuve du résultat annoncé repose sur le théorème important suivant :

**Théorème 7.2** (Lemme de Gauss). *Si  $\mathcal{U}$  est une boule géodésique centrée en  $p \in M$ , le champ de vecteurs radial unitaire  $\frac{\partial}{\partial r}$  est orthogonal aux sphères géodésiques de  $\mathcal{U}$ .*

*Démonstration.* Soit  $q \in \mathcal{U}$  et  $X \in T_qM$  un vecteur tangent à la sphère géodésique passant par  $q$ . Comme  $\exp_p$  est un difféomorphisme sur  $\mathcal{U}$ , il existe  $V, W \in T_pM$  tel que  $\exp_p(V) = q$  et  $(\exp_p)_*W = X$ . Alors si  $R = d(p, q)$ ,  $V \in \partial B_R(0)$  et  $W$  est tangent à  $\partial B_R(0)$ . On considère la géodésique radiale de  $p$  à  $q$ , qui a pour équation  $\gamma(t) = \exp_p(tV)$ , avec pour vecteur vitesse  $\dot{\gamma}(t) = R \frac{\partial}{\partial r}$ . On doit donc montrer  $\langle X, \dot{\gamma}(1) \rangle = 0$ .

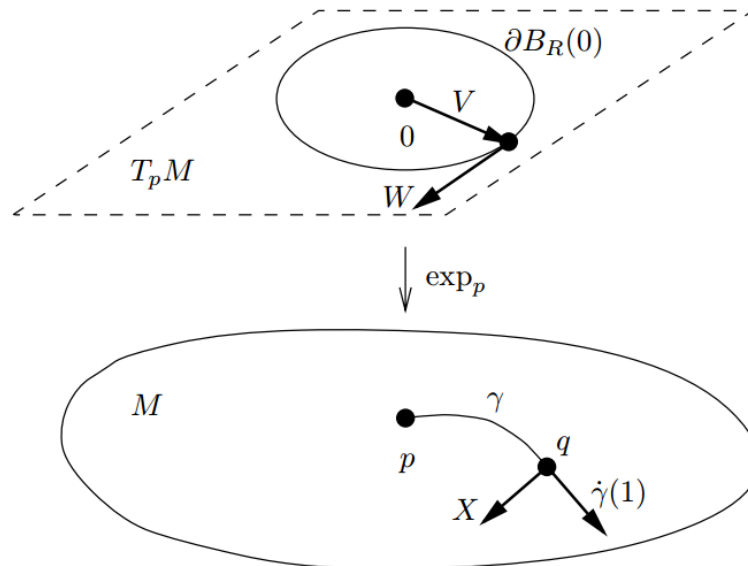


FIGURE 7.6 – Preuve du lemme de Gauss : notations.



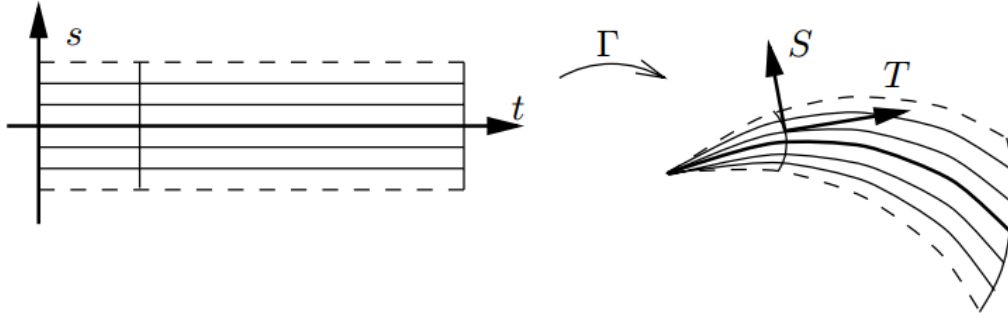


FIGURE 7.7 – Variation  $\Gamma$ , champs de vecteurs  $S$  et  $T$  associés, utilisés lors de la preuve du Lemme de Gauss.

On élargit la géodésique en considérant la variation propre  $\Lambda(s, t) = \exp_{\gamma(t)}(t\sigma(s))$  où  $\sigma : ]-\epsilon, \epsilon[$  d'image incluse dans  $\partial B_R(0)$  et telle que  $\sigma(0) = V$  et  $\dot{\sigma}(0) = W$ . Pour  $s$  dans  $]-\epsilon, \epsilon[$ ,  $\Gamma_s$  est une géodésique de vitesse constante égale à  $R$ . On a naturellement  $\langle S(0), T(0) \rangle = 0$  (car  $S(0) = 0$  vu que la variation est propre), et on a  $\langle S(1), T(1) \rangle = \langle X, \dot{\gamma}(1) \rangle$ . Montrons donc que  $\langle S(t), T(t) \rangle$  ne dépend pas de  $t$ . Les propriétés de la connexion (compatibilité avec la métrique et symétrie) impliquent :

$$\frac{\partial}{\partial t} \langle S, T \rangle = \langle D_t S, T \rangle + \langle S, D_t T \rangle \quad (\text{par compatibilité avec la métrique}) \quad (7.28)$$

$$= \langle D_t S, T \rangle \quad (\text{car } \Gamma_s \text{ est une géodésique}) \quad (7.29)$$

$$= \langle D_s T, T \rangle \quad (\text{par le lemme de symétrie}) \quad (7.30)$$

$$= \frac{1}{2} \frac{\partial}{\partial s} |T|^2 \quad (\text{par compatibilité avec la métrique}) \quad (7.31)$$

$$= 0 \quad (\text{car } T \text{ est unitaire}) \quad (7.32)$$

On a bien  $\langle X, \dot{\gamma}(1) \rangle = 0$ . □

Le lemme de Gauss était la clé pour passer des géodésiques “différentielles”, obtenues via les connexions, aux géodésiques “métriques”, les rayons de boules. En utilisant de manière *cruciale* les propriétés de symétrie et de compatibilité, on a pu caractériser par un énoncé purement géométrique (l’orthogonalité aux surfaces de niveau de la fonction distance au point  $p$ ) les géodésiques tirées via le théorème de Cauchy-Lipschitz à partir du point  $p$ . On peut alors développer, de manière purement géométrique, les propriétés locales des géodésiques :

**Lemme 7.7.** *Si  $(x^i)$  sont les coordonnées normales sur une boule géodésique  $\mathcal{U}$  centrée en  $p \in M$ , et si  $r$  est la fonction distance radiale, alors  $\text{grad } r = \frac{\partial}{\partial r}$  sur  $\mathcal{U} \setminus \{p\}$ .*

**Proposition 7.2.** *Si  $p \in M$  et  $q$  est dans une boule géodésique centrée en  $p$ , alors, à reparamétrisation près, la géodésique radiale de  $p$  à  $q$  est l’unique courbe minimisante de  $p$  à  $q$  dans  $M$ .*

**Proposition 7.3.** *Au sein d’une boule géodésique autour de  $p \in M$ , la distance radiale  $r(x)$  coïncide avec la distance riemannienne entre  $p$  et  $x$ .*

Eva Philippe,  
Arthur Touati,  
Grégoire Macqueron,  
Maxime Sylvestre  
Équilibré

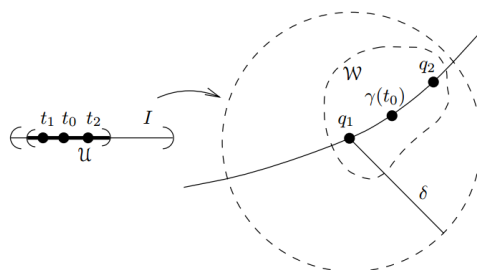


FIGURE 7.8 – Les géodésiques sont localement minimisantes

**Théorème 7.3.** *Toute géodésique riemannienne est localement minimisante.*

*Démonstration.* Soit  $\gamma : I \rightarrow M$  une géodésique, qu'on peut supposer définie sur un intervalle ouvert. Soit  $t_0 \in I$ , et soit  $\mathcal{W}$  un voisinage uniformément normal de  $\gamma(t_0)$  (i.e. il existe  $\delta > 0$  tel que  $\mathcal{W}$  est inclus dans toute boule géodésique de rayon  $\delta$  centrée autour de l'un de ses points – le caractère lisse du flot géodésique dans l'espace des phases garantit en effet que le rayon de bonne définition de la carte exponentielle ne s'effondre pas instantanément, puisque  $(p, v) \mapsto \exp_p(v)$  est lisse), et  $\mathcal{U}$  la composante connexe de  $\gamma^{-1}(\mathcal{W})$  contenant  $t_0$ . Soit  $t_1, t_2 \in \mathcal{U}$  et  $q_i = \gamma(t_i)$ . Alors  $q_2$  est contenu dans une boule géodésique autour de  $q_1$ . D'après la proposition 7.2, la géodésique radiale entre  $q_1$  et  $q_2$  est l'unique courbe minimisante entre ces deux points. Comme la restriction de  $\gamma$  est aussi une géodésique entre  $q_1$  et  $q_2$  dans la même boule géodésique, on en déduit que  $\gamma$  coïncide avec cette géodésique minimisante.  $\square$

### 7.3 Complétude : Le théorème de Hopf-Rinow

**Définition 7.8.** Une variété riemannienne est dite *géodésiquement complète* si toute géodésique maximale est définie pour tout  $t \in \mathbb{R}$ .

Par exemple, un ouvert non trivial de  $\mathbb{R}^n$  muni de la distance euclidienne n'est pas géodésiquement complet car il y a des géodésiques qui atteignent la limite en un temps fini.

**Théorème 7.4** (Théorème de Hopf-Rinow). *Une variété riemannienne connexe est géodésiquement complète si et seulement si elle est complète en tant qu'espace métrique.*

Faute de temps, nous n'aborderons pas la preuve de ce théorème, qui repose en grande partie sur les propriétés de la carte exponentielle. On retiendra simplement les corollaires suivants, qui nous rassurent quand à la bonne définition des géodésiques dans un espace où *les bords sont inaccessibles, à l'infini* :

**Proposition 7.4.** *S'il existe un point  $p$  tel que la carte exponentielle  $\exp_p$  est définie sur tout  $T_p M$ , alors  $M$  est complète.*

**Proposition 7.5.**  *$M$  est complète si et seulement si deux points de  $M$  peuvent être reliés par un segment géodésique minimisant.*

**Proposition 7.6.** *Si  $M$  est compacte, alors toute géodésique peut être définie pour tout temps.*

Deuxième partie

Application à l'étude des Espaces de  
Formes



## Chapitre 8

# Un article fondateur : les triangles de Kendall

*Séance 8*

*[Ken84]  
Basile Coron,  
Paul Fraux,  
Massil Hihat,  
Lino Benedetto  
Élémentaire  
et élégant*

### 8.1 Introduction et motivation à l'étude des triangles

L'étude des formes a motivé un bel effort de recherche dans le courant des années 80 avec en point d'orgue l'étonnant résultat de David G. Kendall : la compréhension définitive de l'espace des formes de triangles, isométrique à une sphère de rayon  $1/2$ .

Ce résultat exceptionnel, s'inscrit dans une démarche de classification topologique des espaces de formes mais a aussi eu pour but l'étude de ceux-ci en tant qu'espaces probabilisés.

Illustrée en Figure 8.1, voici le genre de questions que l'on se posait :

« Étant donnée une collection arbitraire de points du plan, à partir de quels critères peut-on affirmer qu'elle présente "trop" d'alignements, indices d'une structuration sous-jacente ? »

L'*alignement* étant une question indépendante de l'échelle ou de l'orientation des données, il est naturel de travailler sur cette question "à similitude près". Comme elle peut s'exprimer en terme de triplets de points, elle nous incite à concentrer nos efforts sur le cas des triangles . À partir d'une collection de points, on peut construire l'ensemble de tous les triangles y prenant leurs sommets, et donc se ramener à une question bien posée :

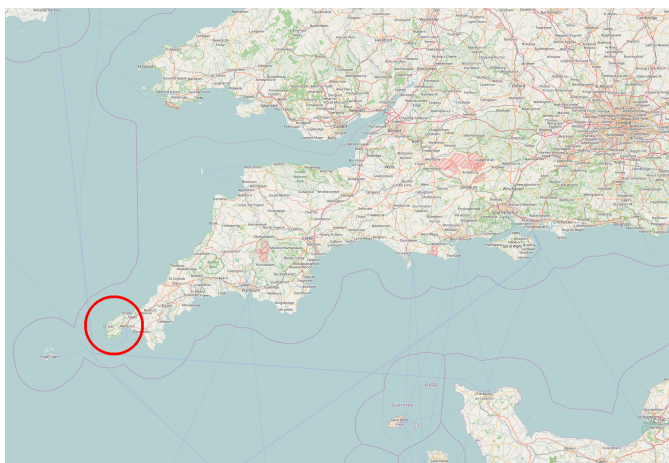
« Quelle est la distribution "standard" des formes de triangles – à similitude près – reliant des points tirés au hasard selon une loi de probabilité simple ?

Plus spécifiquement, si des points sont (par exemple) tirés de manière indépendante et uniforme dans le disque unité, puis reliés trois à trois, quelle sera la proportion observée de triangles "presque plats" avec un angle inférieur à  $1^\circ$  ? »

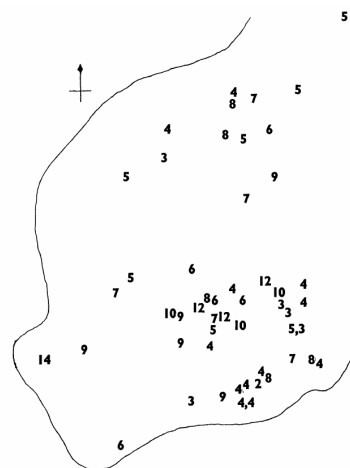
Durant cette séance nous allons d'abord nous concentrer sur les triangles, on pourra ainsi se faire une bonne intuition d'un espace de forme particulier et présenter dans un cadre assez simples les premiers principes et outils servant dans l'étude de ces espaces comme l'analyse procrustéenne. Par ailleurs, nous allons réussir à montrer le surprenant résultat de Kendall à propos de la sphère des triangles en utilisant des outils géométriques et trigonométriques assez simples. Nous verrons ensuite les résultats de généralisation de ces espaces, où nous devons utiliser un bagage beaucoup plus conséquent. Enfin, nous étudierons les questions relatives aux probabilités et aux statistiques sur ces espaces.

[Ken84]

Basile Coron,  
Paul Fraux,  
Massil Hihat,  
Lino Benedetto  
Élémentaire  
et élégant



(a) À l'extrémité ouest des îles britanniques, la pointe des Cornouailles recèlerait-elle un secret ?  
Carte OpenStreetMap.



(b) Sur cette péninsule de 10km sur 15km, on observe une concentration élevée de monolithes.

FIGURE 8.1 – Motivation initiale de David Kendall : les monolithes de Land’s End ont-ils été placés au hasard, ou peut-on voir dans la distribution ci-dessus les traces d’alignements significatifs ? Chaque site est ici représenté par un nombre, celui des “triangles presque plats” dont il fait partie. On reformule alors notre question de manière statistique :

« Sur la carte (b), y a-t-il un nombre anormalement élevé de *triades* de mégalithes alignés ? »

Dans un article d’une grande importance historique, *Simulating the Ley Hunter* paru en 1980, Simon Broadbent laisse à penser que non – ce qui est bien entendu sujet à de sérieuses controverses archéologiques.

## 8.2 Un espace de formes : la sphère des triangles

Pour nous, une forme est un nuage de  $k$  points de  $\mathbb{R}^m$ , les triangles dans le plan sont donc associés aux valeurs  $k = 3$  et  $m = 2$ . Restreignons nous au cas des formes dans le plan ( $m = 2$ ) et considérons que l’on dispose d’une collection  $P^1, \dots, P^N$  de telles formes.

**Action des similitudes** Ces nuages de points  $P^n$ , on peut les transformer à l’aide de translations, homothéties, rotations : si  $(\bar{x}, \bar{y})$  est un vecteur du plan,  $s$  est une échelle et  $\theta$  un angle, on définit l’application

$$S_{\bar{x}, \bar{y}, s, \theta} : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} s \cos(\theta) \cdot x - s \sin(\theta) \cdot y + \bar{x} \\ s \sin(\theta) \cdot x + s \cos(\theta) \cdot y + \bar{y} \end{pmatrix} \quad (8.1)$$

qui est la composée d’une translation de vecteur  $(\bar{x}, \bar{y})$ , d’une rotation d’angle  $\theta$  et d’une homothétie de rapport  $s$  strictement positif. À vrai dire, puisque qu’on se restreint aux formes du plan, on peut se simplifier la vie en adoptant le formalisme complexe. On caractérisera nos points  $(x, y)$  par des affixes  $z = x + iy$  et une similitude du plan est la donnée de deux complexes

$$\tau = \bar{x} + i\bar{y} \quad \text{et} \quad v = se^{i\theta}, \quad (8.2)$$

et on pourra écrire nos similitudes comme des applications

$$S_{\tau, v} : z \mapsto v \cdot z + \tau. \quad (8.3)$$

Si  $P^n = (P_1^n, \dots, P_I^n)$  est l'une de nos "formes", on définira sans ambages son transformé :

$$S_{\tau,v}(P^n) = (S_{\tau,v}(P_1^n), \dots, S_{\tau,v}(P_I^n)). \quad (8.4)$$

**Analyse procrustéenne** Si  $P^m$  et  $P^n$  sont deux individus distincts, ne pourrait-on alors tenter de recaler l'un sur l'autre à l'aide d'une similitude ?

Cela permettrait de décomposer la différence  $P^m - P^n$  entre deux nuages en une composante rigide  $P^m - S_{\tau,v}(P^m)$ , donnée de  $(\bar{x}, \bar{y})$ ,  $s$  et  $\theta$ , et un résiduel irréductible  $S_{\tau,v}(P^m) - P^n$ .

Illustré Figure 8.2, le procédé le plus simple est de trouver le choix de  $(\tau, v)$  qui minimise

$$\|S_{\tau,v}(P^m) - P^n\|_2^2 = \sum_{i=1}^I |S_{\tau,v}(P_i^m) - P_i^n|^2 \quad (8.5)$$

$$= |v \cdot P_1^m + \tau - P_1^n|^2 + \dots + |v \cdot P_I^m + \tau - P_I^n|^2. \quad (8.6)$$

Nous noterons ce minimum  $d_{\text{Procuste}}(P^m \rightarrow P^n)$ . Cette méthode a été popularisée sous le nom d'analyse *Procrustéenne* en référence à Procuste, brigand apparaissant dans la légende de Thésée qui, d'après Diodore de Sicile

« contraignait les voyageurs à se jeter sur un lit ; il leur coupait les membres trop grands et qui dépassaient du lit ; et étirait les pieds de ceux qui étaient trop petits. »

### 8.2.1 Un système de coordonnées adaptées

On va présenter le cadre théorique permettant de répondre de manière élégante à de telles questions. Avant tout, il importe de *représenter* nos triangles sous une forme adaptée aux calculs de **distances** et de **similitudes**.

On adoptera ici les notations de l'article de référence de David Kendall, *Shape Manifolds, Procrustean Metrics and Complex Projective Spaces* (1984), sur lequel est basé la présente discussion. Pour une exposition plus élémentaire, on pourra aussi consulter *Exact Distributions for Shapes of Random Triangles in Convex Sets* du même auteur, publié en 1985 à destination d'un public non spécialisé.

**Paramétrisation naïve** À première vue, un triangle  $ABC$  n'est rien d'autre que la donnée de trois affixes complexes associées aux sommets :  $z_1^*$ ,  $z_2^*$  et  $z_3^*$ . Conformément à l'équation (8.4), pour tout "vecteur"  $\tau$  et rapport  $v$ , on peut écrire

$$S_{\tau,v}(ABC) = S_{\tau,v}(z_1^*, z_2^*, z_3^*) = (v \cdot z_1^* + \tau, v \cdot z_2^* + \tau, v \cdot z_3^* + \tau). \quad (8.7)$$

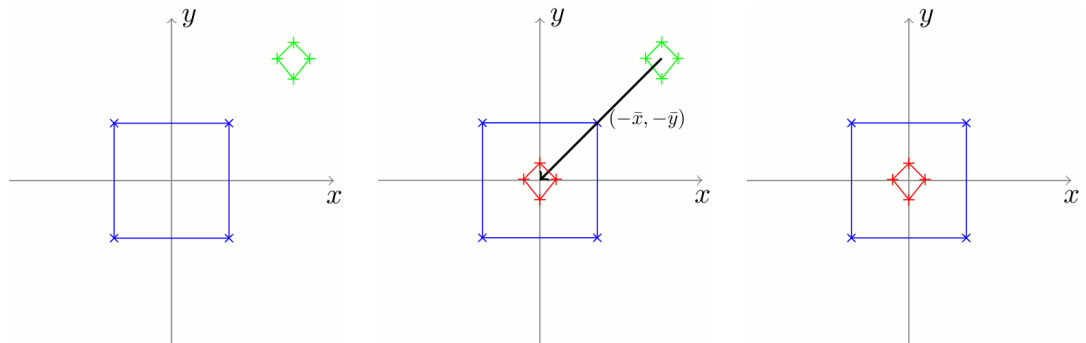
De plus, pour  $(z_1^*, z_2^*, z_3^*)$  et  $(w_1^*, w_2^*, w_3^*)$  deux triangles quelconques, le critère minimisé par l'analyse procrustéenne s'écrit simplement :

$$\|z^* - w^*\|_2^2 = |z_1^* - w_1^*|^2 + |z_2^* - w_2^*|^2 + |z_3^* - w_3^*|^2. \quad (8.8)$$

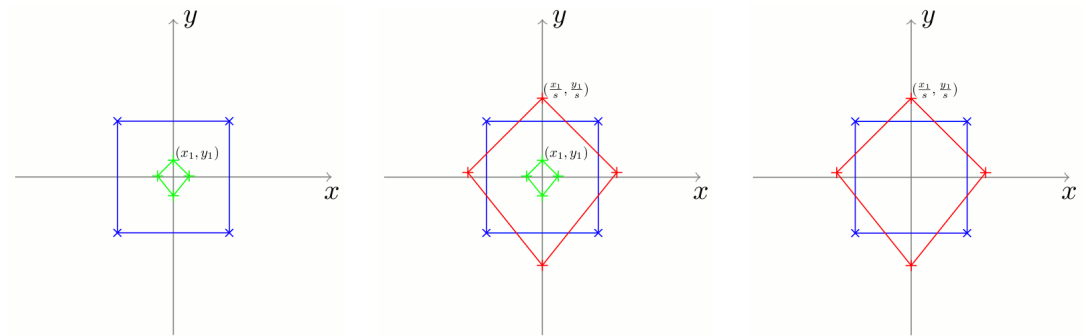
**Coordonnées barycentriques** Grâce à la multiplication complexe, la représentation naïve des triangles par un triplet d'affixes n'est donc pas un mauvais choix. Mais il est possible d'aller plus loin en considérant le nouveau triplet de coefficients

$$(z_0 \quad z_1 \quad z_2) = \left( \frac{1}{\sqrt{3}}(z_1^* + z_2^* + z_3^*) \quad \frac{1}{\sqrt{2}}(z_2^* - z_1^*) \quad \frac{1}{\sqrt{6}}(2z_3^* - z_1^* - z_2^*) \right) \quad (8.9)$$

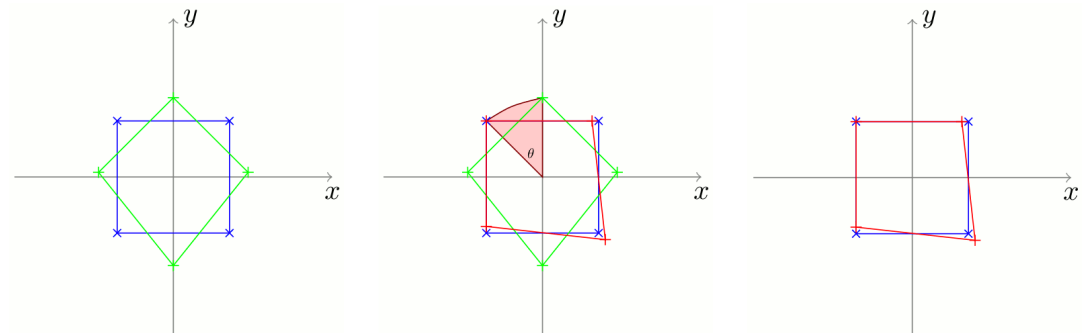
$$= (z_1^* \quad z_2^* \quad z_3^*) \cdot \begin{pmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & +1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & +2/\sqrt{6} \end{pmatrix}. \quad (8.10)$$



(a) Données brutes : la source  $P^m$  est en vert (clair), la cible  $P^n$  en bleu (foncé).  
 (b) La meilleure translation est celle qui aligne les barycentres.  
 (c) Après recalage des barycentres.



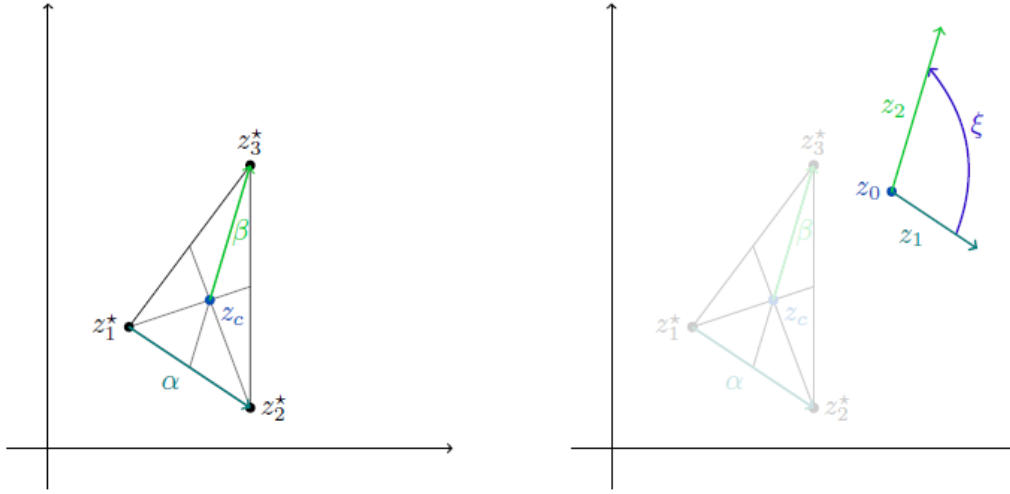
(d) Les deux formes, centrées.  
 (e) On peut calculer le rapport de dilatation optimal.  
 (f) Après application de la translation + homothétie.



(g) Les deux formes sont centrées et normalisées : reste à recalculer les orientations.  
 (h) La meilleure rotation s'obtient elle aussi par un calcul analytique simple.  
 (i) Après recalage par une similitude : le résidu peut être analysé indépendamment.

FIGURE 8.2 – L'analyse Procrustéenne, ou la recherche pas-à-pas de la meilleure similitude pour recaler deux formes – ici, un quadrilatère vert sur le carré bleu. Les paramètres optimaux pour la similitude sont obtenus par un calcul élémentaire, détaillé dans le cas des triangles aux équations (8.20-8.49). Après recalage (en bas à droite), les deux quadrilatères diffèrent uniquement par une fine différence de *forme* (au sens usuel) : la différence qui existe entre un cerf-volant et un carré, indépendante de tout problème d'échelle ou d'orientation. Images tirées de Wikipédia, par Linschn.





(a) Des affixes aux coordonnées barycentriques. (b) Des coordonnées barycentriques au ratio projectif.

FIGURE 8.3 – On peut caractériser un triangle  $(z_1^*, z_2^*, z_3^*)$  par son barycentre  $z_c = \frac{1}{3}(z_1^* + z_2^* + z_3^*)$  et deux vecteurs directionnels  $\alpha = z_2^* - z_1^*$  et  $\beta = \frac{1}{3}(2z_3^* - z_1^* - z_2^*)$ . En normalisant ceux-ci pour obtenir un jeu de coordonnées  $(z_0, z_1, z_2) = (\sqrt{3}z_c, \alpha/\sqrt{2}, \beta\sqrt{3}/\sqrt{2})$ , on dispose d'une représentation adaptée aux similitudes *et* aux calculs de distances – équations (8.11-8.12). Finalement, le ratio  $\xi = z_2/z_1$  encode complètement la forme du triangle aux similitudes près.

Illustré Figure 8.3, ce changement linéaire de coordonnées permet d'écrire les similitudes de manière plus compacte

$$S_{\tau,v}(ABC) = S_{\tau,v}(z_0, z_1, z_2) = (v \cdot z_0 + \sqrt{3}\tau, v \cdot z_1, v \cdot z_2), \quad (8.11)$$

les translations n'ayant d'influence que sur le seul barycentre. Dans le même temps, grâce aux propriétés d'orthogonalité de la matrice de changement de repère, on a toujours une expression simple de la distance quadratique entre deux triangles du plan :

$$\|z^* - w^*\|_2^2 = \|z - w\|_2^2 = |z_0 - w_0|^2 + |z_1 - w_1|^2 + |z_2 - w_2|^2. \quad (8.12)$$

**Relation d'équivalence** Si  $(z_0, z_1, z_2)$  et  $(w_0, w_1, w_2)$  sont deux triangles donnés en coordonnées "barycentriques", on peut maintenant déterminer s'ils sont ou non similaires l'un à l'autre :

$$\exists \tau, v \in \mathbb{C}, S_{\tau,v}(z) = w \iff \exists \tau, v \in \mathbb{C}, \begin{cases} vz_0 + \sqrt{3}\tau = w_0 \\ vz_1 = w_1 \\ vz_2 = w_2 \end{cases} \quad (8.13)$$

$$\iff z_1 w_2 = z_2 w_1. \quad (8.14)$$

Autrement dit, si notre triangle n'est pas dégénéré ( $A = B = C$ , cas que l'on exclut de notre analyse), on peut caractériser sa forme "à similitude près" par le ratio

$$\xi = \frac{z_2}{z_1} = \frac{\frac{1}{\sqrt{6}}(2z_3^* - z_1^* - z_2^*)}{\frac{1}{\sqrt{2}}(z_2^* - z_1^*)} \in \mathbb{C} \cup \{\infty\}. \quad (8.15)$$

[Ken84]

Basile Coron,

Paul Fraux,

Massil Hihat,

Lino Benedetto

Élémentaire

et élégant

Deux triangles non dégénérés  $(z)$  et  $(w)$  seront semblables si et seulement si leurs “ratios” respectifs  $\xi = z_2/z_1$  et  $\xi' = w_2/w_1$  sont égaux. On a réussi à décrire la forme d’un triangle par un unique nombre complexe “projectif” (i.e. qui peut être infini), avec *deux* paramètres réels (module et argument). À vrai dire, comme illustré Figures 8.4 et 8.5, considérer le ratio  $\xi$  revient ni plus ni moins à utiliser une similitude pour recaler  $(A, B) = (z_1^*, z_2^*)$  sur  $(-1, 1)$ , puis à identifier la forme du triangle avec la coordonnée libre restante, issue de  $z_3^*$ .

**Premiers exemples** Certaines valeurs remarquables de  $\xi$  correspondent à des propriétés simples du triangle  $ABC$ ; on trouve :

- $\xi = \infty \iff A = B.$
- $\xi = 0 \iff C$  est le milieu de  $[AB].$
- $\xi \in \mathbb{R} \cup \{\infty\} \iff ABC$  est plat.
- $\xi \in i\mathbb{R} \cup \{\infty\} \iff ABC$  est isocèle en  $C.$
- $\xi \in \{-i, +i\} \iff ABC$  est équilatéral.

Malheureusement, manipuler  $\xi$  manque d’élégance : pourquoi mettre à part le triangle avec  $A = B$ , correspondant à l’infini ? L’espace des triangles devrait refléter une certaine invariance par permutation des sommets, ce qui n’est clairement pas le cas ici. Pour trouver une représentation satisfaisante de l’espace des formes de triangles, il faut donc travailler encore un peu.

**La sphère des triangles** Souvenons-nous : dans le film *Dimensions* repris Figure ??, nous avons appris qu’une “droite projective complexe” n’est rien d’autre qu’une sphère. Par la projection stéréographique, il est donc possible d’identifier un ratio complexe  $\xi \in \mathbb{C} \cup \{\infty\}$  avec un point de la boule unité, caractérisé par deux angles en coordonnées sphériques  $\theta$  et  $\varphi$ . En s’inspirant de cette idée, David Kendall propose d’associer un couple d’angles à chaque triangle  $(z^*) \sim (z) \sim \xi$  par la formule :

$$\theta = 2 \arctan(|\xi|) \quad \text{et} \quad \varphi = \arg(\xi), \quad (8.16)$$

$$\text{de sorte que} \quad \xi = \tan(\theta/2) e^{i\varphi}. \quad (8.17)$$

Chaque classe de triangle à similitude près est alors associée à un point de la sphère, que l’on choisit de rayon  $1/2$  pour une raison qui deviendra claire au Théorème 8.1 :

$$f(\xi) = \frac{1}{2}(\cos \theta, \sin \theta \cos \varphi, \sin \theta \sin \varphi) \quad (8.18)$$

en coordonnées  $(X, Y, Z)$ . Le résultat est illustré Figure 8.7, avec un représentant de classe pour chaque point du globe. Suivant la page précédente, on trouve :

- Le segment  $[A = B, C]$  au point  $(-1/2, 0, 0)$ .
- Le triangle plat où  $C$  est au milieu de  $[AB]$  au point  $(+1/2, 0, 0)$ .
- Les triangles plats sur l’équateur  $Z = 0$ .
- Les triangles isocèles en  $C$  sur le méridien  $Y = 0$ .
- Les deux triangles équilatéraux direct (Nord) et indirect (Sud) aux pôles  $(0, 0, \pm 1/2)$ .

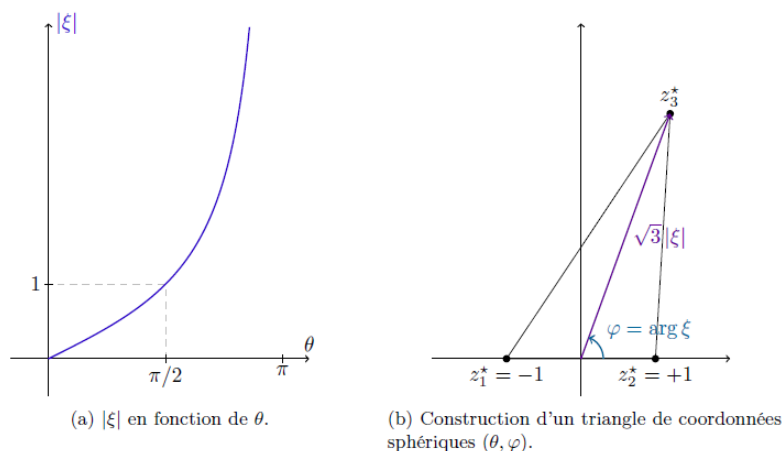


FIGURE 8.4 – Étant donné un point de la sphère de coordonnées sphériques  $(\theta, \varphi) \in [0, \pi] \times [0, 2\pi[$ , comment construire un triangle lui correspondant ? Il suffit de considérer le triplet  $(z_1^*, z_2^*, z_3^*) = (-1, 1, \sqrt{3} \tan(\theta/2) e^{i\varphi})$ .

## 8.2.2 Sphère des triangles et distance procustéenne

La sphère des triangles ne pourrait être qu'un moyen astucieux de rendre compte de la *topologie* de l'ensemble des triangles considérés à similitude près. Un moyen commode de se souvenir que pour passer d'un triangle équilatéral direct à un triangle équilatéral indirect, il faut passer par l'équateur des triangles plats. Mais cette représentation de l'espace de triangles est beaucoup plus profonde qu'un simple moyen mnémotechnique. Avant tout, elle rend parfaitement compte de l'analyse procustéenne sur les triangles, comme indiqué par le théorème suivant.

**Théorème 8.1** (Sphère de Kendall et moindres carrés). *Soit  $ABC$  et  $DEF$  deux triangles du plan aux sommets étiquetés, identifiés à des triplets d'affixes complexes  $z^* = (z_1^*, z_2^*, z_3^*)$  et  $w^* = (w_1^*, w_2^*, w_3^*)$ .*

*Suivant la discussion des pages précédentes, on associe à  $z^*$  (respectivement  $w^*$ ) un couple de complexes "barycentriques"  $z = (z_1, z_2)$  (resp.  $w = (w_1, w_2)$ ), puis un ratio projectif complexe  $\xi = z_2/z_1$  (resp.  $\xi' = w_2/w_1$ ), un couple d'angles  $(\theta, \varphi) \in [0, \pi] \times [0, 2\pi[$  (resp.  $(\theta', \varphi')$ ) et un point de la sphère centrée en 0 de rayon 1/2 dans  $\mathbb{R}^3$ ,  $f(\xi)$  (resp.  $f(\xi')$ ).*

*Alors en notant  $W = \sqrt{|w_1|^2 + |w_2|^2}$  l'échelle de  $DEF$ , on a :*

$$d_{\text{Procuste}}(ABC \rightarrow DEF) = \min_{\tau, \nu} \|S_{\tau, \nu}(z^*) - w^*\|_2 = W \cdot \|f(\xi) - f(\xi')\|_{\mathbb{R}^3}. \quad (8.19)$$

*Autrement dit : la distance procustéenne entre triangles est directement proportionnelle à la distance euclidienne entre points de la sphère. L'ensemble des triangles du plan peut donc être vu comme le produit de deux facteurs : le groupe des similitudes " $S_{\tau, \nu}$ " (de dimension réelle 4), et la sphère de Kendall de dimension 2, qui caractérise les résiduels après recalage rigide.*

La preuve repose sur un simple calcul de minimum par annulation de la dérivée, avec utilisation de moult identités trigonométriques.

[Ken84]

Basile Coron,  
 Paul Froux,  
 Massil Hihat,  
 Lino Benedetto  
 Élémentaire  
 et élégant

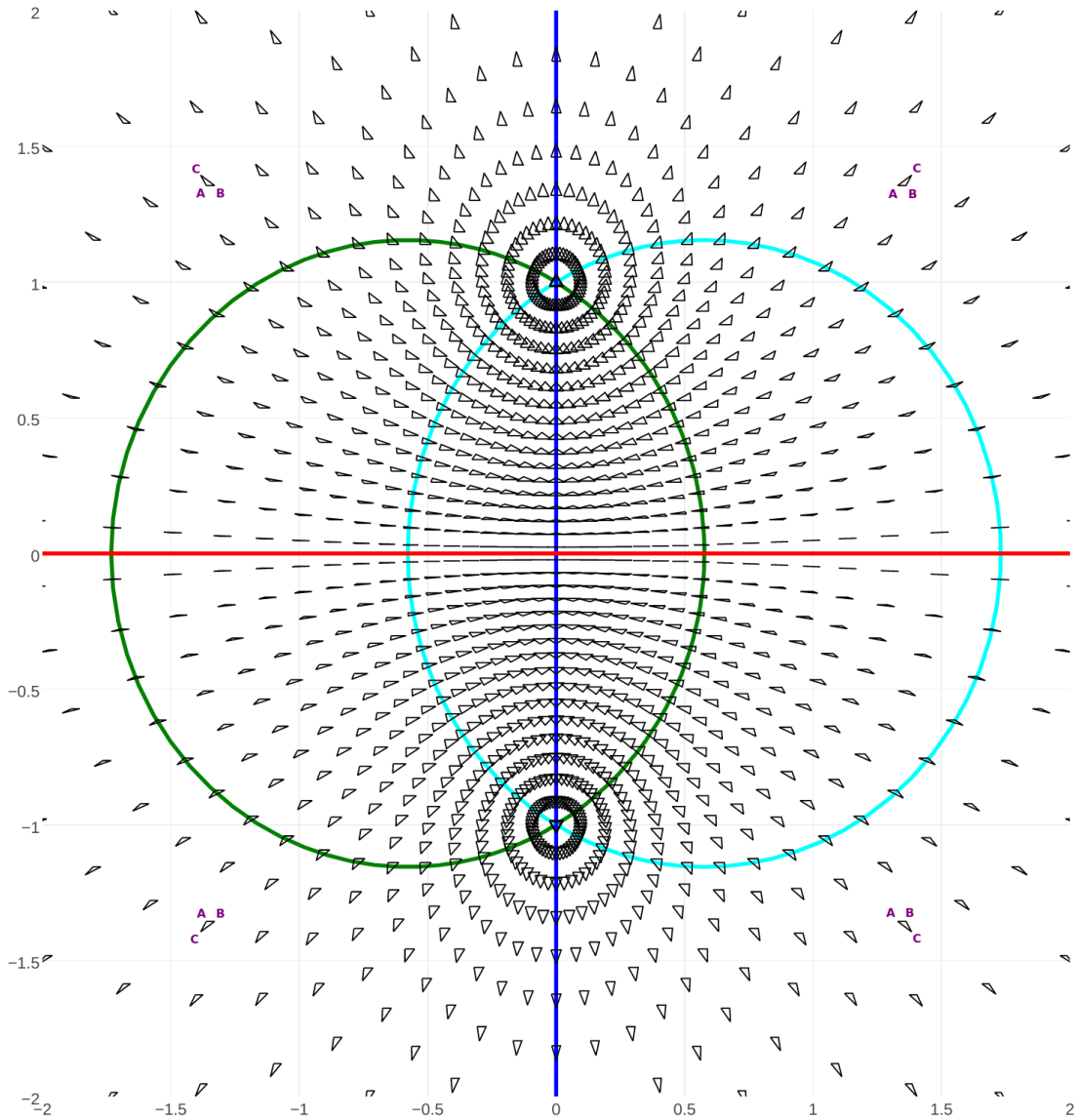


FIGURE 8.5 – L'espace des triangles à similitude près, vu dans le plan des ratios  $\xi \in \mathbb{C} \cup \{\infty\}$ . Dessiné en rouge, l'axe des abscisses correspond aux triangles plats. Les triangles isocèles se répartissent selon trois courbes, qui dépendent du sommet privilégié : en bleu, l'axe des ordonnées correspondant aux triangles isocèles en  $C$  ; en vert la courbe des triangles isocèles en  $A$ , en cyan celle des triangles isocèles en  $B$ . Le point correspondant au ratio  $\xi = 0$  est ici représenté à l'origine du repère.

*Démonstration.* Reprenons les membres de notre équation. À gauche et au centre, un réel défini de manière implicite :

$$d_{\text{Procuste}}(ABC \rightarrow DEF) = \min_{\tau, v \in \mathbb{C}} \sqrt{\sum_{i=1}^3 |v \cdot z_i^* + \tau - w_i^*|^2}. \quad (8.20)$$

À droite, une distance :

$$W \cdot \|f(\xi) - f(\xi')\|_{\mathbb{R}^3} \quad (8.21)$$

$$= W \cdot \left\| \frac{1}{2}(\cos \theta, \sin \theta \cos \varphi, \sin \theta \sin \varphi) - \frac{1}{2}(\cos \theta', \sin \theta' \cos \varphi', \sin \theta' \sin \varphi') \right\|_{\mathbb{R}^3} \quad (8.22)$$

$$= W \cdot \frac{1}{2} \sqrt{(\cos \theta - \cos \theta')^2 + (\sin \theta \cos \varphi - \sin \theta' \cos \varphi')^2 + (\sin \theta \sin \varphi - \sin \theta' \sin \varphi')^2} \quad (8.23)$$

où  $W$ ,  $\theta$ ,  $\theta'$ ,  $\varphi$  et  $\varphi'$  sont définis à partir de  $z^*$  et  $w^*$ .

Pour identifier ces deux quantités, on procède en deux temps : d'abord, on explicite le terme de gauche pour aboutir à une fonction de  $z$  et  $w$  ; ensuite, on développe cette expression dans les coordonnées sphériques  $(\theta, \varphi)$  jusqu'à voir apparaître le terme de droite.

**Passage aux distances quadratiques** Commençons par rappeler que  $x \mapsto \sqrt{x}$  est une fonction positive et strictement croissante de  $x$ . Il suffit donc de démontrer notre égalité sur les carrés des distances mises en jeu, i.e. montrer que :

$$\min_{\tau, v \in \mathbb{C}} \left( \sum_{i=1}^3 |v \cdot z_i^* + \tau - w_i^*|^2 \right) \quad (8.24)$$

$$= W^2 \cdot \frac{1}{4} \left( (\cos \theta - \cos \theta')^2 + (\sin \theta \cos \varphi - \sin \theta' \cos \varphi')^2 + (\sin \theta \sin \varphi - \sin \theta' \sin \varphi')^2 \right) \quad (8.25)$$

**Calcul du minimum procustéen** Explicité à gauche de l'équation précédente,  $\|S_{\tau, v}(z^*) - w^*\|_2^2$  est une fonction positive et continue des variables complexes  $\tau$  et  $v$ , à  $z^*$  et  $w^*$  fixés. Or on voit qu'elle diverge vers  $+\infty$  lorsque  $|\tau|$  ou  $|v|$  s'éloigne à l'infini. Par un raisonnement de compacité on sait qu'il existe un couple de paramètres optimaux  $(\tau_0, v_0)$  qui réalise le minimum.

**La similitude optimale recalcule les barycentres** Reste à le caractériser. Pour cela, remarquons que pour tout accroissement  $d\tau$ , on a :

$$\|S_{\tau_0 + d\tau, v_0}(z^*) - w^*\|_2^2 \geq \|S_{\tau_0, v_0}(z^*) - w^*\|_2^2. \quad (8.26)$$

[Ken84]

Basile Coron,

Paul Froux,

Massil Hihat,

Lino Benedetto

Élémentaire

et élégant

Or on peut calculer que :

$$\|S_{\tau_0+d\tau, v_0}(z^*) - w^*\|_2^2 = \sum_{i=1}^3 |v_0 \cdot z_i^* + \tau_0 + d\tau - w_i^*|^2 \quad (8.27)$$

$$= \sum_{i=1}^3 (v_0 \cdot z_i^* + \tau_0 + d\tau - w_i^*) \overline{(v_0 \cdot z_i^* + \tau_0 + d\tau - w_i^*)} \quad (8.28)$$

$$= \|S_{\tau_0, v_0}(z^*) - w^*\|_2^2 + 3|d\tau|^2 \quad (8.29)$$

$$+ \sum_{i=1}^3 d\tau \cdot \overline{(v_0 \cdot z_i^* + \tau_0 - w_i^*)} + \overline{d\tau} \cdot (v_0 \cdot z_i^* + \tau_0 - w_i^*) \quad (8.30)$$

$$= \|S_{\tau_0, v_0}(z^*) - w^*\|_2^2 + 3|d\tau|^2 \quad (8.31)$$

$$+ 2\Re d\tau \cdot \sum_{i=1}^3 \overline{(v_0 \cdot z_i^* + \tau_0 - w_i^*)}, \quad (8.32)$$

qui est le développement en  $d\tau$  du coût minimisé au voisinage de l'optimum  $(\tau_0, v_0)$ . Par suite de l'inégalité (8.26), le terme linéarisé d'ordre 1 doit être positif pour toute valeur assez faible de  $d\tau$  : ceci n'est possible que si on a *annulation de la dérivée*,

$$\sum_{i=1}^3 (v_0 \cdot z_i^* + \tau_0 - w_i^*) = 0. \quad (8.33)$$

Autrement dit, la condition de minimisation en  $\tau_0$  s'écrit à l'ordre 1 :

$$v \cdot \frac{1}{3} \sum_{i=1}^3 z_i^* + \tau_0 = \frac{1}{3} \sum_{i=1}^3 w_i^*, \quad (8.34)$$

i.e. «  $S_{\tau_0, v_0}$  recale les deux barycentres ». C'est un résultat que nous avons utilisé de manière intuitive dans la Figure 8.2b.

**Passage aux coordonnées barycentriques** On a vu aux équations (8.11-8.12) que l'on pouvait écrire la dissimilarité procustéenne en coordonnées barycentriques :

$$\|S_{\tau, v}(z^*) - w^*\|_2^2 = \|S_{\tau, v}(z) - w\|_2^2 \quad (8.35)$$

$$= |v \cdot z_0 + \sqrt{3}\tau - w_0|^2 + |v \cdot z_1 - w_1|^2 + |v \cdot z_2 - w_2|^2. \quad (8.36)$$

Or nous venons de montrer que la similitude optimale recale les barycentres :

$$v_0 \cdot z_0 + \sqrt{3}\tau_0 = w_0. \quad (8.37)$$

Pour trouver le coût procustéen  $d_{\text{Procuste}}^2(ABC \rightarrow DEF)$ , il suffit donc de minimiser en  $v$  l'expression :

$$\|S_{\tau_0, v}(z^*) - w^*\|_2^2 = |v \cdot z_1 - w_1|^2 + |v \cdot z_2 - w_2|^2. \quad (8.38)$$

**Un coût explicite** Comme au paragraphe précédent, on écrit :

$$\|S_{\tau_0, v_0 + dv}(z^*) - w^*\|_2^2 = \sum_{i=1}^2 ((v + dv) \cdot z_i - w_i) \cdot \overline{((v + dv) \cdot z_i - w_i)} \quad (8.39)$$

$$= \|S_{\tau_0, v_0}(z^*) - w^*\|_2^2 + |dv|^2 \cdot (|z_1|^2 + |z_2|^2) \quad (8.40)$$

$$+ \sum_{i=1}^2 dv \cdot z_i \cdot \overline{(v \cdot z_i - w_i)} + \overline{dv \cdot z_i} \cdot (v \cdot z_i - w_i). \quad (8.41)$$

La condition d'annulation de la dérivée en  $v_0$  est donc :

$$\sum_{i=1}^2 \overline{z_i} \cdot (v_0 \cdot z_i - w_i) = 0, \quad \text{i.e.} \quad v_0 \cdot \sum_{i=1}^2 |z_i|^2 = \sum_{i=1}^2 \overline{z_i} w_i. \quad (8.42)$$

À l'optimum, on a alors :

$$\sum_{i=1}^2 \overline{(v_0 z_i)} \cdot (v_0 z_i) = \sum_{i=1}^2 \overline{(v_0 z_i)} \cdot w_i, \quad (8.43)$$

ce qui nous permet de conclure :

$$d_{\text{Procuste}}^2(ABC \rightarrow DEF) = \sum_{i=1}^2 \overline{(v_0 z_i - w_i)} \cdot (v_0 z_i - w_i) \quad (8.44)$$

$$= \sum_{i=1}^2 \overline{(v_0 z_i)} \cdot (v_0 z_i) - \overline{(v_0 z_i)} \cdot w_i - \overline{w_i} \cdot (v_0 z_i) + \overline{w_i} \cdot w_i \quad (8.45)$$

$$= \sum_{i=1}^2 \overline{w_i} \cdot (w_i - v_0 z_i) \quad (8.46)$$

$$= \sum_{i=1}^2 \overline{w_i} w_i - v_0 \sum_{i=1}^2 \overline{w_i} z_i \quad (8.47)$$

$$= \sum_{i=1}^2 \overline{w_i} w_i - \frac{|\sum_{i=1}^2 \overline{w_i} z_i|^2}{\sum_{i=1}^2 |z_i|^2} \quad (8.48)$$

$$= \left( \sum_{i=1}^2 |w_i|^2 \right) \cdot \left( 1 - \frac{|\sum_{i=1}^2 \overline{w_i} z_i|^2}{\left( \sum_{i=1}^2 |z_i|^2 \right) \cdot \left( \sum_{i=1}^2 |w_i|^2 \right)} \right). \quad (8.49)$$

**Un coût fonction des ratios complexes  $\xi$  et  $\xi'$**  À l'aide de deux conditions d'annulation de la dérivée, nous avons pu résoudre notre problème de minimisation : trouver la similitude optimale pour recaler  $ABC$  sur  $DEF$  n'était finalement pas si difficile.

Maintenant, nous allons chercher à exprimer le coût irréductible du résiduel donné équation (8.49) en fonction des coordonnées sphériques  $(\theta, \varphi)$  et  $(\theta', \varphi')$ . On a convenu de se doter de ratios  $\xi$  et  $\xi'$  tels que :

$$z_2 = \xi z_1 \quad \text{et} \quad w_2 = \xi' w_1. \quad (8.50)$$

[Ken84]

Basile Coron,

Paul Fraux,

Massil Hihat,

Lino Benedetto

Élémentaire

et élégant

On peut donc écrire :

$$d_{\text{Procuste}}^2(ABC \rightarrow DEF) = \left( \sum_{i=1}^2 |w_i|^2 \right) \cdot \left( 1 - \frac{|\sum_{i=1}^2 \bar{w}_i z_i|^2}{\left( \sum_{i=1}^2 |z_i|^2 \right) \cdot \left( \sum_{i=1}^2 |w_i|^2 \right)} \right) \quad (8.51)$$

$$= (|w_1|^2 + |w_2|^2) \cdot \left( 1 - \frac{|\bar{w}_1 z_1 + \bar{w}_2 z_2|^2}{(|z_1|^2 + |z_2|^2) \cdot (|w_1|^2 + |w_2|^2)} \right) \quad (8.52)$$

$$= (|w_1|^2 + |w_2|^2) \cdot \left( 1 - \frac{|\bar{w}_1 z_1 + \bar{w}_1 z_1 \bar{\xi}' \xi|^2}{|z_1|^2 (1 + |\xi|^2) \cdot |w_1|^2 (1 + |\xi'|^2)} \right) \quad (8.53)$$

$$= (|w_1|^2 + |w_2|^2) \cdot \left( 1 - \frac{|1 + \bar{\xi}' \xi|^2}{(1 + |\xi|^2) \cdot (1 + |\xi'|^2)} \right). \quad (8.54)$$

En notant  $W^2 = |w_1|^2 + |w_2|^2$  la variance du triangle  $DEF$ , facteur d'échelle, on trouve :

$$\frac{1}{W^2} d_{\text{Procuste}}^2(ABC \rightarrow DEF) = 1 - \frac{|1 + \bar{\xi}' \xi|^2}{(1 + |\xi|^2) \cdot (1 + |\xi'|^2)}. \quad (8.55)$$

**En coordonnées sphériques** Pour conclure, il suffit alors de s'armer de patience, d'une fiche de formules trigonométriques, et de se souvenir que

$$\xi = \tan(\theta/2) e^{i\varphi}, \quad \xi' = \tan(\theta'/2) e^{i\varphi'}. \quad (8.56)$$

On trouve :

$$\frac{1}{W^2} d_{\text{Procuste}}^2(ABC \rightarrow DEF) = 1 - \frac{|1 + \tan(\theta/2) \tan(\theta'/2) e^{i(\varphi - \varphi')}|^2}{(1 + \tan^2(\theta/2)) \cdot (1 + \tan^2(\theta'/2))} \quad (8.57)$$

$$= \frac{\tan^2(\theta/2) + \tan^2(\theta'/2) - 2 \tan(\theta/2) \tan(\theta'/2) \cos(\varphi - \varphi')}{(1 + \tan^2(\theta/2)) \cdot (1 + \tan^2(\theta'/2))} \quad (8.58)$$

$$= \frac{1}{2} \cdot (1 - \cos \theta \cos \theta' - \sin \theta \sin \theta' \cos(\varphi - \varphi')) \quad (8.59)$$

$$= \frac{1}{4} \cdot ((\cos \theta - \cos \theta')^2 + (\sin \theta \cos \varphi - \sin \theta' \cos \varphi')^2 + (\sin \theta \sin \varphi - \sin \theta' \sin \varphi')^2) \quad (8.60)$$

$$= \|f(\xi) - f(\xi')\|_{\mathbb{R}^3}^2. \quad (8.61)$$

En passant aux racines carrées, on a bien retrouvé l'énoncé du théorème :

$$d_{\text{Procuste}}(ABC \rightarrow DEF) = W \|f(\xi) - f(\xi')\|_{\mathbb{R}^3}. \quad (8.62)$$

À la variance du triangle d'arrivée près, la sphère de rayon 1/2 est bien le bon espace pour calculer les distances procustéennes entre triangles.  $\square$



## 8.3 Généralisation

### 8.3.1 Définition formelle

Nous allons maintenant essayer de généraliser la construction précédente. Considérons une forme à  $k$  points déterminée par les vecteurs de  $\mathbb{R}^m$   $(z_j^*)_{j \in 1 \dots k}$  de barycentre  $z_c^*$ . On autorise les dégénérescences à l'exception du polygone effondré en un point, (tous les  $z_j$  égaux).

Nous allons tout d'abord négliger l'effet des translations. Pour cela, comme dans le cas des triangles, on commence par chercher une base orthonormale qui décrit bien le problème. Puisque seules nous intéressent les positions relatives au barycentre, on considère les coordonnées obtenues par la transformation orthogonale suivante :

$$z_0 = z_c^* \sqrt{k} \quad z_j = \frac{jz_{j+1}^* - (z_1^* + \dots + z_j^*)}{\sqrt{j^2 + j}}. \quad (8.63)$$

Ainsi à partir de maintenant on ignorera  $z_0$  et on essaye de trouver une construction pour  $z = (z_1, \dots, z_{k-1})$  indépendante des autres changements de bases orthonormales.

On définit alors

$$\Sigma_m^k = \{(\mathbb{R}^m)^{k-1} \setminus \{0\}\} / \text{Sim} \quad (8.64)$$

où Sim est le groupe des similitudes engendré par les rotations et les dilatations (non nulles). Le  $0 \in (\mathbb{R}^m)^{k-1}$  est exclu car nous ne considérons pas la forme du  $k$ -polygone totalement dégénéré.

### 8.3.2 Topologie et norme

Pour  $z \in (\mathbb{R}^m)^{k-1}$ , on définit sa norme par :

$$\|z\|^2 = \sum_{j=1}^{k-1} |z_j|^2 = \sum_{i=1}^k |z_i^*|^2 - k|z_c^*|^2 = \sum_{i=1}^k |z_i^* - z_c^*|^2 \quad (8.65)$$

où  $|\cdot|$  est la norme euclidienne de  $\mathbb{R}^m$ . On remarque que cette norme est canonique, au sens où elle ne dépend plus de la transformation orthogonale définie en 8.63.

L'espace quotient  $\Sigma_m^k$  muni la topologie quotient héritée de cette norme, sera appelé *l'espace des formes d'ordre  $(k, m)$*

Pour des raisons pratiques, on considérera le quotient définissant  $\Sigma_m^k$ , comme effectué d'abord par les dilatations, ramenant  $(\mathbb{R}^m)^{k-1}$  à  $S^{m(k-1)-1}$  noté  $S_m^k$ , ensuite par les rotations  $SO(m)$  agissant à gauche de  $(\mathbb{R}^m)^{k-1}$  (identifié à  $M_{m, k-1}$ ), on a donc :

$$\Sigma_m^k = S_m^k / SO(m) \quad (8.66)$$

On note  $p$  la projection de  $S_m^k \rightarrow \Sigma_m^k$ , alors un argument de théorie des groupes montre que la métrique de  $\Sigma_m^k$  est exactement celle engendrée par la distance  $\rho$  :

$$\rho(p(W), p(W_0)) = \inf_{R \in SO(m)} (d(RW, W_0)) \quad (8.67)$$

et  $d$  est la métrique induite par les calculs de géodésique sur  $S_m^k$  (en particulier,  $d$ , et donc  $\rho$ , est majoré par  $\pi$ ).

[Ken84]

Basile Coron,

Paul Fiaux,

Massil Hihat,

Lino Benedetto

Élémentaire

et élégant

**Groupe de transformations** On veut construire un groupe  $G$  d'isométrie sur  $\Sigma_m^k$ . Vu comme sous-espace de dans  $M_{m,k-1}$ , on a une action à gauche sur  $S_m^k$ , celle de  $SO(m)$ , par rotation. Néanmoins  $O(k-1)$  agit aussi à droite (par permutation des points entre autres, ou comme dans la formule 8.63), ces deux actions commutent et l'action de  $O(k-1)$  engendre sur  $\Sigma_m^k$  un groupe d'isométrie  $G$ .

Si  $m$  est pair, on peut augmenter le groupe  $G$  de symétries en

$$G^+ = G \otimes \{1, \tau\}, \quad (8.68)$$

où  $\tau$  représente la multiplication à gauche par  $\text{diag}(1, 1, \dots, 1, -1)$  et on appelle l'action de  $\tau$  sur  $\Sigma_m^k$  la conjugaison, nous reverrons bientôt cette conjugaison dans un cadre plus familier.

**Une idée sur la classification** Les espaces de la forme  $\Sigma_m^{m+1}$  que l'on appellera *espace de forme simplicial*, sont plus important qu'il n'y parait. Une telle forme détermine un simplexe de  $\mathbb{R}^m$ . A. J. Casson à prouver que tous ces espaces sont homéomorphes à des sphères.

En fait tous, les espaces  $\Sigma_m^k$  avec  $k < m + 1$  que l'on appellera *espaces de formes sur-dimensionné* se ramènent aux cas des espaces simpliciaux  $\Sigma_{k-1}^k$ . En effet, modulo  $SO(m)$  dans un tel espace sur-dimensionné, on peut transformer nos formes (vues dans  $M_{m,k-1}$ ), de sorte à ce qu'elles ait les  $m - (k - 1)$  dernières lignes nulles. L'action de conjugaison  $\tau$  sur les formes de  $\Sigma_{k-1}^k$  fait correspondre celles-ci aux formes de  $\Sigma_m^k$ , par un processus de déconjugaison, on peut donc revenir à  $\Sigma_{k-1}^k$ .

## 8.4 Etude des formes du plan, $\Sigma_2^k$

On s'intéresse maintenant, au cas des formes du plan, donc à décrire  $\Sigma_2^k$ . Il nous vient alors la même idée que dans l'étude des triangles, on peut représenter  $(\mathbb{R}^2)^{k-1}$  par  $\mathbb{C}^{k-1}$  et les similitudes sont alors la multiplication par un complexe non nul. L'espace résultant est  $\mathbb{C}^{k-1} - 0$  quotienté par les homothétie non nulles, c'est à dire l'espace projectif  $\mathbb{C}P^{k-2}$ .

On a alors (à un changement d'échelle près, que l'on reverra très bientôt) :

$$\Sigma_2^k = \mathbb{C}P^{k-2} \quad (8.69)$$

**Plus de transformations** Remarquons déjà que par cette identification, la conjugaison par  $\tau$  que l'on a défini précédemment correspond bien à la conjugaison complexe. En fait, voir notre espace de forme, comme un espace complexe nous permet d'enrichir encore plus notre groupe de similitudes  $G^+$ .

En effet, on peut l'étendre en un groupe  $G^{++}$  en ajoutant les éléments de  $U(k-1)/U(1)$  (quotient du groupe unité par son centre, les homothéties de rapport non nul de  $\mathbb{C}^{k-1}$ ). Ainsi, on a que  $\Sigma_2^k$  est homogène car  $G^{++}$  agit alors transitivement. C'est un résultat qui ne se vérifiera pas en dimension  $m \geq 3$ .

**Remarques** Notons, au passage, qu'il vient naturellement que  $\Sigma_2^2 = \mathbb{C}P^0 = \{.\}$ , cette dernière identification peut, par ailleurs, se voir tout aussi bien géométriquement. Aussi,  $\Sigma_2^3 = \mathbb{C}P^1 = S^2$ , ce qui semble déjà nous indiquer que l'espace des triangle est une sphère, nous avons déjà vu une preuve de la nature métrique de cette identification, nous développerons dans la suite d'autres arguments pour cela.

### 8.4.1 Analyse procrustéenne et distance sur $\Sigma_2^k$

Soit  $z$  (ou  $z^*$  si on utilise les coordonnées initiales) représentant un  $k$ -polygone dans  $\mathbb{R}^2$ . On note  $[z]$  ou  $[z^*]$  le point correspondant de  $\Sigma_2^k$ . On suppose  $k \geq 3$ . On introduit la mesure d'écart entre deux  $k$ -uplets de complexes  $(z, w)$  de barycentre  $b$  :

$$\frac{\sum_i |z_i - w_i|^2}{\sum_i |z_i - b|^2 + \sum_i |w_i - b|^2} \quad (8.70)$$

On veut obtenir le minimum de la mesure d'écart entre les  $k$ -polygones  $\alpha z^* + \lambda$  et  $\beta y^* + \mu$ . On peut supposer sans perte de généralité que  $z^*$  et  $w^*$  sont centrés en 0. On cherche donc à minimiser :

$$\frac{\sum_i |\alpha z_i - \beta w_i + (\lambda - \mu)|^2}{\sum_i |\alpha z_i + (\lambda - \mu)/2|^2 + \sum_i |\beta w_i - (\lambda - \mu)/2|^2} \quad (8.71)$$

et on trouve finalement :

$$\Delta([z^*], [w^*]) = 1 - \frac{|\sum_i (z_i^* - z_c^*)|}{\sqrt{\sum_i |z_i^* - z_c^*|^2 \sum_j |w_j^* - w_c^*|^2}} \quad (8.72)$$

Où encore avec  $Z$  et  $W$  normalisés :

$$\Delta([Z], [W]) = 1 - \left| \sum_1^{k-1} Z_j \overline{W_j} \right| \quad (8.73)$$

**Analyse procrustéenne :** On peut reconnaître dans ce calcul un exemple d'analyse procrustéenne, à la différence qu'ici nous ne supprimons pas les réflexions ce qui introduit d'importantes différences dans la formule.

On vérifie facilement que  $\Delta$  est définie positive et symétrique. On pose :

$$\rho = \arccos(1 - \Delta) \quad (8.74)$$

et on vérifie que c'est une distance avec un maximum atteint de  $\frac{\pi}{2}$  et que dans le cas où  $m = 2$  c'est bien la fonction définie dans la partie 8.67.

On a alors de nombreuses distances équivalentes à  $\rho$  comme  $\frac{\rho}{1+\rho}$  ou encore  $c \sin(\frac{\rho}{c})$  avec  $2 \leq c$  (le cas  $c = 2$  conduit même à  $\sqrt{2\Delta}$ ). Mais la fonction  $\rho$  est la bonne puisqu'il s'agit d'une définition alternative de la distance quotient.

Pour obtenir la métrique Riemannienne associée il suffit alors de faire un développement asymptotique de  $\cos(\rho)^2$  en utilisant la carte  $p$  de  $S_2^k$  dans  $\Sigma_2^k$ .

On trouve alors :

$$ds^2 = \frac{(\sum_1^{k-1} z_j \overline{z_j})(\sum_1^{k-1} dz_j d\overline{z_j}) - (\sum_1^{k-1} z_j d\overline{z_j})(\sum_1^{k-1} \overline{z_j} dz_j)}{(\sum_1^{k-1} z_j \overline{z_j})^2}, \quad (8.75)$$

où encore sur une surface telle que  $z_1 \neq 0$ , avec  $\zeta_j = \frac{z_j}{z_1}$  :

$$ds^2 = \frac{(1 + \sum_2^{k-1} \zeta_j \overline{\zeta_j})(\sum_2^{k-1} d\zeta_j d\overline{\zeta_j}) - (\sum_2^{k-1} \zeta_j d\overline{\zeta_j})(\sum_2^{k-1} \overline{\zeta_j} d\zeta_j)}{(1 + \sum_2^{k-1} \zeta_j \overline{\zeta_j})^2}. \quad (8.76)$$

[Ken84]

Basile Coron,

Paul Fraux,

Massil Hihat,

Lino Benedetto

Élémentaire

et élégant

Intéressons nous aux triangles dans le plan ( $m = 2, k = 3$ ). D'après (8.76) nous savons qu'il existe un système de coordonnées sur  $\Sigma_2^3$  privé d'un point tel que :

$$ds^2 = \frac{(1 + |\zeta|^2)|d\zeta|^2 - |\zeta d\zeta|^2}{(1 + |\zeta|^2)^2}. \quad (8.77)$$

Si on pose  $\zeta = re^{i\varphi}$  et  $r = \tan(\frac{\theta}{2})$ , on obtient sur  $\Sigma_2^3$  privé d'un point :

$$ds^2 = \frac{1}{4}(d\theta^2 + \sin^2(\theta)d\varphi^2). \quad (8.78)$$

**Théorème 8.2.**  $\Sigma_2^3$  est isométrique à  $S^2(\frac{1}{2})$

### 8.4.2 Détails géométriques de la carte

Si on se donne ABC un triangle, on cherche le point de  $S^2(\frac{1}{2})$  qui correspond à sa forme.

On se place dans le cas où  $\frac{|AB|}{2} = 1$  et  $\vec{AB}$  est la direction de référence. On note alors le vecteur entre le centre de AB et C pas  $Me^{i\psi}$ . Alors en posant  $r = \frac{M}{\sqrt{3}}$ , on a :

$$X = \frac{1}{2} \cos(\theta) = \frac{1 - r^2}{2(1 + r^2)} \quad (8.79)$$

$$Y = \frac{1}{2} \sin(\theta) \cos(\psi) = \frac{2r \cos(\psi)}{2(1 + r^2)} \quad (8.80)$$

$$Z = \frac{1}{2} \sin(\theta) \sin(\psi) = \frac{2r \sin(\psi)}{2(1 + r^2)} \quad (8.81)$$

Pour faciliter les calculs on retire le facteur  $\frac{1}{2}$

Inversement, si on se donne un point de la sphère, on peut facilement calculer  $(r, \psi)$  et alors un triangle représentant la forme sera :

$$A = -\sqrt{3} - re^{i\psi} \quad (8.82)$$

$$B = \sqrt{3} - re^{i\psi} \quad (8.83)$$

$$C = 2re^{i\psi} \quad (8.84)$$

### 8.4.3 Action de $G^{++}$ sur la sphère

Dans  $G^{++}$  on a un sous groupe d'ordre 6  $G_6$  qui permute les noms des points : Si on veut regarder la forme de BCA à la place de celle de ABC, on doit faire une rotation sur la sphère autour de l'axe OZ d'angle  $+\frac{2\pi}{3}$ . Si on veut regarder celle de BAC à la place de ABC, on doit passer de  $(X, Y, Z)$  à  $(X, -Y, -Z)$

Or dans la plupart des cas pratiques ou théoriques les permutations de sommets sont sans importances.  $S^2(\frac{1}{2})$  est alors découpé en 6 lunes équivalentes. Il peut donc être utile de ne considérer les formes que modulo une réindexation des sommets, ie  $S^2(\frac{1}{2})/G_6$  qui est *topologiquement* équivalent à une sphère. Néanmoins nous sommes obligés de continuer à travailler dans  $S^2(\frac{1}{2})$  si nous voulons une isométrie. Si en plus les effets d'une réflexion ne sont

## Kendall sphere

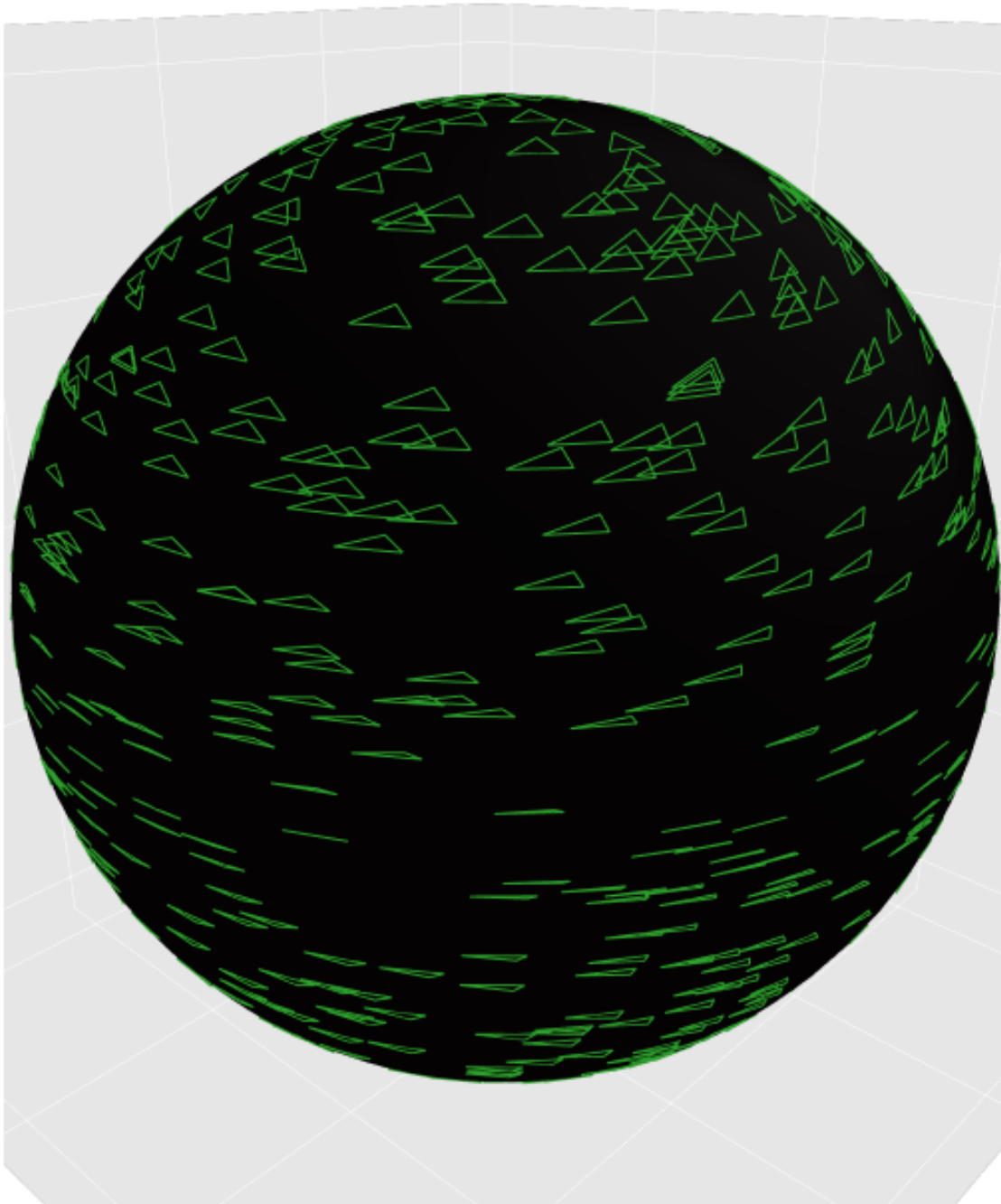


FIGURE 8.6 – L'espace des triangles à similitude près, muni de la distance quotient / procrustéenne. Il est isométrique à la sphère de rayon  $1/2$ . Les triangles équilatéraux se trouvent aux pôles, et les triangles plats à l'équateur. Notez que l'on n'a pas ici quotienté par les symétries ou l'étiquetage des points (dans le groupe symétrique  $S_3$ ) : la même "forme" au sens des rapports de longueurs se retrouve donc 12 fois, sauf exception. L'espace quotient résultant, le "tableau noir sphérique" est montré dans la figure suivante.

[Ken84]

Basile Coron,  
Paul Froux,  
Massil Hihat,  
Lino Benedetto  
Élémentaire  
et élégant

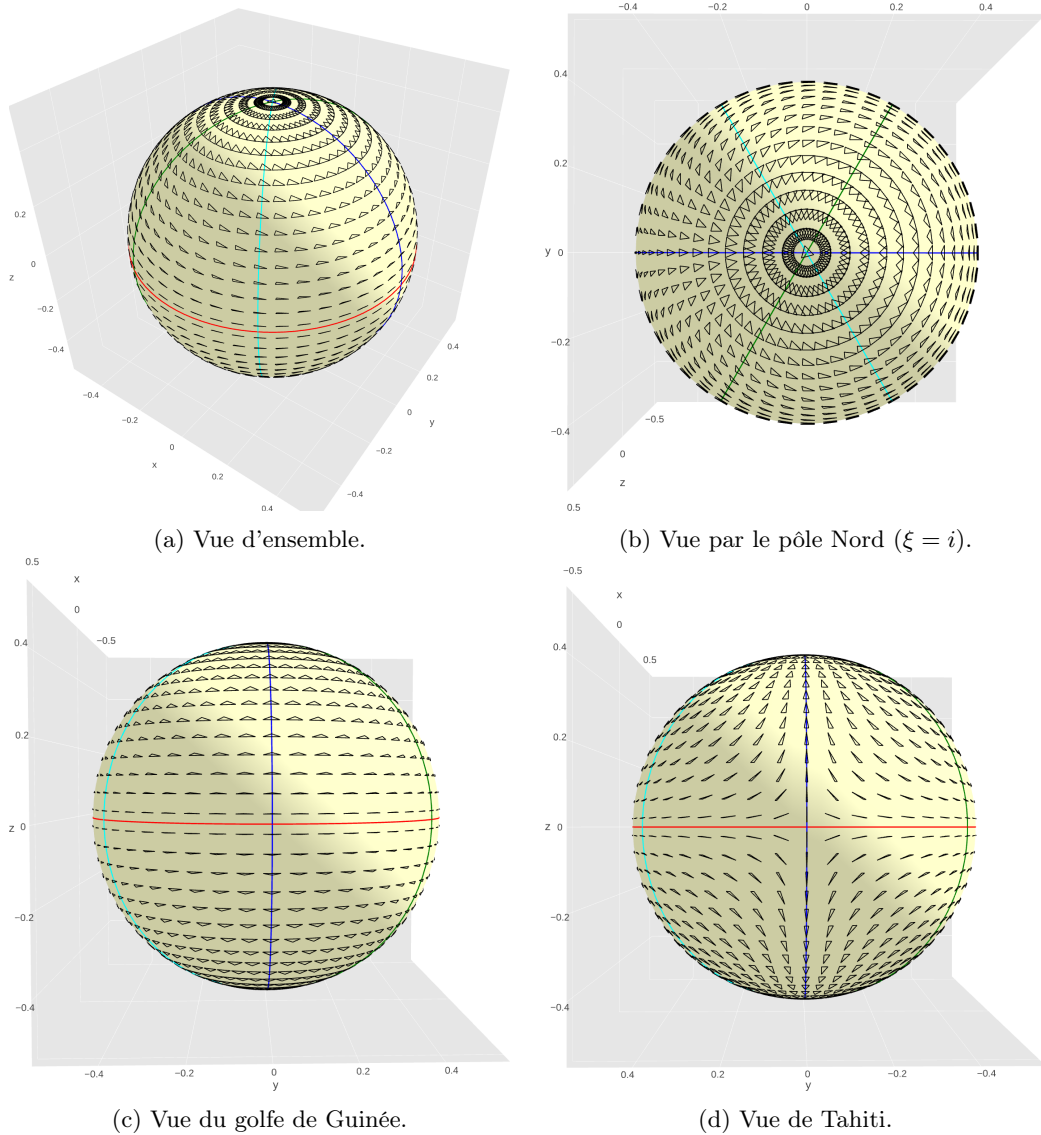


FIGURE 8.7 – L'espace des triangles à similitude près, isométrique à la sphère de rayon  $1/2$ . Homogène, cette représentation permet de rendre compte des symétries intrinsèques de l'espace des triangles liées au ré-étiquetage des points et aux réflexions. Ici, chaque "forme" au sens des rapports de longueurs se retrouve dans les 12 quartiers découpés par l'équateur des triangles plats et les méridiens des triangles isocèles.

pas significatif, on peut considérer le sous groupe d'ordre 12  $G_{12}$  de  $G^{++}$  obtenu par adjonction de l'opérateur de conjugaison à  $G_6$ . on considère alors la demi-lune supérieure :  $S^2(\frac{1}{2})/G_{12}$  qui sera par la suite appelée *tableau sphérique (spherical blackboard)*

On se place alors dans la demi-lune délimitée par le triangle sphérique LMN qui sont les résultats des projections de :

$$L = (r = \frac{1}{\sqrt{3}}, \psi = \pi), \text{ ie } A = C \quad (8.85)$$

$$M = (r = 0, \psi \text{ quelconque}), \text{ ie } C = \frac{A+B}{2} \quad (8.86)$$

$$N = (r = 1, \psi = \frac{\pi}{2}), \text{ ie } AB = AC = BC \quad (8.87)$$

Alors les formes colinéaires sont celles sur l'arc LM et les formes isocèles celles sur les arcs MN et LN.

Avec une forme ABC, on peut alors s'intéresser à la fonction max-angle qui donne l'angle maximal du triangle et regarder quand est-ce qu'elle est proche de  $180^\circ$  (pour alors être presque alignée).

## 8.5 Statistiques sur la sphère

La sphère de Kendall a donc un statut *géométrique* privilégié, en montrant qu'elle capture parfaitement la notion de distance (procustéenne) entre triangles. Cette "canonicité" rejaillit sur la structure intrinsèque de la sphère (distance géodésique, surface, symétries) qui se retrouve tout à coup légitimée : elle n'est pas le fruit du choix arbitraire telles les formules de l'équation (8.16), ou plus généralement des projections stéréographiques issus de l'identification de  $\mathbb{C}P^1$  à  $S^2$ , mais bien du passage au quotient par les similitude de la distance "des moindres carrées", canonique et structurée sur l'espace des coordonnées de points.

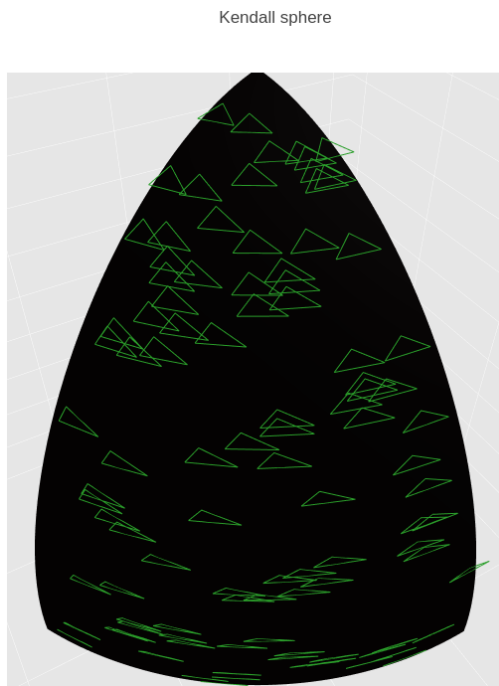
Toutes ces propriétés restent néanmoins cantonnées aux seules questions de distances, de géométrie pure. Essayons de voir ce qui se passe quand les probabilités s'en mêle.

### 8.5.1 Mesures sur l'espace des formes

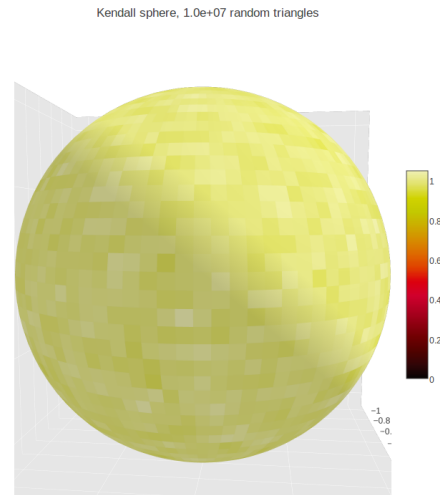
On cherche à munir  $\Sigma_m^k$  d'une mesure de probabilité et pour cela on pense au transfert. On remarque en effet que la projection définie plus haut (8.66),  $p : S_m^k \rightarrow \Sigma_m^k$  est mesurable sur les tribus boréliennes.

Si on se donne  $P_1, \dots, P_k \in \mathbb{R}^m$  qui forme un polygone aléatoire et tel que  $\text{proba}(P_1 = P_2 = \dots = P_k) = 0$  alors la forme  $\sigma$  du k-polygone sera une variable aléatoire de  $\Sigma_m^k$  distribuée suivant une loi des formes déterminée par la loi conjointe de  $(P_1, \dots, P_k)$ .

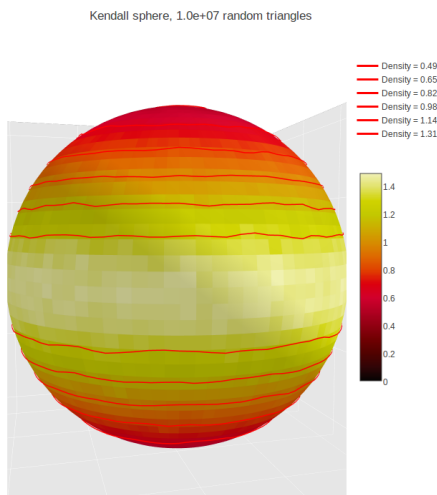
Un des cas les plus intéressant est celui où les points  $(P_i)$  sont indépendamment et identiquement distribués dans  $\mathbb{R}^m$  de loi individuelle  $\mu$ . On note  $\mu^k$  la loi conjointe, elle induit sur  $\Sigma_m^k$  une mesure image  $\mu_*^k = \mu^k p^{-1}$  qui va être inchangée si l'on modifie les paramètres de  $\mu$  contrôlant la position, la taille ou les rotations autour d'un axe. On note  $[\mu]$  la classe d'équivalence de  $\mu$  suivant de telles transformations.



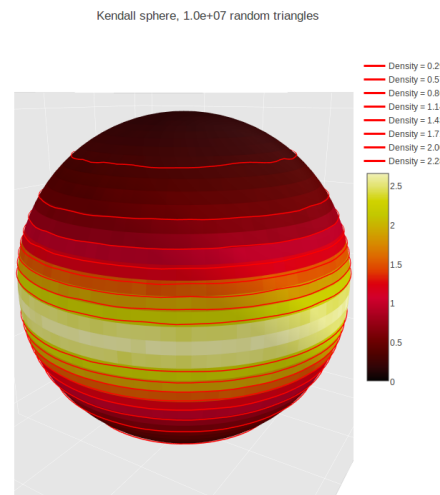
(a) Le douzième de sphère, qui contient tous les triangles à similitude, permutation des points et symétrie près. Les triangles plats sont sur le bord du bas, et les triangles isocèles, sur les deux côtés.



(b) Histogramme empirique sur la sphère, lorsque les points des triangles sont tirés de manière i.i.d. selon une gaussienne isotrope dans le plan. Comme prévu, la densité empirique est uniforme en bonne approximation.



(c) En remplaçant la gaussienne isotrope par une gaussienne anisotrope (ici,  $\sigma_1 = 3, \sigma_2 = 1$ ), on perd l'uniformité. Par contre, et c'est un fait remarquable, la mesure image est une fonction de la seule altitude  $Z$ .



(d) Enfin, lorsque l'anisotropie devient importante (ici,  $\sigma_1 = 6, \sigma_2 = 1$ ), la masse se concentre autour de l'équateur des triangles plats, ce qui correspond bien à notre intuition : les triangles équilatéraux, aux pôles, se font rares !



Beaucoup de questions se pose après avoir définis ces objets pour une vision plus exhaustive on consultera l'article original de Kendall; *Shape Manifolds, Procrustean Metrics and Complex Projective Spaces* (1984).

**La mesure gaussienne des formes** Pour  $m = 2$ , la mesure  $\gamma_*^k$  associée à la distribution gaussienne peut être calculée. Il en ressort une propriété remarquable, celle-ci est invariante suivant  $G^{++}$ , qui par transitivité de son action nous donne le résultat suivant.

**Théorème 8.3.** *Il existe une unique mesure de probabilité invariante par l'action de  $G^{++}$  sur l'ensemble des boréliens de  $\Sigma_2^k$  qui coïncide avec la mesure du volume correspondant à la métrique quotient.*

Pour les triangles, le théorème précédent montre que  $\gamma_*^3$  est la mesure de l'aire classique sur  $S^2(\frac{1}{2})$  normée de manière à avoir une masse totale égale à 1.

Le résultat ci-dessus est remarquable : il fait le lien entre la loi normale isotrope dans le plan (qui est en un sens la loi de probabilité la plus simple, la moins structurée dans  $\mathbb{R}^2$ ) et la loi uniforme sur la sphère. Autrement dit, voir les triangles sur la sphère est aussi naturel que de tirer des sommets indépendamment selon une loi normale isotrope. Sur la sphère, chaque élément de surface est d'une importance proportionnée à sa représentativité, ce qui est un progrès flagrant par rapport au plan des  $\xi$  présenté Figure 8.5.

**Intérêt pratique du résultat** Si nous tirons les points du plan non plus selon une loi normale, mais de manière uniforme dans un domaine fixé – par exemple, la péninsule de Land's End dans les Cornouailles –, la répartition des formes de triangles sur la sphère s'en trouve sensiblement modifiée. À la dernière ligne de la Figure 8.9, on représente les densités empiriques sur la sphère pour des points tirés uniformément dans le disque unité, puis dans des ellipses de rapports d'anisotropie 2 et 4 pour 1. Ici, point de distribution uniforme sur la sphère : là où la loi normale et sa "bosse" maintenaient à flot les triangles "à deux petits côtés + un grand", ceux-ci sont maintenant clairement désavantagés au profit des triangles à "deux grands côtés + un petit".

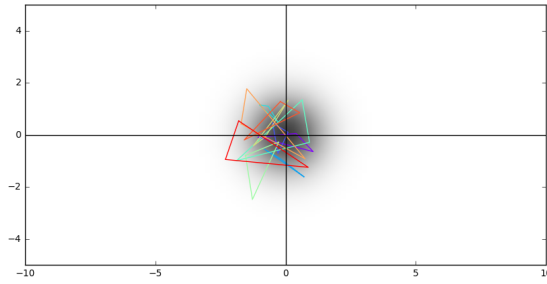
Par contre, le phénomène de tassement sur l'équateur des triangles plats se confirme à mesure que la loi de tirage devient anisotrope. In fine, la comportement précis de la distribution des angles d'un triangle en fonction de la loi de tirage de ses sommets peut être bien comprise, et des tables mises à dispositions des archéologues et des biologistes. Tester une hypothèse d'indépendance dans la génération des sommets par cette statistique est donc devenu une opération de routine.

**Quantification des erreurs statistiques** On peut aussi obtenir un résultat statistique sur l'erreur commise lors de la prise de mesures empiriques par exemple.

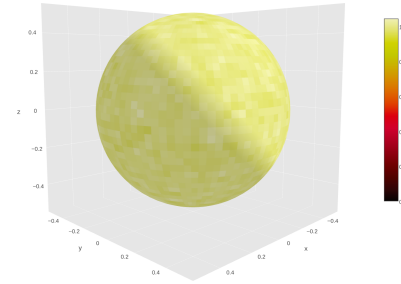
**Théorème 8.4.** *Si on mesure la précision de la position de la forme en terme de longueur de la géodésique entre valeur réelle et mesurée, asymptotiquement pour une erreur initiale  $d$  petite, on aura une erreur de  $\sqrt{\frac{k-2}{k-1}} \frac{d}{\tau}$  avec  $\frac{d}{\tau}$  qui mesure la taille des erreurs relativement à la taille du  $k$ -polygone.*

[Ken84]

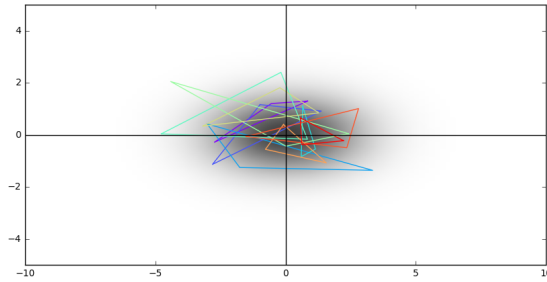
Basile Coron,  
Paul Fraux,  
Massil Hihat,  
Lino Benedetto  
Élémentaire  
et élégant



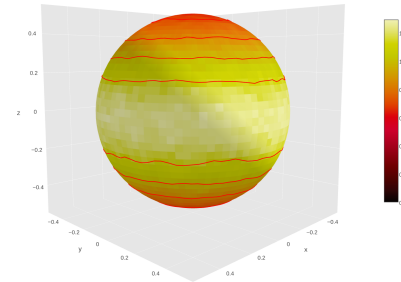
(a) 10 triangles,  $s = 1$ .



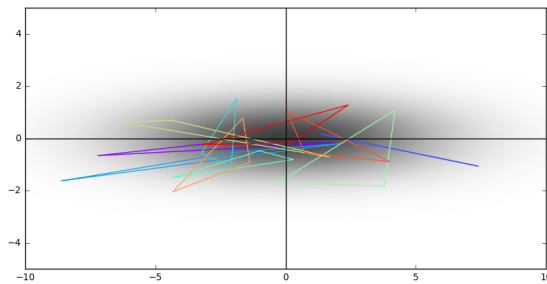
(b)  $10^7$  triangles,  $s = 1$ .



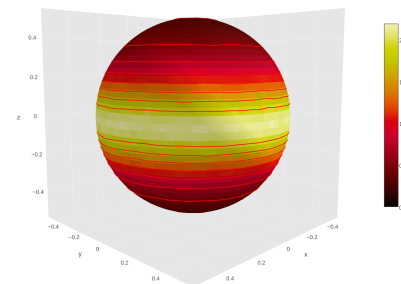
(c) 10 triangles,  $s = 2$ .



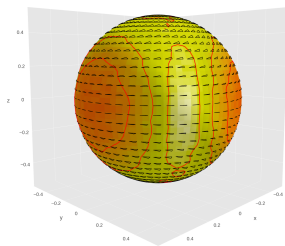
(d)  $10^7$  triangles,  $s = 2$ .



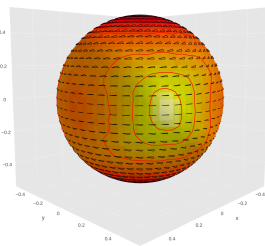
(e) 10 triangles,  $s = 4$ .



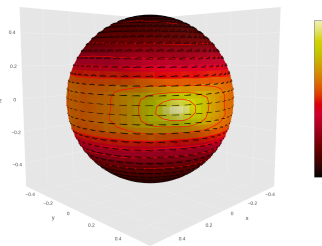
(f)  $10^7$  triangles,  $s = 4$ .



(g)  $10^7$  triangles,  $r = 1$ .



(h)  $10^7$  triangles,  $r = 2$ .



(i)  $10^7$  triangles,  $r = 4$ .

FIGURE 8.9 – Histogrammes empiriques sur la sphère de Kendall. Sur les trois premières lignes (a-f), les points sont tirés de manière i.i.d. selon une gaussienne isotrope (a-b) ou anisotrope (c-d) et (e-f). À gauche, on représente la densité de tirage des points et une dizaine de triangles. Comme indiqué par le théorème 8.3, la densité sur la sphère (représentée à droite par rapport à la mesure de surface) ne dépend que de l’altitude  $Z$ .

Sur la ligne du bas (g-i), des histogrammes analogues sont tracés pour des points tirés de manière uniforme dans une ellipse de grands axes  $(r, 1)$ . Les triangles isocèles “deux petits côtés + un grand” perdent en influence au profit des triangles “deux grands côtés + un petit”, ce qui conduit à la formation de trois îlots d’attraction espacés régulièrement le long de l’équateur.

# Chapitre 9

## Étude statistique dans un espace de formes simple

*Séance 9*

*[Ken84]  
Grégoire Macqueron,  
Maxime Sylvestre  
Très visuel*

### 9.1 Généralisation de l'analyse de Kendall à des polygones quelconques

**En ligne.** Cette séance repose sur une collection de trois notebooks Python, accessible à l'adresse suivante :

[http://www.math.ens.fr/~feydy/Teaching/GeomData\\_2018/GeomData\\_2018\\_TP.zip](http://www.math.ens.fr/~feydy/Teaching/GeomData_2018/GeomData_2018_TP.zip)

Une version statique de ces pages, avec le code, peut être lue aux adresses suivantes :

- [http://www.math.ens.fr/~feydy/Teaching/GeomData\\_2018/Kendall1\\_1.html](http://www.math.ens.fr/~feydy/Teaching/GeomData_2018/Kendall1_1.html)
- [http://www.math.ens.fr/~feydy/Teaching/GeomData\\_2018/Kendall1\\_2.html](http://www.math.ens.fr/~feydy/Teaching/GeomData_2018/Kendall1_2.html)
- [http://www.math.ens.fr/~feydy/Teaching/GeomData\\_2018/Kendall1\\_3.html](http://www.math.ens.fr/~feydy/Teaching/GeomData_2018/Kendall1_3.html)

#### 9.1.1 Analyse procustéenne

Considérons deux triangles A et B non dégénérés, non égaux. On peut les représenter par des triplets de nombres complexes. Pour la programmation on se place cependant dans  $\mathbb{R}^2$ . Comme vu dans le chapitre 8 on s'intéresse à minimiser la distance entre les sommets à l'aide de similitudes.

$$\tau = \frac{1}{3} \sum_{i=1}^3 (B_i - A_i)$$

$$\arg \min_{v, \tau \in \mathbb{C}^2} \sum_{i=1}^3 |v \cdot (A_i + \tau) - B_i|^2 \quad (9.1)$$

$$= \arg \min_{\lambda, \theta, \tau \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^2} \sum_{i=1}^3 |e^{\lambda + i\theta} \cdot (A_i + \tau) - B_i|^2 \quad (9.2)$$

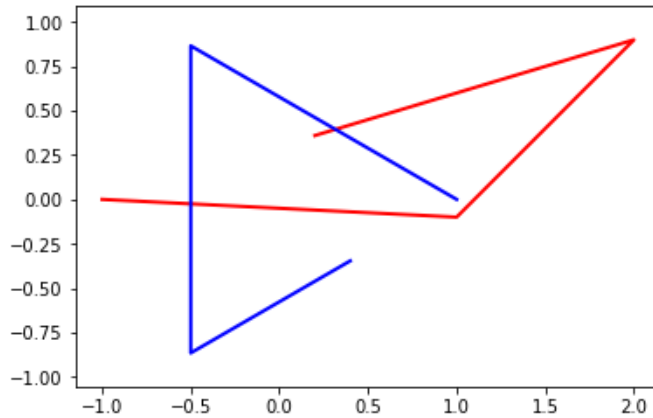
A l'aide du module PyTorch il est facile de résoudre ce problème.

Mais en réalité il suffit de remarquer qu'il faut commencer par aligner les barycentres des triangles à l'aide d'une translation puis d'une variation des longueur et une rotation. Ainsi on a :

[Ken84]

Grégoire Macqueron,  
Maxime Sylvestre  
Très visuel

```
A = tensor([ [-1, 0], [ 1, -.1], [ 2,.9] ])
B = tensor([ [ 1, 0], [-.5, np.sqrt(3)/2], [-.5,-np.sqrt(3)/2] ])
plt.figure()
plot(A, "red")
plot(B, "blue")
plt.axis("equal")
plt.show()
```



```
class RigidRegistration(Model) :
    "Find the optimal translation, scaling and rotation."
    def __init__(self, source) :
        "Defines the parameters of a rigid deformation of source."
        super(Model, self).__init__()
        self.x = source
        self.λ = Parameter(tensor( 0. )) # log-Scale
        self.θ = Parameter(tensor( 0. )) # Angle
        self.τ = Parameter(tensor( [0.,0.] )) # Position

    def __call__(self, x) :
        "Applies the model on some point cloud, an N-by-2 tensor."
        # At the moment, PyTorch does not support complex numbers...
        return self.λ.exp() * rot(x + self.τ, self.θ)

    def cost(self, target) :
        ""Returns a cost to optimize.
        Here, the squared L2 distance to the target""
        return ((self(self.x) - target)**2).sum()
```

$$\tau = \frac{1}{3} \sum_{i=1}^3 (B_i - A_i)$$

Si on étudie alors les triangles centrés  $a$ ,  $b$ , le produit hermitien contient alors les informations permettant d'ajuster l'échelle et la rotation :

$$a^*b = \langle a, b \rangle_{\mathbb{C}^3} = \sum_{i=1}^3 \bar{a}_i b_i$$

Ainsi on en déduit, en dérivant la somme des modules au carré précédentes :

$$v = e^{\lambda+i\theta} = \frac{a^*b}{\|a\|^2}.$$

En prenant en compte cela on a enfin réussi à expliciter la distance procustéenne à l'aide des affixes des sommets des triangles avec leurs barycentres superposés. On a alors :

$$\|b - \frac{a^*b}{\|a\|^2}a\|_{\mathbb{C}^3}^2 = \|b\|_{\mathbb{C}^3}^2 \cdot \left|1 - \left|\frac{a}{\|a\|} \star \frac{b}{\|b\|}\right|^2\right|^2.$$

Le facteur  $\|b\|^2$  est le facteur d'ajustement d'échelle, et le reste de cette expression ne dépend que des triangles normalisés centrés  $a/\|a\|$  et  $b/\|b\|$ .

Ainsi il est possible de généraliser, si  $A$  et  $B$  sont deux polygones non-dégénérés du plan complexe, on peut alors définir la distance optimale  $L^2$  en normalisant  $A$  et  $B$  en polygones centrés normalisés  $a$  et  $b$  pour définir, à l'aide du produit hermitien de  $\mathbb{C}^n$  :

$$d_{L^2}([A], [B]) = d_{L^2}(a, b) = \sqrt{1 - |a^*b|^2}$$

### 9.1.2 Sphère de Kendall

La distance définie précédemment est invariante par transformation par similitude. Ainsi on se permet d'étudier les classes d'équivalence par transformation par similitudes. En réalité Kendall a remarqué que l'on peut même étudier l'espace quotient comme un espace métrique. On identifie cette espace quotient avec une sphère comme on l'a fait dans le chapitre 8.

**Théorème 9.1.** *Il existe une bijection entre la sphère de  $\mathbb{R}^3$  de rayon 1/2 et la représentation dans le plan des triangles non-dégénérés  $A$  et  $B$  dans  $\mathbb{C}^3$  vérifiant :*

$$d_{L^2}([A], [B]) = \|f(A) - f(B)\|_{\mathbb{R}^3}.$$

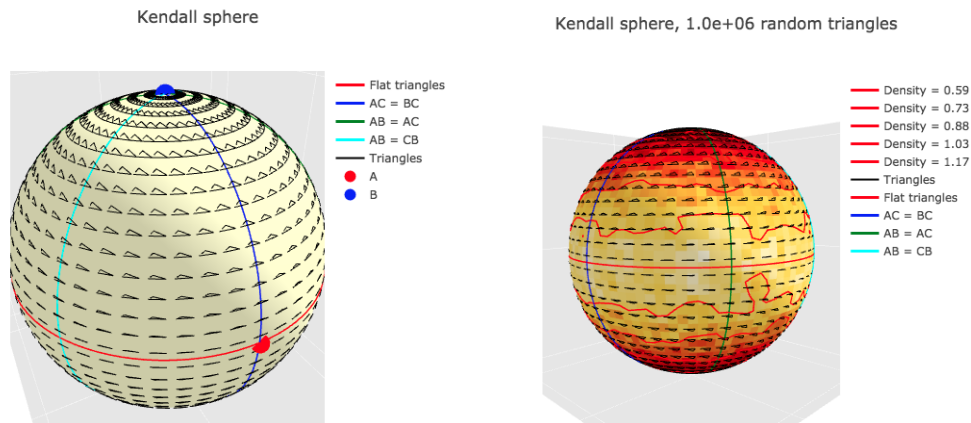
### 9.1.3 Espace Riemannien des formes

Si la distance  $L^2$  correspond bien à la distance de corde dans la sphère de Kendall ce n'est pas une distance de chemin. On ne peut donc pas définir proprement de loi de répartition à l'aide de la topologie induite.

Cependant on peut introduire la distance par arc assez facilement. Si on pose à l'aide de  $a$  et  $b$  les transformations normalisées de  $A$  et  $B$ , on a

$$d_{\text{arc}}([A], [B]) = \arccos(|a^*b|).$$

Cette distance est même géodésique ce qui va nous permettre d'introduire des manipulations statistiques à l'aide de la méthode PCA.



```

M = KendallTriangles()

npoints = 1000000
s = 3.
points = s*random.randn(npoints,3) + 1j * random.randn(npoints,3)

M.set_data(points)

```

FIGURE 9.1

**Théorème 9.2.** *Si on pioche dans le plan complexe selon une gaussienne des sommets et ainsi des triangles, alors dans l'espace quotient on obtient une distribution de triangles qui ne dépend que du caractère anisotrope de la distribution Gaussienne.*

Si la distribution est isotrope on obtient alors une loi uniforme sur la sphère de Kendall. Sinon il y a une concentration de triangle plats.

## 9.2 Statistiques sur la sphère algorithme d'ACP

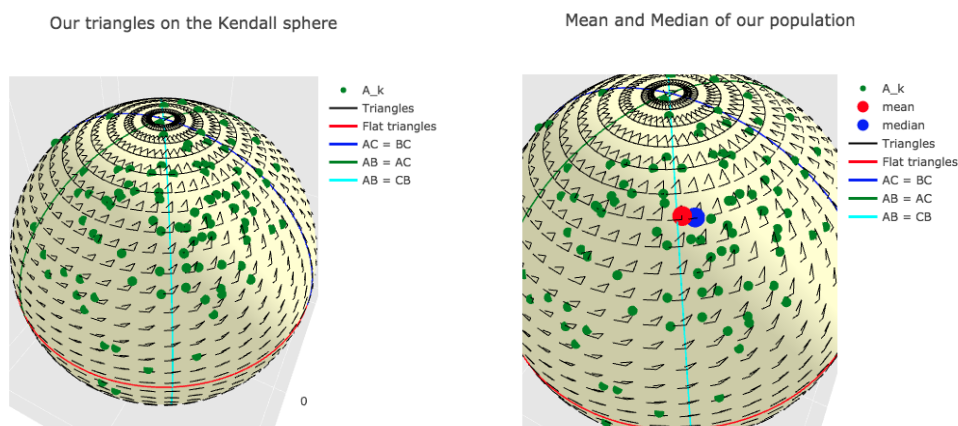
### 9.2.1 Calculer une moyenne de Fréchet

Nous mettons maintenant en œuvre des opérations statistiques (simples) sur la sphère de Kendall. L'idée à retenir est que les outils classiques du statisticien peuvent être généralisés aux variétés riemanniennes, à la fois d'un point de vue théorique mais aussi d'un point de vue pratique. Ici on utilise l'exemple de la sphère de Kendall.

Étant donné une distribution empirique  $A_k$ , la première mesure statistique que l'on peut chercher à calculer est la moyenne. Mais sur la sphère, il n'existe pas d'opérateur "+".

L'un des moyens les plus simples pour généraliser la moyenne à un espace métrique arbitraire  $(X, d)$  est la définition des moindres carrés. Pour tout ordre  $p > 0$ , nous définissons donc la moyenne  $p$ -Fréchet des  $A_k$  par :

$$F_p(A_1, \dots, A_N) = \arg \min_{x \in X} \sum_{k=1}^N d(x, A_k)^p$$



Dans un espace vectoriel euclidien, les moyennes de Fréchet d'ordre 1 et 2 coïncident respectivement avec la médiane et la moyenne arithmétique de la distribution.

### 9.2.2 Calculer une géodésique

Comme on l'a déjà vu on a pu définir les géodésiques à l'aide d'équation différentielles d'ordre 2. Si  $a$  et  $b$  sont des points de notre variété et si  $v$  est un vecteur tangent en  $a$ , alors  $\text{Exp}_a(v)$  désignera la position à l'instant  $t = 1$  de l'unique courbe géodésique  $\gamma_t$  telle que :

$$\gamma_0 = a \quad \text{et} \quad \dot{\gamma}_0 = v.$$

La longueur d'une telle courbe géodésique est simplement égale à la norme du vecteur tangent à la position  $a$ .

Pour tout point de référence  $a$ , la carte exponentielle définit un isomorphisme entre un voisinage de 0 (sur le plan tangent) sur un voisinage de  $a$  sur notre variété.

Nous pouvons ainsi définir la carte  $\text{Log}_a$  comme l'inverse de  $\text{Exp}_a$  sur un voisinage de  $a$ , et l'utiliser pour représenter la variété courbe dans le plan tangent.

Le plus souvent, les opérateurs  $\text{Exp}$  et  $\text{Log}$  sont difficiles à implémenter. Mais dans le cas très spécifique (et important) des formes polygonales à similitude près, la théorie classique de Kendall fournit des formules explicites (Figure 9.7).

### 9.2.3 ACP tangente : Analyse en composante principale

Nous avons vu comment calculer une moyenne et cartographier les données de notre variété sur l'espace tangent. La prochaine étape est d'analyser la variabilité au voisinage de la moyenne par une analyse statistique linéaire classique. Elle consiste à étudier les axes de variabilité caractéristique de la distribution. Cela revient à étudier sur la carte exponentielle la distribution et à placer une ellipse centrée en la moyenne qui représente le mieux la distribution.

[Ken84]

Grégoire Macqueron,

Maxime Sylvestre

Très visuel

```

def log(a,b):
    """
    Assuming that a and b are both normalized, returns
    the Riemannian logarithm of b at point a.
    """
    # Optimally rotate b onto a
    bsa = herm(b, a)
    b = rot(b, angle(bsa))

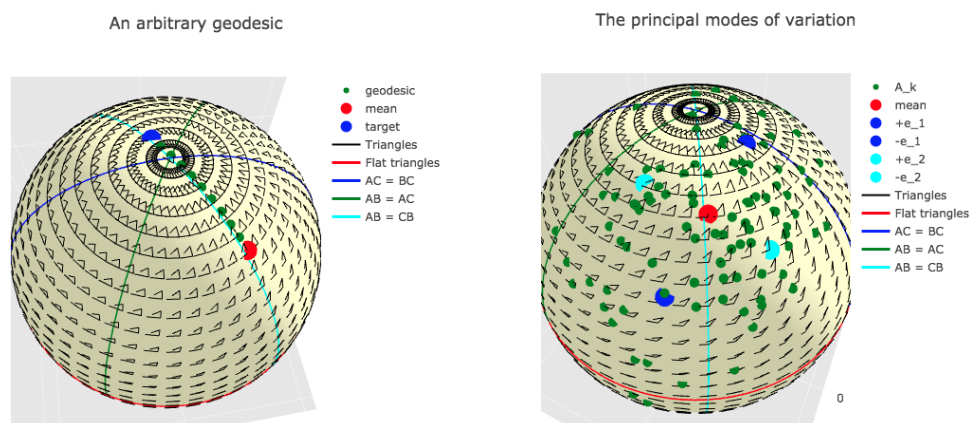
    # Compute the normal vector
    c = b - (a*b).sum()*a
    c = c / (c**2).sum().sqrt()
    L = arc(a,b)
    return L*c

def exp(a,v):
    "Shoots along the tangent vector v at location a."
    L = (v**2).sum().sqrt()
    return a*torch.cos(L) + (v/L)*torch.sin(L)

def geodesic(a,b):
    """
    Returns the geodesic between two *normalized* polygons a and b,
    sampled on 10 timesteps with the normal tangent vector at a.
    """
    v = log(a,b)
    g_t = [ exp(a, t*v) \
            for t in torch.linspace(0,1,11) ]
    return torch.stack( g_t )

```

FIGURE 9.2





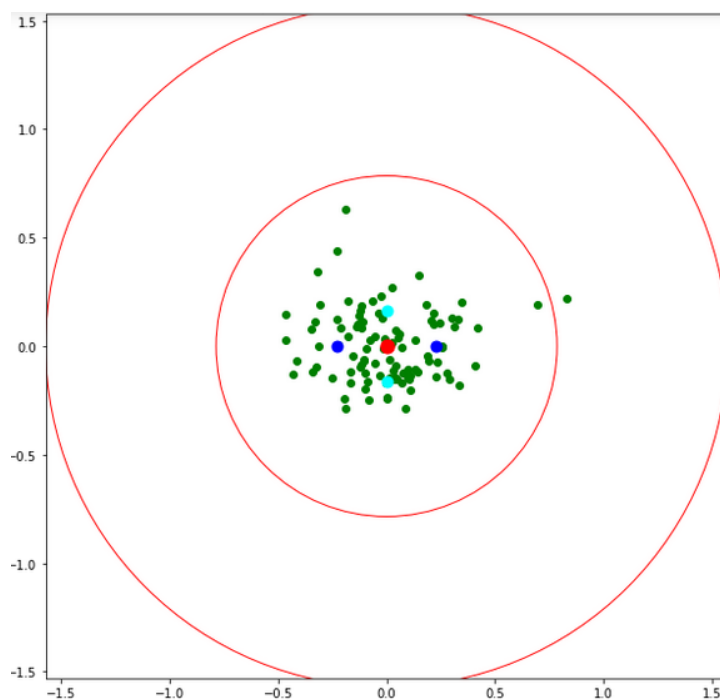


FIGURE 9.3

## 9.3 Application sur les petits vers

### 9.3.1 L'ensemble des données

Nous avons vu comment effectuer une ACP tangente dans la variété des triangles à similitudes près. Généralisons maintenant cette analyse aux espaces de landmarks de dimension supérieure, ou nuages de points étiquetés à similitudes près. Notre base de données : une petite vidéo d'un ver *C. elegans*, tirée de la base de données comportementales *C.elegans*.

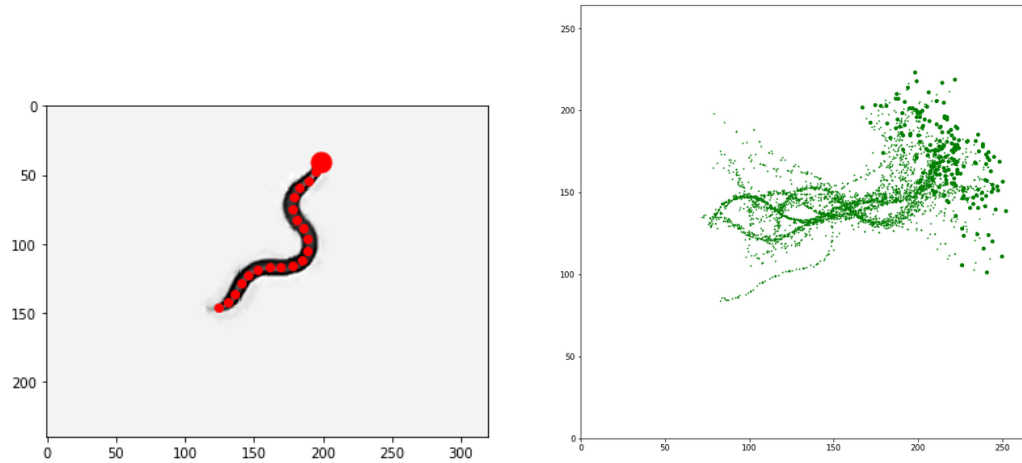
En utilisant un algorithme de segmentation barebone (seuillage de gradient + opérations morphologiques fournies par la bibliothèque d'images scikit), nous avons enregistré la forme du ver sur 200 images sous forme d'un nuage de points  $A_k$  dans  $\mathbb{R}^{20 \times 2}$ .

### 9.3.2 Les formules de Kendall

Si le nombre  $N$  de points par image dépasse trois, l'espace quotient des formes ne peut pas s'interpréter aussi simplement que la sphère des triangles. Cependant, la plupart des résultats de Kendall restent valables : l'espace des formes de  $N$ -points dans le plan à similitudes près, est un quotient de la sphère de dimension  $2N - 3$  qui peut être dotée d'une distance arc définie par la même formule que dans le cas des triangles mais avec le produit hermitien de  $\mathbb{C}^N$ .

### 9.3.3 Études statistique du déplacement des vers

Dès lors on cherche à étudier les mouvements du ver donc en observant seulement les dispositions successives de son corps à similitudes près (sans variation d'échelle). On peut alors effectuer une ACP après avoir calculé la moyenne qui dans ce cas se résume à la position du ver



inactif. Puis les axes de variabilité sont les formes qui permettent de décomposer efficacement le mouvement. On observe ici des ondulations à des fréquences différentes, c'est une décomposition de Fourier.

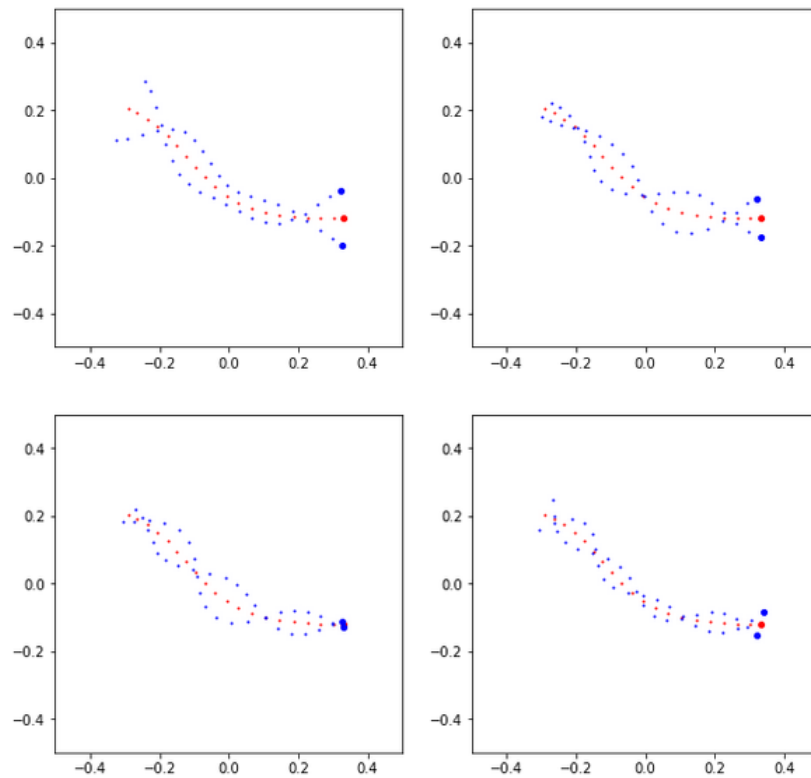


FIGURE 9.4 – En rouge la moyenne et en bleu les axes de variabilité caractéristiques

Dans la séance d'aujourd'hui consacrée à la théorie classique des formes, nous avons détaillé

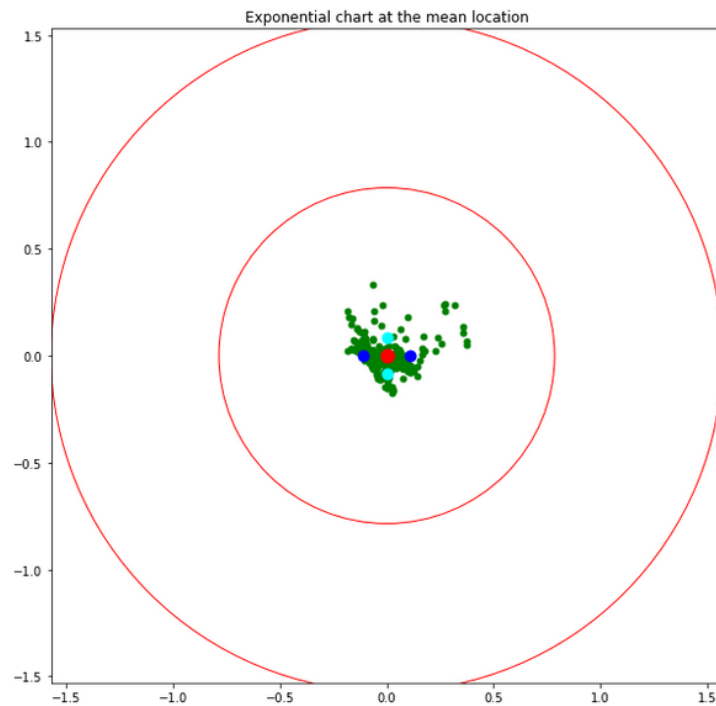


FIGURE 9.5

une méthode de base pour décomposer la variabilité d'une population anatomique. Les algorithmes ci-dessus permettent d'étudier séparément le vecteur de similitude (position, orientation et échelle, codé par un couple de nombres complexes) et la variété des formes résiduelles. Dans des contextes très simples (faible variance, données étiquetées), une analyse des formes résiduelles peut suffire. Mais dans un contexte médical, la plupart des problèmes d'anatomie sont mal posés. La conception de méthodes posant un cadre topologique plus fort fera l'objet d'une prochaine séance.

Séance 9 126

[Ken84]

Grégoire Macqueron,

Maxime Sylvestre

Très visuel

Chapitre 9. Étude statistique dans un espace de formes simple

# Chapitre 10

## Travailler avec des données non-étiquetées

*Séance 10*

*[CT13]  
Valentin Carlier,  
Louis-Pierre  
Chaintron  
Retour aux  
fondamentaux*

### 10.1 Introduction

Les séances précédentes ont montré comment comparer des nuages de points étiquetés  $(x_i)_{i \in \llbracket 1;n \rrbracket}$  de  $(\mathbb{R}^d)^n$ , par la construction d'une distance canonique. Néanmoins dans le cas général, les échantillons de données à comparer ne contiennent pas tous le même nombre de points, en particulier lorsqu'ils proviennent de la discrétisation de phénomènes continus ou de formes tridimensionnelles. On s'intéresse ici à la construction d'une distance plus générale sur un espace de forme quelconque, adaptée au cas de nuages de points non étiquetés.

### 10.2 Distances sur un espace de formes

#### 10.2.1 Représenter les formes comme des mesures

Les nuages de points n'étant plus étiquetés, on cherche une façon de représenter les données indépendamment de toute numérotation. À un nuage de points  $(x_i)_{i \in \llbracket 1;n \rrbracket}$  de  $\mathbb{R}^d$ , on choisit ainsi d'associer la mesure canonique

$$\nu = \frac{1}{n} \sum_{i=1}^n \delta_{x_i} \quad (10.1)$$

où  $\delta_x$  représente la mesure de Dirac en  $x$ . Cette représentation a l'avantage d'effacer le nombre et l'ordre des points conservés par la discrétisation. En revanche, le nouvel espace considéré est l'ensemble des mesures, qui peut être vu comme le dual des fonctions de  $\mathbb{R}^d$ . Cet espace de dimension infini est nettement plus gros que  $\mathbb{R}^d$ , ce qui complexifie a priori l'étude. Néanmoins la structure duale permet des représentations plus efficaces, indépendantes de la paramétrisation. Par exemple

$$\int f d\nu = \frac{1}{n} \sum_{i=1}^n f(x_i) \quad (10.2)$$

Si l'importance de certaines données doit prédominer sur d'autres, il est possible de l'exprimer en pondérant les mesures :

$$\nu = \sum_{i=1}^n \alpha_i \delta_{x_i} \quad (10.3)$$

[CT13]

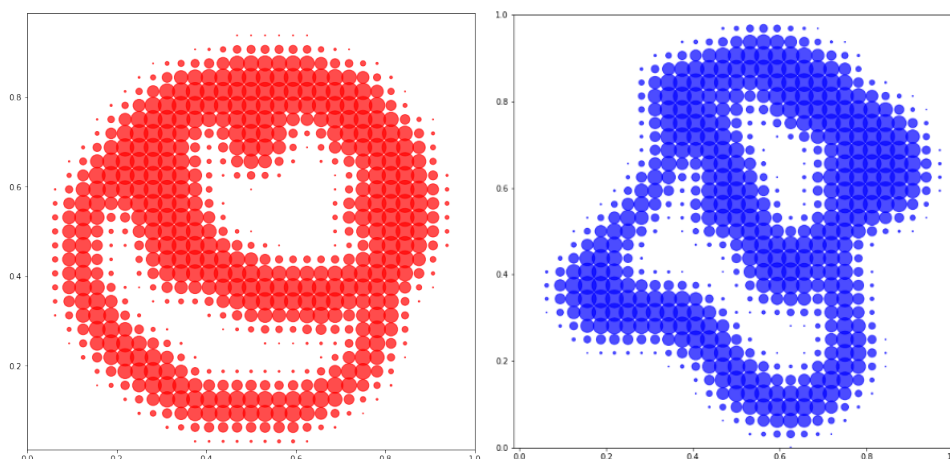
Valentin Carlier,

Louis-Pierre

Chaintron

Retour aux fondamentaux

avec  $\sum_{i=1}^n \alpha_i = 1$ . Le problème initial consistant à caractériser une forme à partir d'un échantillon de points, se ramène donc à la détermination d'une mesure en se donnant ses valeurs en seulement certains points. Pour mesurer l'erreur commise dans ce travail d'approximation, on va s'atteler à la construction de distances sur des espaces de mesures bien choisies.



## 10.2.2 Transformée de Fourier et produits de convolution

On introduit dans cette partie les bases de la transformée de Fourier, qui permettra plus loin de construire une distance répondant au problème. La transformée de Fourier est une généralisation continue de la décomposition d'une fonction en série de Fourier. Pour une fonction  $f$  suffisamment régulière, une telle décomposition s'écrit

$$f(x) = \sum_{n \in \mathbb{Z}} c_n e^{in x} \quad (10.4)$$

On a ainsi décomposé  $f$  comme somme de fonctions périodiques en extrayant les coefficients de Fourier  $c_n$ . Dans le cas continu, la somme discrète est remplacée par une intégrale, et la suite  $(c_n)_{n \in \mathbb{N}}$  des coefficients par une fonction continue  $\hat{f}$ , qui vérifie l'identité

$$f(x) = \frac{1}{2\pi} \int \hat{f}(\omega) e^{i\omega x} d\omega \quad (10.5)$$

$\hat{f}$  est la *transformée de Fourier* de  $f$ , et on vérifie que  $\hat{f}$  est donnée par

$$\hat{f}(\omega) = \int e^{-i\omega x} f(x) dx \quad (10.6)$$

*Remarque 1.* Les fonctions  $x \mapsto e^{i\omega x}$  forment une base Hilbertienne orthonormale de  $L^2(\mathbb{R}, \mathbb{C})$ , tout comme les fonctions de Dirac  $\delta_x$ . On peut alors interpréter la transformée de Fourier comme le changement de base correspondant, qui à une fonction associe ses composantes en prenant les produits scalaires avec les  $x \mapsto e^{i\omega x}$ . Il s'agit de l'analogie en dimension infinie d'un simple changement de base en dimension finie. Par exemple dans  $\mathbb{C}^n$ , on peut préférer travailler dans la base orthonormale donnée par les vecteurs  $u_j = \left( \frac{1}{\sqrt{n}} e^{\frac{2i\pi k l}{n}} \right)_{0 \leq l \leq n-1}$  où  $j$  est dans  $\llbracket 0; n-1 \rrbracket$ . La matrice de passage de la base canonique à la base des  $f_j$  est circulante (et de Van der Monde).

Comme cette base est orthonormale, un vecteur  $x = (x_0, \dots, x_{n-1})$  dans la base canonique s'écrit directement

$$x = \sum_{j=0}^{n-1} c_j f_j \quad \text{avec} \quad c_j = \langle x, f_j \rangle = \sum_{l=0}^{n-1} e^{\frac{-2i\pi kl}{n}} x_l \quad (10.7)$$

La donnée de la suite  $(c_j)_{0 \leq j \leq n-1}$  constitue la *transformée de Fourier discrète* de  $x$ , très utile en Informatique notamment dans le domaine de la compression d'image.

**Proposition 1.** La transformée de Fourier satisfait les propriétés suivantes

1. L'application  $f \mapsto \widehat{f}$  est linéaire
2. Pour  $(a, b)$  dans  $\mathbb{R}^2$ , la transformée de  $x \mapsto f(ax + b)$  est  $\omega \mapsto \frac{1}{|a|} \widehat{f}\left(\frac{\omega}{a}\right) e^{ib\omega}$
3. Pour  $f$  à valeurs réelles, notant  $\widetilde{f} : x \mapsto f(-x)$ , on a  $\widehat{\widetilde{f}} = \overline{\widehat{f}}$
4.  $\widehat{f \star g} = \widehat{f} \times \widehat{g}$
5.  $\langle f, g \rangle = \langle \widehat{f}, \widehat{g} \rangle$
6. Si  $f$  est de classe  $C^1$ , alors  $\widehat{f}' : \omega \mapsto i\omega \widehat{f}(\omega)$
7. De même la transformée de  $x \mapsto xf(x)$  est  $i\widehat{f}'$
8. Si  $\widehat{f}$  est dans  $L^1$  alors  $f$  est continue.

le produit scalaire considéré étant  $\langle f, g \rangle = \int_{-\infty}^{+\infty} \overline{f(x)} g(x) dx$

*Démonstration.* On ne détaille ici que certains points :

1. Résulte de la linéarité de l'intégrale, et du produit scalaire hermitien en sa seconde variable.
2. Changement de variable.
3. On calcule directement

$$\widehat{\widetilde{f}}(\omega) = \int_{-\infty}^{+\infty} e^{-i\omega x} f(-x) dx \quad (10.8)$$

$$= \int_{-\infty}^{+\infty} e^{i\omega x} f(x) dx \quad (10.9)$$

$$= \int_{-\infty}^{+\infty} \overline{e^{-i\omega x} f(x)} dx \quad (10.10)$$

$$= \overline{\int_{-\infty}^{+\infty} e^{-i\omega x} f(x) dx} = \overline{\widehat{f}(\omega)} \quad (10.11)$$

4. Là encore, grâce au Théorème de Fubini

$$\widehat{f \star g}(\omega) = \int_{-\infty}^{+\infty} e^{-i\omega x} (f \star g)(x) dx \quad (10.12)$$

$$= \int_{-\infty}^{+\infty} e^{-i\omega x} \int_{-\infty}^{+\infty} f(y) g(x-y) dy dx \quad (10.13)$$

$$= \int_{-\infty}^{+\infty} f(y) \int_{-\infty}^{+\infty} e^{-i\omega x} g(x-y) dx dy \quad (10.14)$$

$$= \int_{-\infty}^{+\infty} f(y) \int_{-\infty}^{+\infty} e^{-i\omega z} e^{-i\omega y} g(z) dz dy \quad (10.15)$$

$$= \widehat{f}(\omega) \widehat{g}(\omega) \quad (10.16)$$

Valentin Carlier,  
Louis-Pierre  
Chaintron  
Retour aux fondamentaux

5. Calcul analogue.
6. On dérive sous le signe somme.
7. Intégration par parties.
8. Théorème de continuité sous l'intégrale.

□

*Remarque 2.* Ces notions se généralisent à des espaces multidimensionnels, et les propriétés sont conservés

$$\widehat{f}(\vec{\omega}) = \int e^{-i\vec{\omega} \cdot \vec{x}} f(\vec{x}) d\vec{x} \quad (10.17)$$

La transformée de Fourier conserve le produit scalaire, donc pour  $f$  dans  $L^2$ , on a la formule de Parseval-Plancherel

$$\int_{-\infty}^{+\infty} |f(x)|^2 dx = \int_{-\infty}^{+\infty} |\widehat{f}(\omega)|^2 d\omega \quad (10.18)$$

*Remarque 3.* La transformée de Fourier d'une dérivée donne un résultat remarquable : la dérivation est un *opérateur diagonal* pour la transformation de Fourier. En particulier, si  $f$  est suffisamment régulière, la transformée de  $f^{(n)}$  est  $\omega \mapsto i^n \omega^n \widehat{f}$ . Si la régularité permet de repasser à  $f^{(n)}$  par transformation de Fourier inverse, la fonction  $\omega \mapsto i^n \omega^n \widehat{f}$  est alors intégrable : imposer des conditions de régularité à  $f$ , c'est imposer à sa transformée de Fourier de décroître d'autant plus vite en  $\pm\infty$ . Cette propriété correspond au dernier point de la proposition précédente.

De la même manière, on définit la transformée de Fourier d'une mesure finie  $\mu$  sur  $\mathbb{R}^d$ , comme la fonction  $\widehat{\mu} : \mathbb{R}^d \mapsto \mathbb{R}$  donné par

$$\widehat{\mu}(\omega) = \int_{\mathbb{R}^d} e^{-i\omega \cdot x} d\mu(x) \quad (10.19)$$

Si  $\mu$  est une mesure de densité  $f$  par rapport à la mesure de Lebesgue, alors  $\widehat{f} = \widehat{\mu}$  car

$$\widehat{\mu}(\omega) = \int_{\mathbb{R}^d} e^{-i\omega \cdot x} d\mu(x) = \int_{\mathbb{R}^d} e^{-i\omega \cdot x} f(x) dx = \widehat{f}(\omega) \quad (10.20)$$

Le produit de convolution d'une mesure et d'une fonction de  $\mathbb{R}^d$  est également défini par

$$g \star \mu(x) = \int_{\mathbb{R}^d} g(x-y) d\mu(y) \quad (10.21)$$

**Proposition 2.** On a les propriétés

1.  $\widehat{g \star \mu} = \widehat{g} \times \widehat{\mu}$
2.  $\langle \mu, f \rangle = \langle \widehat{\mu}, \widehat{f} \rangle$

Les mesures étant vues comme des formes linéaires, il s'agit du crochet de dualité  $\langle \mu, f \rangle = \int f d\mu$ .



### 10.2.3 Une première norme duale

Une mesure peut être vue comme une forme linéaire continue sur un espace de fonctions continues. Le plus simple pour construire une distance sur les mesures est alors d'utiliser une norme duale. Les mesures étudiées sont continues au sens de  $\|\cdot\|_\infty$ , ce qui donne la norme duale

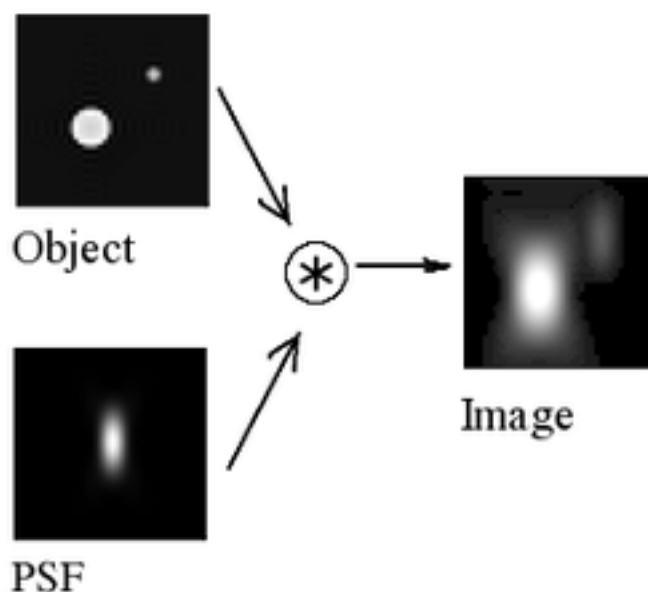
$$\|\mu\|_\infty = \sup_{\|f\|_\infty=1} \langle f, \mu \rangle \quad (10.22)$$

Néanmoins cette norme est complètement inadaptée : il suffit que deux nuages de point  $(x_i)$  et  $(y_j)$  soient à support disjoints pour avoir  $\|\mu - \nu\|_\infty = 2$ , en prenant par exemple une fonction valant 1 sur les  $x_i$  et -1 sur les  $y_j$ . Pour les mesures concernées ici (combinaisons linéaires pondérées de mesures de Dirac), la distance induite est à valeurs dans  $[0; 2]$ , et peut valoir 2 sur des formes a priori très proches. A l'inverse, elle est faible sur des formes intuitivement très différentes...

## 10.3 Normes à noyau

### 10.3.1 Définitions

La norme précédente s'est avérée inadéquate car les mesures de Dirac ponctuelles ne prennent pas à compte les caractéristiques de la forme au voisinage d'un point. Une solution consiste à "flouter l'image échantillonnée" afin d'inclure cette information. On réalise cela en effectuant le produit de convolution de la mesure avec une fonction dite de *floutage*  $g : \mathbb{R}^d \rightarrow \mathbb{R}$  : il en résulte une mesure floutée  $g \star \mu(x)$ , qui est en fait la fonction de densité de la mesure floutée.



Une fois les mesures transformées en fonctions plus maniables, on définit une nouvelle norme sur notre espace de mesures par la formule

$$\|\mu\|^2 = \langle g \star \mu, g \star \mu \rangle \quad (10.23)$$

Posant  $\tilde{g}(x) = g(-x)$  et  $k = \tilde{g} \star g$ , on a l'égalité remarquable ( $g$  étant à valeurs réelles)

$$\|\mu\|_k^2 = \langle g \star \mu, g \star \mu \rangle = \langle \widehat{g} \times \widehat{\mu}, \widehat{g} \times \widehat{\mu} \rangle = \langle \widehat{\mu}, \widehat{\tilde{g}} \times \widehat{g} \times \widehat{\mu} \rangle = \langle \widehat{\mu}, \widehat{\tilde{g} \star g \star \mu} \rangle = \langle \mu, k \star \mu \rangle \quad (10.24)$$

Cette norme est donc naturellement associée à un produit scalaire

$$\langle \mu, \nu \rangle_k = \langle \mu, k \star \nu \rangle \quad (10.25)$$

A chaque fonction *noyau*  $k$  de la forme  $\tilde{g} \star g$  est donc associé un produit scalaire. Si de plus les mesures sont combinaison de mesures de Dirac (cas des nuages de points), on a l'expression très simple

$$\langle \mu, \nu \rangle_k = \sum_i \sum_j k(x_i - y_j) \mu_i \nu_j \quad \text{où} \quad \mu = \sum_{i=1}^n \mu_i \delta_{x_i} \text{ et } \nu = \sum_{j=1}^n \nu_j \delta_{y_j} \quad (10.26)$$

### 10.3.2 Choix du noyau

Dans la recherche d'une distance adaptée aux mesures étudiées, on a introduit la notion de norme à noyau associée à une fonction  $k$ . On s'intéresse à présent au choix spécifique de  $k$ , ainsi qu'à la nature duale de ces normes. L'étude des bonnes fonctions  $k$  est liée à la théorie *RKHS*, et privilégie les  $k$  telles que

$$\forall \omega \in \mathbb{R}, \quad \widehat{k}(\omega) > 0 \quad (10.27)$$

**Théorème 10.1.** *Pour toute fonction  $k$  vérifiant la condition ci-dessus, il existe une fonction de floutage  $g$  telle que  $k = g \star g$ . Elle est donnée par*

$$\forall \omega \in \mathbb{R}, \quad \widehat{g}(\omega) = \sqrt{\widehat{k}(\omega)} \quad (10.28)$$

### 10.3.3 Espaces de Sobolev

La compréhension des normes à noyaux va permettre l'interprétation des mesures considérées comme des formes linéaires, donc le retour à une norme duale. Le bon cadre pour cela est la théorie des espaces de Sobolev. Les espaces de Sobolev ont été introduits dans le but d'apporter un cadre général aux équations aux dérivées partielles. Ce sont des espaces de fonctions dont les dérivées successives doivent vérifier des conditions de régularité particulières. En tout généralité, on considère  $\Omega$  un ouvert de  $\mathbb{R}^n$ ,  $m$  entier naturel, et  $p$  dans  $[1, +\infty]$ . L'espace de Sobolev associé est

$$W^{m,p}(\Omega) = \{f \in L^p(\Omega); \forall \alpha \text{ liste d'au plus } m \text{ indices}, D^\alpha f \in L^p(\Omega)\} \quad (10.29)$$

On impose simplement que toutes les dérivées partielles successives jusqu'à l'ordre  $p$  de  $f$  soient elles aussi dans  $L^p$ . Il existe une norme naturelle sur  $W^{m,p}(\Omega)$  :

- Lorsque  $p = \infty$ , il s'agit de  $\|f\|_{W^{m,\infty}} = \max_{|\alpha| \leq m} \|D^\alpha f\|_\infty$ .
- Pour  $p$  fini, ce qu'on supposera dorénavant, on pose

$$\|f\|_{W^{m,p}} = \left( \sum_{|\alpha| \leq m} \|D^\alpha f\|_p^p \right)^{1/p} \quad (10.30)$$

Le cas lié à notre étude est celui où  $p = 2$ , et on notera plus simplement  $H^m$  pour  $W^{m,2}$ . La norme des espaces de Sobolev est alors induite par le produit scalaire

$$\langle f, g \rangle_m = \sum_{|\alpha| \leq m} \langle D^\alpha f, D^\alpha g \rangle \quad (10.31)$$

où comme précédemment  $\langle u, v \rangle = \int_{\mathbb{R}^n} \bar{u}v$ .  $H^m$  est un espace de Hilbert. Pour  $f$  dans  $L^2$  de classe

$C^\infty$  et  $\alpha$  liste d'au plus  $m$  indices, la condition  $\|f\|_{H^m} < \infty$  se réécrit  $\|\widehat{f}\|_{H^m} < \infty$  d'après l'égalité de Parseval, c'est-à-dire

$$\int_{\mathbb{R}^n} |\widehat{f}(\omega)|^{|\alpha|} d\omega \tag{10.32}$$

car la dérivation est un opérateur diagonal pour la transformée de Fourier. On a donc

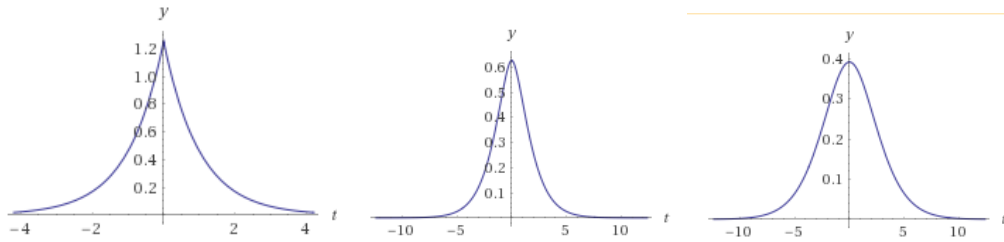
$$H^m = \{f \in L^2(\mathbb{R}^n); \int_{\mathbb{R}^n} |\widehat{f}(\omega)|^2 (1 + |\omega|^2)^m d\omega < \infty\} \tag{10.33}$$

avec

$$(f, g)_m = \int_{\mathbb{R}^n} \overline{\widehat{f}(\omega)} \widehat{g}(\omega) (1 + |\omega|^2)^m d\omega \tag{10.34}$$

produit hermitien équivalent à  $\langle \cdot, \cdot \rangle_m$ .

*Remarque 4.* Le choix du polynôme d'ordre  $2m$  qui apparaît importe un peu. On peut démontrer que les normes ainsi obtenues sont toutes équivalentes, d'où le terme générique de *normes de Sobolev d'ordre  $m$* . On représente ci-après des noyaux de Sobolev, à savoir les transformées de Fourier inverse de  $\omega \mapsto \frac{1}{(\omega^2+1)^m}$  pour  $m = 1, 2, 4$ .



### 10.3.4 Normes duales

Considérons à présent une fonction  $k$  telle que  $\widehat{k} > 0$ . Par analogie avec les espaces de Sobolev, on souhaite introduire la norme de Hilbert

$$\|f\|_V^2 = \int_{\mathbb{R}^n} \frac{|\widehat{f}(\omega)|^2}{\widehat{k}(\omega)} d\omega = \langle k^{(-1)} \star f, f \rangle \tag{10.35}$$

Le bon cadre pour cela est l'espace de Hilbert

$$V = \{f \in L^2(\mathbb{R}^n); \|f\|_V < \infty\} \tag{10.36}$$

dont la norme  $\|\cdot\|_V$  est induite par le produit hermitien

$$\langle f, g \rangle_V = \int_{\mathbb{R}^n} \frac{\overline{\widehat{f}(\omega)} \widehat{g}(\omega)}{\widehat{k}(\omega)} d\omega = \langle k^{(-1)} \star f, g \rangle \tag{10.37}$$

*Remarque 5.* Le lien avec la section précédente apparaît clairement : l'espace de Sobolev  $H^m$  n'est rien d'autre que l'espace  $V$  ci-dessus pour la fonction de noyau  $k$  telle que  $\widehat{k}(\omega) = \frac{1}{(1+|\omega|^2)^m}$ .

[CT13]

Valentin Carlier,

Louis-Pierre

Chaintron

Retour aux fondamentaux

Dans ce contexte, il est naturel de se restreindre aux fonctions  $k$  telles que l'évaluation en un point  $x$  de  $V$  soit une forme linéaire continue lorsqu'on la voit comme

$$\delta_x = \begin{cases} (V, \|\cdot\|_V) \rightarrow (\mathbb{R}, |\cdot|) \\ f \mapsto f(x) \end{cases} \quad (10.38)$$

Une condition suffisante à cela est que  $\widehat{k}$  soit à la fois continue et dans  $L^1$ . Le Théorème de Représentation de Riesz dans un espace de Hilbert permet alors d'identifier  $\delta_x$  à la fonction continue  $k \star \delta_x : y \mapsto k(y - x)$ , grâce à l'identité

$$\forall f \in V, \quad f(x) = \langle \delta_x, f \rangle = \langle k \star \delta_x, f \rangle_V \quad (10.39)$$

Cette approche a permis une fois de plus d'identifier les mesures à des formes linéaires continues sur des espaces de fonctions continues. La norme naturelle pour une mesure  $\mu$  est alors la norme duale  $\|\mu\| = \sup_{\|f\|_V \leq 1} \langle \mu, f \rangle$ . On l'explicite par la méthode habituelle des multiplicateurs de Lagrange. On se ramène à maximiser la quantité  $\langle \mu, f \rangle$  sous la contrainte  $\|f\|_V = 1$ . La condition de colinéarité des gradients assure l'existence d'un réel  $\lambda$  tel que la fonction maximisante  $f$  vérifie  $f = \lambda k \star \mu$ . De

$$1 = \|f\|_V^2 = \langle \lambda k \star \mu, \lambda k^{(-1)} \star k \star \mu \rangle = \lambda^2 \langle \mu, \lambda k \star \mu \rangle \quad (10.40)$$

on déduit  $\lambda = \frac{1}{\sqrt{\langle \mu, \lambda k \star \mu \rangle}}$ , puis  $f = \frac{k \star \mu}{\sqrt{\langle \mu, \lambda k \star \mu \rangle}}$ . Finalement

$$\|\mu\| = \left\langle \mu, \frac{k \star \mu}{\sqrt{\langle \mu, \lambda k \star \mu \rangle}} \right\rangle = \frac{\langle \mu, \lambda k \star \mu \rangle}{\sqrt{\langle \mu, \lambda k \star \mu \rangle}} = \sqrt{\langle \mu, \lambda k \star \mu \rangle} = \|\mu\|_k \quad (10.41)$$

et la norme  $\|\cdot\|_k$  s'identifie naturellement à la **norme duale** sur  $V$

$$\forall \mu \in V^*, \quad \|\mu\|_k = \langle \mu, k \star \mu \rangle = \sup_{\|f\|_V \leq 1} \langle \mu, f \rangle \quad (10.42)$$

*Remarque 6.* On s'est ramené à une situation analogue à celle qu'on avait pour  $\|\cdot\|_\infty$  dans l'espace de Sobolev associé à  $L^\infty(\mathbb{R}^n)$ . A ceci près que désormais, on dispose d'une distance capable de comparer les formes.

## 10.4 Application : recalculer des formes

On s'intéresse ici au recalcul de formes dans le cas du choix particulier de  $k : x \mapsto -\|x\|$ . Ce noyau n'est pas de Sobolev, et l'annulation pour la fréquence 0 restreint la comparaison à des mesures de même masse. Il s'agit à présent de trouver la transformation rigide qui minimise la distance entre deux formes données. On considère ici des nuages de points bidimensionnels, nos objets tests étant les deux cœurs de la page 1 représentées par les mesure  $\alpha$  et  $\beta$ . La première étape consiste à implémenter les fonctions auxiliaires permettant le calcul de la fonction de coût (distance entre les mesures), encodée dans *fidelity*.

```

def scal( f, g ) :
    "Scalar product between two vectors."
    return torch.dot( f.view(-1), g.view(-1) )

def sqdistances(x_i, y_j) :
    "Matrix of squared distances, C_ij = |x_i-y_j|^2."
    return ( (x_i.unsqueeze(1) - y_j.unsqueeze(0)) ** 2).sum(2)

def distances(x_i, y_j) :
    "Matrix of distances, C_ij = |x_i-y_j|."
    return (x_i.unsqueeze(1) - y_j.unsqueeze(0)).norm(p=2,dim=2)

def fidelity(alpha, beta) :
    "Energy Distance between two sampled probability measures."
    alpha_i, x_i = alpha
    beta_j, y_j = beta
    K_xx = -distances(x_i, x_i)
    K_xy = -distances(x_i, y_j)
    K_yy = -distances(y_j, y_j)
    cost = .5*scal( alpha_i, K_xx @ alpha_i ) \
          - scal( alpha_i, K_xy @ beta_j ) \
          + .5*scal( beta_j, K_yy @ beta_j )
    return cost

class RigidRegistration(Model) :
    "Find the optimal translation, scaling and rotation."
    def __init__(self, alpha) :
        "Defines the parameters of a rigid deformation of alpha."
        super(Model, self).__init__()
        self.alpha, self.x = alpha[0].detach(), alpha[1].detach()
        self.lam = Parameter(tensor( 0. )) # Log-Scale
        self.theta = Parameter(tensor( 0. )) # Angle
        self.tau = Parameter(tensor( [0.,0.] )) # Position

    def __call__(self, t=1.) :
        # At the moment, PyTorch does not support complex numbers...
        x_t = (t*self.lam).exp()*rot(self.x, t*self.theta) + t*self.tau
        return self.alpha, x_t

    def cost(self, target) :
        "Returns a cost to optimize."
        return fidelity( self(), target)

```

On implémente ensuite les transformations rigides, caractérisées par trois paramètres :  $\lambda$  coefficient de dilatation,  $\theta$  angle de la rotation, et  $\tau$  translation. La transformation s'écrit alors

$$f_{\lambda, \theta, \tau} : x \mapsto \lambda R_{\theta}(x) + \tau \quad (10.43)$$

[CT13]

Valentin Carlier,

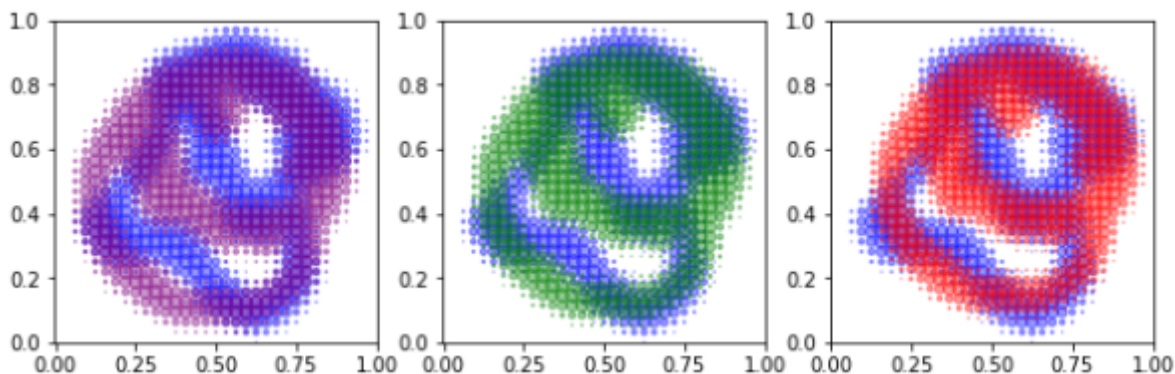
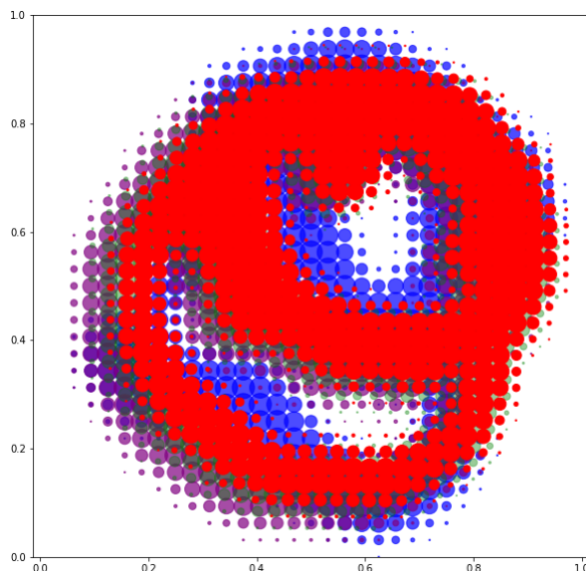
Louis-Pierre

Chaintron

Retour aux fondamentaux

Il reste alors à minimiser la distance entre  $\beta$  et  $f_{\lambda,\theta,\tau}(\alpha)$ , vue comme une mesure image mais également mesure associée aux points de la première figure à laquelle on a appliqué la transformation rigide. On calcule donc le minimum de

$$g : (\lambda, \theta, \tau) \mapsto \|f_{\lambda,\theta,\tau}(\alpha) - \beta\|_k \quad (10.44)$$



## 10.5 Conclusion

La comparaison des formes nous a d'abord entraîné à changer de point de vue en considérant nos nuages de points comme des mesures. Le problème s'est alors ramené à construire une distance permettant de comparer efficacement deux mesures. La solution présentée ici consiste à utiliser des normes à noyau, en voyant nos mesures comme des formes linéaires sur l'analogie d'un espace de Sobolev. Une telle norme correspond à la donnée d'une fonction  $k$  suffisamment régulière, qu'on choisira en pratique avec des dérivées nulles en zéro est un « rayon de floutage » suffisamment large pour préserver l'information que le nuage de points donne sur la forme. D'autres méthodes fructueuses pour construire des distances sur les mesures consisteraient à utiliser des estimations probabilistes, ou des techniques de transport optimal.

# Chapitre 11

## Géodésiques Riemanniennes, équations Hamiltoniennes

Séance 11

Antoine Le Calvez  
Raphaël Barboni  
Deux résultats  
remarquables

### 11.1 Métrique riemannienne, distance de chemins

#### Rappel

On définit en tout point d'une variété riemannienne  $M$  une métrique locale  $g$  qui sert à évaluer les distances dans le plan tangent.

**Définition 11.1.** On appelle métrique sur  $\mathbb{R}^n$  une matrice symétrique définie positive  $g$  définissant un nouveau produit scalaire  $\langle gv, w \rangle$  et donc une nouvelle norme :

$$\|v\|_g^2 = \langle gv, v \rangle$$

On définit alors sur le dual la cométrique  $K := g^{-1}$

*Remarque 11.1.* Si  $p = gv$ ,  $\|v\|_g^2 = \langle gv, v \rangle = \langle p, v \rangle = \langle p, Kp \rangle = \|p\|_K^2$

**Définition 11.2** (Métrique Riemannienne). On appelle métrique riemannienne sur une variété  $M$  une application qui associe à tout point de la variété une métrique sur l'espace tangent en ce point.

$$g : q \in M \mapsto g_q \in T_q M$$

Pour des raisons pratiques, on supposera de plus que les métriques sont  $\mathcal{C}^1$ .

**Définition 11.3** (Coût d'un chemin). Etant donné  $q : [0; 1] \rightarrow M$ ,  $\mathcal{C}^1$  par morceaux, un chemin sur une variété, on définit le coût et la longueur de  $q$  par :

$$C(q) = l(q)^2 = \int_0^1 \|q'(t)\|_{q(t)}^2 dt$$

Cela permet de munir  $M$  d'une structure métrique en posant :

$$d(a, b) = \inf_{\substack{q: [0,1] \rightarrow M \\ q(0)=a, q(1)=b}} l(q)$$

**Définition 11.4** (Plus court chemin). On appelle plus court chemin de  $a$  à  $b$  un chemin dont la longueur réalise  $d(a, b)$ . Les géodésiques sont localement des plus courts chemins, elles minimisent le coût entre deux points proches.

Dans la suite notre but sera de caractériser par une équation les géodésiques sur une variété.

## Equation hamiltonienne des géodésiques

Dans un chapitre précédent nous avons pu décrire explicitement les géodésiques de trois espaces homogènes : le plan euclidien, la sphère et le plan hyperbolique (ou disque de Poincaré). Grâce aux nombreuses isométries de ces espaces, nous avons en effet réussi à nous ramener à des cas simples, accessibles au calcul et au raisonnement direct.

En toute généralité, ce n'est plus possible : la Figure 11.1 montre bien que sur une surface quelconque, on ne peut espérer de description exhaustive simple des courbes géodésiques. D'ailleurs, ce n'est pas parce qu'une courbe est *géodésique*, ou "droite" (ce qui est une notion locale : entre deux points proches, elle minimise la distance) qu'elle est nécessairement un plus court chemin entre ses deux extrémités : on peut souvent trouver d'autres géodésiques plus astucieuses, qui coupent à travers les "vallées" !

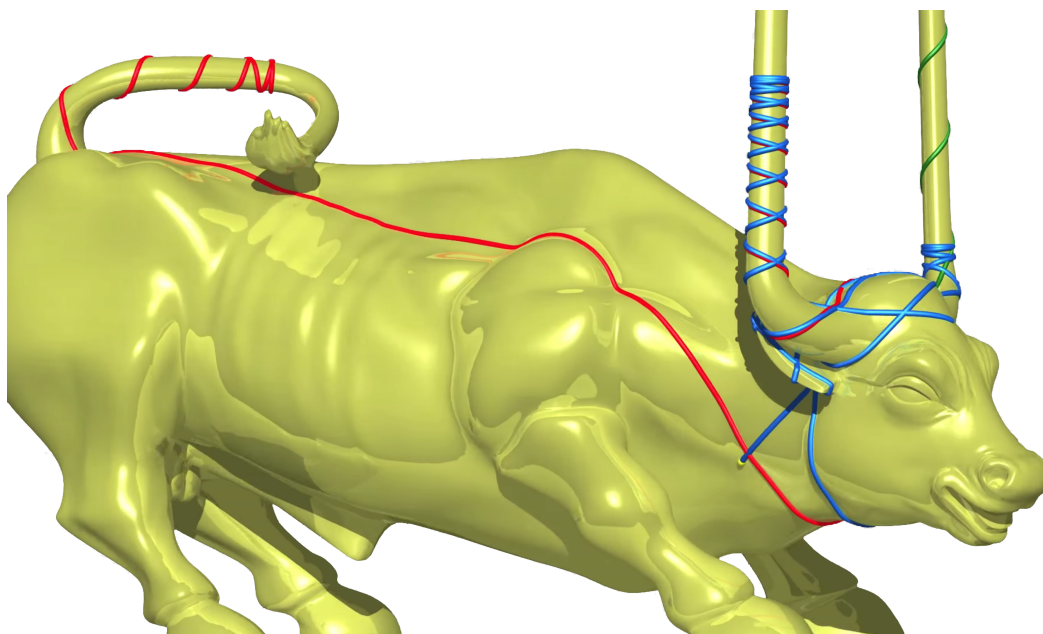


FIGURE 11.1 – Trois exemples de trajectoires géodésiques – i.e. qui minimisent localement la longueur – sur une statue de taureau. On l'aura compris, l'ensemble des "lignes droites" est ici bien plus riche que sur la sphère homogène !

Image tirée du très bon film *Chaos* de Jos Leys, Étienne Ghys et Aurélien Alvarez, accessible librement à l'adresse suivante : [www.chaos-math.org](http://www.chaos-math.org).

Nous allons établir une équation différentielle vérifiée par les plus courts chemins, ce qui permettra de les calculer algorithmiquement (à l'aide d'une méthode d'Euler par exemple).

Ici on travaille sur  $U$ , ouvert de  $\mathbb{R}^n$ .

**Proposition 11.1.** Soit  $q : [0; 1] \rightarrow U$  un chemin paramétré de classe  $C^1$ . On note  $v = \dot{q}$ . Les assertions suivantes sont équivalentes :

$$i \quad q \text{ minimise } \int_0^1 \|v(t)\|_{q(t)}^2 dt$$

$$ii \quad q \text{ minimise } \int_0^1 \|v(t)\|_{q(t)} dt \text{ et } v \text{ est de norme constante.}$$

*Démonstration.* C'est une conséquence de l'inégalité de Cauchy-Schwartz. □



**Définition 11.5** (Energie cinétique). On pose, par analogie avec la physique, l'énergie cinétique du couple  $(q, v)$  par :

$$E : (q, v) \in \mathbb{R}^n \times \mathbb{R}^n \longmapsto \frac{1}{2} \|v\|_q^2 = \frac{1}{2} \langle g_q v, v \rangle$$

Comme on a considéré la métrique  $\mathcal{C}^1$ ,  $E$  l'est aussi. On a :

$$C(q) = 2 \int_0^1 E(q(t), \dot{q}(t)) dt$$

**Lemme 11.1** (Formule de Christoffel). Soit  $q$  un plus court chemin sur  $U$ , on a :

$$\frac{d}{dt} \partial_v E(q, \dot{q}) = \partial_q E(q, \dot{q})$$

*Démonstration.* Si  $q$  est un plus court chemin entre  $a$  et  $b$  dans  $U$ . Par définition,  $q$  est un point critique de la fonction de coût. On se ramène alors à une dérivation en dimension 1 en perturbant  $q$  par  $xQ$  où  $Q$  est un chemin s'annulant en 0 et en 1 et où  $x \in \mathbb{R}$  est un paramètre proche de 0. On pose  $f(x) := C(q + xQ)$  définie et dérivable au voisinage de 0 avec  $f'(0) = 0$ . La densité des chemins s'annulant en 0 et en 1 dans l'ensemble des chemins permet ensuite de conclure.  $\square$

*Exercice 11.1* (Géodésique de  $\mathbb{R}^n$ ). On retrouve avec la formule précédente que les géodésiques de  $\mathbb{R}^n$ , muni de la métrique euclidienne, sont les courbes d'accélération nulle, c'est à dire les lignes droites.

Plutôt que d'essayer de développer cette expression, on pose :

$$p := g_q v = \partial_v E(q, v)$$

On effectue alors le changement de variable :  $(q, p) := (q, g_q v)$

*Remarque 11.2.* Le couple de variable vit désormais dans le fibré cotangent, que l'on appelle aussi l'espace des phases.

**Définition 11.6** (Hamiltonien). Avec ces nouvelles variables, on redéfinit :

$$H(q, p) = \frac{1}{2} \langle p, K_q p \rangle = E(q, K_q p)$$

On a alors :

$$\partial_q H(q, p) = -\partial_q E(q, K_q p) \partial_p H(q, p) = K_q p$$

Cette reparamétrisation du problème nous permet alors d'avoir la caractérisation suivante :

**Théorème 11.1** (Equations de Hamilton). On a équivalence entre les deux propositions suivantes :

1.  $\gamma$  est une courbe géodésique, qui minimise (localement) la distance.
2. La trajectoire relevée  $(q_t, p_t)$  dans l'espace des phases suit le flot hamiltonien correspondant au gradient symplectique (tourné de  $-\frac{\pi}{2}$ ) de  $H$  :

$$(\text{Ham}) : \begin{cases} \dot{q}_t &= +\frac{\partial H}{\partial p}(q_t, p_t) &= +K_{q_t} p_t \\ \dot{p}_t &= -\frac{\partial H}{\partial q}(q_t, p_t) &= -\partial_q(p_t, K_q p_t)(q_t) \end{cases} \quad (11.1)$$

Remarque 11.3. Pour  $q$  un plus court chemin, on a :

$$\begin{pmatrix} \dot{q} \\ p \end{pmatrix} = R_{-\frac{\pi}{2}} \nabla H \begin{pmatrix} q \\ p \end{pmatrix}$$

$\begin{pmatrix} \dot{q} \\ p \end{pmatrix}$  est orthogonal à  $\nabla H$  tout le long du chemin donc  $q$  suit une équipotentielle : on retrouve que  $H(q, p)$  est constant tout le long du chemin.

Remarque 11.4. La réciproque est fautive : un chemin vérifiant les équations de Hamilton n'est pas forcément un plus court chemin entre ses deux extrémités. On peut seulement affirmer que c'est localement un plus court chemin, c'est à dire un géodésique.

## Lien avec les Christoffels

On a défini précédemment l'équation des géodésiques dans un système de coordonnées locales. En effet, au lieu d'effectuer un changement de variables dans la formule de Christoffel, on peut choisir de développer les termes :

Propriété 11.1 (Equation des géodésiques). Si  $q$  est un chemin sur la variété  $M$ , on a que  $q$  est une géodésique si et seulement si elle vérifie "l'équation des géodésiques avec symboles de Christoffel" : dans un système de coordonnées locales  $q_t = (q_t^1, \dots, q_t^D)$ ,

$$\forall d \in 1, \dots, D, \quad \ddot{q}_t^d = - \sum_{i,j=1}^D \Gamma_{ij}^d(q_t) \cdot \dot{q}_t^i \dot{q}_t^j, \quad (11.2)$$

où le symbole de Christoffel  $\Gamma_{ij}^d(q)$  est donné par :

$$\Gamma_{ij}^d(q) = \frac{1}{2} \sum_{l=1}^D g^{dl}(q) \cdot (\partial_i g_{jl}(q) + \partial_j g_{il}(q) - \partial_l g_{ij}(q)) \quad (11.3)$$

avec  $g_{ij}$  les coefficients de la métrique, et  $g^{dl}$  ceux de son inverse, la cométrique.

L'accélération d'une particule géodésique  $\gamma_t$  est donc fonction lisse de sa position et de sa vitesse, avec un contrôle quadratique en cette dernière.

La description ci-dessus a un caractère déterministe : elle permet d'assurer que pour tout couple position-vitesse  $(q, v)$ , **il existe une unique trajectoire géodésique sur la variété passant en  $t = 0$  par le point  $q$  avec la vitesse  $v$** . C'est une conséquence du Théorème de Cauchy-Lipschitz, mais il semble difficile d'aller plus loin : l'équation 11.2 manque de structure, et est trop générique pour pouvoir être vraiment travaillée au corps. En particulier, la nécessité du calcul des coefficients de la métrique  $g$  et de la cométrique  $K$  rend l'implémentation numérique compliquée.

## 11.2 Application numérique : un code de tir géodésique

On retrouve l'objectif des séances précédentes : étant donné deux nuages de points sur le plan, on cherche à définir de manière cohérente une distance entre ces deux nuages et à établir une transformation continue d'un nuage vers l'autre. On utilise pour cela la caractérisation des géodésiques par l'équation avec le hamiltonien et l'on résout le problème à l'aide d'une implémentation numérique de la méthode d'Euler pour les équations différentielles. Les géodésiques sont représenté dans la carte exponentielle. Etant donné une origine et un vecteur vitesse initiale, la carte exponentielle donne la géodésique correspondante.

# Chapitre 12

## Cométries à noyaux, méthode LDDMM

Séance 12

Emmanuel Kammerer  
Léo Hahn-Lecler

### 12.1 Introduction

Rappelons notre motivation initiale : on cherche des géodésiques sur l'espace des nuages de points muni de la métrique de transport corrélé, c'est-à-dire trouver un chemin  $q$  minimisant

$$c(q) = \int_0^1 \|\dot{q}(t)\|_{q(t)}^2 dt$$

où  $\|\dot{q}(t)\|_{q(t)}^2 = \langle g_{q(t)}\dot{q}(t), \dot{q}(t) \rangle$ . Cependant, dans les équations hamiltonniennes, ce sont les moments et la cométrie  $K_q = g_q^{-1}$  qui interviennent. À partir de la métrique  $g$ , il faudrait inverser sa matrice pour pouvoir appliquer les algorithmes, d'où l'intérêt de se donner dès le départ la cométrie. Ensuite se pose une deuxième question : quelle cométrie adopter ? On cherche à la fois une cométrie facilitant les calculs et une cométrie qui limite les déformations.

### 12.2 Cométries à noyaux : l'algorithme

#### 12.2.1 Coût réduit des cométries à noyaux

Rappelons le schéma algorithmique utilisé pour le calcul de géodésiques.

On avait posé  $H(q, p) = \frac{1}{2}p^T K_q p$  et une condition nécessaire pour être une géodésique était de vérifier l'équation différentielle suivante :

$$(1) \quad \frac{dq}{dt} = + \frac{dH}{dp}$$

$$(2) \quad \frac{dp}{dt} = - \frac{dH}{dq}$$

On pose alors  $\text{cost}(p) = \|\text{shoot}(q_i, p) - q_f\|_2^2$  et l'on cherche à minimiser cette fonction par une descente de gradient.

Pour limiter la complexité de notre algorithme il faut donc que  $H(q, p)$  soit facile à calculer en tout point et nous devons disposer d'une procédure de différentiation rapide et précise.

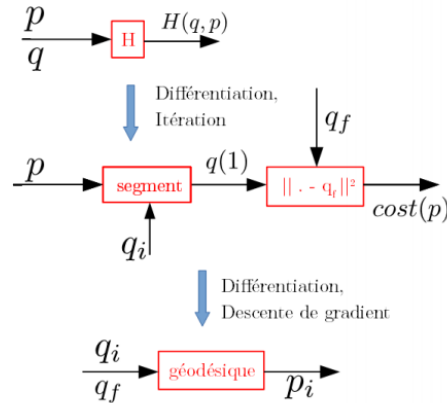


FIGURE 12.1 – Schéma pour obtenir une géodésique

Pour ce qui est du calcul de  $H(q, p) = \frac{1}{2}p^T K_q p$ , la multiplication par  $p$  va être rapide avec PyTorch car on va pouvoir effectuer de nombreux calculs simples en parallèle en utilisant la carte graphique. Reste donc le calcul de  $K_q = g_q^{-1}$  qui dépend de la métrique choisie.

Pour la différentiation, il existe deux méthodes : le calcul approché et la différentiation automatique.

**Calcul approché.** Si  $f : \mathbb{R}^n \rightarrow \mathbb{R}^d$  on a  $\frac{df_i}{dx_j}(x) = \lim_{h \rightarrow 0} \frac{f_j(x + h e_i) - f_j(x)}{h}$  on peut donc calculer la dérivée de manière approchée par la formule  $\frac{f_j(x + h_0 e_i) - f_j(x)}{h_0}$  pour un  $h_0$  petit. Cependant, l'expression de  $f$  est souvent (et ce sera le cas pour nous) trop compliquée pour que ce calcul soit parallélisable.

**Différentiation automatique.** On calcule la dérivée de manière exacte, comme le ferait un humain. Cette méthode est plus adaptée à PyTorch.

Une remarque importante que l'on peut faire est que la métrique  $g_q$  n'intervient pas dans nos calculs qui utilisent uniquement  $K_q = g_q^{-1}$ . Il est donc intéressant de se donner directement la cométrie  $K_q$  pour ne pas devoir inverser (coûteux algorithmiquement) une matrice à chaque fois que l'on veut calculer  $H(q, p)$ .

Maintenant que l'on a décidé de se donner la cométrie  $K_q$ , il faut la construire rapidement. Cela fait  $(Nd)^2$  coefficients à calculer, où  $d$  est la dimension de l'espace et  $N$  le nombre de points dans le nuage. Pour cela on peut utiliser la cométrie à noyaux

$$K_q = \begin{pmatrix} k(q_1 - q_1)I_d & k(q_1 - q_2)I_d & \cdots & k(q_1 - q_N)I_d \\ k(q_2 - q_1)I_d & k(q_2 - q_2)I_d & \cdots & k(q_2 - q_N)I_d \\ \vdots & \vdots & \ddots & \vdots \\ k(q_N - q_1)I_d & k(q_N - q_2)I_d & \cdots & k(q_N - q_N)I_d \end{pmatrix}, \quad (12.1)$$

où  $k$  est une fonction de noyau (par exemple une gaussienne  $k(x) = e^{-\frac{\|x\|^2}{2\sigma^2}}$ ). Elle est symétrique et on vérifiera plus tard qu'elle est définie positive. Cette matrice est égale au produit tensoriel de la matrice  $I_d$  et de la matrice de noyau réduite  $k_q = [k(q_i - q_j)]_{i,j}$ . On peut calculer rapidement

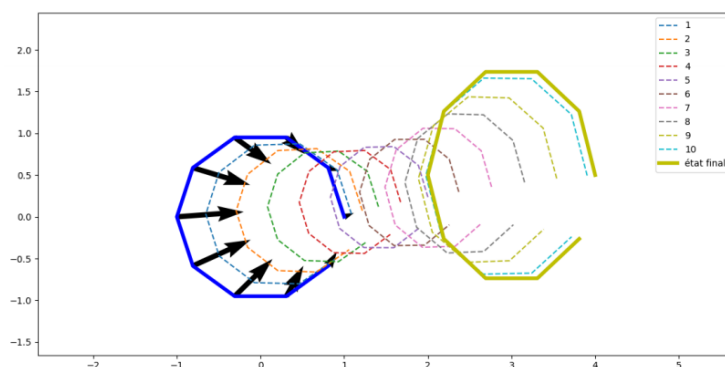


FIGURE 12.2 – Géodésique entre un cercle et une ellipse

la matrice car la formule est parallélisable. De même pour le produit par un vecteur : on verra qu'on peut réécrire le produit  $K_q v$  sous la forme d'un produit de convolution.

Ainsi, dans le TP, l'algorithme cherche un chemin minimisant le coût qui dépend du moment initial  $p_0$  (et non de la vitesse initiale). Le coût est composé de deux termes : un terme contrôlant si l'on est proche du nuage ciblé et un terme correspondant au coût du chemin.

Remarquons que la cométrie à noyaux permet bien d'empêcher les déchirures. En effet, elles associent un poids fort aux points proches. Cependant, on peut remarquer que les points proches ont même tendance à se resserrer durant le trajet pour s'écarter à la fin. Comment expliquer ce phénomène de "covoiturage" ?

## 12.3 Comprendre les cométries à noyaux

### 12.3.1 Le cas discret

**Calcul de la métrique associée sur un exemple jouet** Les cométries à noyaux sont extrêmement simples à calculer : il suffit d'une application de la fonction noyau pour chaque terme  $(q^i - q^j)$ . La matrice de noyau réduite pourra alors être vue comme une matrice de corrélation entre les points du nuage, qui associe un poids fort aux paires de points proches.

On a vu que la métrique était rapide à calculer, mais reste à voir si la métrique associée fait sens ! Prenons le temps de détailler les calculs dans un cas simple, avec  $k(x) = e^{-\frac{\|x\|^2}{2\sigma^2}}$  le noyau gaussien d'échelle  $\sigma$ ,  $d = 2$  et  $N = 6$ .

On considère un état  $q \in \mathbb{R}^{12}$ , donné par six points  $(q_1, q_2, q_3, q_4, q_5, q_6)$  du plan sur lesquels on fait les hypothèses simplificatrices suivantes :

- Les points se répartissent en trois groupes  $(q_1)$ ,  $(q_2, q_3, q_4)$ , et  $(q_5, q_6)$ , qui sont éloignés les uns des autres à une distance très grande devant l'échelle  $\sigma$  du noyau.
- $q_5$  et  $q_6$  sont à une distance  $l$  donnée l'un de l'autre.
- Le groupe  $(q_2, q_3, q_4)$  est un triangle équilatéral de côté  $l$ .

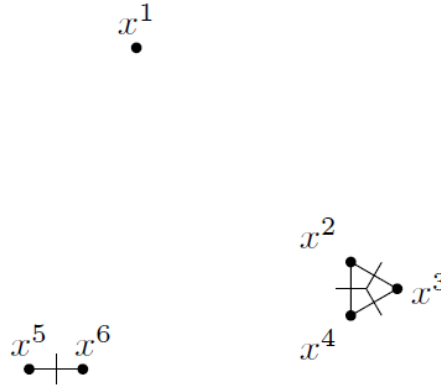


FIGURE 12.3 – Nuage de six points

Sous ces hypothèses, on peut écrire très simplement la matrice de noyau réduite en le nuage de points  $q$  :

$$k_q = \begin{pmatrix} 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & a & a & \cdot & \cdot \\ \cdot & a & 1 & a & \cdot & \cdot \\ \cdot & a & a & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & a \\ \cdot & \cdot & \cdot & \cdot & a & 1 \end{pmatrix}, \tag{12.2}$$

où  $a = e^{\frac{-t^2}{2\sigma^2}} \in [0, 1]$ , et où l'on a remplacé les termes négligeables par des points.

À la limite (quand les groupes de points sont très éloignés), on obtient un matrice diagonale par blocs de la forme :

$$B_n(a) = (1 - a) \cdot I_n + a \cdot (1) \ ^t(1) = \begin{pmatrix} 1 & & & & & \\ & 1 & & (a) & & \\ & & \ddots & & & \\ & (a) & & 1 & & \\ & & & & 1 & \\ & & & & & 1 \end{pmatrix}. \tag{12.3}$$

Pour trouver la métrique associée, c'est-à-dire inverser  $k_q$ , inversons les blocs  $B_n(a)$  associés aux amas de  $n$  points.

Pour inverser  $B_n(a)$ , on peut remarquer qu'elle est symétrique et la diagonaliser pour trouver son inverse. Un vecteur propre évident est  $e = \frac{(1)}{\|(1)\|}$ , associé à la valeur propre  $1 + (n - 1)a$  et on remarque que  $B_n(a)$  agit comme  $(1 - a)$  fois l'identité sur l'orthogonal. Ainsi,

$$B_n(a)^{-1} = \frac{1}{1 + (n - 1)a} e^t e + \frac{1}{1 - a} (I_n - e^t e)$$

Si on décompose un vecteur  $v$  sous la forme

$$v = e^t(ev) + (v - e^t(ev)) = v_{\text{moy}} + v_{\text{var}}, \tag{12.4}$$

une partie ‘‘moyenne’’ constante et une partie de somme nulle, la variance. On a

$${}^t v(B_n(a))^{-1} v = \frac{1}{1 + (n-1)a} \|v_{\text{moy}}\|_2^2 + \frac{1}{1-a} \|v_{\text{var}}\|_2^2 \quad (12.5)$$

**Interprétation** Tout d’abord, un fait rassurant : le rôle particulier joué par les champs de vitesses constants, ‘‘colinéaires’’. Lorsque  $n$  est grand, le coût associé à une translation est petit par rapport à celui des autres transformations. Le coût associé à un champ de vitesses sur l’amas est la somme d’un terme de *translation*, proportionnel à  $\|v_{\text{moy}}\|_2^2$ , et d’un terme de *régularisation* pénalisant la non-uniformité en  $\|v_{\text{var}}\|_2^2$ .

Lorsque  $l \gg \sigma$ ,  $a$  est petit devant 1. On a alors

$$\frac{1}{1 + (n-1)a} \simeq 1 \simeq \frac{1}{1-a}, \quad (12.6)$$

l’équilibre entre les deux pénalisations. Le coût  ${}^t v(B_n(a))^{-1} v$  est simplement égal au coût  ${}^t v v = \|v\|_2^2$ , cela se ramène bien au cas où on a  $n$  blocs de taille 1.

À l’inverse, si le diamètre  $l$  de l’amas est petit devant  $\sigma$ , si le noyau voit les points de l’amas comme quasiment confondus, on aura  $a \simeq 1^-$  et par suite

$${}^t v(B_n(a))^{-1} v \simeq \frac{1}{n} \|v_{\text{moy}}\|_2^2 + \infty \|v_{\text{var}}\|_2^2. \quad (12.7)$$

Ainsi, les déchirures sont surpénalisées alors que le poids des translations est atténué en  $\frac{1}{n}$ . Tout se passe donc comme si notre amas de  $n$  particules se réduisait à un seul atome, très difficile à éclater mais aussi facile à transporter qu’une particule seule.

Dans notre exemple, la métrique (réduite) est :

$$g_q = (k_q)^{-1} = \left( \begin{array}{ccc|cc} 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & & & & & \\ \cdot & B_3(a)^{-1} & & \cdot & \cdot & \cdot \\ \cdot & & & \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & & \\ \cdot & \cdot & \cdot & \cdot & B_2(a)^{-1} & \cdot \end{array} \right) \quad (12.8)$$

Les trois amas sont donc complètement indépendants, ce qui n’est pas une surprise puisqu’ils sont décorrélés au sens de  $k$ .

**Interprétation** Étant donné un champ de vitesses  $v = (v_1, v_2, v_3, v_4, v_5, v_6)$ , comment se trouve-t-il pénalisé par  $g_q$ ? Les comportements limites se déduisent de ce que l’on a établi pour un amas simple. Aussi, lorsque  $l \gg \sigma$  et donc  $a \simeq 0$ , on a :

$${}^t v g_q v \simeq \|v_1\|_2^2 + \|v_2\|_2^2 + \|v_3\|_2^2 + \|v_4\|_2^2 + \|v_5\|_2^2 + \|v_6\|_2^2. \quad (12.9)$$

Les points sont donc indépendants.

Par contre, si  $l \ll \sigma$ , alors un champ de coût fini s’écrit nécessairement :

$$(v_1, v_2, v_3, v_4, v_5, v_6) = (w_1, w_2, w_2, w_2, w_3, w_3), \quad (12.10)$$

avec  $w_i$  la ‘‘vitesse de groupe’’ de l’amas  $i$ . Autrement dit, on n’a le droit que de translater indépendamment chaque amas. On a :

$${}^t v g_q v = \|v_1\|_2^2 + \frac{1}{3} (\|v_2\|_2^2 + \|v_3\|_2^2 + \|v_4\|_2^2) + \frac{1}{2} (\|v_5\|_2^2 + \|v_6\|_2^2) \quad (12.11)$$

$$= \|w_1\|_2^2 + \|w_2\|_2^2 + \|w_3\|_2^2. \quad (12.12)$$

C'est comme si l'on avait affaire à trois particules indépendantes.

Sur cet exemple, la cométrie à noyaux réalise bien ce que l'on voulait : pénaliser les déchirures.

### 12.3.2 Rappel sur les normes de Sobolev

Avant les vacances, on avait défini (à peu de choses près) l'espace de Hilbert :

$$H = \left\{ f \in L^2(\mathbb{R}^d); \int_{\mathbb{R}^d} \frac{\|\hat{f}(\omega)\|^2}{\hat{k}(\omega)} d\omega < +\infty \right\}$$

muni de du produit scalaire hermitien

$$\langle f, g \rangle_k = \int_{\mathbb{R}^d} \frac{\langle \hat{f}(\omega), \hat{g}(\omega) \rangle_{\mathbb{C}^d}}{\hat{k}(\omega)} d\omega = \left\langle \hat{f}, \frac{\hat{g}}{\hat{k}} \right\rangle_{L^2}$$

duquel dérive la norme

$$\|f\|_k^2 = \int_{\mathbb{R}^d} \frac{\|\hat{f}(\omega)\|^2}{\hat{k}(\omega)} d\omega$$

On fait en sorte que les mesures  $\delta_x$  appartiennent au dual  $H^*$ .  $H$  est alors appelé un RKHS.

Si  $f \in H$  et  $\mu \in H^*$ , on a défini  $\langle \mu, f \rangle := \langle k * \mu, f \rangle_k$ .

Enfin, la norme duale d'une mesure  $\mu$  dans  $H^*$  est égale à

$$\|\mu\|_k^2 = \int_{\mathbb{R}^d} \|\hat{\mu}(\omega)\|^2 \hat{k}(\omega) d\omega$$

et  $H^*$  est muni du produit scalaire auquel on pense.

### 12.3.3 Principe de réduction

Dans cette partie, nous allons voir qu'à un plus court chemin dans l'espace des nuages de points  $\mathbb{R}^{Nd}$  correspond une déformation de l'espace ambiant (l'espace des champs de vecteurs de  $\mathbb{R}^d$ ) de coût minimal.

Déjà, remarquons que si  $q \in \mathbb{R}^{Nd}$  est un nuage de points,  $k(q_i - q_j) = \langle \delta_{q_i}, k * \delta_{q_j} \rangle$  pour tous  $i, j$ .

Si  $p$  est un nuage de moments et  $v = K_q p$ , notons  $\pi = \sum_{i=1}^N p_i \delta_{q_i} \in H^*$ . Alors on a

$$v_i = (K_q p)_i = \sum_{j=1}^N k(q_i - q_j) p_j = \sum_{j=1}^N \langle \delta_{q_i}, k * \delta_{q_j} \rangle p_j = \left\langle \delta_{q_i}, k * \sum_{j=1}^N p_j \delta_{q_j} \right\rangle = \langle \delta_{q_i}, k * \pi \rangle = (k * \pi)(q_i) = u(q_i)$$

en posant  $u = k * \pi \in H$ . Le champ (constant) de vecteurs  $u$  interpole alors le nuage de points  $v$ . Il peut s'écrire aussi sous la forme

$$u(x) = \sum_{i=1}^N p_i k(x - q_i)$$



Il est grand temps de vérifier que la métrique  $K_q$  est définie positive. On a :

$$\langle p, K_q p \rangle = \sum_i p_i (K_q p)_i = \sum_i p_i (k * \pi)(q_i) = \langle \pi, k * \pi \rangle_{L^2} = \left\langle \hat{\pi}, \widehat{k * \pi} \right\rangle_{L^2} = \left\langle \hat{\pi}, \hat{k} \hat{\pi} \right\rangle_{L^2} > 0$$

si  $p \neq 0$  car  $\hat{k} > 0$  et  $k$  est continue.

Remarquons que la norme définie sur  $H$  permet de voir un nouvel aspect de la métrique à noyaux : elle pénalise les hautes fréquences dans les champs de vitesses.

Maintenant que l'on a étendu la métrique aux champs de vitesses, il faut définir les transformations de champs vectoriels et leur coût.

**Définition 12.1.** Si  $u \in L^1([0, 1], H)$  est un champ de vecteurs variable, si  $q_0$  et  $q_1$  sont deux nuages de points, on dit que  $u$  transporte  $q_0$  sur  $q_1$  si  $q_1$  est l'image de  $q_0$  par le flot en  $t = 1$  de  $u$ . On appelle coût de  $u$  le réel  $c(u) = \int_0^1 \|u(t)\|_k^2 dt$ .

Soit  $V = \text{Vect}(\delta_{q_i})$  (de dimension finie). Alors  $H = V \oplus^\perp V^\perp$ . On montre alors que  $H^* = (k * V) \oplus^\perp (k * V)^\perp = (k * V) \oplus^\perp \{u; \forall i u(q_i) = 0\}$ .

Par Pythagore on obtient que, étant donné un nuage de points  $q$ , un champ de vitesses  $v$ ,  $p = g_q v$ ,  $\pi = \sum_{i=1}^N p_i \delta_{q_i}$ , et  $u = k * \pi$ , parmi tous les champs de vitesses interpolant  $v$  en  $q$ ,  $u$  est le champ de norme minimale.

Venons en au théorème important de cette section :

**Théorème 12.3.1.** Soit  $Q$  un plus court chemin entre  $q_0$  et  $q_1$  dans l'espace des nuages de points. Soit  $p(t) = g_q Q'(t)$ . Soit  $\pi(t) = \sum_{i=1}^N p_i \delta_{Q_i(t)}$  et  $u(t) = k * \pi(t)$  le champ de vitesse que correspond. Alors, pour tout  $t \in [0, 1]$ ,  $u(t)$  interpole  $Q'(t)$  en  $Q(t)$  et  $u$  est le champ de vitesses de coût minimal qui transporte  $q_0$  sur  $q_1$ . Il a le même coût que le chemin  $Q$ .

En d'autres termes, toute trajectoire géodésique dans l'espace des nuages de points peut se relever en une trajectoire géodésique dans l'espace des difféomorphismes de  $\mathbb{R}^d$  muni de la métrique riemannienne.

Initialement développé pour décrire la dynamique des *solitons* (ou "vagues solitaires"), ce résultat lie la cométrique  $K_q$  sur les espaces de landmarks à une action de déformation de l'espace ambiant. C'est, en un sens, le **résultat fondamental** de la théorie qui fait le lien entre transport optimal et analyse procustéenne : le **cadre LDDMM, pour Large Deformation Diffeomorphic Metric Mapping**.

## 12.4 conclusion

Les métriques à noyaux sont, en un sens, les moins chères sur un espace de nuages de points grâce à la programmation parallèle. L'analyse de la section précédente permet de comprendre leur comportement, et de garantir le caractère "difféomorphique" du recalage obtenu (car à un recalage sur les nuages de points correspond un recalage sur les difféomorphismes de coût minimal). Le choix du noyau permet de contrôler la régularité de la déformation.

Néanmoins, un effet indésirable subsiste : le "covoiturage" dont on avait parlé en première partie. On peut l'expliquer de la façon suivante : les phases de rapprochement et d'éloignement sont coûteuses (car les vitesses ne sont pas colinéaires) mais la phase centrale est avantageuse (car on a vu que translater  $n$  points à distance nulle coûte autant qu'en translater un seul). Ainsi, si la distance à parcourir est suffisamment longue, le covoiturage devient avantageux.

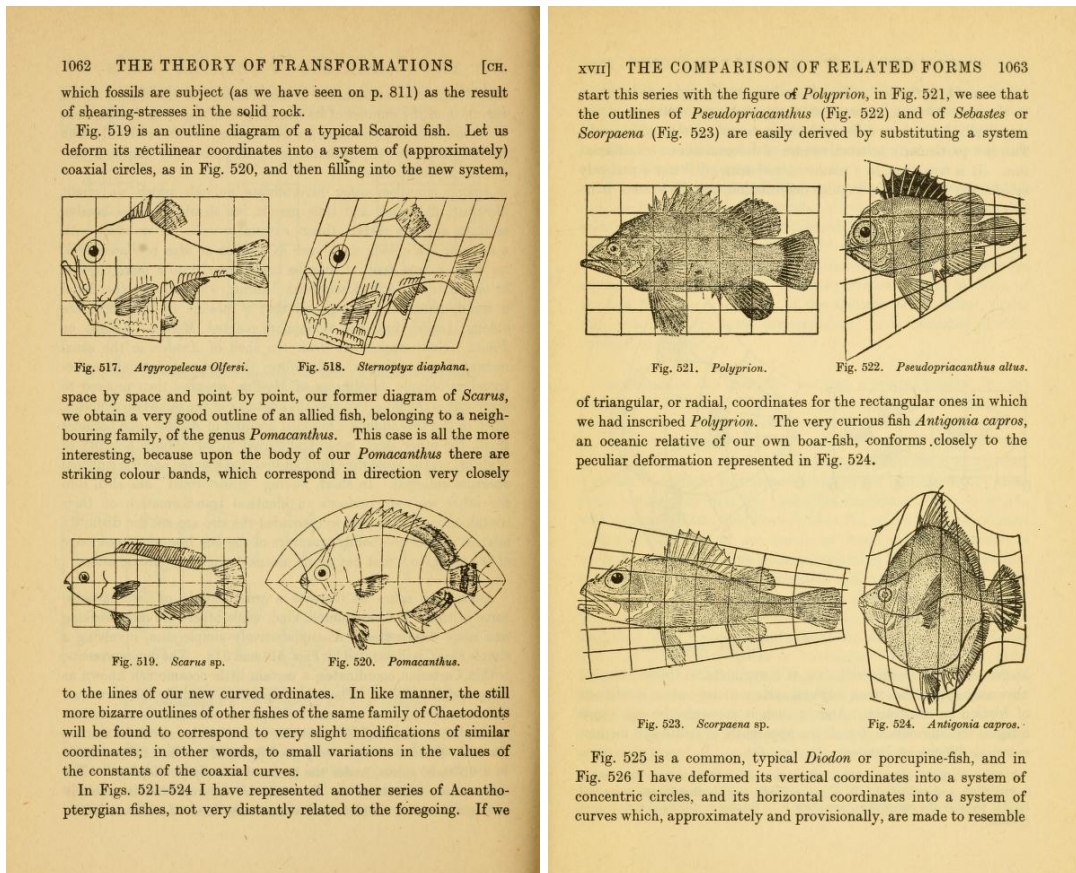


FIGURE 12.4 – Dans un ouvrage en tous points remarquable, le lettré, biologiste et mathématicien D’Arcy Wentworth Thompson (1860-1948) souligna l’importance des facteurs environnementaux et physiques (en opposition à la seule hérédité) dans la morphogenèse des êtres vivants. La forme des poissons étant peu ou prou optimale, il n’y a pas une infinité de “plans” différents les uns des autres mais bien une poignée seulement de patrons originaux, qui permettent par des déformations (non-triviales) d’engendrer toutes les formes privilégiées par l’évolution. Pour décrire la variabilité anatomique d’une famille ou d’une population observée, il suffit donc de donner un *template* de référence (arbitrairement complexe, mais commun à toutes les observations) et les déformations qui permettent de passer dudit template aux individus. La complexité se retrouve alors découpée en deux composantes intelligibles : une *image de référence*, complexe mais fixe ; des *déformations* propres aux sujets observées, souvent assez simples pour être décrites avec peu de paramètres. Fait remarquable : les schémas ci-dessus présentent la variabilité des formes de poissons non pas comme des déplacements arbitraires de squelettes, mais comme **des changements de coordonnées, des déformations de l’espace ambiant**. C’est sur cette idée mathématique de déformation *extrinsèque* de l’espace (en opposition aux mouvements intrinsèques des particules de poisson) que repose l’analyse procustéenne et la théorie “LDDMM” présentée dans ce chapitre. Grâce aux algorithmes détaillés dans les pages qui suivent, on peut aujourd’hui produire des figures de ce type de manière *automatique*. *On Growth and Forms*, 1946.

# Chapitre 13

## Méthodes à noyaux pour l'appariement de nuages de points

*Bonus*

*[Gla05]*

*Vadim Lebovici  
et Arthur Touati  
Fondamental*

### 13.1 Objectifs

#### 13.1.1 Cadre concret

Dans le domaine du médical, il est nécessaire de pouvoir comparer deux formes pour pouvoir détecter des maladies, du cœur ou du cerveau par exemple. Pour pouvoir le faire dans de grandes proportions et à une vitesse élevée, le numérique est indispensable. Il nous faut donc trouver un moyen de comparer deux formes. Pour ceci, nous déformons la première en une forme proche de la seconde et associons un coût – qui reflètera une distance – à une telle transformation.

Ceci consiste en fait en ce qu'on appelle un *problème d'appariement*, c'est-à-dire, à partir de deux objets mathématiques, déterminer le difféomorphisme optimal pour que le "transport" du premier soit "aussi proche du second que possible." Plus précisément,

1. On considère un ensemble  $E$  dans lequel vivent nos objets.
2. On considère un groupe de difféomorphismes  $G$  et une action de  $G$  sur  $E$  qui modélisera le transport.
3. On considère une distance  $d_E$  sur  $E$  pour mesurer l'appariement des objets après l'action d'un difféomorphisme sur le premier, c'est l'*attache aux données*.

#### 13.1.2 Solutions possibles

Une première solution serait de ne considérer qu'un groupe de difféomorphismes constitué de *déformations rigides*, c'est-à-dire des déformations linéaires, rotations, translations, dilatations. C'est ce que l'on a fait dans la séance précédente, et ceci nous apporte des résultats satisfaisant : on a même réussi à expliciter l'espace quotient des triangles à similitude près. Cependant, le transport d'une forme complexe vers une autre est peu satisfaisant. Un cerveau ne diffère pas d'un autre linéairement, et ceci ne permet pas de faire correspondre un bras tendu avec un bras plié par exemple.

On propose ici un modèle nettement plus riche. Pour cette séance, nous ne considérerons pas des formes continues, mais des formes discrétisées, afin de n'avoir qu'un nombre fini de points à traiter :  $E$  sera un espace de landmarks. Pour discrétiser, il est pertinent de définir à l'avance

Vadim Lebovici et Arthur Touati des points de références sur les formes que nous serons amenés à comparer (artère pulmonaire, artère aorte pour le cœur par exemple).  
Fondamental

## 13.2 Difféomorphismes

Supposer les déformations linéaires est trop restrictif pour faire de la reconnaissance de formes, c'est pourquoi il est nécessaire de se placer en dimension plus grande et même infinie pour s'autoriser des déformations plus complexes et ainsi accéder à des résultats plus fins. On se placera donc dans un espace de difféomorphismes.

Le schéma général est le suivant : on construit d'abord un espace de vecteurs dont la norme définira le coût d'une déformation infinitésimale et les difféomorphismes permettant de déformer nos formes seront obtenus par intégration d'une équation de flot. On doit pour ceci travailler avec un contrôle  $L^1$  ou  $L^2$  sur nos champs de vecteurs.

Construisons les champs utiles à la construction de nos déformations élastiques.

**Définition 13.1.** Un espace vectoriel  $V$  de champs de vecteurs sur  $\mathbb{R}^d$  sera dit *admissible* s'il vérifie les conditions suivantes :

1.  $V$  est un espace de Hilbert. On note sa norme  $\|\cdot\|_V$  et son produit scalaire  $\langle \cdot, \cdot \rangle_V$ .
2.  $(V, \|\cdot\|_V)$  s'injecte continûment dans  $(C_0^1(\mathbb{R}^d, \mathbb{R}^d), \|\cdot\|_{1,\infty})$  des champs de classe  $C^1$  sur  $\mathbb{R}^d$  s'annulant à l'infini, ainsi que leurs dérivées partielles. Il existe donc  $c_V > 0$  telle que :

$$\forall v \in V, \|v\|_{1,\infty} \leq c_V \|v\|_V \quad (13.1)$$

En fait, les difféomorphismes que nous allons définir sont vus comme solutions d'une équation de flot. On voit en fait un champ de vecteur  $v_t$  (modélisant des déformations infinitésimales), déformer notre espace au cours du temps. Arrivé au temps  $t$ , la déformation définit un difféomorphisme. On dispose d'un théorème qui nous assure l'existence et l'unicité de telles solutions. Cependant, il nous faut des hypothèses supplémentaires sur nos champs de vecteurs, un contrôle  $L^1$ . On définit alors

**Définition 13.2.** On définit les espaces suivant :

$$L_V^1 = L^1([0, 1], V) \quad (13.2)$$

$$L_V^2 = L^2([0, 1], V) \quad (13.3)$$

ainsi que les normes :

$$\|v\|_{L_V^1} = \int_0^1 \|v_t\|_V dt. \quad (13.4)$$

$$\|v\|_{L_V^2} = \sqrt{\int_0^1 \|v_t\|_V^2 dt}. \quad (13.5)$$

*Remarque 13.1.* L'espace  $(L_V^1, \|v\|_{L_V^1})$  est un Banach et  $(L_V^2, \|v\|_{L_V^2})$  est un Hilbert. De plus notons aussi que  $L_V^2$  s'injecte dans  $L_V^1$  par inégalité de Cauchy-Schwarz.

On peut alors, si l'on prend un champ de vecteur adéquat, énoncer le théorème suivant et construire nos difféomorphismes :

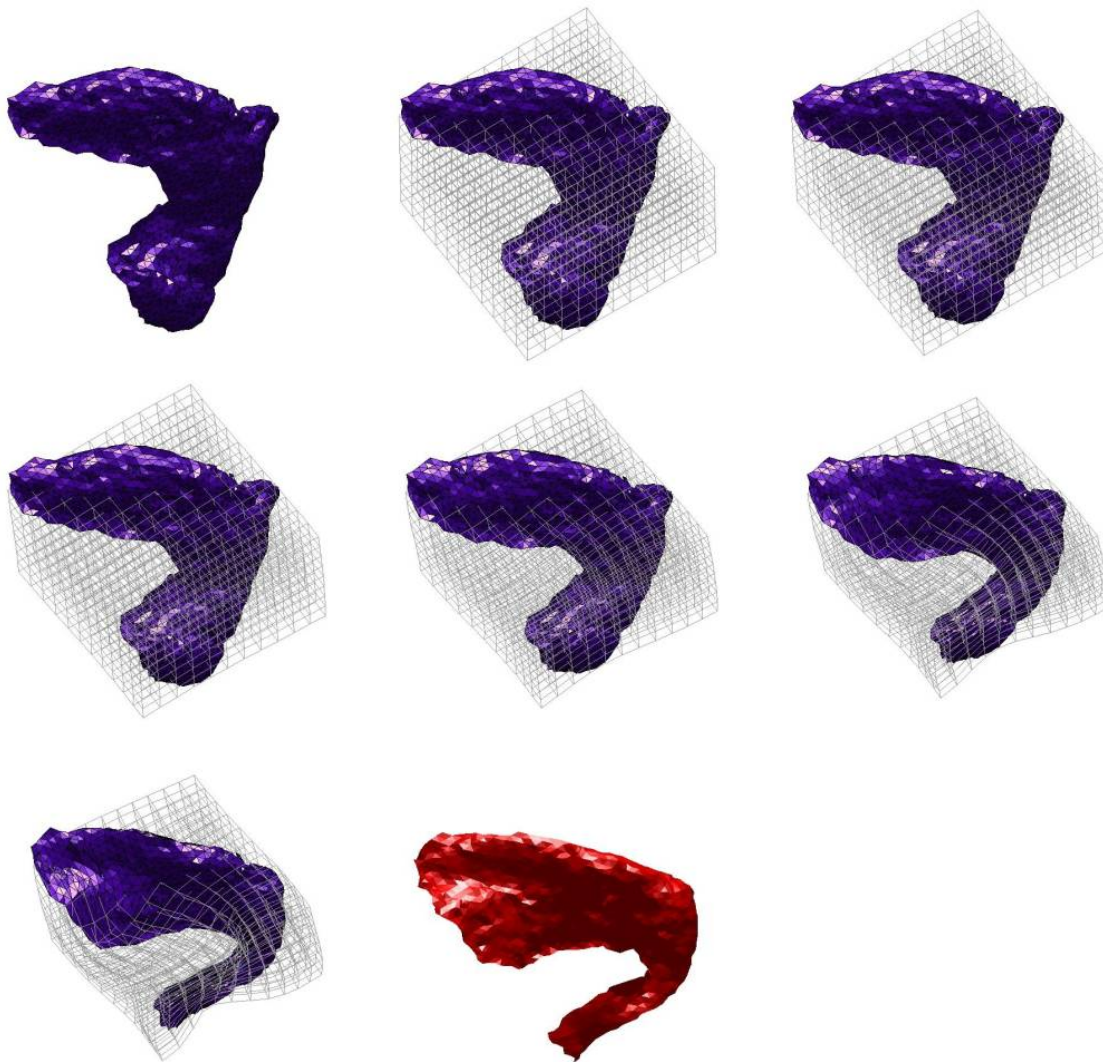


FIGURE 13.1 – Déformation d'une surface en une autre, sous l'action d'un difféomorphisme. Étape par étape, on voit la grille se déformer sous l'action d'un champs de vecteur  $v_t$ .

Vadim Lebovici  
et Arthur Touati  
Fondamental

**Théorème 13.1.** Soit  $v \in L_V^1$ . Pour tout  $x \in \mathbb{R}^d$ , il existe une unique application  $t \mapsto \Phi_t^v(x)$  continue de  $[0, 1]$  dans  $\mathbb{R}^d$  vérifiant l'équation de flot

$$\Phi_t^v(x) = x + \int_0^t v_s \circ \Phi_s^v(x) ds. \quad (13.6)$$

Autrement dit,

$$\Phi_0^v(x) = x \quad \dot{\Phi}_t^v(x) = v_t(\Phi_t^v(x)) \quad (13.7)$$

On peut en fait montrer que l'application  $x \mapsto \Phi_t^v(x)$  est un homéomorphisme de  $\mathbb{R}^d$  dans  $\mathbb{R}^d$ . Les difféomorphismes qui vont nous servir sont donc les éléments de l'ensemble

$$\mathcal{A}_V = \{\Phi_1^v \mid v \in L_V^1\} \quad (13.8)$$

En effet, il suffit de considérer les flots au temps 1 – renormalisation en temps –, quitte à changer le vecteur  $v$  en  $\tilde{v}$  tel que  $\tilde{v}_s = v_s$  si  $s \leq t$  et  $\tilde{v}_s = 0$  sinon.

Cette définition des difféomorphismes sur notre espace nous permet d'établir une métrique, et ainsi va nous permettre d'évaluer le coût associé à l'utilisation d'un difféomorphisme pour déformer notre espace.

**Proposition 13.1.**  $\mathcal{A}_V$  est un groupe et un espace complet pour la métrique :

$$d_V(id, \Phi) = \inf \{ \|v\|_{L_V^1} \mid v \in L_V^1, \Phi_V^1 = \Phi \}. \quad (13.9)$$

que l'on peut étendre par invariance à droite :

$$d_V(\Phi, \Psi) = d_V(id, \Psi \circ \Phi^{-1}). \quad (13.10)$$

En fait nous disposons d'un lemme qui nous permet de nous restreindre à l'étude des champs de vecteurs avec un contrôle  $L^2$  :

**Lemme 13.1.** Pour tout  $v \in L_V^1$ , il existe  $u \in L_V^2$  tel que  $\Phi_1^u = \Phi_1^v$  et  $\|u\|_{L_V^2} \leq \|v\|_{L_V^1}$ .

Ce qui nous permet de réduire la description de  $\mathcal{A}_V$  :

**Proposition 13.2.**

$$\mathcal{A}_V = \{\Phi_1^v \mid v \in L_V^2\}, \quad (13.11)$$

et pour tout  $\Phi \in \mathcal{A}_V$ ,

$$d_V(id, \Phi) = \inf \{ \|v\|_{L_V^2} \mid v \in L_V^2, \Phi_1^v = \Phi \}. \quad (13.12)$$

Si l'on prend deux difféomorphismes de  $\mathcal{A}_V$ , il existe une géodésique entre eux pour la métrique que l'on a définie.

**Proposition 13.3.** Soient  $\Phi, \Psi \in \mathcal{A}_V$ .

Il existe  $v \in L_V^2$  tel que  $\Phi_1^v = \Phi \circ \Psi^{-1}$  et  $d(\Phi, \Psi) = \|v\|_{L_V^2} = \|v\|_{L_V^1}$ . En fait pour un tel champ  $v$ , la norme  $\|v_t\|_V$  est constante pour presque tout  $t$ .

On dira que  $\mathcal{A}_V$  est le groupe de difféomorphismes modelé sur  $V$  à partir de l'identité.

*Exemple 13.1.* Dans le cas "limite" où  $V$  est l'espace des champs constant, identifié à  $\mathbb{R}^m$  muni de la métrique euclidienne,  $\mathcal{A}_V$  est tout simplement l'espace des translations muni de la distance euclidienne naturelle.

*Exemple 13.2.* De même, si  $V$  est l'espace des translations + rotations infinitésimales, alors  $\mathcal{A}_V$  est l'espace des isométries affines qui préservent l'orientation.

En prenant un espace  $V$  plus souple, on fait exploser le nombre de degrés de libertés, et donc la complexité de l'espace des difféomorphismes  $\mathcal{A}_V$ . Tout l'objet de cet exposé sera de voir comment choisir  $V$  pour garder un problème *raisonnable*, résoluble numériquement.

### 13.3 Retour sur le problème d'appariement

Revenons sur le problème d'appariement. Nous avons défini dans les sections précédentes :

1. L'espace  $E$  dans lequel vivent les formes est l'ensemble des nombres finis de points de  $\mathbb{R}^d$  avec souvent  $d = 2$  ou  $3$ .
2. Notre groupe déformant l'espace est  $\mathcal{A}_V$  qui agit de la manière suivante :

$$\Phi.O = \Phi(O) = \{\Phi(x), x \in O\} \quad (13.13)$$

C'est bien ça ??

3. Pour la distance à définir sur les formes pour quantifier la différence entre  $O_2$  et  $\Phi.O_1$  il nous faut une distance  $d_E$  sur  $E$ .

Le point 3 est assez subtil dans le cas de formes continues, de surfaces, etc, mais ce n'est pas l'objet de cet exposé.

Nous supposons ici disposer d'un nombre fini de *landmarks*, qui sont des points de références numérotés : pensez à un biologiste qui aurait annoté à la main une série d'images de poissons, en cliquant sur le bout du nez, la commissure des lèvres, les bouts de nageoires... On peut alors définir un terme d'attache aux données naïf, si  $O_1$  et  $O_2 \in E$  sont de la forme :  $O_1 = \{x_i, i \in \llbracket 1, n \rrbracket\}$  et  $O_2 = \{y_i, i \in \llbracket 1, n \rrbracket\}$  décrit par leurs points de références, on définit :

$$d_E = \sum_{i=1}^n |x_i - y_i|^2. \quad (13.14)$$

Tout ceci nous permet alors de poser un coût pertinent à la transformation de  $O_1$  à  $O_2$ , qu'on notera  $J$  et on peut définir proprement le problème d'appariement :

**Définition 13.3.** Soit  $A : \mathcal{A}_V \mapsto [0, +\infty[$  une fonctionnelle (appelée attache aux données), et  $\lambda > 0$  fixé. Le *problème d'appariement* pour  $A$  consiste à minimiser sur  $\mathcal{A}_V$

$$J(\Phi) = \lambda d_V(id, \Phi) + A(\Phi). \quad (13.15)$$

ou encore de manière équivalente à minimiser sur  $L_V^2$  la fonctionnelle :

$$J(v) = \lambda \int_0^1 \|v_t\|_V^2 dt + A(\Phi_1^v). \quad (13.16)$$

*Remarque 13.2.* L'équivalence entre les deux formulations provient du théorème d'existence de géodésiques sur  $\mathcal{A}_V$ . Attention : en dimension infinie, c'est un résultat compliqué !

Pour résoudre ce problème d'appariement, on possède un premier théorème d'existence de solution, qui nous assure que ceci n'est pas vain et qu'il existe effectivement des difféomorphismes qui vont nous permettre en pratique de passer d'une forme à une autre de façon optimale.

En pratique, il s'agira de trouver un difféomorphisme  $\Phi$ , qui déforme une forme  $X$  en une forme  $\Phi(X)$  proche de  $\tilde{X}$  : on prendra typiquement pour  $A$  une attache quadratique du type  $\|\Phi(X) - \tilde{X}\|^2$  avec une norme à préciser sur l'espace  $E$  des observations. La forme optimale  $\Phi(X)$  pourra alors être vue comme un *barycentre* des points  $X$  – de poids  $\lambda$  – et  $\tilde{X}$  – de poids 1.

Le réel  $\lambda$  est un poids de *régularisation*, qui force  $\Phi(X)$  à rester dans un voisinage de  $X$  : on prendra souvent  $0 < \lambda \ll 1$ , de façon à avoir un modèle  $\Phi(X)$  proche de l'observation  $\tilde{X}$ , mais qui reste à distance "finie" de la forme  $X$ . Comme nos normes RKHS – définies ci-dessous – pénalisent de manière exponentielle les discontinuités de  $\Phi$  à une échelle inférieure à une longueur caractéristique  $\sigma$ , résoudre le problème d'appariement sera alors un moyen de trouver, parmi les déformations  $\sigma$ -lisses de  $X$ , la forme la plus proche de  $\tilde{X}$ .



Vadim Lebovici  
et Arthur Touati  
Fondamental

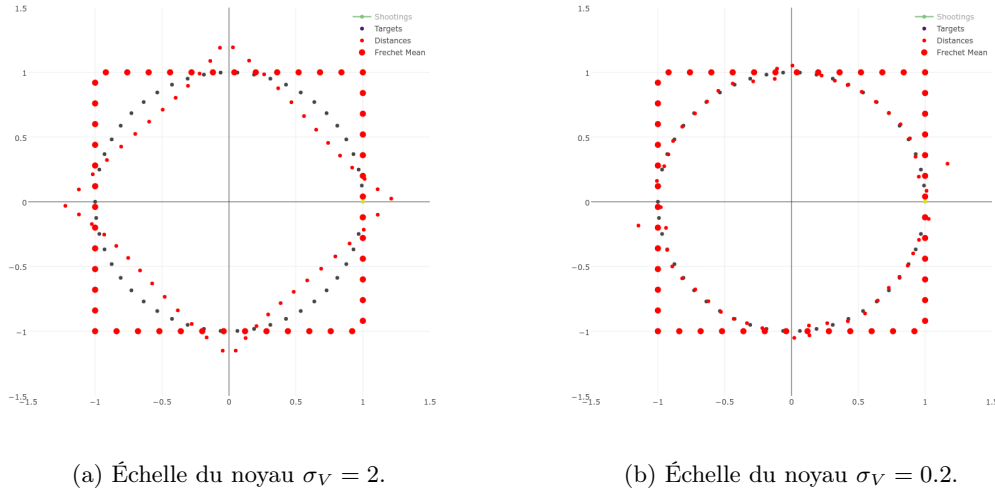


FIGURE 13.2 – Matching d'un carré sur un cercle, à l'aide d'un noyau gaussien ( $k(x, y) = \exp(-\|x - y\|^2 / 2\sigma^2)$ ). Un tel noyau force la préservation des structures plus fines que  $\sigma$ , en corrélant fortement les points tels que  $k(x, y) \simeq 1$ .

### 13.4 Élargissement du modèle aux déformations rigides

Il ne s'agit pas d'oublier que l'espace des champs de vecteurs que nous considérons pour construire nos difféomorphismes sont des champs de vecteurs nuls à l'infini et donc qu'ils ne permettent pas de prendre en compte les *déformations rigides*, c'est-à-dire les translations, les rotations et éventuellement les dilatations.

On construit ici un modèle qui traite indépendamment les déformations élastiques et les déformations rigides. Sinon, on peut plutôt que de quotienter l'espace par toutes les déformations rigides, identifier des axes de références sur les formes que nous cherchons à comparer et les faire coïncider avant la comparaison à l'aide des difféomorphismes. Ceci donne une souplesse supplémentaire.

### 13.5 Méthodes à noyaux

On l'a vu, pour arriver à des logiciels riches et pratiques, il nous faut considérer (et pénaliser) des déformations infinitésimales plus complexes que des simples translations/rotations/dilatations. On présente ici une modélisation possible de notre espace de champ de vecteur, qui se révélera extrêmement féconde du point de vue théorique, et efficace au point de vue algorithmique.

#### 13.5.1 Espaces de Hilbert

Si  $H$  est un espace de Hilbert, on note  $\langle \cdot, \cdot \rangle_H$  son produit scalaire et  $\|\cdot\|_H$  la norme associée. On rappelle que le théorème de représentation de Riesz nous donne une isométrie bijective  $K_H$  entre  $H'$  le dual topologique de  $H$  et  $H$  qui vérifie :

$$\forall \eta \in H', \forall h \in H, \eta(h) = \langle K_H \eta, h \rangle_V \tag{13.17}$$



### 13.5.2 Espace reproduisant et noyaux

**Définition 13.4.** Soit  $S$  un ensemble, et  $H$  un espace de Hilbert de champs de vecteurs  $S \rightarrow \mathbb{R}^m$  où  $m \geq 2$ . Si  $S = \mathbb{R}^m$ , on parlera de champs de vecteurs *denses*, et *discrets* si  $S = \llbracket 1, n \rrbracket$ .

$H$  est dit espace à noyau reproduisant (ou RKHS) lorsque les formes linéaires  $\delta_x^\alpha : f \in H \mapsto f(x) \cdot \alpha$  où  $x \in S$  et  $\alpha \in \mathbb{R}^m$  sont continues.

On peut alors appliquer le théorème de Riesz aux  $\delta_x^\alpha$  et définir le noyau reproduisant d'un RKHS :

**Définition 13.5.** Si  $H$  est un RKHS, le noyau reproduisant de  $H$  est l'application  $k_H : S^2 \mapsto \mathcal{L}(\mathbb{R}^m)$  telle que :

$$\forall x \in S, k_H(x, \cdot)\alpha = K_H \delta_x^\alpha \quad (13.18)$$

c'est-à-dire que  $k_H(x, \cdot)\alpha$  est l'unique élément de  $H$  tel que :

$$\forall h \in H, \delta_x^\alpha(h) = \langle k_H(x, \cdot)\alpha, h \rangle_H = h(x) \cdot \alpha \quad (13.19)$$

**Définition 13.6.** On appelle noyau vectoriel positif de dimension  $m$  sur  $S$  toute application  $k : S^2 \mapsto \mathcal{L}(\mathbb{R}^m)$  telle que

$$\forall (x, y) \in S^2, k(x, y) = \overline{k(y, x)} \quad (13.20)$$

$$\forall (x^1, \dots, x^n) \in S^n, \forall (\alpha^1, \dots, \alpha^n) \in (\mathbb{R}^m)^n, \sum_{1 \leq i, j \leq n} \alpha^j \cdot k(x^i, x^j) \alpha^i \geq 0 \quad (13.21)$$

En terme de matrice, c'est demander que la matrice des noyaux

$$K_x = \begin{pmatrix} k(x_1, x_1) & k(x_1, x_2) & \cdots & k(x_1, x_n) \\ k(x_2, x_1) & k(x_2, x_2) & \cdots & k(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_n, x_1) & k(x_n, x_2) & \cdots & k(x_n, x_n) \end{pmatrix} \quad (13.22)$$

définie par blocs soit symétrique positive.

Le noyau est dit strictement positif si lorsque les  $x^i$  sont distincts, l'expression ci-dessus est nulle si et seulement si les  $\alpha^i$  sont tous nuls. On montre l'équivalence entre la notion de noyau reproduisant et la notion de noyau vectoriel positif :

**Théorème 13.2.** 1. *le noyau reproduisant d'un RKHS est un noyau positif.*

2. *A tout noyau positif  $k$  sur  $S$  de dimension  $m$ , il existe un unique RKHS inclus dans  $(\mathbb{R}^m)^S$  dont le noyau reproduisant est  $k$ .*

Dans le cas d'un RKHS séparable, si on se donne une base orthonormée  $(e_n)_{n \in \mathbb{N}}$ , un simple calcul nous donne que :

$$\forall (x, y) \in S^2, \forall \alpha \in \mathbb{R}^m, k_H(x, y)\alpha = \lim_{n \rightarrow +\infty} \sum_{k=0}^n (\alpha \cdot e_n(x)) e_n(y) \quad (13.23)$$

Vadim Lebovici  
et Arthur Touati  
Fondamental

### 13.5.3 Application au problème d'appariement

L'intérêt principal des noyaux reproduisant est de fournir une méthode simple pour la construction de l'espace des champs des déformations  $S$ . On choisit d'abord le noyau  $k$  puis grâce au théorème 13.2, on définit  $V$  comme l'unique RKHS correspondant à  $k$ . On rappelle l'expression de la fonctionnelle :

$$J(v) = \lambda \int_0^1 \|v_t\|_V^2 dt + A(\Phi_1^v).$$

Le premier terme correspond au coût de la déformation, évalué par la norme du champ  $v$  dans  $L_V^2$ . Le plus souvent les objets que l'on va vouloir déformer se situent dans une région de l'espace  $S \subset \mathbb{R}^d$ . On suppose donc que le terme d'attache aux données n'est plus fonction de toute la déformation  $\Phi_1^v$  mais seulement de  $\Phi_{1|S}^v$ . Si  $V$  est un espace de champs de déformation dans  $\mathbb{R}^n$  et  $k_V$  son noyau, on peut alors énoncer la proposition cruciale suivante :

**Proposition 13.4.** Soit  $S \subset \mathbb{R}^d$ .

1.  $V_S := \{v \in V, v|_S = 0\}$  est un sous-ensemble fermé de  $V$  et

$$V_S^\perp = \overline{\text{Vect}\{k_V(x, \cdot)\alpha, x \in S, \alpha \in \mathbb{R}^d\}}. \quad (13.24)$$

2. Pour tout  $v_0 \in V$  il existe un champ  $v_0^S \in V_S^\perp$  tel que  $v_0|_S = v_0^S|_S$ . Le champ  $v_0^S$  est le champ de norme minimale dans  $V$  interpolant  $v_0$  sur  $S$ .

*Démonstration.* Un champ  $v$  est dans  $V_S$  si et seulement si  $\alpha \cdot v(x) = 0$  pour tout  $x \in S$  et  $\alpha \in \mathbb{R}^d$  donc si et seulement si  $\langle k_V(x, \cdot)\alpha, v \rangle_V = 0$  pour tout  $x \in S$  et  $\alpha \in \mathbb{R}^d$ . Ainsi  $V_S = \text{Vect}\{k_V(x, \cdot)\alpha, x \in S, \alpha \in \mathbb{R}^d\}^\perp$ . La deuxième assertion découle juste de la définition du projeté orthogonal.  $\square$

On montre maintenant comment le noyau  $k_V$  nous permet d'interpoler n'importe quel champ de vecteur en minimisant la norme.

**Théorème 13.3.** Supposons que  $k_V$  soit strictement positif. Soient  $x^1, \dots, x^n$  des points tous distincts de  $\mathbb{R}^d$  et  $\gamma^1, \dots, \gamma^n$  des vecteurs de  $\mathbb{R}^d$ . Il existe une unique solution  $v_* \in V$  au problème d'interpolation suivant :

$$\begin{cases} v(x^i) = \gamma^i \\ \|v\|_V \text{ minimal} \end{cases} \quad (13.25)$$

La solution  $v_*$  s'écrit

$$v_* = \sum_{i=1}^n k_V(x^i, x) \alpha^i \quad (13.26)$$

où les  $\alpha^i \in \mathbb{R}^d$  sont solutions du système linéaire suivant :

$$\sum_{i=1}^n k_V(x^i, x^j) \alpha^i = \gamma^j, 1 \leq j \leq n. \quad (13.27)$$

*Démonstration.* Le système linéaire précédent admet une unique solution par stricte positivité du noyau. Le champ  $v_*$  satisfait par construction la condition d'interpolation. On est dans la situation de la proposition précédente avec  $S = \{x^i, 1 \leq i \leq n\}$  et  $v_0 = v_*$ . Par construction  $v_* \in V_S^\perp$  donc c'est l'unique solution du problème d'interpolation et  $v_*$  est bien de norme minimale.  $\square$

On s'attaque maintenant au principe de réduction. On entend par là l'opération qui consiste à prouver qu'un problème de minimisation d'une fonctionnelle dans un espace revient au même problème de minimisation dans un espace plus petit. Dans la suite le noyau est toujours supposé strictement positif. On va s'intéresser au cas où le terme d'attache aux données n'est fonction que du déplacement des points de  $S$  :

$$J(v) = \lambda \int_0^1 \|v_t\|_V^2 dt + A((\Phi_1^v(x))_{x \in S}). \quad (13.28)$$

La minimisation se fait a priori sur  $L_V^2$ . Notons  $L_{V,S}^2$  le sous-ensemble fermé de  $L_V^2$  suivant :

$$L_{V,S}^2 = \{v \in L_V^2, \forall t \in [0, 1], v_t \in V_{\Phi_t^v(S)}^\perp\}. \quad (13.29)$$

On a le résultat suivant :

**Théorème 13.4.** *Pour tout champ  $v \in L_V^2$ , il existe  $v^S \in L_{V,S}^2$  tel que  $\Phi_{1|S}^v = \Phi_{1|S}^{v^S}$  et  $\|v_t^S\| \leq \|v_t\|_V$  pour tout  $t \in [0, 1]$ .*

*Démonstration.* Pour  $t \in [0, 1]$ , on définit  $v_t^S$  comme le champ minimal interpolant  $v_t$  sur  $\Phi_t^v(S)$ . Par construction on a  $\|v_t^S\| \leq \|v_t\|_V$  pour tout  $t \in [0, 1]$ . De plus, par unicité du flot,  $\Phi_{1|S}^v = \Phi_{1|S}^{v^S}$  pour tout  $t \in [0, 1]$ . D'après la proposition 13.4,  $v_t^S \in V_{\Phi_t^v(S)}^\perp = V_{\Phi_t^{v^S}(S)}^\perp$  donc  $v^S \in L_{V,S}^2$ .  $\square$

On énonce alors le principe de réduction sous la forme du corollaire suivant :

**Corollaire 13.1.** *La solution du problème optimal*

$$J(v) = \lambda \int_0^1 \|v_t\|_V^2 dt + A((\Phi_1^v(x))_{x \in S})$$

*peut être cherchée dans  $L_{V,S}^2$ .*

*Les  $v_t$  optimaux sont donc typiquement nuls sauf au voisinage des points de  $\Phi_t^v(S)$ .*

*Démonstration.* Si  $v \in L_V^2$ , et si  $v^S$  est défini par le théorème précédent, les termes d'attache aux données dans  $J(v)$  et  $J(v^S)$  sont égaux vu que  $\Phi_{1|S}^v = \Phi_{1|S}^{v^S}$ . Les inégalités  $\|v_t^S\| \leq \|v_t\|_V$  impliquent par croissance de l'intégrale que  $J(v^S) \leq J(v)$ .  $\square$

### 13.5.4 Cas des Landmarks

Dans le cadre de l'appariement de nuage de points,  $S = \{x^i, 1 \leq i \leq n\}$ , avec des points  $x^i$  distincts. Nous allons décrire l'ensemble  $L_{V,S}^2$ . Si  $v^S \in L_{V,S}^2$ , alors pour  $t$  dans  $[0, 1]$ ,  $v_t \in V_{\Phi_t^v(S)}^\perp = \overline{\text{Vect}\{k_V(\Phi_t^v(x^i), \cdot)\alpha, 1 \leq i \leq n, \alpha \in \mathbb{R}^d\}}$ , donc il existe des vecteurs  $\alpha_t^i$  dans  $\mathbb{R}^d$  tels que :

$$v_t(x) = \sum_{i=1}^n k_V(\Phi_t^v(x^i), x) \alpha_t^i. \quad (13.30)$$

où les trajectoires  $x_t^i = \Phi_t^v(x^i)$  vérifient :

$$\begin{cases} x_0^i = x^i \\ \partial_t x_t^i = v_t(x_t^i) = \sum_{j=1}^n k_V(x^j, x_t^i) \alpha_t^j. \end{cases} \quad (13.31)$$

Vadim Lebovici et Arthur Touati  
Fondamental

On voit que l'on peut paramétrer l'espace de minimisation  $L^2_{V,S}$  par les vecteurs  $\alpha_t^i$ . Un des intérêts majeurs de la formule définissant  $v_t$  est de pouvoir calculer facilement la norme de  $v_t$  :

$$\|v_t\|_V^2 = \sum_{i,j=1}^n \alpha_t^j \cdot k_V(x^j, x^i) \alpha_t^j \quad (13.32)$$

Nous pouvons énoncer le résultat final suivant :

**Proposition 13.5.** *Supposons que le noyau  $k_V$  soit strictement positif. Alors lorsque  $S = \{x^i, 1 \leq i \leq n\}$ , les points  $x^i$  étant distincts, la fonctionnelle d'appariement peut être écrite comme fonction des paramètres  $\alpha^i$ , éléments de  $L^2([0, 1], (\mathbb{R}^d)^n)$  :*

$$J((\alpha_t^i)) = \lambda \int_0^1 \sum_{i,j=1}^n \alpha_t^j \cdot k_V(x^j, x^i) \alpha_t^j dt + A((x_1^i)) \quad (13.33)$$

où les trajectoires  $x_t^i \in \mathbb{R}^d$  sont calculées par résolution du système différentiel 13.31.

## 13.6 Équations Hamiltoniennes

En considérant des espaces RKHS de champs de vecteurs, nous avons donc ramené le problème de trouver une trajectoire optimale  $v_t \in V$  de dimension infinie à un simple problème de contrôle en dimension finie, sur les coefficients  $\alpha_t$ . A priori, cela semble être un simple tour de passe-passe : après avoir complexifié un modèle – des nuages de points aux difféomorphisme –, le voilà de retour en dimension finie sous une forme légèrement différente. En réalité, il s'agit d'un résultat extrêmement profond sur la nature des géodésique, connu par les physiciens depuis 1833, mais qui n'était pas encore complètement assimilé par la communauté du traitement d'image au moment de l'écriture de l'article, en 2005 : élaborée à l'origine par des probabilistes (oui oui!), cette théorie des déformations difféomorphes a mis plusieurs années à assimiler tous les progrès réalisés depuis cent-cinquante ans par les théoriciens du contrôle optimal, spécialistes des trajectoires géodésiques.

Dans cette section, on présente donc un théorème de structure très fort sur les équations géodésiques, dû à Hamilton, qui n'est que rarement présenté en cours de géométrie Riemannienne "mathématique", pour des raisons culturelles.

**Cotangent plane and cometric** Consider a Riemannian manifold  $(M, g)$  of dimension  $m$ . For every pair  $(q, v)$  in its tangent bundle  $TM$ , we can consider the quadratic length

$$\|v\|_{g_q}^2 = (v, g_q v). \quad (13.34)$$

Thinking in terms of *coordinates*, we are often used to say that  $g_q$  is an  $m$ -by- $m$  positive definite symmetric matrix, thus working with the usual dot product  $(u, v) = u^T v$ . However, as we will now discover the prominent role played by the *cotangent* bundle  $T^*M$ , we have to be more careful and disambiguate the distinction between  $T_q M$  and  $T_q^* M = (T_q M)^*$ . The so-called bilinear "dot product" shall then be reserved to couples  $(u, v) \in T_q^* M \times T_q M$ , with

$$(u, v) = u(v) = (v, u). \quad (13.35)$$

At each location  $q \in M$ , the metric tensor  $g_q$  can now be seen as a linear isomorphism

$$g_q : T_q M \rightarrow T_q^* M, \quad (13.36)$$

and we will pay special attention to its inverse, the *kernel* or co-metric tensor

$$K_q : T_q^*M \rightarrow T_qM. \quad (13.37)$$

The symmetry of  $g_q$  ensures that  $K_q$  is also symmetric :  $K_q^* = K_q$ , i.e.

$$\forall p_1, p_2 \in T_q^*M, \quad (p_1, K_q p_2) = (K_q p_1, p_2). \quad (13.38)$$

**Dual norm** A remarkable fact about the positive definite symmetric tensor  $K_q$  is that it encodes on  $T_q^*M$  the dual norm of  $g_q$  :

$$\forall p \in T_q^*M, \quad \|p\|_{g_q^*}^2 = \left( \sup_{(v, g_q v)=1} (p, v) \right)^2 = (p, K_q p). \quad (13.39)$$

Eventually, if  $v = K_q p$  – that is, if  $p = g_q v$  – one has :

$$\|v\|_{g_q}^2 = (v, g_q v) = (K_q p, p) = \|p\|_{K_q}^2 \quad (13.40)$$

### 13.6.1 The reduced Hamiltonian

Assume we are given a smooth path  $q_t = q(t)$  in  $M$ . Considering the lifted smooth path

$$(q_t, v_t) = (q_t, \dot{q}_t) \quad (13.41)$$

in the tangent bundle  $TM$  makes no difficulty : undergrad physics courses have taught us how insightful it can be to study dynamical systems in the “phase plane”.

Pushing even further, we can define the smooth trajectory

$$(q_t, p_t) = (q_t, g_{q_t} v_t) = (q_t, g_{q_t} \dot{q}_t) \quad (13.42)$$

in the *cotangent* bundle  $T^*M$ . Kinetic energy can now be expressed smoothly as a function – the so-called *Hamiltonian* – of the phase points  $(q, v)$  or  $(q, p)$  :

$$E(q, v) = \frac{1}{2}(v, g_q v), \quad H(q, p) = \frac{1}{2}(p, K_q p). \quad (13.43)$$

At first sight, those two formalisms are completely equivalent, and we barely see the need for a non-intuitive “co-tangent” phase space...

**Théorème 13.5** (Hamiltonian formulation of the geodesic equations).

*Given the above notations,  $(q_t)$  is a solution of the geodesic equation on  $M$  if and only if  $(q_t, p_t)$  is a solution of the Hamiltonian equation :*

$$(\text{Ham}) : \begin{cases} \dot{q}_t &= +\frac{\partial H}{\partial p}(q_t, p_t) &= +K_{q_t} p_t \\ \dot{p}_t &= -\frac{\partial H}{\partial q}(q_t, p_t) &= -\partial_q(p_t, K_q p_t)(q_t) \end{cases} \cdot \quad (13.44)$$

Otherwise said, the geodesics of  $M$  are exactly the projection on the space coordinate “ $q$ ” of the integral trajectories – in phase space  $T^*M$  – of the hamiltonian dynamical system

$$(\dot{q}, \dot{p}) = \text{“Rot}_{-90^\circ}” \nabla H(q, p), \quad (13.45)$$

Vadim Lebovici and we have an equivalence between the two *shootings* Cauchy problems, associated respectively  
et Arthur Touati to the geodesic and hamiltonian equations :  
Fondamental

$$(\text{Shoot}_{q_0, v_0}) : \begin{cases} q(0) = q_0 \\ \dot{q}(0) = v_0 = K_{q_0} p_0 \end{cases}, \quad (\text{Shoot}_{q_0, p_0}^*) : \begin{cases} q(0) = q_0 \\ p(0) = p_0 = g_{q_0} v_0 \end{cases}. \quad (13.46)$$

The above theorem is a stunning result : it reduces the geodesic equation – a priori, an order 2 ODE given through highly non-trivial Christoffel coefficients – to a sole order 1 equation in phase space, whose vector field happens to be the *rotated* gradient of the “kinetic energy”.

The standard proof of this result relies on the calculus of variations, and is very much related to the fact that geodesics locally minimize the length. In this report, for a change, we will give an elementary proof : The only “hard to prove” lemma will be the fact that our  $m$ -manifold  $M$  can be isometrically embedded in the euclidean space  $\mathbb{R}^n$ , at least locally – the global statement is true, and has been proved by Nash, but is way too hard to be detailed here. This fits well within the spirit of the lecture group “Introduction to Riemannian Geometry through the study of shapes spaces”, given to undergrad students at the Ecole Normale Supérieure in the autumn of 2016.

*Main proof : Hamiltonian formulation of the geodesic equations.*

**Lemme 13.2** (Locally,  $M$  can be isometrically embedded in a euclidean space  $\mathbb{R}^n$ ).

Let  $q_0$  be a base point in  $M$ . Then there exists a neighborhood  $\Omega$  of  $q_0$  in  $M$ , a euclidean space  $\mathbb{R}^n$  – with  $n \geq m$  – and a smooth diffeomorphism

$$\varphi : (\Omega, g) \rightarrow (\mathbb{R}^n, g_{\text{eucl}}) \quad (13.47)$$

which is an isometry – i.e. such that  $g = \varphi^* g_{\text{eucl}}$ .

*Démonstration.*

A well-written proof of the global Nash embedding theorem can be found on Terence Tao’s blog : <http://terrytao.wordpress.com/2016/05/11/notes-on-the-nash-embedding-theorem/>.

I shall provide here a brief sketch within a few days. □

**Simplifications** Now, consider a smooth trajectory  $q_t = q(t)$  in  $M$  such that  $q(0) = q_0$ . For all time  $t$ , we can define  $p_t = g_{q_t} \dot{q}_t$ , and the above lemma gives us a neighborhood  $\Omega$  of  $q_0$ , and an embedding  $\varphi : \Omega \rightarrow \mathbb{R}^n$ .

We want to characterize on  $p_t$  the property «  $q_t$  is a solution of the geodesic equation » : as this is a local problem, we can restrict our study to the case where  $M = \Omega$ . Restricting the neighborhood even further, we can even assume that  $\Omega$  is covered by a single coordinate chart : we can safely assume that  $M = \Omega$  is an open set of  $\mathbb{R}^m$ , endowed with a smooth metric  $g$ .

**Notations** Our diffeomorphism  $\varphi$  is an isometry :

$$\forall q \in M, \forall u, v \in T_q M, \quad (u, g_q v) = \langle d\varphi_q u, d\varphi_q v \rangle_{\text{eucl}}. \quad (13.48)$$

Denoting  $d\varphi_q^*$  the euclidean adjoint of  $d\varphi_q$  – in coordinates, it’s nothing but the transpose –, we then have :

$$g_q = d\varphi_q^* d\varphi_q, \quad K_q = (d\varphi_q^* d\varphi_q)^{-1}. \quad (13.49)$$

The well-known linear algebra formula

$$\forall A \in M_{n, m}(\mathbb{R}), \quad \pi_{\text{Vect}(A)}^\perp = A(A^\top A)^{-1} A^\top \quad (13.50)$$

then becomes :

$$\forall q \in M, \quad \pi_{T_{\varphi(q)}\varphi(M)}^\perp = d\varphi_q(d\varphi_q^*d\varphi_q)^{-1}d\varphi_q^* \quad (13.51)$$

$$= d\varphi_q K_q d\varphi_q^*. \quad (13.52)$$

Finally, we remind the reader that since  $\varphi$  is of class  $C^2$ , it has an order two differential at every point  $q$ ,

$$d^2\varphi_q : T_q M \rightarrow \mathcal{L}(T_q M, \mathbb{R}^n), \quad (13.53)$$

which is such that, according to Schwarz's theorem,

$$\forall q \in M, \forall u, v \in T_q M, \quad (d^2\varphi_q \cdot u) \cdot v = (d^2\varphi_q \cdot v) \cdot u. \quad (13.54)$$

### First calculation : acceleration of the moving point $\varphi(q_t)$

The diffeomorphism  $\varphi$  is an isometry, hence we know that

$$\llcorner q_t \text{ is a geodesic in } (M, g) \llcorner \quad (13.55)$$

$$\iff \llcorner \varphi(q_t) \text{ is a geodesic in } (\varphi(M), g_{\text{eucl}}) \llcorner \quad (13.56)$$

$$\iff \forall t, \quad \pi_{T_{\varphi(q)}\varphi(M)}^\perp \widehat{\varphi(q_t)} = 0 \quad (13.57)$$

$$\iff \forall t, \quad d\varphi_{q_t} K_{q_t} d\varphi_{q_t}^* \widehat{\varphi(q_t)} = 0 \quad (13.58)$$

$$\iff \forall t, \quad d\varphi_{q_t}^* \widehat{\varphi(q_t)} = 0 \quad (13.59)$$

since  $d\varphi_q$  is injective and  $K_q$  is invertible for all  $q$ .

It is therefore necessary to compute an expression for the acceleration of the moving point  $\varphi(q_t)$  as a function of the phase coordinates  $(q_t, p_t)$ . We have, using the chain rule :

$$\widehat{\dot{\varphi}(q_t)} = d\varphi_{q_t} \cdot \dot{q}_t \quad (13.60)$$

$$\widehat{\ddot{\varphi}(q_t)} = (d^2\varphi_{q_t} \cdot \dot{q}_t) \cdot \dot{q}_t + d\varphi_{q_t} \cdot \ddot{q}_t \quad (13.61)$$

Replacing the time derivative  $\dot{q}_t$  by its phase-space expression  $K_{q_t} p_t$ , we also know that :

$$\ddot{q}_t = \widehat{K_{q_t} p_t} \quad (13.62)$$

$$= (d_q K_q \cdot \dot{q}_t) p_t + K_{q_t} \dot{p}_t \quad (13.63)$$

$$= (d_q K_q \cdot K_{q_t} p_t) p_t + K_{q_t} \dot{p}_t. \quad (13.64)$$

Computing  $d_q K_q$  is not that hard : we know that

$$\forall \delta q \in T_q M, \quad d_q g_q \cdot \delta q = d_q (d\varphi_q^* d\varphi_q) \cdot \delta q \quad (13.65)$$

$$= (d^2\varphi_q^* \cdot \delta q) d\varphi_q + d\varphi_q^* (d^2\varphi_q \cdot \delta q), \quad (13.66)$$

and, since  $g_q K_q = \text{Id}_{T_q^* M}$  for all  $q \in M \subset \mathbb{R}^m$ ,

$$0 = (d_q g_q) K_q + g_q (d_q K_q), \quad (13.67)$$

$$\text{i.e. } d_q K_q = -K_q (d_q g_q) K_q. \quad (13.68)$$

Combining all of the above, we have

$$d_q K_q \cdot K_{q_t} p_t = -K_{q_t} ((d^2\varphi_{q_t}^* \cdot K_{q_t} p_t) d\varphi_{q_t} + d\varphi_{q_t}^* (d^2\varphi_{q_t} \cdot K_{q_t} p_t)) K_{q_t}. \quad (13.69)$$

Vadim Lebovici  
et Arthur Touati  
Fondamental

Eventually,

$$\begin{aligned} \widehat{\varphi(q_t)} &= (d^2\varphi_{q_t} \cdot K_{q_t}p_t) \cdot K_{q_t}p_t \\ &\quad + d\varphi_{q_t} \cdot [-K_{q_t}((d^2\varphi_{q_t}^* \cdot K_{q_t}p_t) d\varphi_{q_t} + d\varphi_{q_t}^* (d^2\varphi_{q_t} \cdot K_{q_t}p_t)) K_{q_t}p_t \\ &\quad \quad + K_{q_t}\dot{p}_t] \end{aligned} \quad (13.70)$$

$$\begin{aligned} &= (d^2\varphi_{q_t} \cdot K_{q_t}p_t) \cdot K_{q_t}p_t \\ &\quad + d\varphi_{q_t} K_{q_t} \cdot [ -((d^2\varphi_{q_t}^* \cdot K_{q_t}p_t) d\varphi_{q_t} + d\varphi_{q_t}^* (d^2\varphi_{q_t} \cdot K_{q_t}p_t)) K_{q_t}p_t \\ &\quad \quad + \dot{p}_t] \end{aligned} \quad (13.71)$$

**Null tangential acceleration condition** This expression is pretty frightening, as no symmetry is assumed on  $d^2\varphi_{q_t}$ . Thankfully, most terms will soon cancel out as we are only interested in the projection of the above acceleration vector on the *tangent plane*. Indeed, the condition (13.59) now reads

$$0 = d\varphi_{q_t}^* \widehat{\varphi(q_t)} \quad (13.72)$$

$$\text{i.e. } 0 = d\varphi_{q_t}^* (d^2\varphi_{q_t} \cdot K_{q_t}p_t) \cdot K_{q_t}p_t + d\varphi_{q_t}^* d\varphi_{q_t} K_{q_t} \cdot [\dots] \quad (13.73)$$

$$\text{i.e. } 0 = d\varphi_{q_t}^* (d^2\varphi_{q_t} \cdot K_{q_t}p_t) \cdot K_{q_t}p_t + [\dots] \quad (13.74)$$

$$\begin{aligned} \text{i.e. } 0 &= d\varphi_{q_t}^* (d^2\varphi_{q_t} \cdot K_{q_t}p_t) \cdot K_{q_t}p_t \\ &\quad - ((d^2\varphi_{q_t}^* \cdot K_{q_t}p_t) d\varphi_{q_t} + d\varphi_{q_t}^* (d^2\varphi_{q_t} \cdot K_{q_t}p_t)) K_{q_t}p_t \\ &\quad + \dot{p}_t \end{aligned} \quad (13.75)$$

$$\text{i.e. } \dot{p}_t = (d^2\varphi_{q_t}^* \cdot K_{q_t}p_t) d\varphi_{q_t} K_{q_t}p_t \quad (13.76)$$

**Interpretation of  $\dot{p}_t$  as a gradient** The last step of our proof is to see that the above expression is nothing but the opposite of the Hamiltonian's  $q$ -gradient. Indeed, according to the very definition of what is a gradient and the expression computed equation (13.68), we know that for any displacement  $\delta q \in T_{q_t}M$ ,

$$-\frac{1}{2} \left( \partial_q (p_t, K_q p_t) \right) \cdot \delta q = \frac{1}{2} \left( p_t, K_{q_t} [(d^2\varphi_{q_t}^* \cdot \delta q) d\varphi_{q_t} + d\varphi_{q_t}^* (d^2\varphi_{q_t} \cdot \delta q)] K_{q_t} p_t \right) \quad (13.77)$$

Now, developping by linearity :

$$-\frac{1}{2} \left( \partial_q (p_t, K_q p_t) \right) \cdot \delta q = \frac{1}{2} \left( p_t, K_{q_t} (d^2\varphi_{q_t}^* \cdot \delta q) d\varphi_{q_t} K_{q_t} p_t \right) + \frac{1}{2} \left( p_t, K_{q_t} d\varphi_{q_t}^* (d^2\varphi_{q_t} \cdot \delta q) K_{q_t} p_t \right) \quad (13.78)$$

$$= \frac{1}{2} \left( K_{q_t} p_t, (d^2\varphi_{q_t}^* \cdot \delta q) d\varphi_{q_t} K_{q_t} p_t \right) + \frac{1}{2} \left( K_{q_t} p_t, d\varphi_{q_t}^* (d^2\varphi_{q_t} \cdot \delta q) K_{q_t} p_t \right) \quad (13.79)$$

$$= \frac{1}{2} \left( (d^2\varphi_{q_t} \cdot \delta q) K_{q_t} p_t, d\varphi_{q_t} K_{q_t} p_t \right) + \frac{1}{2} \left( d\varphi_{q_t} K_{q_t} p_t, (d^2\varphi_{q_t} \cdot \delta q) K_{q_t} p_t \right) \quad (13.80)$$

$$= \left( (d^2\varphi_{q_t} \cdot \delta q) K_{q_t} p_t, d\varphi_{q_t} K_{q_t} p_t \right) \quad (13.81)$$

since  $K_{q_t}$  is symmetric (i.e.  $K_{q_t}^* = K_{q_t}$ ) and  $d\varphi_{q_t}^{**} = d\varphi_{q_t}$ . The last step is then to remember that, according to Schwarz's theorem,

$$\left( (d^2\varphi_{q_t} \cdot \delta q) K_{q_t} p_t, d\varphi_{q_t} K_{q_t} p_t \right) = \left( (d^2\varphi_{q_t} \cdot K_{q_t} p_t) \delta q, d\varphi_{q_t} K_{q_t} p_t \right) \quad (13.82)$$

$$= \left( \delta q, (d^2\varphi_{q_t}^* \cdot K_{q_t} p_t) d\varphi_{q_t} K_{q_t} p_t \right) \quad (13.83)$$



Otherwise said,

$$-\frac{1}{2} \left( \partial_q \left( p_t, K_q p_t \right) \right) = (d^2 \varphi_{q_t}^* \cdot K_{q_t} p_t) d\varphi_{q_t} K_{q_t} p_t \quad (13.84)$$

in the sense of the space of linear forms  $T_q^* M$ .

**Conclusion** All in all, we have proved that given  $p_t = g_{q_t} \dot{q}_t$  the only trajectory in the space of co-vectors (« momentums ») such that

$$\dot{q}_t = + \frac{\partial H(q, p)}{\partial p} (q_t, p_t) = K_{q_t} p_t, \quad (13.85)$$

then  $\varphi(q_t)$  is a geodesic of the surface  $\varphi(M)$  embedded in the the euclidean space  $\mathbb{R}^n$  if and only if the acceleration of this kinematic point is purely normal, i.e.

$$\forall t, \quad \pi_{T_{\varphi(q)} \varphi(M)}^{\perp} \widehat{\ddot{\varphi}(q_t)} = 0, \quad (13.86)$$

which is verified if and only if

$$\forall t, \quad \dot{p}_t = - \frac{\partial H(q, p)}{\partial q} (q_t, p_t) = (d^2 \varphi_{q_t}^* \cdot K_{q_t} p_t) d\varphi_{q_t} K_{q_t} p_t. \quad (13.87)$$

The theorem has therefore been proved.  $\square$

Now, a word about our proof. It is filled with elementary but cumbersome calculations, no doubt about it. But it also allowed us to interpret the co-metric in a rather unusual way : denoting  $X = \varphi(M)$  the embedded  $m$ -hypersurface and  $\psi = \varphi^{-1} : X \rightarrow M$  our isometry, we have seen that

$$\forall x \in X, \quad K_{\psi(x)} = d\psi_x \pi_{T_x X}^{\perp} d\psi_x^*, \quad (13.88)$$

which means that the action of  $K_q = K_{\psi(x)}$  can be decomposed as follow :

1. Given a linear-form/momentum  $p \in T_q^* M$ , consider its isometric representant  $d\psi_x^* p$  in the dual euclidean space  $(\mathbb{R}^n)^*$ .
2. Taking advantage of the canonical identification  $(\mathbb{R}^n, g_{\text{eucl}}) \simeq ((\mathbb{R}^n)^*, g_{\text{eucl}}^*)$ , orthogonally *project* this representant onto  $T_x X \simeq T_x^* X$  – in all of the above, we have omitted to write explicitly the Riesz mapping  $I_n : (\mathbb{R}^n)^* \rightarrow \mathbb{R}^n$ .
3. In the end, push(“-back”) the projected tangent vector towards  $T_q M$ .

$K_q$  could therefore be seen as a *projection* operator, which forgets about irrelevant/inaccessible/inefficient directions to concentrate on the sole tangent plane.

### 13.6.2 Point de vue hamiltonien

Comme expliqué ci-dessus, il y a donc équivalence entre l'équation des géodésiques et les équations d'Hamilton. Nous montrons ici comment faire le lien entre ce que nous venons de faire et ce formalisme, pour obtenir une nouvelle interprétation de la minimisation de la fonctionnelle  $J$  en terme de géométrie riemannienne et surtout une nouvelle étape de réduction, qui justifie à elle seule tout le travail fait dans la première moitié du cours. Nous modifions les notations pour coller avec celles du formalisme hamiltonien. On pose alors :

$$p_t^i = \alpha_t^i \quad q_t^i = x_t^i \quad (13.89)$$

Vadim Lebovici et Arthur Touati  
Fondamental

On considère aussi les éléments de  $(\mathbb{R}^d)^n$  suivants :

$$p_t = \begin{pmatrix} p_t^1 \\ \vdots \\ p_t^n \end{pmatrix} \quad q_t = \begin{pmatrix} q_t^1 \\ \vdots \\ q_t^n \end{pmatrix} \quad (13.90)$$

Enfin on définit :

$$K_{q_t} = (k_V(x_t^i, x_t^j))_{1 \leq i, j \leq n} \quad (13.91)$$

La matrice  $K_{q_t}$  est inversible pour tout  $t$  dans  $[0, 1]$  par stricte positivité du noyau. On a alors  $v_t = \dot{q}_t = K_{q_t} p_t$ , on retrouve bien le formalisme hamiltonien. De plus,  $\|v_t\|_V^2 = (p_t, K_{q_t} p_t)$ . On montre maintenant comment réduire à nouveau le problème de minimisation.

**Etape 1 : on est sur une géodésique** Si la trajectoire  $q_t$  est telle que  $v_t = \dot{q}_t$  minimise la fonctionnelle  $J(v)$ , alors elle minimise  $\int_0^1 \|\dot{q}_t\|_V dt$  c'est-à-dire la longueur du chemin parcouru, c'est donc une courbe minimisante. Comme on l'a vu en géométrie riemannienne, les courbes minimisantes sont des géodésiques. On a montré que si une trajectoire minimise notre fonctionnelle, c'est une géodésique. Le couple  $(q_t, p_t)$  vérifie donc les équations de Hamilton :

$$\begin{cases} \dot{q}_t = K_{q_t} p_t \\ \dot{p}_t = -\partial_q(p_t, K_{q_t} p_t)(q_t) \end{cases} \quad (13.92)$$

**Etape 2 : réduction** Si on est sur une géodésique, la vitesse est de norme constante c'est-à-dire

$$\langle v_t, v_t \rangle_V = (p_t, K_{q_t} p_t) \quad (13.93)$$

est constant égal à  $(p_0, K_{q_0} p_0)$ . Dans les équations de Hamilton c'est un fait qui est reflété par le fait que l'évolution suit un gradient de l'énergie cinétique... tourné de  $90^\circ$  ! La fonctionnelle prend alors la forme très simple suivante :

$$J(p_0) = \lambda(p_0, K_{q_0} p_0) + A(p_0) \quad (13.94)$$

où  $p_0 \in (\mathbb{R}^d)^n$ . On a considérablement simplifié le problème : grâce à nos connaissances mathématiques sur l'équation des géodésiques, on a transformé un problème de minimisation sur la trajectoire  $L_V^2$  en une minimisation "simple" sur le moment initial  $p_0$ , un vecteur de la même dimension que notre forme.

# Chapitre 14

## Appariement de courbes et surfaces polygonales

*Bonus*

*[CT13]*

*Arnaud Eteve et  
Yohan Mandin*

*-Huble  
Aspects  
pratiques*

### 14.1 Introduction

L'objectif de ce chapitre est d'apporter plusieurs solutions au problème de l'appariement non-labellisé, c'est-à-dire entre deux nuages de points considérés globalement, où les points ne sont plus regroupés par paires point source, point cible. On pourra travailler avec un nombre de points source et un nombre de points cible différents, et s'affranchir de la nécessité d'associer les points par paires (opération souvent arbitraire et non pertinente, comme dans le cas de points répartis sur une courbe). Nous présenterons trois modélisations différentes : par les mesures, par les courants et par les varifolds, avant d'en décrire les avantages et limitations respectives.

### 14.2 Transport et appariement de mesures

Une première manière de visualiser un nuage de points (ou plus généralement une courbe, une surface) est de l'exprimer en termes de masse. Pour ce faire, on va se placer dans l'espace des mesures boréliennes signées sur  $\mathbb{R}^d$ . On verra un point comme la mesure de Dirac en ce point, un nuage de points comme la somme des mesures de Dirac de ces points (avec éventuellement certains coefficients, selon l'importance respective accordée aux différents points). On ne s'intéressera qu'au cas des nuages de points, mais on peut de la même manière associer à une courbe ou une surface une mesure. Autrement dit, nous nous plaçons dans l'espace vectoriel des mesures boréliennes signées sur  $\mathbb{R}^d$ , noté  $M_s(\mathbb{R}^d)$ .

#### 14.2.1 Sur le terme d'attache aux données

Comme on l'a vu à la séance précédente, résoudre le problème d'appariement revient à minimiser une fonction constituée de deux termes : un terme représentant le coût de la déformation, sous forme d'une intégrale, et un autre exprimant la distance entre la forme obtenue par déformation et la forme cible. Dans cette section nous conservons le premier terme sous sa forme déjà énoncée, mais nous modifions le second, appelé terme d'attache aux données. Nous cherchons à définir une norme sur  $M_s(\mathbb{R}^d)$  (rappelons que les formes sont maintenant des mesures). Pour cela, on peut simplement utiliser la caractérisation de  $M_s(\mathbb{R}^d)$  par dualité :

$$M_s(\mathbb{R}^d) = C_0(\mathbb{R}^d)', \quad (14.1)$$

Arnaud Eteve et et, étant donné une norme  $\|\cdot\|$  sur un sous-ensemble dense de  $C_0(\mathbb{R}^d)$ , on peut définir la norme  
Yohan Mandin duale associée sur  $M_s(\mathbb{R}^d)$  par

$$\forall \mu \in M_s(\mathbb{R}^d), \|\mu\|^\star = \sup \left\{ \int f d\mu, \|f\| \leq 1 \right\}. \quad (14.2)$$

-Huble  
Aspects  
pratiques

Le problème est alors de trouver une norme sur les fonctions qui réalise un bon compromis entre :  
— la simplicité mathématique/algorithmique.

— la pertinence pour nos problèmes de matching : la distance  $\|\mu - \nu\|^\star$  devrait se réduire au fur et à mesure que l'objet représenté par  $\mu$  se "rapproche" de l'objet représenté par  $\nu$ .

Nous allons comparer, pour l'exemple, deux normes sur  $M_s(\mathbb{R}^d)$  (qui ne seront pas celles utilisées in fine). Si, pour commencer, nous utilisons la norme infinie :

$$\|\mu\|_\infty^\star = \sup \left\{ \int_{\mathbb{R}^d} f d\mu, f \in C_0(\mathbb{R}^d), \|f\|_\infty \leq 1 \right\} \quad (14.3)$$

Cette mesure ne peut que donner de très mauvais résultats. En effet, si  $\mu$  et  $\nu$  correspondent à deux nuages disjoints de points, chaque point ayant pour masse 1, on peut trouver une fonction continue bornée par 1 valant 1 sur les points de  $\mu$  et  $-1$  sur ceux de  $\nu$ . Autrement dit,

$$\|\mu - \nu\|_\infty^\star = |\mu| + |\nu|, \quad (14.4)$$

ce qui n'encode aucune information sur les positions relatives de  $\mu$  et  $\nu$ . Cela vient du fait que la norme infinie n'interdit aucunement à une fonction de sauter du coq à l'âne entre le support de  $\mu$  et celui de  $\nu$ , même s'ils sont géométriquement proches. On va donc tenter d'imposer une certaine lissité à la fonction, en rajoutant un terme régularisant.

Nous définissons par exemple :

$$\|\mu\|_{\infty,\lambda}^\star = \sup \left\{ \int_{\mathbb{R}^d} f d\mu, f \in C_1(\mathbb{R}^d), \|f\|_\infty + \lambda \|df\|_\infty \leq 1 \right\}, \quad (14.5)$$

où  $\lambda$  est un réel strictement positif à choisir. En pratique, ce paramètre  $\lambda$  représente le rayon caractéristique d'une zone autour de  $\mu$  où une proximité géométrique avec  $\nu$  se traduit par une petite valeur de  $\|\mu - \nu\|_{\infty,\lambda}$ , ce qui est souhaitable : une fonction  $f$  telle que  $\|df\|_\infty \leq 1/\lambda$  - i.e. une fonction  $(1/\lambda)$ -lipschitzienne - ne saurait passer de 1 à  $-1$  entre deux points  $\lambda$ -proches. En pratique, il ne faudrait donc pas choisir  $\lambda$  trop petit, sous peine d'isoler  $\mu$  et de retomber dans le cas de la norme infinie, mais pas trop grand non plus, pour ne pas que n'importe quelle mesure soit proche de  $\mu$  (floutage trop important).

### 14.2.2 Transport de masse

A la séance précédente, on avait remarqué qu'un difféomorphisme agissait sur un ensemble de points (en lui associant l'ensemble des images des points). Ici, il faut définir une action similaire et intuitive d'un difféomorphisme sur une mesure, puisque l'on a adopté ce formalisme. Un difféomorphisme  $\varphi$  de  $\mathbb{R}^d$  agit naturellement sur  $M_s(\mathbb{R}^d)$  : à une mesure  $\mu$  on peut associer la mesure  $\varphi.\mu$  définie par :

$$\int_{\mathbb{R}^+} f d(\varphi.\mu) = \int_{\mathbb{R}^+} f \circ \varphi d\mu \quad (14.6)$$

Cette action est linéaire, et si on considère un point  $x$ , on a pour toute  $f$  :

$$\int_{\mathbb{R}^+} f d(\varphi.\delta_x) = \int_{\mathbb{R}^+} f \circ \varphi d\delta_x = f \circ \varphi(x) \quad (14.7)$$

Donc :

$$\varphi.\delta_x = \delta_{\varphi(x)} \quad (14.8)$$

D'où le résultat suivant pour un nuage de points :

$$\varphi.\left(\sum a^i \delta_{x^i}\right) = \sum a^i \delta_{\varphi(x^i)} \quad (14.9)$$

Ainsi, l'action de  $\varphi$  transporte la masse conformément à l'intuition : la masse totale de la mesure est conservée.

### 14.2.3 Normes RKHS sur les mesures

Le problème des normes introduites précédemment sur l'espace des mesures est qu'elles ne sont pas facilement et rapidement calculables. Pour accéder aux applications pratiques, nous allons donc, une fois de plus, avoir recours à la notion d'espace à noyaux reproduisants.

**Proposition 14.1.** *Soit  $I$  un espace reproduisant scalaire sur  $\mathbb{R}^d$ , i.e. un espace de Hilbert qui s'injecte continûment dans  $C_0(\mathbb{R}^d)$ . Alors  $M_s(\mathbb{R}^d)$  (en tant qu'espace vectoriel normé pour  $\|\cdot\|_\infty^*$ ) s'injecte continûment dans le dual topologique de  $I$ , noté  $I'$ .*

*Démonstration.* Soit  $\varphi$  une injection continue de  $I$  dans  $C_0(\mathbb{R}^d)$ , il existe une constante  $C$  telle que pour tout  $f \in I$  :

$$\|\varphi(f)\|_\infty \leq C \|f\|_I. \quad (14.10)$$

Posons alors :

$$\begin{aligned} \theta : M_s(\mathbb{R}^d) &\rightarrow I' \\ \mu &\mapsto \left( f \mapsto \int_{\mathbb{R}^d} \varphi(f) d\mu \right) \end{aligned} \quad (14.11)$$

On a :

$$|\theta(\mu)(f)| \leq \|\varphi(f)\|_\infty \|\mu\|_\infty^* \leq C \|f\|_I \|\mu\|_\infty^* \quad (14.12)$$

Donc :

$$\|\theta(\mu)\|_{I'} \leq C \|\mu\|_\infty^* \quad (14.13)$$

□

Voici maintenant la formule permettant de calculer la norme d'une mesure associée à un nuage de points (i.e. une somme de Dirac) (on identifie désormais  $M_s(\mathbb{R}^d)$  et son plongement dans  $I'$ ).

**Proposition 14.2.** *Soit  $\mu = \sum_i a^i \delta_{x^i}$  une mesure,  $a^i \in \mathbb{R}$ ,  $x^i \in \mathbb{R}^d$ . Alors :*

$$\|\mu\|_{I'}^2 = \sum_{i,j} a^i a^j k_I(x^i, x^j) \quad (14.14)$$

*Démonstration.* Par bilinéarité du produit scalaire, on a :

$$\|\mu\|_{I'}^2 = \sum_{i,j} a^i a^j \langle \delta_{x^i}, \delta_{x^j} \rangle \quad (14.15)$$

Puis, si  $x$  et  $y$  sont deux points, comme  $K_I$  est une isométrie :

$$\langle \delta_x, \delta_y \rangle_{I'} = \langle K_I(\delta_x), K_I(\delta_y) \rangle_I \quad (14.16)$$

Enfin :

$$\langle K_I(\delta_x), K_I(\delta_y) \rangle_I = \delta_y(K_I(\delta_x)) = K_I(\delta_x)(y) = k_I(x, y) \quad (14.17)$$

□

Arnaud Eteve et  
Yohan Mandin  
-Huble  
Aspects  
pratiques

Attardons-nous quelque peu sur cette formule, en rappelant qu'on prendra typiquement  $k_I(x, y) = \exp(-(x - y)^2 / 2\sigma_I^2)$ . Si on a par exemple

$$\mu = a\delta_x \qquad \nu = b\delta_y, \qquad (14.18)$$

on aura alors :

$$\begin{aligned} \|\mu - \nu\|^2 &= \langle \mu, \mu \rangle - 2\langle \mu, \nu \rangle + \langle \nu, \nu \rangle \\ &= a^2 - 2ab \cdot k(x, y) + b^2, \end{aligned} \qquad (14.19)$$

ce qui, dans le cas d'une mesure de comptage – où  $a = b = 1$  – se réduit à

$$\|\mu - \nu\|^2 = 2 \cdot (1 - k(x, y)) \qquad (14.20)$$

**Le point clé** de tout ce chapitre est alors de comprendre que la fonction “noyau” de corrélation entre deux formes élémentaires (ici, des points isolés, représentés comme des diracs) permet en fait de coder une distance, un terme d'attache aux données bien défini. Ainsi, si on suppose que  $y$  est la cible et  $x$  le point à transporter, minimiser  $\|\mu - \nu\|$  reviendra simplement à tenter de maximiser la corrélation  $k(x, y)$ , en suivant le gradient  $\partial_x k(x, y)$ .

Il est important de noter que l'attractivité de  $y$  est en pratique limitée à la zone de l'espace où  $\partial_x k(x, y)$  est “non nul”, i.e. une boule centrée sur  $y$  de rayon  $\simeq 3 \cdot \sigma_I$ . Au delà de ce rayon, les points  $x$  et  $y$  n'interagissent plus (au sens où bouger  $x$  et  $y$  ne change rien à  $\|\mu - \nu\|$ ), et un algorithme de matching utilisant cette attache aux données échouera.

#### 14.2.4 Appariement de mesures

On conserve les mêmes notations qu'à la séance précédente en ce qui concerne les champs de vecteurs et les difféomorphismes associés par intégration.

**Définition 14.1.** Le difféomorphisme optimal sur  $\mathbb{R}^d$  pour l'appariement de  $\mu$  et  $\nu$  est la solution du problème d'optimisation sur  $V$  consistant à minimiser :

$$J(v) = \lambda \|v\|_{L_V^2}^2 + \|\varphi_1^v \mu - \nu\|_{I'}^2 \qquad (14.21)$$

Ce problème admet une solution en toute généralité sous l'hypothèse que  $\mu$  est à support compact, qui est vérifiée car on travaille avec des sommes finies de Dirac. Ceci tient à la présence du  $\lambda > 0$ . En effet, celui-ci assure que les minima de  $J$  peuvent être cherchés dans une boule pour la norme  $L_V^2$  centrée en 0. Celle-ci n'est pas compacte pour la norme sur  $L_V^2$  (c'est un espace de dimension infinie), mais un travail non trivial permet d'obtenir des résultats qui assurent la présence d'un minimum – c'est une conséquence du théorème de compacité faible de Banach-Alaoglu, un des points culminants du cours d'analyse fonctionnelle. Cette solution peut être obtenue par un algorithme de descente de gradient qui sera explicité plus loin dans le cadre des varifolds. Nous allons enfin énoncer un théorème justifiant notre démarche (qui consiste à échantillonner la courbe source et la courbe cible, trouver une déformation optimale d'un nuage de points vers l'autre, et espérer que cette déformation fonctionne encore entre les deux courbes).

**Théorème 14.1.** *Supposons que le noyau  $k_I$  vérifie  $k_I(x, y) > 0$  pour tous  $x, y \in \mathbb{R}^d$  et  $\partial_1 \log(k_I)$  borné sur tout compact. Soient  $\mu$  et  $\nu$  deux mesures sur  $\mathbb{R}^d$  à support borné, et supposons que la solution  $v^*$  associée au problème d'appariement de  $\mu$  vers  $\nu$  est unique. Soient  $(\mu_n)$  et  $(\nu_n)$  deux suites dans  $M_s(\mathbb{R}^d)$  à support uniformément borné convergeant faiblement vers  $\mu$  et  $\nu$  respectivement. Alors les champs de déformation optimaux  $v_n^*$  associés au problème d'appariement de  $\mu_n$  vers  $\nu_n$  convergent presque sûrement dans  $L^2([0, 1], V)$  vers  $v^*$ .*

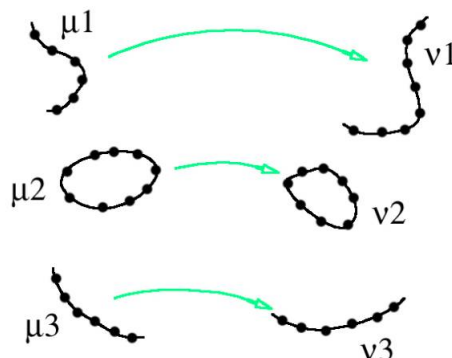


FIGURE 14.1 – Principe de l'étiquetage partiel :  $\mu_i$  sera matché sur  $\nu_i$ , sans information supplémentaires point-à-point.

A titre d'exemple de résultats pouvant être obtenus à l'aide de ce formalisme, mentionnons un lemme utile à la démonstration de ce théorème. Il s'agit de contrôler de manière uniforme sur un certain ensemble de mesures leurs déformations par des champs de vecteurs bornés, sous quelques hypothèses concernant le noyau reproduisant de  $I$ .

**Lemme 14.1.** *Si le noyau scalaire  $k_I$  est tel que  $k_I(x, y) > 0$  pour tous  $x, y \in \mathbb{R}^d$  et  $\partial_1 \log(k_I)$  est borné sur tout compact, alors, pour tout compact  $K$  et  $M > 0$ , il existe  $C > 0$  tel que pour tout  $\mu$  dans  $M_s(\mathbb{R}^d)$  de support inclus dans  $K$  et tout  $v \in L_V^1$  vérifiant  $\|v\|_{L_V^1} \leq M$ , alors  $\|\varphi_1^v \mu\|_{I'} \leq C \|\mu\|_{I'}$ .*

*Autrement dit, une déformation de coût fini ne peut pas faire exploser la norme, quelque soit la mesure prise à support dans  $K$ .*

*Remarque 14.1.* En pratique,  $k_I$  est un noyau gaussien, qui vérifie donc les hypothèses du lemme (et du théorème), et les mesures sont à support fermé borné, donc compact – ce sont des cœurs, des cerveaux, des poissons, pas des hyperboloïdes!

### 14.2.5 Étiquetage partiel

En utilisant le formalisme des mesures, on peut non seulement gérer des données où les éléments identifiés dans l'objet sont des points isolés, mais également celles où ce sont des groupes de points (ce qui peut avoir un intérêt pratique). Chaque groupe de point est alors associé à une somme de masses ponctuelles, et on peut appliquer les résultats et l'algorithme déjà obtenus.

Arnaud Eteve et  
Yohan Mandin

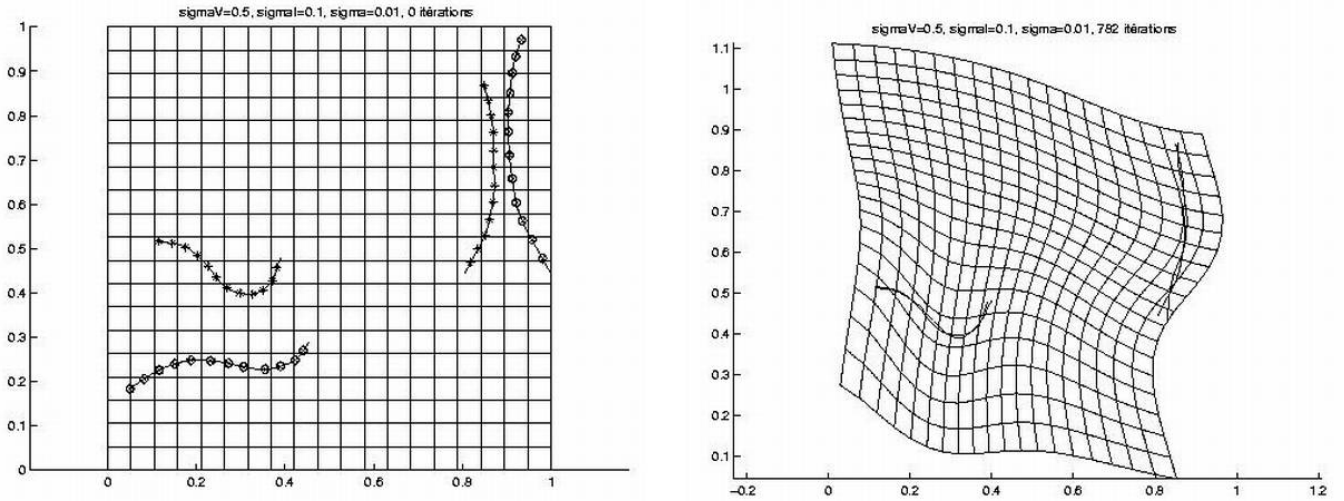


FIGURE 14.2 – Expérience d'appariement de deux paires de courbes – des cercles sur les croix. À gauche, les deux jeux de données ; à droite, le matching obtenu.

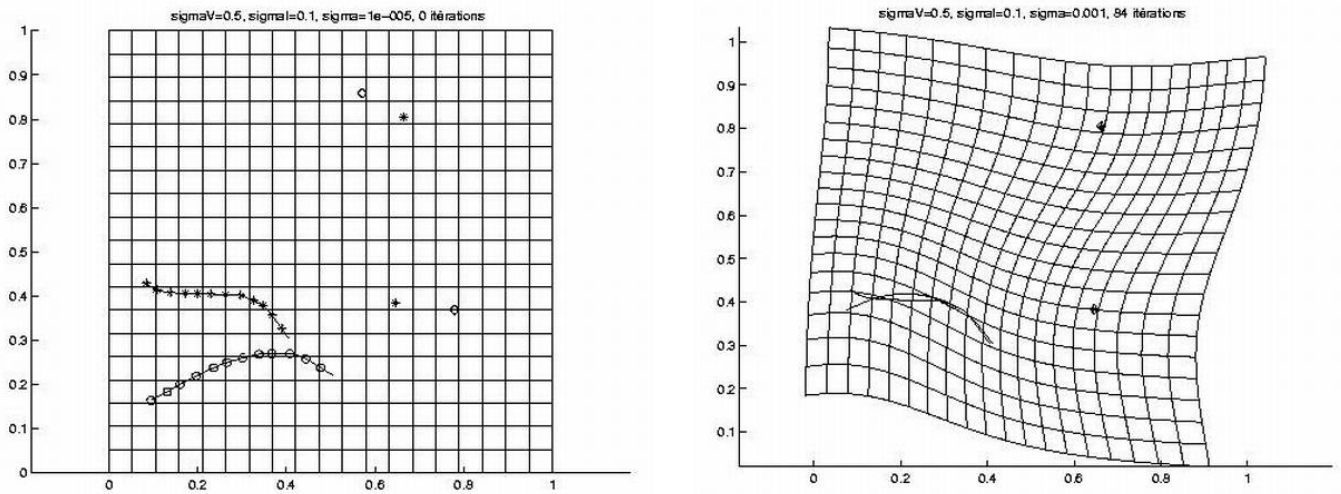


FIGURE 14.3 – Matching d'une courbe et de deux points – des cercles sur les croix.



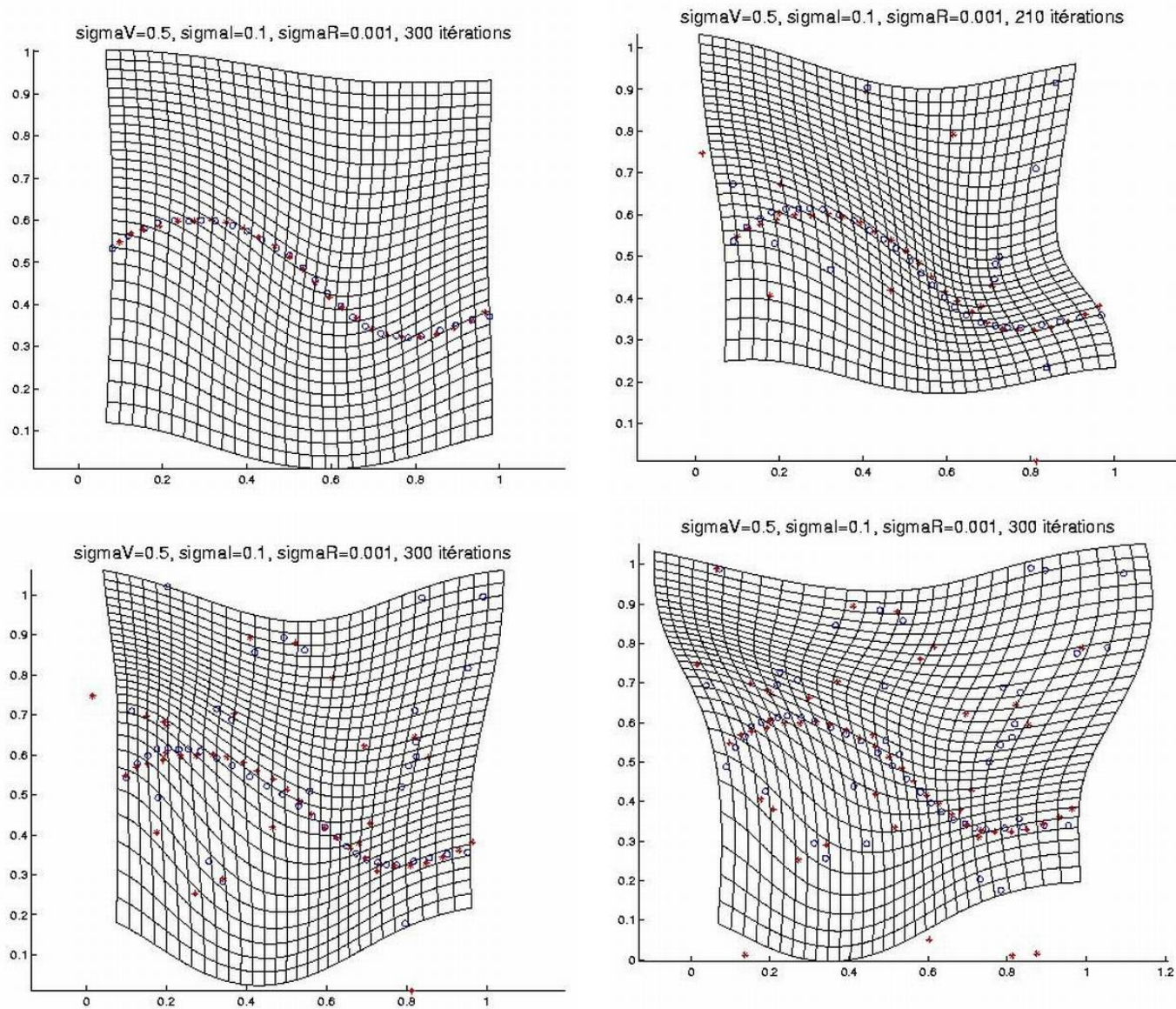


FIGURE 14.4 – Robustesse vis-à-vis des données non significatives (0, 20, 40 et 60 points ajoutés aléatoirement). La régularité imposée par la norme infinitésimale sur les déformations (RKHS gaussien de déviation  $\sigma_V = 0.5$ ) permet d'encaisser un bruitage conséquent : si la déformation est très affectée sur les coins du domaine, le matching de la courbe ne se laisse pas distraire, et reste satisfaisant sur les quatres images.

Arnaud Eteve et  
Yohan Mandin  
-Huble  
Aspects

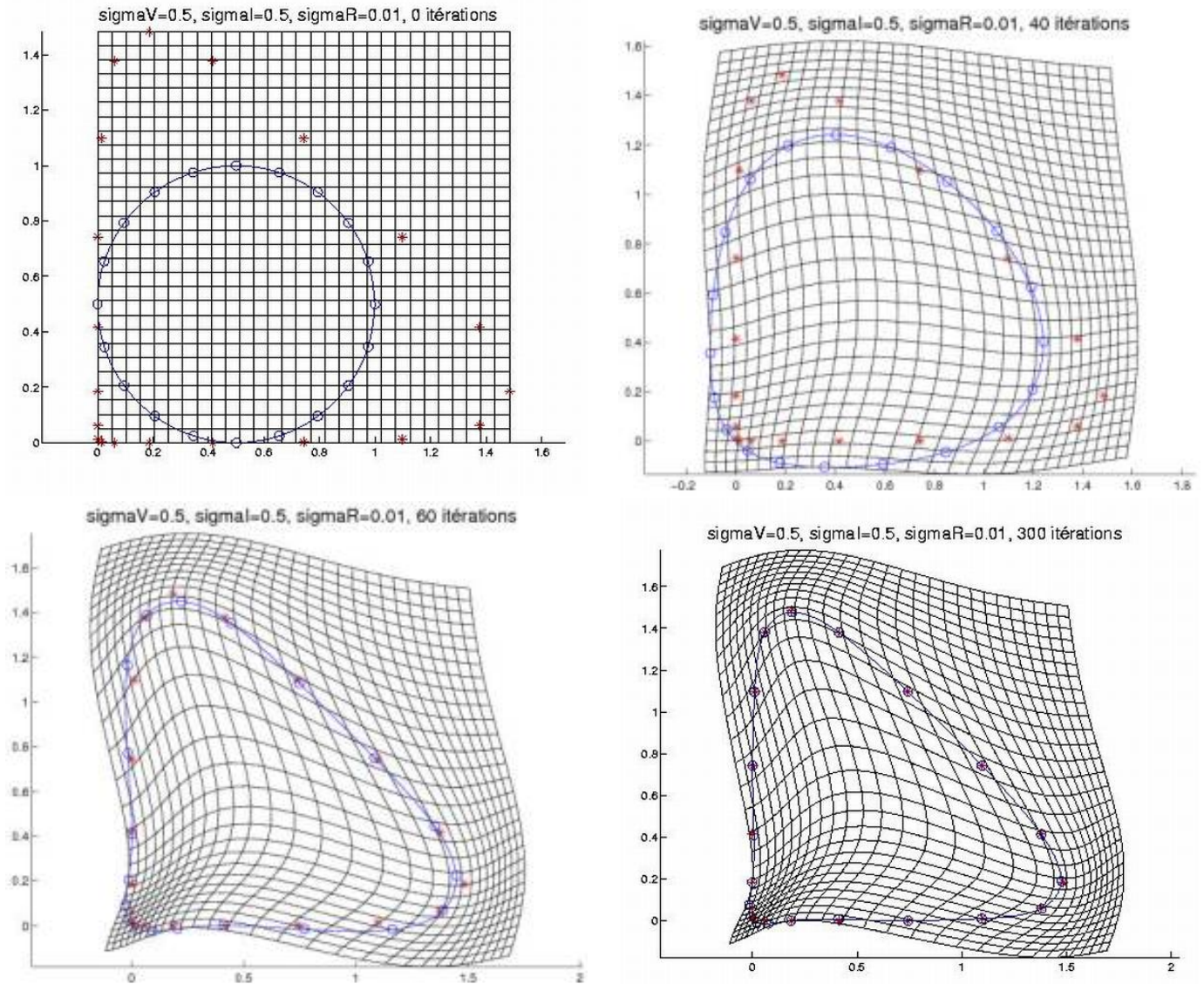


FIGURE 14.5 – Evolution de la descente de gradient : on tente de matcher un cercle sur un “triangle”. Les points vont peu à peu interagir via l’attache aux données, conduisant ainsi le processus de tir géodésique à déformer le cercle en un triangle approximatif. On a donc “habillé” le triangle rouge avec le cercle bleu.



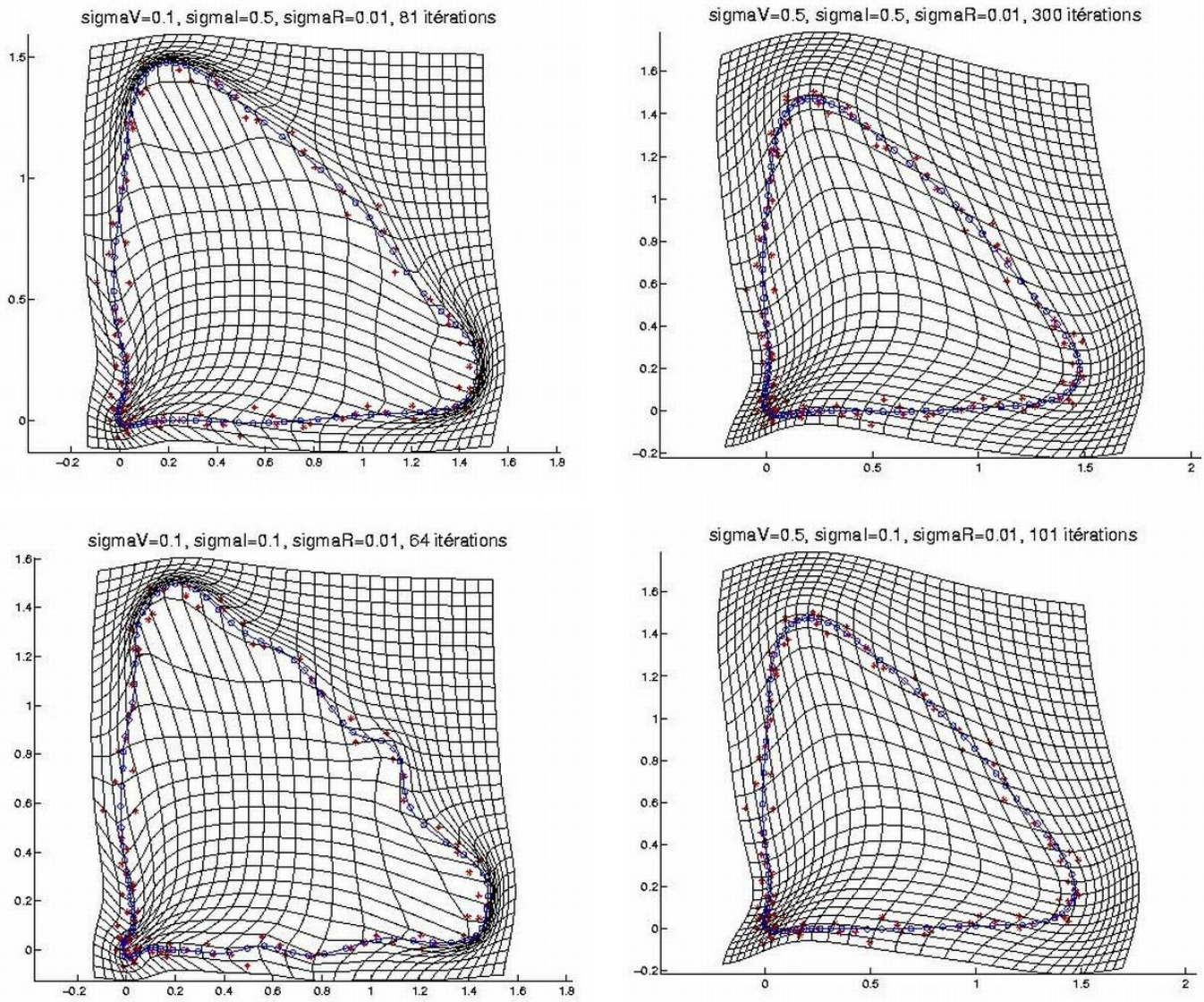


FIGURE 14.6 – Comparaison de plusieurs tailles de noyaux avec données bruitées.  $\sigma_V$  correspond à la taille du noyau de déformation,  $\sigma_I$  à celui de l'attache aux données.

Arnaud Eteve et  
Yohan Mandin  
-Huble  
Aspects

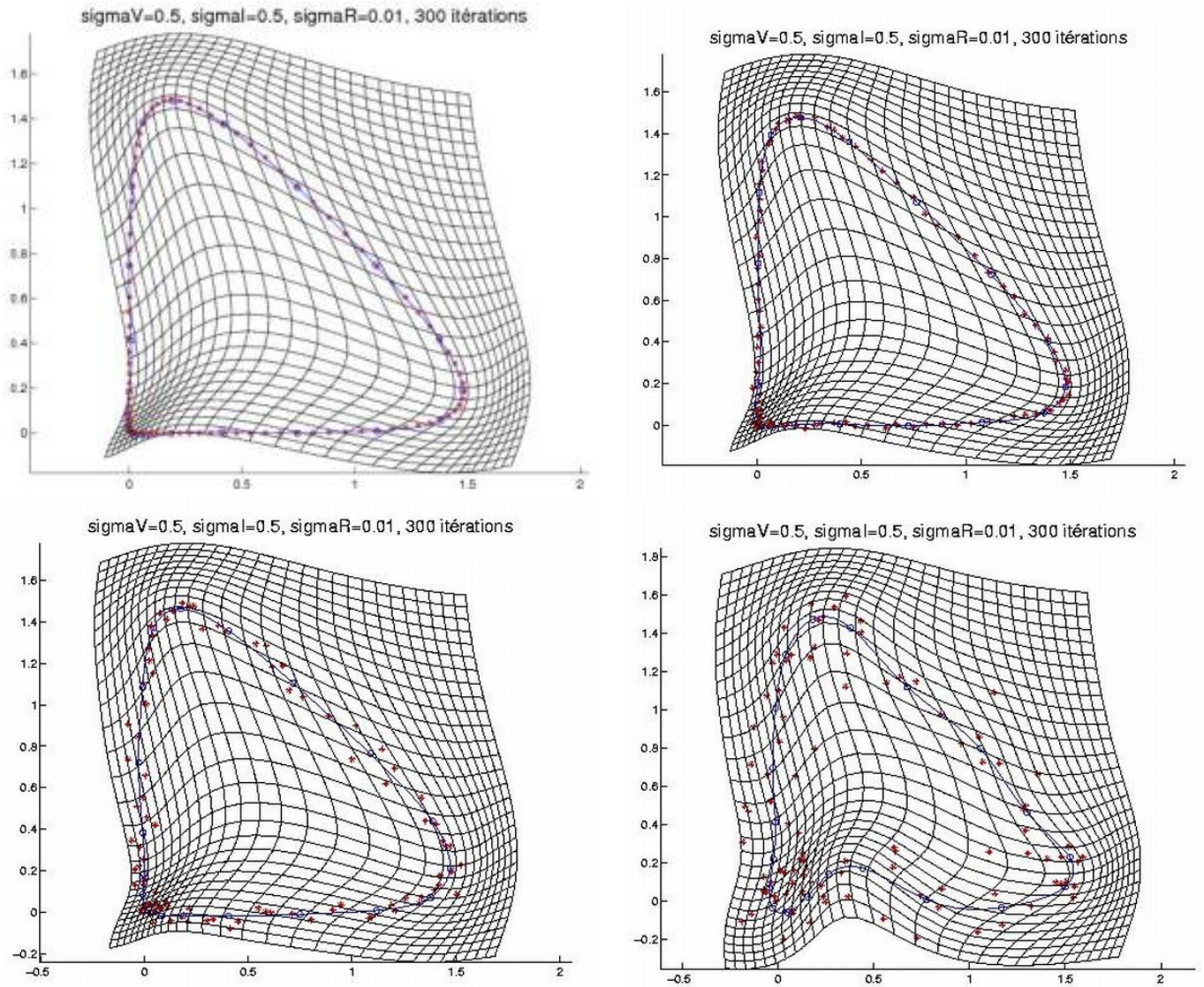


FIGURE 14.7 – Expérience avec données bruitées, avec  $\sigma_V = 0.5$ ,  $\sigma_I = 0.5$ , pour un bruit blanc gaussien de variances du bruit 0, 0.01, 0.03, et 0.1. On le voit, choisir des valeurs importantes pour ces deux paramètres permet de considérablement lisser le matching : en diminuant la précision de nos attaches, on diminue aussi la sensibilité au bruit. En pratique, le choix de paramètres appropriés aux données aura une influence considérable sur la pertinence des résultats obtenus.

## 14.3 Courants

Le transport de masse nous a permis du surmonter le problème de l'étiquetage des landmarks. Malheureusement, pour réussir des appariements de courbes/surfaces, il reste un outil trop élémentaire : en représentant des segments/triangles par des diracs de masse, on perd toute information sur l'orientation. Deux bouts de courbes seront "proches" au sens des mesures s'ils sont géométriquement voisins – à une échelle définie par le rayon  $\sigma_I$  – et qu'ils ont la même longueur, sans plus de précisions.

Pour aller plus loin, et obtenir un recalage plus fin où l'alignement des espaces tangents est mis en avant, plusieurs formalismes ont été proposés. Dans cette section, nous présenterons le plus ancien d'entre eux, celui des courants : on manipulera les espaces tangents aux travers de leurs normales sortantes, traitées de manière *linéaire*.

### 14.3.1 Définitions

**Définition 14.2.** Une forme différentielle de degré  $p$  sur  $\mathbb{R}^d$  est une application continue de  $\mathbb{R}^d$  dans l'espace des formes  $p$ -linéaires alternées sur  $\mathbb{R}^d$ . En d'autres termes, il s'agit d'un champ continu sur  $\mathbb{R}^d$  de forme  $p$ -linéaires alternées – i.e. quelque chose que l'on intègre sur une  $p$ -surface. (Par exemple, si  $f$  est une fonction numérique différentiable, alors  $df$  est une forme différentielle de degré 1.)

**Définition 14.3.** L'espace des  $p$ -courants sur  $\mathbb{R}^d$  est l'espace des formes linéaires continues sur l'espace des formes différentielles continues de dimension  $p$  sur  $E$  qui s'annulent à l'infini.

A toute sous-variété  $X$  de  $E$  de dimension  $d$  de volume fini on peut associer un  $d$ -courant, défini par :

$$C_X : w \mapsto \int_X w$$

Cette définition est l'analogie de celle de la circulation sur une courbe, ou du flux à travers une surface. L'exemple le plus simple de courants est celui des courants de Dirac : si  $x$  est un point et  $\xi$  un  $p$ -uplet, on définit le courant  $\delta_x^\xi : \omega \mapsto \omega(x)(\xi)$ .

La question suivante qui se pose est : comment encoder un nuage de points en termes de courant ? Plaçons-nous par exemple dans le cas d'un nuage de points échantillonnant une surface dans  $\mathbb{R}^3$  : C'est la donnée d'une famille de landmarks, et d'une triangulation présetnée comme une liste de triplets d'indices. A chaque triangle  $(x_1, x_2, x_3)$ , on associe le 2-courant de Dirac pour  $x = \frac{x_1+x_2+x_3}{3}$  et  $\xi = (x_2 - x_1) \wedge (x_3 - x_1)$  donnée comme une normale sortante, un élément de surface  $\vec{n} \cdot dS$ . Par cette méthode, on encode à la fois la localisation, l'aire du triangle et l'**orientation de la surface**. C'est la principale différence pratique entre le formalisme des courants et celui des mesures. Et si maintenant le nuage échantillonne une courbe, alors à chaque couple  $(x_1, x_2)$  de points successifs (la courbe est orientée), on associe le 1-courant de Dirac pour  $x = \frac{x_1+x_2}{2}$  et  $\xi = x_2 - x_1$ .

Tout comme on avait associé à chaque nuage de points (et plus généralement à chaque forme) une mesure, on leur associe maintenant un courant. On cherche alors une norme sur l'espace des courants pour définir le terme d'attache aux données qui soit calculable et qui traduise notre intuition géométrique : de la même manière que dans la section précédente, on plonge continûment l'espace des courants dans le dual d'un RKHS (sous l'hypothèse que le RKHS se plonge continûment dans l'espace des formes différentielles de degré  $p$ ). On peut ainsi utiliser la norme de Hilbert sur le dual du RKHS, qui présente l'avantage d'être explicitement calculable



Arnaud Eteve et dans le cas de courants associés à des nuages de points (sommes finies de courants de Dirac) :

Yohan Mandin  
-Huble  
Aspects  
pratiques

$$\langle \delta_{x_1}^{\xi_1}, \delta_{x_2}^{\xi_2} \rangle_{W'} = k(x_1, x_2) \langle \xi_1, \xi_2 \rangle \quad (14.22)$$

$$= k(x_1, x_2) \mu_1 \mu_2 \cos(\theta) \quad (14.23)$$

où les  $\xi_i$  doivent être compris comme des normales sortantes dans le cas des surfaces, et comme des vitesses dans le cas des courbes, où  $\mu_i = \|\xi_i\|$  la masse associée à l'élément de courbe/surface, et où  $\theta$  est l'angle entre les deux vecteurs  $\xi_i$ . Le formalisme des courants permet donc d'encoder une courbe/surface comme une superposition de courants élémentaires, données d'une position et d'un vecteur "aire · direction orientée". La corrélation donnée par la formule ci-dessus permettra alors de comparer ces nuages de courants, et calculer son gradient nous permettra d'obtenir des "désirs" d'alignements pertinents.

### 14.3.2 Le problème de l'orientation

Les courants ne peuvent modéliser que des objets déjà orientés. Ceci pose plusieurs problèmes en pratique : si un objet a  $N$  composantes connexes (typiquement, un cerveau), il y a  $2^N$  orientations possibles, et il arrive que rien a priori ne donne de raisons d'en privilégier une. Notons par ailleurs d'une orientation aléatoire ne donnera pas de résultats satisfaisants de matching, et qu'il existe des courbes ou des surfaces non orientables.

Un autre problème se pose : celui de l'évanouissement de la masse du à des compensations entre courants avec des orientations opposées dans un petit domaine de l'espace. En effet, si  $X$  est un objet et si  $X^*$  représente le même ensemble avec une orientation opposée, on a  $C_X = -C_{X^*}$ . De plus, dans le cas d'un noyau gaussien d'échelle  $\sigma$  (cas le plus courant) :

$$\|\delta_{x_1}^{\xi} - \delta_{x_2}^{-\xi}\|_{W'} = 2|\xi|^2 (1 - e^{-\frac{|x-y|^2}{\sigma^2}}). \quad (14.24)$$

Cette distance s'annule lorsque la distance entre  $x$  et  $y$  est petite devant  $\sigma$ . Il en découle certaines pathologies : des courbes orientées de longueur plus grande que 2 dont la norme du courant associée est arbitrairement petite, des difficultés à matcher des structures épineuses comme les oreilles de lapin... Ceci motive l'introduction d'un nouvel objet mathématique, modélisant des courbes et surfaces non orientées : les varifolds.

## 14.4 Varifolds

Dans toute la suite  $E$  est l'espace ambiant dans lequel sont plongées nos variétés – en pratique, on a donc  $E = \mathbb{R}^2$  ou  $\mathbb{R}^3$ .

### 14.4.1 Grassmannienne

Pour contourner le problème posé par l'orientation, on introduit un nouvel objet : les varifolds, ceux ci ne prennent en compte que la direction. Précédemment on pouvait représenter une direction par un vecteur – typiquement, la vitesse ou la normale sortante – et l'espace des orientations était un espace vectoriel : ce n'est plus possible avec les directions simplement.

**Définition 14.4** (Grassmannienne). La variété grassmannienne de dimension  $d$  dans  $E$  est l'ensemble des sous-espaces de dimension  $d$  de  $E$ . On la note  $G_d(E)$ .

On peut le voir comme l'ensemble des familles de  $d$  vecteurs libres quotienté par la relation d'équivalence :

$$(x_1, \dots, x_d) \sim (y_1, \dots, y_d) \iff \text{Vect}(x_1, \dots, x_d) = \text{Vect}(y_1, \dots, y_d) \quad (14.25)$$

**Représentation** En pratique, la grassmannienne peut être décrite simplement :

- Si  $d = 1$ , elle correspond au cercle/sphère unité quotienté par  $x \sim -x$ .
- Si  $d = 2$  en dimension 3 (codimension 1), on peut se ramener au cas précédent en représentant un plan par sa droite orthogonale.
- Dans les autres cas (jamais reconstruits pour le matching de formes), il est possible de plonger  $G_d(E)$  de manière continue dans  $\text{End}(E)$  en représentant un espace  $V$  par la projection orthogonale sur  $V$ .

### 14.4.2 Mesure de Hausdorff

Soit  $(X, \rho)$  un espace métrique,  $\delta > 0$ , on pose pour une partie  $S$  quelconque de  $X$

$$H_\delta^d(S) = \inf \left\{ \sum_{i=1}^{\infty} (\text{Diam } U_i)^d, S \subset \bigcap_{i=1}^{\infty} U_i, \text{Diam } U_i < \delta \right\} \quad (14.26)$$

Et on définit alors la mesure de Hausdorff par :

$$H^d(S) = \lim_{\delta \rightarrow 0} H_\delta^d(S) \quad (14.27)$$

On vérifie alors que  $H^d$  est une mesure extérieure, que sa tribu contient les boréliens. De plus elle est bornée sur les parties compactes et est invariante par translations. Ainsi si  $E$  est de dimension  $d$ ,  $H^d = K\lambda$  où  $\lambda$  est la mesure de Lebesgue et  $K$  une constante. La mesure de Hausdorff n'est autre que la généralisation de la  $d$ -aire aux objets non-lisses (voire fractals), comme les collections dénombrables de segments/triangles.

### 14.4.3 Varifolds

**Définition 14.5.** Un varifold de dimension  $d$  sur  $E$  est une mesure borélienne finie sur  $E \times G_d(E)$ . En particulier c'est un élément de  $C_0(E \times G_d(E))'$ , le dual topologique de  $C_0(E \times G_d(E))$ .

Cette notion tirée de la théorie de la mesure géométrique semble a priori hors de propos, trop sophistiquée pour être utilisable. On va pourtant voir qu'elle permet de représenter facilement une surface à l'ordre 1, et qu'elle légitimera un hack bien pratique de la formule donnée équation (14.23). Étant donné  $X$  une variété de dimension  $d$ ,  $\mu_X$  est le varifold associé à  $X$ , et on le définit par :

$$\ll \mu_X(A) = H^d(x \in X, (x, T_x X) \in A) \text{ pour tous les boréliens } A \subset E \times G_d(E) \gg. \quad (14.28)$$

De manière équivalente, on voit cette mesure agir sur  $C_0(E \times G_d(E))$  par :

$$\forall \omega \in C_0(E \times G_d(E)), \mu_X(\omega) = \int_{E \times G_d(E)} \omega(x, V) d\mu_X(x, V) \quad (14.29)$$

Dans une optique plus computationnelle, on disposera de diracs :  $\delta_{(x, V)}(\omega) = \omega(x, V)$ , et on approchera une mesure quelconque par une somme de diracs  $\sum c_i \delta_{(p_i, V_i)}$  :

Pour  $d = 1$ , on considère une courbe, les points sont  $x_1, \dots, x_N$  et les segments  $(f_i^1, f_i^2)$  :

$$p_i = \frac{x_{f_i^1} + x_{f_i^2}}{2} \quad (14.30)$$

$$V_i = (x_{f_i^1}, x_{f_i^2}) \quad (14.31)$$

$$c_i = \|\overrightarrow{x_{f_i^1} x_{f_i^2}}\| \quad (14.32)$$

Arnaud Eteve et  
Yohan Mandin  
-Huble  
Aspects  
pratiques

Pour  $d = 2$ , on considère une surface, de sommets  $x_1, \dots, x_N$ , et de faces triangulaires données sous forme de triplets d'indices  $(f_i^1, f_i^2, f_i^3)$  :

$$p_i = \frac{x_{f_i^1} + x_{f_i^2} + x_{f_i^3}}{3} \quad (14.33)$$

$$V_i = [\overrightarrow{x_{f_i^1} x_{f_i^2}}, \overrightarrow{x_{f_i^1} x_{f_i^3}}] \quad (14.34)$$

$$c_i = \frac{1}{2} \|\overrightarrow{x_{f_i^1} x_{f_i^2}} \wedge \overrightarrow{x_{f_i^1} x_{f_i^3}}\| \quad (14.35)$$

Avec  $V_i$  le plan engendré par les deux vecteurs.

Enfin si  $\varphi$  est un difféomorphisme et  $\mu$  un varifold on définit le transport  $(\varphi_*\mu)(\omega) = \mu(\varphi^*\omega)$  où  $\varphi^*\omega = |\varphi'(x)|\omega \circ \varphi$ .

*Propriété 14.1.* Si  $X$  est une sous variété de dimension  $d$  et  $\varphi$  un difféomorphisme alors

$$\varphi^*\mu_X = \mu_{\varphi(X)}. \quad (14.36)$$

*Démonstration.* Tout repose sur la formule de changement de variable, on écrit

$$\varphi_*\mu_X(\omega) = \mu(\varphi^*\omega) \quad (14.37)$$

$$= \int_X \varphi^*\omega(x, T_x X) dH^d(x) \quad (14.38)$$

$$= \int_X \omega(\varphi(x), d_x\varphi(T_x X)) |d_x\varphi \cdot T_x X| dH^d(x) \quad (14.39)$$

$$(14.40)$$

Comme  $\varphi$  est un difféomorphisme d'après la formule de changement de variable on a :

$$\varphi_*\mu_X(\omega) = \int_{\varphi(X)} \omega(y, T_y\varphi(X)) dH^d(x) = \mu_{\varphi(X)}(\omega) \quad (14.41)$$

□

#### 14.4.4 Noyaux

On sait déjà construire un noyau sur  $E$ , par exemple le noyau gaussien :  $k_E(x, y) = e^{-\frac{|x-y|^2}{2\sigma^2}}$ . L'objectif de cette partie est double : la construction d'un noyau  $k_V$  sur  $G_d(E)$ , puis la construction d'un noyau sur  $E \times G_d(E)$  que l'on obtiendra en considérant le produit tensoriel des deux noyaux sur chacun des deux membres du produits.

##### Construction de deux noyaux

On a vu qu'en pratique, on pouvait représenter deux éléments  $V$  et  $W$  de  $G_d(E)$  par deux vecteurs unitaires  $p_V$  et  $p_W$  définis aux signes près, que l'on peut aussi représenter matriciellement comme des projections orthogonales sur  $V$  et  $W$ , pour obtenir simplement une norme. On définit alors simplement deux noyaux sur cet espace de directions :

$$k_{CB}(V, W) = \langle p_V, p_W \rangle^k = \cos(\theta)^{2k} \quad (14.42)$$

$$k_G(V, W) = e^{-|p_V - p_W|^2/2\sigma^2} = e^{-4 \sin^2(\theta)/\sigma^2} \quad (14.43)$$



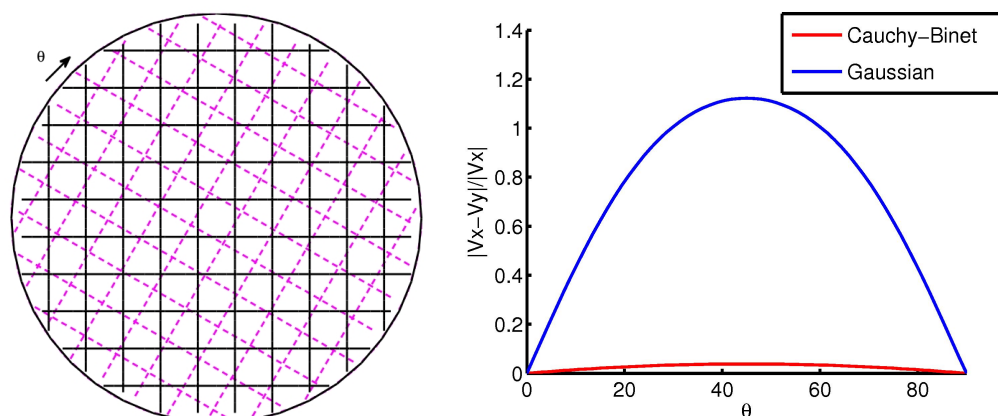


FIGURE 14.8 – Calcul de la distance entre les deux grilles, fonction de l'angle  $\theta$  de rotation de la grille rose par rapport à la noire. Le noyau spatial est d'une taille bien supérieure à celle des figures – i.e. floutage complet en position –, et on utilise sur les orientations un noyau de Cauchy-Binet (rouge) et un noyau gaussien (bleu) de faible déviation angulaire. Pour  $\theta = 0$  ou  $\pi/2$ , les deux grilles coïncident : il n'est donc pas étonnant que la dissimilarité la plus importante soit obtenue avec un décalage angulaire  $\theta = \pi/4$ . Sans surprise, il apparaît que le noyau gaussien sépare beaucoup mieux les orientations que Cauchy-Binet. On peut en effet calculer explicitement les interactions entre une croix de la grille noire et une croix de la grille rose :  $k(\theta, 0) + k(\theta, \pi/2) = \cos^2(t) + \cos^2(\pi/2 - t) = 1$  pour le noyau Cauchy-Binet (d'où l'indiscernabilité, qui n'est qu'approchée numériquement vu que la grille n'est pas découpée de manière régulière par le cercle) ; par contre, ce même terme d'interaction peut prendre des valeurs intéressantes si  $k$  est un pic plus resseré autour de 0.

### Produit tensoriel de deux noyaux

**Lemme 14.2.** Si  $A$  et  $B$  sont des ensembles et  $k_A$  et  $k_B$  sont des noyaux positifs sur  $A$  et  $B$  alors le produit tensoriel :

$$k_A \otimes k_B((a, b), (a', b')) = k_A(a, a')k_B(b, b') \quad (14.44)$$

est un noyau positif sur  $A \times B$ .

*Propriété 14.2.* Soit  $k_e$  un noyau réel sur  $E$ , tel que  $k_e$  est continu et borné, et que la fonction  $k_e(x, \cdot)$  s'annule à l'infini. Soit  $k_t$  un noyau sur  $G_d(E)$  qui est aussi continu, alors le noyau  $k_e \otimes k_t$  génère un RKHS qui s'injecte continuellement dans  $C_0(E \times G_d(E))$

Notons  $W$  le Hilbert généré par les fonctions du type  $k((x, V), \cdot)$ , on a  $W \subset C_0((E \times G_d(E)))$ . Ceci implique qu'il y a une application continue  $C_0((E \times G_d(E)))' \rightarrow W'$ , il en résulte que l'on peut représenter les sous variété  $X$  de  $E$  dans le dual par  $\mu_X$ , et on obtient la norme suivante pour la corrélation entre deux éléments de surface :

$$\langle \delta_{(x_1, V_1)}, \delta_{(x_2, V_2)} \rangle_{W'} = k_e(x_1, x_2)k_t(V_1, V_2) \quad (14.45)$$

Nous avons donc désormais une méthode simple pour construire des noyaux sur les varifolds.\* Cependant il n'est pas certain que la norme obtenue sur  $W'$  soit bien séparante : cela requiert l'injectivité de la fonction  $i^* : C_0((E \times G_d(E)))' \rightarrow W'$ , qui n'est pas toujours vérifiée. Pour le

Arnaud Eteve et Yohan Mandin -Huble  
Aspects pratiques

noyau de Cauchy Binet ce n'est en effet pas le cas, on l'a bien vu lorsque celle-ci ne séparait pas les deux grilles tournées l'une par rapport à l'autre, et ce quelque soit l'angle de rotation. On dira que  $i^*$  est  $C_0$  universelle si elle est injective.

*Propriété 14.3.* Prenons  $k = k_e \otimes k_t$  un noyau défini comme dans la proposition précédente, et supposons que  $k_e$  est  $C_0$  universel. Supposons de plus que  $k_t$  est tel que  $k_t(V, V) > 0$  pour tout  $V \in G_d(E)$ .

Alors la norme induite sur les formes est séparante, au sens où, si  $X = \bigcup_{i=1}^N X_i$  et  $Y = \bigcup_{i=1}^M Y_i$  sont deux unions finies de sous variétés compactes de dimension  $d$  de  $E$ , on a  $\|\mu_X - \mu_Y\| = 0$  si et seulement si  $X = Y$ .

*Propriété 14.4.* Si  $k = k_e \otimes k_G$  et  $k_e$  est  $C_0$  universel et  $k_G$  est le noyau gaussien défini comme précédemment alors  $k$  est  $C_0$  universel sur  $E \times G_d(E)$ .

La conclusion des deux propriétés précédentes est que si on choisit le noyau gaussien sur  $G_d(E)$  alors on est certain que la norme que l'on obtient est bien séparante.

### 14.4.5 Propriété des Varifolds

Premièrement regardons ce qui se passe lorsqu'on brouille complètement l'image, on prend sur  $E$  le noyau gaussien avec un  $\sigma = \infty$  (ou tout simplement,  $\sigma$  bien supérieur au diamètre de nos formes), si bien que  $k_e(x, y) = 1$  uniformément. Si on prend  $X$  une sous variété de dimension  $d$  alors la norme du courant associé s'exprime :

$$\|C_X\|_{W'}^2 = \int \int_{X \times X} \langle \xi(x), \xi(y) \rangle = \left\langle \int_X \xi(x), \int_X \xi(y) \right\rangle \quad (14.46)$$

, ce qui d'après la formule de Stokes ne dépend que du bord de  $X$ . Aussi, toute information sur l'intérieur de  $X$  est-elle perdue... ce qui n'est pas le cas pour le varifold : la norme s'exprime en effet (on remarquera qu'elle ne dépend que de la partie sur  $G_d(E)$  ce qui est normal puisqu'on a tué toute l'information sur la position) :

$$\|\mu_X\|_{W'}^2 = \int \int_{X \times X} k_t(T_x X, T_y X) dH^d(x) dH^d(y) \quad (14.47)$$

Cette dernière intégrale n'est à priori pas nulle même pour une variété sans bord ce qui traduit l'idée que l'on a pas perdu toute l'information sur la structure, il reste encore l'information concernant l'agencement des plans tangents, encodée comme une sorte de norme sur l'histogramme des orientations.

On obtient en particulier l'inégalité suivante :

*Propriété 14.5.* Soit  $X$  une sous variété de  $E$  incluse dans la boule unité, on suppose les noyaux choisis comme dans les propriétés précédentes. Alors il existe une constante  $K$  indépendante de  $X$  telle que

$$\|\mu_X\|_{W'} \geq KH^d(X) \quad (14.48)$$

C'est une très bonne nouvelle : la destruction de masse est ainsi évitée.

Étant donné un shoot sur l'espace des formes on peut se demander comment évolue  $\mu_X$ .

*Propriété 14.6.* Soit  $X$  une sous variété compacte et orientable et  $\mu_X$  le varifold associé. Soit  $v$  un champ de vecteur  $C^1$  à support compact défini sur  $E$ , et considérons le groupe à une dimension généré par le flot de  $v$ ,  $t \mapsto \varphi_t$ . Alors si  $X_t = \varphi(X)$  la variété transportée, pour toute fonction  $\omega$ ,  $C^1$  de  $E \times G_d(E)$  on a

$$\frac{d}{dt}|_{t=0}\mu_X(\omega) = \int_X \left( \frac{\partial\omega}{\partial x} - \text{div}_X \left( \frac{\partial\omega}{\partial x} \right) - \omega H_X |v^\perp \right) + \int_{\partial X} \langle \nu, \omega v^\top + \left( \frac{\partial\omega}{\partial V} |v^\perp \right) \rangle \quad (14.49)$$

où  $v^\perp$  et  $v^\top$  sont les parties tangentielle et normale de  $v$  le long de  $X$ , et  $\nu$  est la normale unité le long de  $\partial X$ .

Nous pouvons faire quelques remarque sur cette expression. Tout d'abord, les variations de la norme du varifold ne dependent que de  $v$  et pas de sa dérivé. Ensuite il y a des termes qui dépendent de  $\partial X$  et cela caractérise l'extension tangentielle de  $X$  (autrement dit si  $X$  est très étiré la norme du varifold associé est assez grande). Enfin sur la partie se trouvant à l'intérieur de  $X$ , la variation ne dépend que de  $v^\top$ , ce qui traduit le fait que le gradient de la distane d'appareillement est orthogonal à  $X$ .

### 14.4.6 Récapitulons

On a vu comment réduire le problème de l'attache aux données à l'obtention d'une formule simple donnant la corrélation entre deux éléments de courbe/surface élémentaires. Les trois formalismes développés, s'ils peuvent sembler éloignés conceptuellement, sont en fait extrêmement proches d'un point de vue algorithmique : nous rappelons ici les principales formules obtenues aujourd'hui, utiles en pratique.

	Mesures	Courants	Varifolds
Repr. d'un segment $[x_1 \rightarrow x_2]$ :	$\ x_2 - x_1\  \delta_{\frac{x_1+x_2}{2}}$	$\delta_{\frac{x_2-x_1}{2}}$	$\ x_2 - x_1\  \delta_{\left(\frac{x_1+x_2}{2}, \frac{x_2-x_1}{\ x_2-x_1\ }\right)}$
Repr. d'un triangle $[x_1, x_2, x_3]$ de normale $\xi = (x_2 - x_1) \wedge (x_3 - x_1)$ :	$\ \xi\  \delta_{\frac{x_1+x_2+x_3}{3}}$	$\delta_{\frac{\xi}{2}}$	$\ \xi\  \delta_{\left(\frac{x_1+x_2+x_3}{3}, \xi/\ \xi\ \right)}$
Corrélation entre deux diracs :	$\langle \mu\delta_x, \nu\delta_y \rangle$ $= \mu\nu k_{\sigma_I}(x, y)$	$\langle \delta_x^\xi, \delta_y^\zeta \rangle$ $= k_{\sigma_I}(x, y) \langle \xi, \zeta \rangle$ $= \mu\nu k_{\sigma_I}(x, y) \cos(\theta)$	$\langle \mu\delta_{(x,u)}, \nu\delta_{(y,v)} \rangle$ $= \mu\nu k_{\sigma_I}(x, y) \cos^2(\theta)$ ou $\mu\nu k_{\sigma_I}(x, y) \exp(-4 \sin^2(\theta)^2 / \sigma^2)$

En rappelant que, pour deux formes données par des collections de segments/triangles, la mesure/courant/varifold totale est la somme des diracs associés, avec :

$$\left\| \sum_i \delta_{x_i} - \sum_j \delta_{y_j} \right\|^2 = \left\langle \sum_i \delta_{x_i}, \sum_i \delta_{x_i} \right\rangle - 2 \left\langle \sum_i \delta_{x_i}, \sum_j \delta_{y_j} \right\rangle + \left\langle \sum_j \delta_{y_j}, \sum_j \delta_{y_j} \right\rangle \quad (14.50)$$

$$= \sum_i \sum_{i'} \langle \delta_{x_i}, \delta_{x_{i'}} \rangle - 2 \sum_i \sum_j \langle \delta_{x_i}, \delta_{y_j} \rangle + \sum_j \sum_{j'} \langle \delta_{y_j}, \delta_{y_{j'}} \rangle \quad (14.51)$$

## 14.5 Retour au problème d'appariement

Maintenant que nous avons bien étudié les propriétés des varifolds, et souligné les garanties que nous offre une telle représentation, il nous faut maintenant un algorithme pour calculer la solution du problème d'appareillement :

Étant donné deux formes :  $O_1$  et  $O_2$  respectivement la source et le but, On cherche la déformation optimale dans le groupe de difféomorphisme engendré par  $L_V^2([0, 1])$  où  $V$  est un espace de

Arnaud Eteve et Hilbert de champs de vecteurs sur  $E$ . Cela revient à résoudre le problème d'optimisation suivant :  
Yohan Mandin  
-Huble  
Aspects  
pratiques

$$\inf_{v \in L^2_V([0,1])} J(v) = \frac{1}{2} \int_0^1 |v_t|^2 dt + \frac{\gamma}{2} d(\varphi_1^v \cdot O_1, O_2) \quad (14.52)$$

où le premier terme est le coût de la déformation et le deuxième est l'écart restant. Ici  $\gamma$  est le paramètre du barycentre de la séance précédente, (ie un  $\gamma$  grand forcera la solution à être proche de  $O_2$ )

On a vu à la séance précédente que ce problème se ramenait à une optimisation sur le moment de tir initial  $p_0$  : la seule difficulté était de trouver un gradient sur  $p_0$ , une direction de descente qui nous permettrait de modifier ce tir de manière intelligente. En prenant  $d(\varphi_1^v \cdot O_1, O_2) = \|\varphi_1^v \cdot O_1 - O_2\|^2$  pour une des trois normes que nous avons défini aujourd'hui, nous sommes capable de calculer rapidement le gradient en  $q_1 = \varphi_1^v \cdot O_1$  du terme d'attaches aux données.

Étant donné une forme  $q_0$  (un nuage de points) que l'on cherche à matcher sur une forme  $\tilde{q}_1$ , on procédera de manière itérative pour trouver un moment de tir optimal  $p_0$ , permettant d'obtenir une déformation  $q_1 = \varphi_1^v \cdot q_0$  proche de  $\tilde{q}_1$ . À chaque itération de l'algorithme de descente sur  $p_0$ , il s'agira donc de répéter les actions suivantes :

1. Étant donné  $p_0$ , tirer la géodésique  $q_t$  par intégration du flot hamiltonien, partant de  $(q_0, p_0)$  pour arriver en  $(q_1, p_1)$ .
2. À l'arrivée, calculer l'opposé du gradient de l'attache aux données  $-\partial_{q_1} \gamma \|\mu_{q_1} - \mu_{\tilde{q}_1}\|^2 / 2$  pour obtenir un désir de mouvement  $\delta q_1$ .
3. "Transporter en arrière" ce désir  $\delta q_1$  pour obtenir un pas de descente  $\delta p_0$ .
4. Régulariser  $\delta p_0$ , en lui ajoutant l'opposé du gradient du terme de longueur,  $-\partial_{p_0} (p_0, K_{q_0} p_0) / 2 = -K_{q_0} p_0$  - il s'agit d'une descente de gradient  $L^2$ .
5. Suivre le pas de descente régularisé :  $p_0 \leftarrow p_0 + \delta p_0$ .

### 14.5.1 Applications

Dans la suite on prendra le noyau gaussien  $k_e(x, y) = e^{-\frac{|x-y|^2}{\sigma_e^2}}$  sur  $E$ . On a présenté que deux noyaux sur  $G_d(E)$  à savoir le noyau gaussien et le noyau de Cauchy-Binet, le deuxième est nettement moins bon en terme de séparation mais il a l'avantage d'être plus rapide à calculer (c'est juste un produit matriciel) donc facilement parallélisable et de ne pas ajouter de paramètre, le noyau gaussien permet quand à lui de bien mieux séparer les éléments selon l'agencement des plans tangents (ie des directions). Toutes les images qui suivent sont tirées du papier de Nicolas Charon et d'Alain Trouvé et ont été calculées en utilisant les noyaux gaussiens sur les deux espaces.

## 14.6 Conclusion

On a ainsi présenté trois manières de résoudre le problème d'appariement entre deux surfaces a priori non étiquetées. L'appariement simple de mesure, la méthode des courants puis les varifolds, rangés par ordre de précision, mais aussi de coût algorithmique : en pratique, il s'agira de trouver un bon compromis entre temps de calcul et précision souhaitée. Dans tous les cas toutes ces méthodes sont facilement parallélisables (il s'agit de simples produits matriciels . calculs de noyaux point par points), ce qui permet de les faire tourner sur des données réelles à l'aide de cartes graphiques - des fermes de quelques centaines/milliers de petits processeurs mis en parallèle, qui ont permis une révolution dans les méthodes de traitement d'images ces dix dernière années.

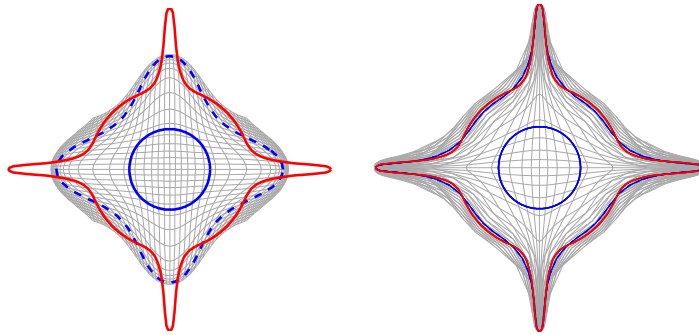


FIGURE 14.9 – On cherche à appairer ici le cercle bleu (au centre) avec l'étoile ninja en rouge, la version avec les courants lisse la structure complète alors que l'algorithme avec les varifolds donne un résultat bien meilleur.

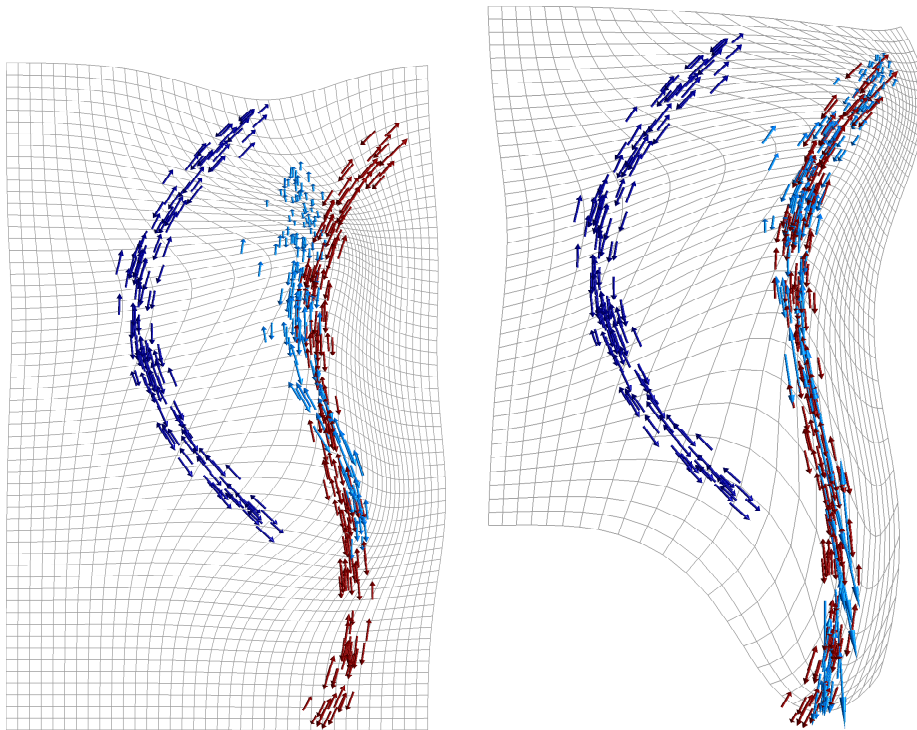


FIGURE 14.10 – Application au suivi de fibres – *fiber tracking* : grâce aux progrès effectués ces dernières années dans le design des séquences IRM, on dispose en chaque point d'une petite flèche (en réalité, c'est bien souvent un ellipsoïde) caractérisant la diffusion de l'eau dans les tissus observés. À partir de ces informations d'un genre nouveau, on peut typiquement repérer des hémorragies cérébrales (importante diffusion isotrope), ou inférer le tracé des fibres neuronales dans la matière blanche et la moelle épinière. On cherche ici à matcher l'ensemble bleu à gauche sur l'ensemble rouge : Les flèches de l'ensemble rouge ne sont pas orientées de manière cohérente et les courants (à gauche) ont donc beaucoup de mal à approcher correctement la structure ; cette absence d'orientation canonique ne pose aucun problème à l'attache varifolds (à droite) : le matching est bien plus satisfaisant, d'une fibre sur une autre.

Arnaud Eteve et  
Yohan Mandin  
-Huble  
Aspects  
pratiques

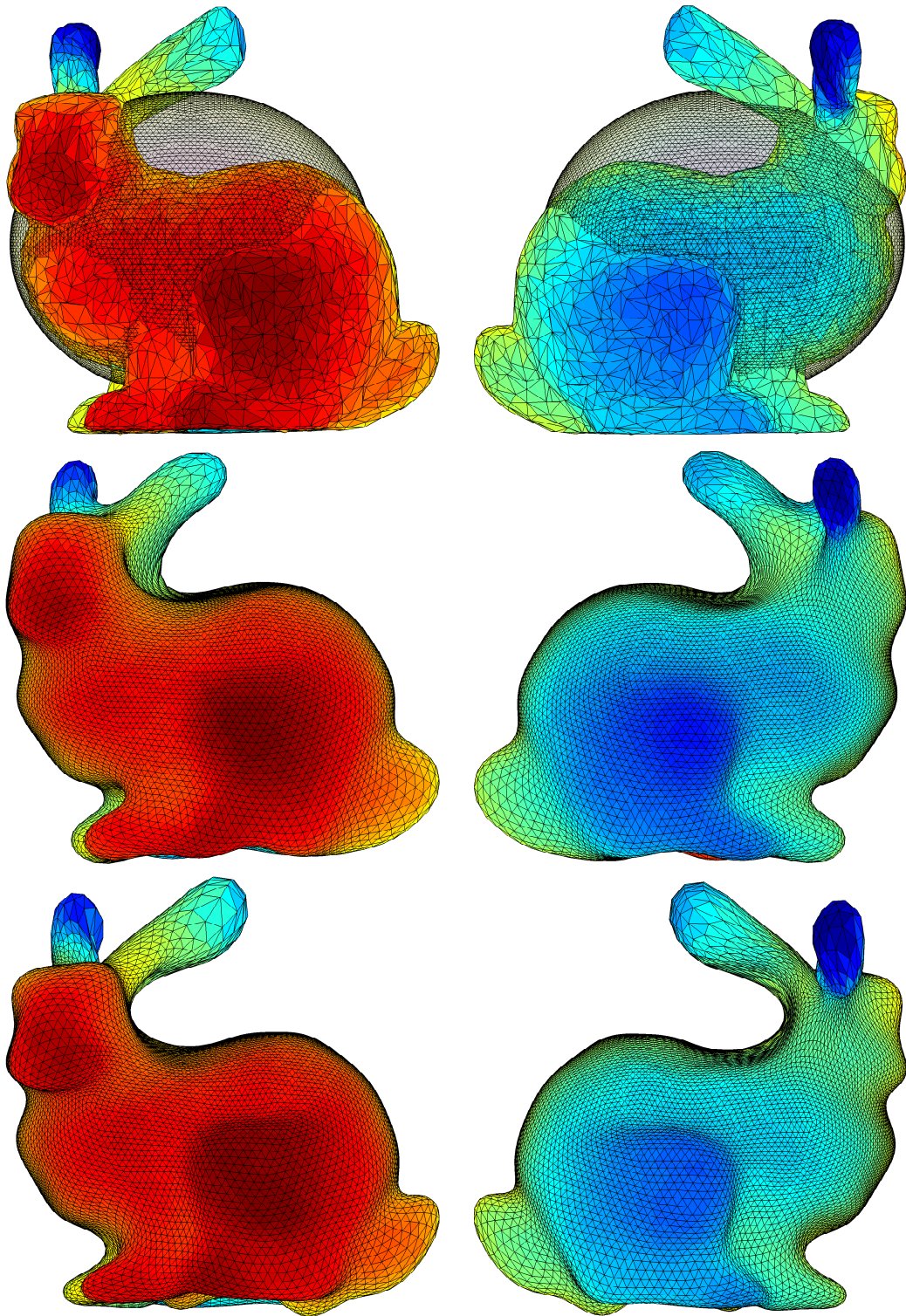


FIGURE 14.11 – On cherche à déformer la boule sur le lapin (ligne du haut, avec deux vues différentes). Là encore, les courants (ligne du milieu) ont un effet lissant tandis que les varifolds (ligne du bas) sont plus précis.



# Chapitre 15

## Construction d'Atlas, étude statistique et aide au diagnostic

*Bonus*

*[CCT14]*

*Eva Philippe  
et Robin Lemaire  
Applications  
réelles*

### 15.1 Introduction

Dans le domaine médical, on a souvent besoin d'étudier un paramètre, tout en s'affranchissant de la variabilité géométrique entre deux sujets. Par exemple, dans une rétine la présence d'un glaucome se traduit par une élévation de la pression au niveau du nerf optique. Pour comparer deux relevés d'épaisseur de la rétine, on veut pouvoir passer outre les différences physiologiques intrinsèques des individus.

L'étude des espaces de formes réalisées dans les chapitres précédents nous fournit de bons outils pour matcher une forme en une autre. On va donc s'intéresser maintenant au transport d'un signal numérique sur ces formes.

### 15.2 Ajout d'un terme fonctionnel

Dans les chapitres précédents, on a vu comment définir un espace de formes géométriques, lui donner une métrique ou y définir des varifolds. On s'intéresse maintenant au cas où en plus de notre forme géométrique  $X$ , on a un champ scalaire sur cette forme  $f : X \rightarrow \mathbb{R}$ . On va donc définir un espace  $\mathcal{F}$  de formes fonctionnelles (fshapes)  $(X, f)$ , qui soit correctement structuré, ainsi qu'une distance qui prenne en compte la composante géométrique et la composante fonctionnelle, qui assureront l'existence de géodésiques entre deux formes fonctionnelles quelconques de  $\mathcal{F}$ . Néanmoins  $\mathcal{F}$  est un espace de formes fonctionnelles idéales qui ne peut pas contenir des formes fonctionnelles trop bruitées observées en réalité. Pour prendre en compte plus de formes fonctionnelles, on étendra la notion de varifolds étudiée la semaine dernière avec des varifolds fonctionnels (fvarifolds).

#### 15.2.1 Définition des fshapes

Soit  $E$  l'espace vectoriel de dimension finie dans lequel sont plongées nos formes.

On a vu à la séance 9 comment construire un espace de difféomorphismes en intégrant des équations de flot sur des espaces à noyau reproduisant (RKHS) : soit donc  $V$  un RKHS de champs de vecteurs sur  $E$  ( $V$  est donc plongé continûment dans  $C_0^1(E, E)$  l'espace des fonctions continu-

Eva Philippe  
et Robin Lemaire  
Applications  
réelles



FIGURE 15.1 – Exemple de métamorphose entre deux visages, tirée du clip *Matilda* du groupe Alt-J ( $\Delta$ ). Observez comme des features géométriques (forme des oreille, coiffure) sont continûment déformées, tandis que des informations fonctionnelles (lunettes, foulard) sont obtenues par modification de l'image, via un simple fondu. Un matching “par métamorphose” permet donc de distinguer ces deux informations et, in fine, de comparer des images définies sur des supports géométriques différents. Notez que ce morphing n'est pas parfait (je ne sais pas comment il a été obtenu, probablement par un étiquetage automatique ou manuel des points saillants du visage + déformations localement affines) : le changement de nez est ici compris comme une information fonctionnelle, alors que nous voudrions l'inclure dans la partie géométrique.

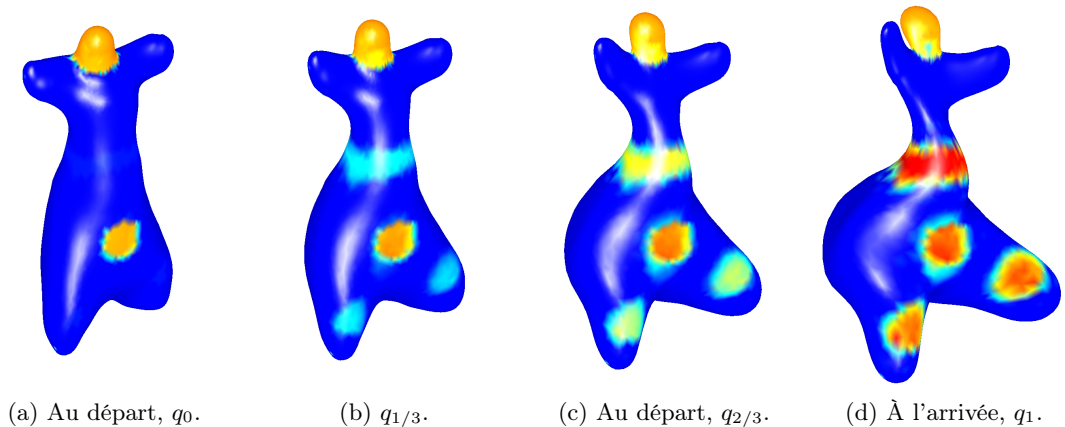
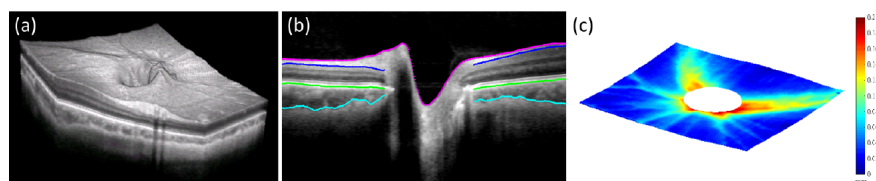
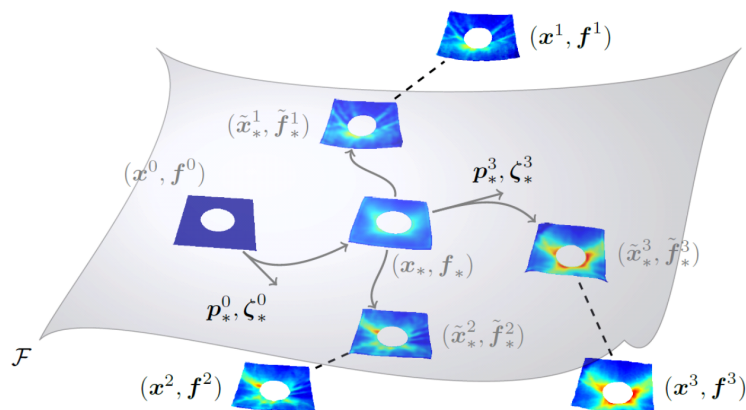


FIGURE 15.2 – Exemple de transformation d'une fshape le long d'un tir géodésique.

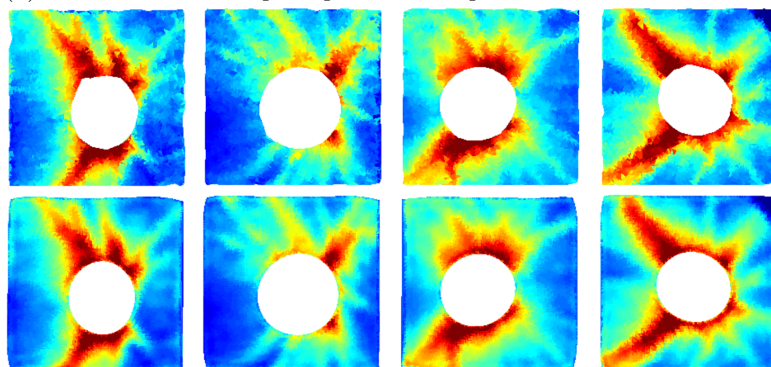




(a) Obtention de nos données : plutôt que de travailler avec un mille-feuille de couches rapprochées, on préférera assimiler nos rétines à des feuillettes simples, sur lesquels un signal donne l'épaisseur en microns.



(b) Illustration de l'idée principale de ce chapitre : la construction d'atlas.



(c) Sur la première ligne, quatre données typiques. Sur la seconde, les modèles obtenus par shooting, caractérisés dans la carte exponentielle par les moments de tir initiaux. Notez comme les zones de répartition de l'épaisseur dans le "carré" dépendent du cadrage et de la vascularisation sous-jacente, propre à chaque individu.

FIGURE 15.3 – Notre jeu de donnée "cas d'école" : scan de rétines optenus par tomographie à cohérence optique. L'objectif est d'arriver à une détection automatique du glaucome (une pathologie dûe à une surpression intra-oculaire, caractérisée par la dégénérescence de la rétine autour du point de branchement du nerf optique selon des patterns qui correspondent typiquement à la vascularisation sous-jacente, variable selon les individus). Ces images sont tirées de l'article d'application des idées exposées ici, *Atlas-based Shape Analysis and Classification of Retinal Optical Coherence Tomography Images using the Functional Shape (fshape) Framework*, par Sieun Lee, Nicolas Charon, Benjamin Charlier, Karteek Popuri, Evgeniy Lebed, Marinko V. Sarunic, Alain Trouvé et Mirza Faisal Beg

Eva Philippe  
et Robin Lemaire  
Applications  
réelles

ment différentiables de  $E$  dans  $E$  qui s'annulent à l'infini) et  $G$  le groupe des difféomorphismes associé obtenu en intégrant les champs de vecteurs dans  $L^2([0, 1], V)$ .

Soit  $\chi_0$  l'espace homogène engendré par un sous-ensemble  $X_0$  compact et de volume  $d$ -dimensionnel fini de  $E$  (i.e  $\chi_0 = \{\varphi(X_0) | \varphi \in G\}$ ).

On note  $L^2(X)$  l'espace des fonctions de carré intégrable sur  $X$  pour la mesure de Hausdorff en dimension  $d$ , i.e

$$L^2(X) = \left\{ f : X \rightarrow \mathbb{R} \mid \int_X f^2(x) d\mathcal{H}^d(x) < +\infty \right\}. \quad (15.1)$$

**Définition 15.1.** On définit alors l'espace des fshapes :

$$\mathcal{F} = \{(X, f) | X \in \chi_0, f \in L^2(X)\} \quad (15.2)$$

## 15.2.2 Distance de métamorphose

Pour une déformation uniquement géométrique  $\varphi \in G$ , un transport naturel d'une fshape  $(X, f)$  serait  $\varphi.(X, f) = (\varphi(X), f \circ \varphi^{-1})$ . Néanmoins, cela ne prend pas en compte la variation du signal elle-même. Nous allons voir comment considérer des transformations qui agissent simultanément sur la composante géométrique et la composante fonctionnelle, ce qui permettra de définir une distance sur  $\mathcal{F}$ .

Soit  $(X, f) \in \mathcal{F}$  et  $(v, h) \in L^2([0, 1], V \times L^2(X))$  des vitesses instantanées (pour  $t \in [0, 1]$ ,  $v_t$  est donc un champ de vecteurs de  $V$  et  $h_t$  est une fonction sur  $X$  de carré intégrable). On peut alors définir un chemin

$$(\varphi^v, \zeta^h) : t \in [0, 1] \mapsto (\varphi_t^v, \zeta_t^h), \quad (15.3)$$

où  $\varphi_t^v$  est le difféomorphisme obtenu en intégrant l'équation de flot de  $v$  à partir de l'identité (pour tout  $x \in \mathbb{R}^d$ ,  $\varphi_0^v(x) = (x)$  et  $\dot{\varphi}_t^v(x) = v_t(\varphi_t^v(x))$ ) et  $t \mapsto \zeta_t^h = \int_0^t h_s ds$  est le chemin de  $L^2(X)$  de vitesse instantanée donnée par  $h$ .

Puis on peut définir  $t \mapsto (X_t, f_t)$  un chemin sur  $\mathcal{F}$  par :

$$(X_t, f_t) = (\varphi_t^v.X, (f + \zeta_t^h) \circ (\varphi_t^v)^{-1}). \quad (15.4)$$

On note  $(\varphi_1^v, \zeta_1^h)$  l'extrémité finale de ce chemin.

On associe à un tel chemin un coût :

$$E_X(v, h) = \frac{\gamma_V}{2} \int_0^1 \|v_t\|_V^2 + \frac{\gamma_f}{2} \int_0^1 \int_X |h_t|^2(x) |d_x \varphi_t^v|_{T_x X} |d\mathcal{H}^d(x) \quad (15.5)$$

où  $|d_x \varphi_t^v|_{T_x X}$  est le Jacobien au point  $x$  de  $\varphi_t^v$  restreint à  $T_x X$  l'espace tangent à  $X$  en  $x$  et  $\gamma_V, \gamma_f$  sont deux paramètres strictement positifs que l'on peut ajuster pour donner plus de poids à la variation géométrique ou à la variation fonctionnelle.

La présence de ce Jacobien est une conséquence de la formule de changement de variables et s'explique par le fait qu'on comptabilise plus ou moins les variations du signal selon "l'aire" (qui varie) de la région concernée.

**Définition 15.2.** On peut maintenant définir pour  $(X, f), (X', f')$  dans  $\mathcal{F}$  :

$$d_{\mathcal{F}}((X, f), (X', f')) = \left( \inf \{ E_X(v, h) | X' = \varphi_1^v.X, f' = (f + \zeta_1^h) \circ (\varphi_1^v)^{-1} \} \right)^{\frac{1}{2}}. \quad (15.6)$$

**Théorème 15.1.** La fonction  $d_{\mathcal{F}}$  définit une distance sur  $\mathcal{F}$ .

*Démonstration.*  $d_{\mathcal{F}}$  est bien définie et est finie car  $\mathcal{F}$  a été bien défini.

La symétrie vient du fait qu'on peut inverser le sens de parcours des chemins.

L'inégalité triangulaire est obtenue par concaténation de chemins.

On vérifie aussi que si  $d_{\mathcal{F}}((X, f), (X', f')) = 0$ , alors  $X = X'$  et  $f = f'$ .  $\square$

### 15.2.3 Existence de géodésiques

Maintenant que le cadre a été bien posé, on a un théorème d'existence d'une géodésique minimisante entre deux fshapes quelconques de  $\mathcal{F}$  :

**Théorème 15.2.** *Pour  $(X, f)$  et  $(X', f')$  dans  $\mathcal{F}$ , il existe  $(v, h) \in L^2([0, 1], V \times L^2(X))$  tel que*

$$E_X(v, h)^{\frac{1}{2}} = d_{\mathcal{F}}((X, f), (X', f')). \quad (15.7)$$

*En particulier, le chemin*

$$t \mapsto (X_t, f_t) = (\varphi_t^v \cdot X, (f + \int_0^t h_s ds) \circ (\varphi_t^v)^{-1}), \quad (15.8)$$

*peut être considéré comme une géodésique minimisante entre  $(X, f)$  et  $(X', f')$ .*

### 15.2.4 Modèle tangentiel

La distance de métamorphose définie précédemment, qui calcule l'évolution fonctionnelle  $t \mapsto \zeta_t$  selon une métrique  $L^2(X_t)$  qui dépend du temps :

$$\int_0^1 \int_X |\dot{\zeta}_t(x)|^2 |d_x \varphi_t^v|_{T_x X} |d\mathcal{H}^d(x)| dt = \int_0^1 \int_{X_t} |\dot{\zeta}_t \circ (\varphi_t^v)^{-1}|^2(y) d\mathcal{H}^d(y) dt, \quad (15.9)$$

pose le problème suivant : si on ne pénalise pas du tout la variation géométrique et qu'on pénalise la variation du signal selon "le volume" (ce qui correspond à la multiplication par le Jacobien), une manière de minimiser le coût pourrait être de compresser la forme géométrique, faire varier le signal à moindres frais sur un petit volume puis redilater.

Pour éviter cette solution peu satisfaisante, on peut geler la métrique à sa valeur initiale sur  $X_0 = X$  et négliger la variation de la mesure.

On obtient alors le coût suivant :

$$\int_0^1 \int_X |\dot{\zeta}_t \circ (\varphi_t^v)^{-1}|^2(y) d\mathcal{H}^d(y) dt, \quad (15.10)$$

soit après application du théorème de Fubini :

$$\int_X \int_0^1 |\dot{\zeta}_t(x)|^2 d\mathcal{H}^d(x) dt. \quad (15.11)$$

Or, à  $x \in X$  fixé et étant données les extrémités  $\zeta_0 = 0$  et  $\zeta_1$ , le plus court chemin est la droite  $t \mapsto t\zeta_1(x)$ . On obtient donc le coût habituel sur  $L^2(X)$  :

$$|\zeta_1|_X^2 = \int_X |\zeta_1(x)|^2 d\mathcal{H}^d(x). \quad (15.12)$$

Eva Philippe  
et Robin Lemaire  
Applications  
réelles

Avec ce nouveau coût on peut à nouveau définir une application  $\tilde{d}_{\mathcal{F}}$  sur  $\mathcal{F}$  :

$$\tilde{d}_{\mathcal{F}}((X, f), (X', f')) = \left( \inf \left\{ \frac{\gamma_V}{2} \int_0^1 |v_t|_V^2 dt + \frac{\gamma_f}{2} |\zeta|_X^2 |X' = \varphi_1^v \cdot X, f' = (f + \zeta) \circ (\varphi_1^v)^{-1} \right\} \right)^{\frac{1}{2}} \quad (15.13)$$

$\tilde{d}_{\mathcal{F}}$  n'est plus une distance mais on a quand-même un résultat d'existence d'un chemin optimal pour matcher deux fshapes quelconques.

La méthode tangentielle est également plus rapide à calculer mais présente l'inconvénient qu'il faut "se souvenir de la paramétrisation au temps initiale" à chaque instant de la métamorphose pour la pénaliser de manière adéquate.

### 15.2.5 Fvarifolds (varifolds fonctionnels)

Jusque-là on a considéré des comparaisons de fshapes qui appartiennent à un même espace  $\mathcal{F}$ , mais cette hypothèse est trop forte pour les données concrètes qu'on voudra comparer dans des situations pratiques (elle impose par exemple que les supports géométriques puissent être matchés par une déformation appartenant au groupe  $G$ ). Pour résoudre ce problème, on va procéder comme lors des séances précédentes : réaliser un matching imparfait et introduire une mesure de dissimilarité entre la fshape obtenue et celle désirée grâce au terme d'attache aux données. On peut adapter l'approche des varifolds présentée lors de la séance 10 pour prendre en compte la composante fonctionnelle.

Comme précédemment, on considère  $E$  l'espace vectoriel de dimension  $n$  où sont plongées les supports géométriques des fshapes, et comme la semaine dernière,  $G_d(E)$  est la Grassmannienne de  $E$  : la variété de tous les sous-espaces de dimension  $d$  de  $E$ .

**Définition 15.3.** On dit que  $\mu$  est un *varifold fonctionnel* (ou *fvarifold*) de dimension  $d$  si  $\mu$  est une mesure borélienne finie sur l'espace  $E \times G_d(E) \times \mathbb{R}$ , ou de manière équivalente :  $\mu \in C_0(E \times G_d(E) \times \mathbb{R})'$ .

La corrélation entre deux fshapes (par exemple des triangles colorés)  $\delta_{(x,V,c)}$  et  $\delta_{(x',V',c')}$  (où  $x$  code la position,  $V$  l'orientation (la normale pour un triangle) et  $c$  le signal) est alors donnée par une formule à noyau :

$$\langle \delta_{(x,V,c)}, \delta_{(x',V',c')} \rangle = k_e(x, x') k_t(V, V') k_f(c, c') \quad (15.14)$$

pour des diracs, où  $k_e$  est un noyau positif sur  $E$  continu, borné et tel que  $k_e(x, \cdot)$  s'annule à l'infini pour tout  $x$ ,  $k_t$  est un noyau continu sur la variété  $G_d(E)$  et  $k_f$  est un noyau sur  $\mathbb{R}$  continu, borné et tel que  $k_f(c, \cdot)$  s'annule à l'infini pour tout  $c$ .

De la même manière que les métriques à noyaux poussent les points à s'aligner, on pousse ici les signaux à s'ajuster.

## 15.3 Etude de population avec moyenne à estimer

Lors d'une expérience réelle, on obtient un jeu de données éventuellement bruitées et pour pouvoir réaliser les appariements, on cherche à définir un modèle moyen (template)  $(X, f)$  qui va permettre les comparaisons de fshapes. On s'intéresse dans cette partie à l'optimisation d'un tel template. En formalisant l'estimation d'atlas comme un problème variationnel, on obtient des résultats d'existence. Les deux parties suivantes seront consacrées à deux algorithmes qui permettent le calcul effectif d'un template.

On approximera dans cette section une forme fonctionnelle  $(X_c, f_c)$  par un ensemble discret de points  $\mathbf{x}$  reliés, associé à la fonction restreinte à cet ensemble de points. On obtient ainsi un polyèdre fonctionnel  $(X, f)$ , qui est la seule approximation discrète raisonnable, bien qu'on ne discutera pas de la convergence de ce modèle. On approximera de même la mesure  $\mu_{X,f}$  par la somme de Diracs  $\mu_{\mathbf{x},\mathbf{y}}$ .

Comme on souhaite minimiser la fonctionnelle du coût, les trajectoires de la déformation sont des géodésiques. Ainsi nous avons vu que les géodésiques sont déterminées de manière unique par le vecteur vitesse initiale. (Nous sommes ici en présence d'un vecteur de dimension  $dP$  où  $d$  est la dimension de l'espace et  $P$  le nombre de points de notre approximation.)

Il suffit alors de minimiser la fonctionnelle

$$\frac{\gamma_V}{2} \int_0^1 |v_t|_V^2 dt + \frac{\gamma_f}{2} \int_0^1 (D(\mathbf{x}_t) \mathbf{h}_t | \mathbf{h}_t) dt \quad (15.15)$$

Avec  $(\mathbf{x}_0, \mathbf{f}_0)$  et  $(\mathbf{x}_1, \mathbf{f}_1)$  fixés et

$$\begin{cases} \dot{\mathbf{x}}_t = v_t \cdot \mathbf{x}_t \\ \dot{\mathbf{f}}_t = \mathbf{h}_t \end{cases} \quad (15.16)$$

en notant  $v \cdot \mathbf{x} = (v(x_k))_{1 \leq k \leq P}$ .

Ainsi en introduisant les quantités de mouvements et hamiltonien, on obtient

$$\begin{cases} \dot{\mathbf{x}} = \partial_{\mathbf{p}} H_r^{\text{geo}}(\mathbf{x}, \mathbf{p}) \\ \dot{\mathbf{p}} = -\partial_{\mathbf{x}} H_r^{\text{geo}}(\mathbf{x}, \mathbf{p}) \\ \mathbf{f}_t = \mathbf{f} + t\zeta \end{cases} \quad (15.17)$$

Une fois ces bases de modèle posées, tentons de voir comment obtenir notre template moyen.

### 15.3.1 Hypertemplate

Le premier algorithme présenté part d'un hypertemplate, que l'on déformera afin d'obtenir un template moyen entre les observations.

La méthode est telle que présentée dans la figure :

L'algorithme prend en entrée  $N$  observations  $(\mathbf{x}^i, \mathbf{f}^i)$ , un hypertemplate  $(\mathbf{x}_0, \mathbf{f}_0)$  où  $\mathbf{f}_0 = 0$ , les quantités de mouvements  $\mathbf{p}_0$  et les  $\mathbf{p}^1, \dots, \mathbf{p}^N$  ainsi que les fonctionnelles  $\mathbf{f}$  et  $\zeta^1, \dots, \zeta^N$ .

On écrit d'abord un algorithme calculant l'énergie  $\mathbf{J}_0^{\text{tan}}$ .

$$\mathbf{J}_0^{\text{tan}}(\mathbf{p}_0, \mathbf{f}, (\mathbf{p}^i, \mathbf{f}^i)_{1 \leq i \leq N}) = \frac{\gamma_{V_0}}{2} |v^{\mathbf{p}_0}|_{V_0}^2 + \frac{\gamma_f^0}{2} |\mathbf{f}|_{\mathbf{x}}^2 + \sum_{i=1}^N \left( \frac{\gamma_V}{2} |v^{\mathbf{p}^0}|_{V_0}^2 + \frac{\gamma_f}{2} |\zeta^i|_{\mathbf{x}}^2 + \frac{\gamma_W}{2} g((\tilde{\mathbf{x}}^i, \tilde{\mathbf{f}}^i), \mathbf{x}^i, \mathbf{f}^i) \right) \quad (15.18)$$

Où  $g$  calcule le terme d'attache aux données pour aller du template moyen en cours de construction vers les observations.

L'algorithme calcule  $\mathbf{x}$  en intégrant la quantité de mouvement initiale  $\mathbf{p}_0$ , puis les  $(\tilde{\mathbf{x}}^i, \tilde{\mathbf{f}}^i)$  en intégrant les  $\mathbf{p}^i$  et par  $\tilde{\mathbf{f}}^i = \mathbf{f} + \zeta^i$ .

On ajoute ensuite les coûts  $|v^{\mathbf{p}^i}|_V^2$  et  $|v^{\mathbf{p}_0}|_{V_0}^2$  et les coûts  $|\zeta^i|_{\mathbf{x}}^2$  et  $|\mathbf{f}|_{\mathbf{x}}^2$ .

On écrit ensuite un algorithme calculant le gradient de cette énergie en calculant simplement le  $\mathbf{f}$  gradient de chaque terme en chaque variable.

L'algorithme final optimisera des déformations de notre hypertemplate pour se rapprocher au plus d'un template moyen entre les observations. Nous le verrons après la deuxième méthode.

Eva Philippe  
et Robin Lemaire  
Applications  
réelles

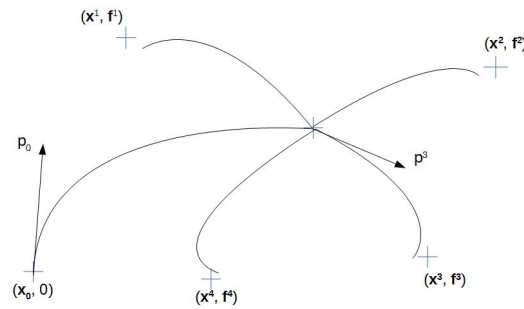


FIGURE 15.4 – Méthode de l'hypertemplate

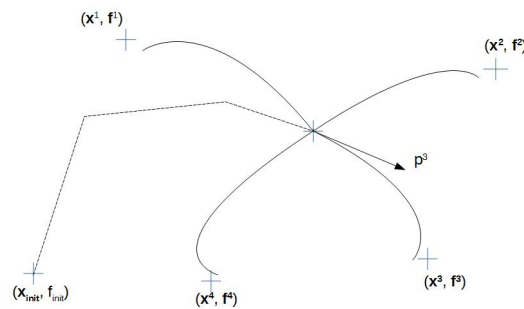


FIGURE 15.5 – Méthode de la moyenne libre

La sortie est alors une estimation du template moyen, les moments  $\mathbf{p}^1, \dots, \mathbf{p}^N$  et les fonctionnelles résiduelles  $\zeta^1, \dots, \zeta^N$ .

Cet algorithme pose le problème du choix de l'hypertemplate. En effet il faut utiliser une forme topologiquement proche des observations, sous peine de se retrouver avec quelque chose d'absurde et très loin des observations. Néanmoins il n'est pas conseillé d'utiliser une des observations sous peine de biaiser l'atlas, le template "moyen" que l'on obtient étant plus proche d'une des observations, choisie arbitrairement.

### 15.3.2 Forme moyenne "libre"

La deuxième méthode proposée est proche de la première, mais au lieu d'avoir un hyper-template que l'on déforme, nous allons mettre à jour le template moyen lors de la descente de gradient. A chaque pas, le template moyen se modifiera un peu.

Cette méthode a le mérite d'avoir moins de contraintes sur l'évolution du template que la première, en effet il est "libre" de se déplacer lors du processus d'optimisation. Cela lui permet probablement d'être moins dépendant aux conditions initiales.

Néanmoins, si l'on peut montrer la correction de la méthode précédente, celle-ci n'a pas les mêmes garanties théoriques.

Elle prend les mêmes entrées que la première méthode, moins la quantité de mouvement  $\mathbf{p}_0$ , et telle que le template initial est  $(\mathbf{x}_{init}, \mathbf{f}_{init})$  où  $\mathbf{f}_{init}$  n'est pas forcément identiquement nulle.

L'énergie pour cette méthode est la même que celle de la méthode précédente, moins le terme dû au passage de l'hyper-template au template moyen estimé temporaire. On la calcule de la même manière, ainsi que son gradient.

Cette deuxième méthode renvoie les mêmes sorties que la première, à savoir une estimation du template moyen, les moments  $\mathbf{p}^1, \dots, \mathbf{p}^N$  et les fonctionnelles résiduelles  $\zeta^1, \dots, \zeta^N$ .

### 15.3.3 Optimisation

L'algorithme d'optimisation est une descente de gradient dite adaptative. En effet, les variables sont nombreuses et de types différents : ce sont les  $\mathbf{f}, \mathbf{p}, \mathbf{f}^i, \mathbf{p}^i$ . Ces variables vivant dans des espaces différents, ça n'aurait que peu de sens de les incrémenter de la même manière. On utilise donc quatre tailles de pas différentes. L'algorithme consiste à voir les quantités de mouvement comme des désirs d'avancer. Ici, on utilise seulement le gradient de l'énergie, laquelle pénalise les écarts quadratiques aux observations (ceux-ci incluant les coûts de déformation et les coûts d'attache aux données). Quand on l'implémente, on peut de plus ajouter un coefficient sur ses deux composantes (fonctionnelle et géométrique) afin d'accorder plus ou moins d'importance à chacune, selon le travail que l'on veut effectuer avec.

On remarquera avec profit qu'appliquer cet algorithme plusieurs fois avec des noyaux de plus en plus fins permet d'améliorer la pertinence du template obtenu, si l'on applique à chaque itération l'algorithme en le template d'arrivée.

Dans tous les cas l'algorithme se fait ainsi :

1. **Entrées** : La fonctionnelle  $\mathbf{J}$ , énergie initiale, et son gradient initial  $\nabla \mathbf{J} = (\nabla_{\mathbf{u}_i} \mathbf{J})_{i=1}^K$ .  
Les variables  $\mathbf{f}^{init}, \mathbf{p}^{init}, \mathbf{f}_i^{init}, \mathbf{p}_i^{init}$  et les tailles de pas  $(\delta_{\mathbf{f}}^{init}, \delta_{\mathbf{p}}^{init}, \delta_{\mathbf{f}_i}^{init}, \delta_{\mathbf{p}_i}^{init})$
2. **Début**
3. Initialisation : On initialise les  $\mathbf{f}, \mathbf{p}, \mathbf{f}_i, \mathbf{p}_i$  et  $(\delta_{\mathbf{f}}, \delta_{\mathbf{p}}, \delta_{\mathbf{f}_i}, \delta_{\mathbf{p}_i})$ , et breakLoop à False
4. **Tant que** "Nombre d'itérations inférieur au nombre maximal d'itérations"  
. et "les pas ne sont pas trop petits" et "la descente n'est pas trop petite" :
5. . Calculer  $J^{actuel} \leftarrow \mathbf{J}$  et son gradient  $\nabla_{\mathbf{u}} J^{actuel} \leftarrow \nabla \mathbf{J}$ . Où  $\mathbf{u}$  est  $\mathbf{f}, \mathbf{p}, \mathbf{f}_i$  ou  $\mathbf{p}_i$ .
6. . Mettre à jour toutes les variables  $\mathbf{u}_i^{new} \leftarrow \mathbf{u} + \delta_{\mathbf{u}} \nabla_{\mathbf{u}} J^{actuel}$  et calculer  $J^{new}$ .
7. . Adapter les pas
8. . **Tant que** "breakLoop = False" et "les pas ne sont pas trop petits" :
9. . Calculer l'énergie en mettant à jour chaque variable séparément :  
.  $(\hat{J}_i^{new})$  et la calculer en mettant les variables à jour en même temps :  $\hat{J}^{new}$
10. . Garder la meilleure configuration (pour minimiser l'énergie) :  
. on obtient des  $\mathbf{f}^*, \mathbf{p}^*, \mathbf{f}_i^*, \mathbf{p}_i^*$  et  $(\delta_{\mathbf{f}}^*, \delta_{\mathbf{p}}^*, \delta_{\mathbf{f}_i}^*, \delta_{\mathbf{p}_i}^*)$ .

Eva Philippe  
et Robin Lemaire  
Applications  
réelles

```

11. .      Si  $J(\mathbf{f}^*, \mathbf{p}^*, \mathbf{f}_i^*, \mathbf{p}_i^*) < J^{actuel}$ 
12. .      Alors
13. .          Mettre à jour les  $\mathbf{u} \leftarrow \mathbf{u}^*$  et  $\delta_{\mathbf{u}} \leftarrow \delta_{\mathbf{u}}^*$ 
14. .          breakLoop  $\leftarrow$  True
15. .      Sinon
16. .          Diminuer les tailles de pas
17. .      Fin Si
18. .  Fin Tant que
19. Fin Tant que
20. Fin
21. Sortie : Les variables  $\mathbf{f}, \mathbf{p}, \mathbf{f}_i, \mathbf{p}_i$ 

```

### 15.3.4 Pourquoi deux algorithmes ?

La différence entre les deux algorithmes est que la méthode utilisant l'hypertemplate impose des contraintes au template moyen calculé. Il reste relié à l'hypertemplate par un difféomorphisme, qui l'empêche de partir à l'infini. Ainsi la méthode est bien fondée mathématiquement, car les templates provisoires obtenus à chaque itération vont converger. Néanmoins elle implique d'être un peu attachée à l'initialisation.

En ce qui concerne le template "libre", aucune contrainte ne nous permet d'affirmer que les templates obtenus à chaque itération vont converger (souvenons nous qu'en dimension infinie, une boule n'est pas compacte). Même si en pratique cet algorithme tourne bien et plus vite que le premier, il peut créer des problèmes, notamment en bordure de forme comme nous le verrons. Sa limitation principale vient du fait qu'il tente de répondre à une question mal posée, car il cherchera une "vraie" forme moyenne entre les différentes observations, alors que rien n'indique qu'elle existe. De son côté, le premier algorithme restreindra sa recherche aux formes qui sont en un sens proches de l'hypertemplate, ce qui assure l'existence de l'optimum.

## 15.4 Limitations et contraintes d'implémentation

### 15.4.1 Perte de masse

Le fait d'avoir ici des formes fonctionnelles, et la norme associée, induit des problèmes lorsqu'on veut matcher des formes de tailles différentes. En effet nous nous intéressons au produit scalaire de diracs, or nous avons

$$|\langle \delta_{x_1, V_1, f_1}, \delta_{x_2, V_2, f_2} \rangle_{W'}| \leq |k_e(\cdot, \cdot)|_{\infty} |k_t(\cdot, \cdot)|_{\infty} k_f(f_1, f_2) \quad (15.19)$$

Ainsi comme  $k_f(f_1, \cdot)$  est continue et tend vers 0 en l'infini, on obtient que ce produit scalaire peut être arbitrairement petit si l'on choisit une fonction variant beaucoup. Cela entraîne des comportements problématiques quand on laisse une grande liberté à l'algorithme pour changer le signal (paramètre  $\gamma_f$ ).

Par exemple, si l'on souhaite envoyer un grand carré de signal uniforme 0 sur un petit carré de signal uniforme 1

Si l'on laisse une liberté totale à l'algorithme ( $\gamma_f = 0$ ), il créera des oscillations déraisonnables de  $f$  dans le complémentaire du but, ce qui l'annule artificiellement d'après lui. En augmentant un peu  $\gamma_f$ , les oscillations sont plus faibles, mais le signal du but est moins précis et un peu plus faible.



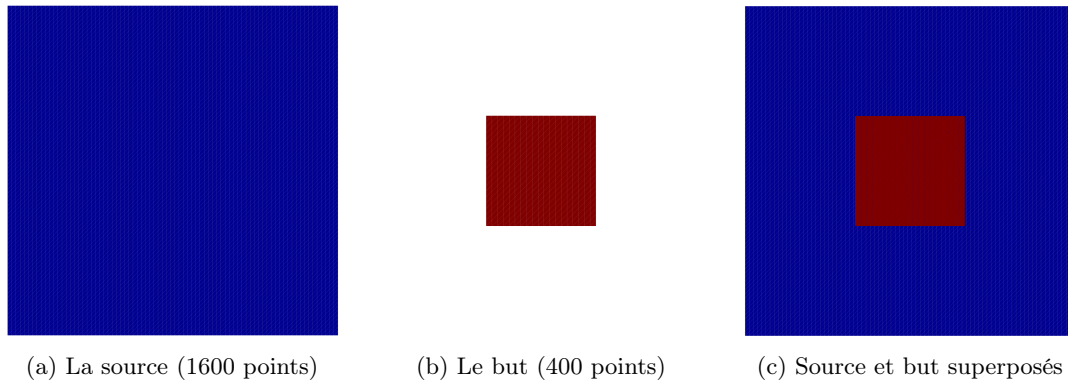


FIGURE 15.6 – Nous allons tenter d'envoyer le grand carré sur le petit, en refusant toute transformation géométrique.

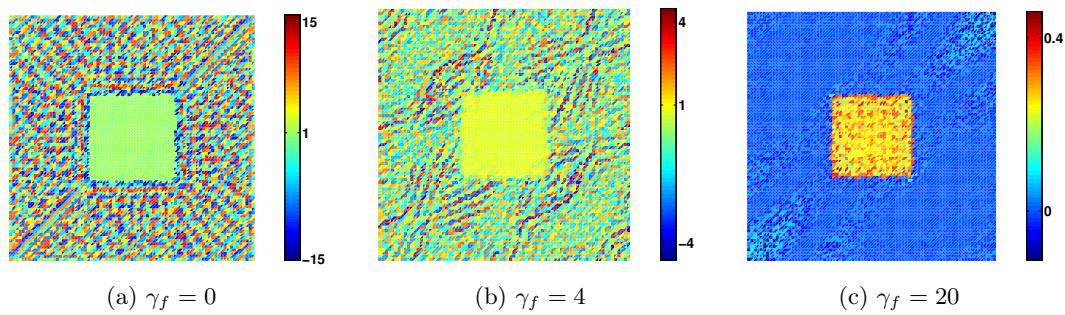


FIGURE 15.7 – On obtient en sortie de descente de gradient ces formes pour différentes valeurs de  $\gamma_f$ .

Eva Philippe  
et Robin Lemaire  
Applications  
réelles

Enfin si l'on augmente encore  $\gamma_f$ , les oscillations disparaissent presque, et le complémentaire du but garde un signal faible, mais le signal du but est d'intensité faible (ici environ 0.25 et non plus 1)

Une solution envisageable serait de régulariser le gradient de  $\mathbf{J}$  par rapport à  $\mathbf{f}$ , de manière analogue à ce que nous ferons au paragraphe suivant, avec la partie géométrique de l'énergie.

## 15.4.2 Problèmes de bordures

Lorsqu'on laisse le template moyen se déplacer librement lors de sa création, on observe expérimentalement des problèmes d'ordre topologique (le signal n'a ici pas d'influence) lorsqu'on manipule des formes possédant des bordures.

Dans notre approximation discrète, nous avons un polyèdre formant un réseau entre les points. On peut donc définir un point qui serait sur un bord comme suit :

Le problème est que si l'on regarde les désirs de se déplacer de chaque point, les points intérieurs de la forme étudiée n'ont que peu de désirs. Alors le gradient de l'énergie n'est important que proche des bords, où il explose.

Par exemple on souhaite matcher les formes ci-dessous :

Alors l'algorithme proposé donne cette descente de gradient :

Une solution proposée est de définir un nouveau gradient grâce à une nouvelle métrique, obtenue grâce à un noyau gaussien  $K_{reg}$ . On note  $K_{\mathbf{x},\mathbf{x}}^{reg} = (K_{reg}(x_k, x_{k'}))_{1 \leq k, k' \leq P}$  et on pose

$$\tilde{\nabla}_{\mathbf{x}} \mathbf{J}_{geo}(\mathbf{x}) = K_{\mathbf{x},\mathbf{x}}^{reg} \nabla_{\mathbf{x}} \mathbf{J}_{geo}(\mathbf{x}) \quad (15.20)$$

Ce nouveau gradient est alors plus régulier, ce qui donne une descente de gradient plus progressive et de meilleure facture.

## 15.4.3 Problèmes dûs aux observations

Cette technique permet notamment de repérer des anomalies dans des organes humains en établissant un organe "moyen", qui servira d'étalon pour discriminer les organes sains de ceux présentant une anomalie. Néanmoins on ne peut pas rentrer un organe dans l'algorithme, et les images réalisées à partir des organes (observations) peuvent présenter des divergences dues au praticien ou au hasard. Dans le cas des images de rétines, le cadrage, la taille du modèle, etc. peuvent varier d'une observation à l'autre. Malheureusement ce modèle ne prend pas en compte de telles différences dans les observations, ce qui fausse invariablement les résultats.

Cette technique permet notamment de repérer des anomalies dans des organes humains en établissant un organe "moyen", qui servira d'étalon pour discriminer les organes sains de ceux présentant une anomalie. Néanmoins on ne peut pas rentrer un organe dans l'algorithme, et les images réalisées à partir des organes (observations) peuvent présenter des divergences dues au praticien ou au hasard. Dans le cas des images de rétines, le cadrage, la taille du modèle, etc. peuvent varier d'une observation à l'autre. Malheureusement ce modèle ne prend pas en compte de telles différences dans les observations, ce qui fausse invariablement les résultats.

## 15.5 Conclusion

Ainsi, on a des outils très bien théorisés permettant de construire un template moyen à partir de ces observations ainsi que des cartes renvoyant à ces observations. On peut donc effectivement comparer des signaux sur des formes (par exemples des rétines) en s'affranchissant de la variabilité géométrique entre individus. La méthode développée est assez coûteuse en calcul mais elle est plus

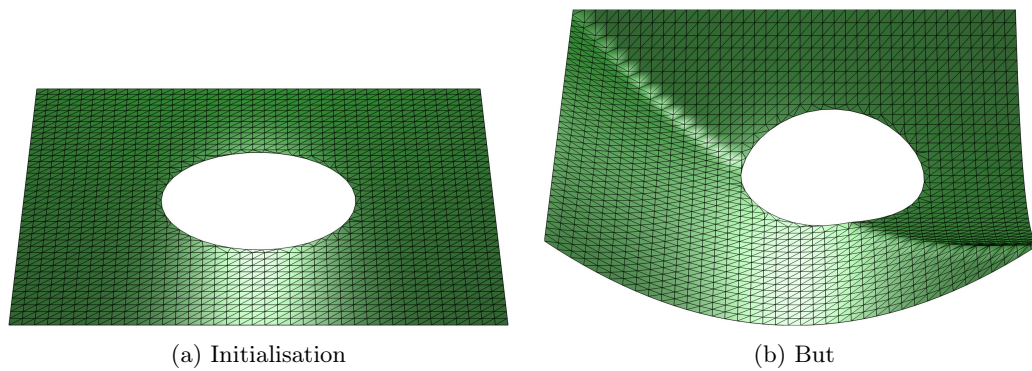
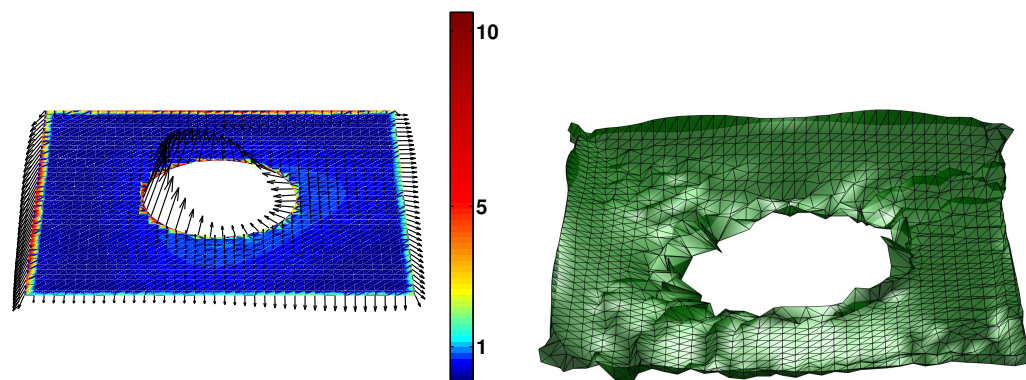


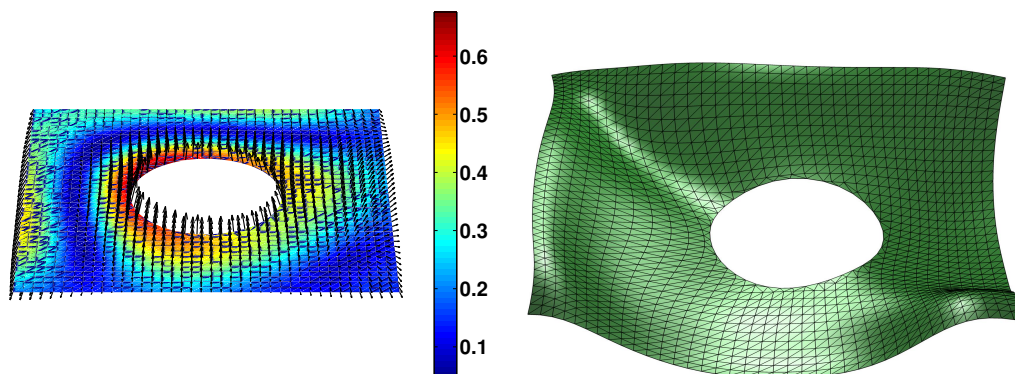
FIGURE 15.8 – On prend ici une forme ressemblant à celle d’une coupe de rétine.



(a) Les flèches représentent le gradient à l’initialisation, la couleur représente la norme du gradient avec une échelle logarithmique.

(b) Après quelques pas de descente de gradient, on obtient ceci.

FIGURE 15.9 – Descente de gradient sur l’énergie par la méthode du template moyen libre.



(a) Les flèches représentent le nouveau gradient à l’initialisation, la couleur représente la norme du gradient avec une échelle logarithmique.

(b) Après quelques pas de descente de gradient, on obtient cette forme.

FIGURE 15.10 – Descente de gradient régularisée.

Eva Philippe  
et Robin Lemaire  
Applications  
réelles

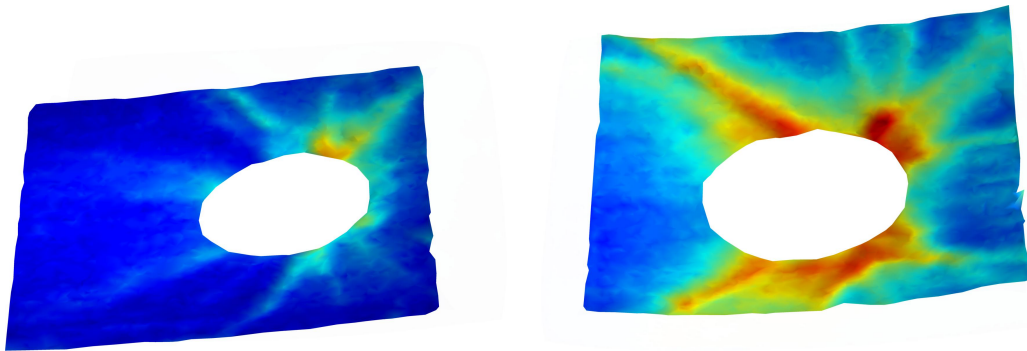


FIGURE 15.11 – Ces deux observations diffèrent en leur géométrie (taille et position du trou central, qui dénote le nerf optique). Pourtant le cadrage différent sera considéré comme étant une différence de forme.

efficace que mettre à la main des étiquettes sur les points pertinents pour réaliser un matching car elle prend en compte l'information globale (sur un cerveau par exemple, on ne peut pas reconnaître localement les régions à matcher). De plus, la construction d'atlas donne également de l'information pertinente. Néanmoins, on ne réussit pas à s'affranchir des biais dus aux manières différentes dont les données ont été collectées.

Troisième partie

Critiques et perspectives



# Chapitre 16

## Conséquences pratiques du choix des métriques à noyaux

Séance 12

[MMM12]

Manolis Perrot

### 16.1 Introduction.

Il est nécessaire de pouvoir comparer deux formes, dans le domaine médical, afin de détecter des pathologies par exemple. Pour ce faire, on considère nos formes comme des nuages de points, puis on définit une distance entre les deux formes comme le coût de la déformation minimale de l'une sur l'autre. On travaille donc avec des méthodes à noyau pour l'appariement de nuages de points (chapitre 9), mais en imposant que l'appariement soit exact. Dans ce chapitre, nous allons montrer comment calculer numériquement cette déformation minimale, qui revient à un calcul de géodésiques, puis nous verrons différents comportements des géodésiques en lien avec la courbure de cet espace de formes, lesquels peuvent biaiser l'interprétation des données.

### 16.2 Le modèle de déformation et son exploitation.

#### 16.2.1 Notations et rappels.

Nous allons regarder nos formes comme des ensembles landmarks, i.e. des éléments de  $\mathcal{L}^N(\mathbb{R}^D) = \{x^1, \dots, x^N : x^a \in \mathbb{R}^D\}$ . Soit  $V$  un RKHS de champs de vecteurs (applications de  $\mathbb{R}^D$  dans  $\mathbb{R}^D$ ), de produit scalaire  $\langle \cdot, \cdot \rangle$ , et de noyau scalaire  $k(x, y)$   $x, y \in \mathbb{R}^D$ .

*Remarque 16.1.* En toute rigueur,  $k$  doit être une application à valeurs dans  $\mathbb{R}^D$ , mais comme on le prendra de la forme  $f(x, y)I_D$ , on se restreint directement à un noyau scalaire. Prendre un noyau de cette forme signifie qu'il n'y a pas de direction de déformation privilégiée.

Typiquement, on aura  $k(x, y) = \gamma(\|x - y\|)$  ( $\gamma$  noyau de Gauss, Bessel, Cauchy...), donc on notera souvent  $k(x - y)$ . On note également l'ensemble des courbes de landmarks

$$Q := \{t \rightarrow q(t) = (q^1(t), \dots, q^N(t))\} = C^1([0, 1], \mathcal{L}^N(\mathbb{R}^D)). \quad (16.1)$$

Pour  $v \in L^2([0, 1], V)$ , le système dynamique :

$$\begin{cases} \dot{z} = v_t(z) \\ z(0) = x \end{cases} \quad (16.2)$$

Manolis Perrot admet une unique solution du type  $\psi(t, x)$ . Si on regarde tous les points obtenus au temps 1,  $v$  induit alors un difféomorphisme  $\varphi^v$  de  $\mathbb{R}^D$ , défini par :

$$\varphi^v(x) = \psi(1, x). \quad (16.3)$$

Ceci nous permet de définir une distance entre deux ensembles de landmarks  $X = (x^1, \dots, x^N)$  et  $Y = (y^1, \dots, y^N)$  comme :

$$d(X, Y) = \inf \left\{ E[v] = \int_0^1 \|v(t, \cdot)\|_V dt : v \in L^2([0, 1], V), \forall a \varphi^v(x^a) = y^a \right\}. \quad (16.4)$$

*Remarque 16.2.* Il n'y a pas de terme d'attache aux données dans cette énergie, contrairement à ce qui a été vu dans les chapitres antérieurs. Cela indique que pour notre modèle, on considère un appariement exact. Ainsi,  $\varphi^v$  est un difféomorphisme de  $\mathbb{R}^D$ , donc son injectivité implique qu'il ne peut y avoir de collision dans les trajectoires des points (car dans ce cas les trajectoires seraient confondues), et aussi que tous les landmarks sont distincts à tout temps.

Si on fixe un chemin  $q \in Q$ , et qu'on se place à  $t$ , cela définit  $N$  couples  $(q^1(t), \dot{q}^1(t)), \dots, (q^N(t), \dot{q}^N(t))$  que l'on peut interpoler, d'après un théorème vu au chapitre 9, par  $v_t^* \in V$ , de norme minimale. En regardant ce minimiseur pour tout  $t$ , on a donc construit un unique  $v^* \in L^2([0, 1], V)$  vérifiant  $\dot{q}(t) = v_t(q(t))$  et tel que  $\int_0^1 \|v(t, \cdot)\|_V^2 dt$  soit minimale, donné par :

$$v^*(t, x) := \sum_{a=1}^N p_a(t) k(x - q^a(t)), \quad (16.5)$$

où les moments  $p_a$  sont donnés par

$$p_a(t) = \sum_{b=1}^N \left[ K^{q(t)} \right]_{ab}^{-1} \dot{q}^b(t) \quad (16.6)$$

avec  $K^{q(t)} = [k(q^a(t) - q^b(t))]_{a,b=1,\dots,N}$ .

Ainsi on peut poser  $\tilde{E}[q] := E[v_{[q]}^*]$ . Nous allons maintenant calculer explicitement cette énergie en fonction de  $k$  et de  $q$ . La définition (16.5) de  $v^*$  donne :

$$\|v^*\|_V^2 = \sum_{a,b} \langle k(q^a, \cdot) p_a, k(q^b, \cdot) p_b \rangle_V, \quad (16.7)$$

or la définition d'un noyau reproduisant (scalaire) donne que

$$\forall v \in V \quad \forall x, p \in \mathbb{R}^D \quad \langle p, k(x, \cdot), v(\cdot) \rangle_V = \langle p, v(x) \rangle_{\mathbb{R}^D} \quad (16.8)$$

d'où

$$\|v^*\|_V^2 = \sum_{a,b} k(q^a, q^b) \langle p_a, p_b \rangle_{\mathbb{R}^D}, \quad (16.9)$$

ce qui s'écrit :

$$\|v^*\|_V^2 = \langle p, \mathbf{K}_q p \rangle_{\mathbb{R}^{ND}}, \quad (16.10)$$

si on note

$$p = \begin{pmatrix} p_1 \\ \vdots \\ p_N \end{pmatrix} \in (\mathbb{R}^N)^D \quad (16.11)$$



et

$$\mathbf{K}_q = \begin{pmatrix} K_q & & (0) \\ & \ddots & \\ (0) & & K_q \end{pmatrix} \quad (16.12)$$

$\mathbf{K}_q$  la matrice  $ND \times ND$  bloc diagonale, dont les blocs diagonaux sont  $K_q$ . Enfin, il suffit de remarquer que (16.6) se réécrit :

$$p = \mathbf{K}_q^{-1} \dot{q}, \quad (16.13)$$

ce qui donne finalement :

$$\|v^*\|_V^2 = \langle \dot{q}, \mathbf{g}_q \dot{q} \rangle_{\mathbb{R}^{ND}} \quad (16.14)$$

où  $\mathbf{g}_q = \mathbf{K}_q^{-1}$  est la métrique associée à la cométrique  $\mathbf{K}_q^{-1}$ .  $E$  est donc une énergie riemannienne, qui est minimale si  $q$  est une géodésique reliant  $I$  à  $I'$ . Trouver cette géodésique revient à résoudre le système hamiltonien :

$$\begin{cases} \dot{q}^a = \sum_{b=1}^N k(q^a - q^b) p_b \\ \dot{p}_a = - \sum_{b=1}^N \nabla k(q^a - q^b) \langle p_a, p_b \rangle \end{cases} \quad (16.15)$$

*Remarque 16.3.* Pour comprendre le comportement de nos points, on peut garder à l'esprit que les  $p_a$  sont les intentions de bouger de chaque point, et  $\dot{q}^a = \sum_{b=1}^N k(q^a - q^b) p_b$  indique que le déplacement réel de  $q^a$  dépend de l'intention de bouger de tous les autres points.

La réduction du problème à un calcul de géodésiques simplifie grandement la tâche, car on peut aisément implémenter leur calcul en intégrant le système hamiltonien (avec un schéma d'Euler par exemple).

## 16.2.2 Analyse statistique sur les espaces de formes.

Le modèle ci-dessus est notamment utilisé en analyse d'IRM. Dans ce domaine, le but est d'arriver à créer, à partir d'un jeu de données, des modèles pour un cerveau sain typique et pour un cerveau à un certain stade d'une pathologie, qui a modifié sa géométrie et sa structure (par exemple la maladie d'Alzheimer). L'objectif étant bien sûr de créer un logiciel de diagnostic. La démarche est la suivante : étant donné un jeu d'observations  $\mathcal{B} = \{J_1, \dots, J_M\} \subset \mathcal{L}^N(\mathbb{R}^D)$ , on calcule la moyenne intrinsèque, ou moyenne de Karcher (on remarquera l'analogie avec la méthode des moindres carrés) :

$$m := \operatorname{argmin}_{I \in \mathcal{I}} \sum_{J \in \mathcal{B}} d(I, J)^2. \quad (16.16)$$

Puis on détermine les vecteurs  $w_i \in T_m \mathcal{I}$  tels que  $\exp_m(w_i) = J_i$  (i.e. la géodésique partant de  $m$  avec une vitesse initiale  $w_i$  arrive en  $J_i$  au bout d'un temps 1). L'analyse statistique se déroule alors sur les  $w_i$ , c'est-à-dire qu'on a une interprétation des données dans un espace plat. Par conséquent, nous allons voir que les comportements des géodésiques et la courbure non nulle de l'espace des formes donnent des interprétations erronées des données dès que l'on s'éloigne un peu trop de la position initiale du tir. Ainsi l'analyse des données à partir des tirs géodésiques n'est pertinente que pour de faibles déformations, ce qui, fort heureusement, est le cas en pratique (par exemple pour des cerveaux, il ne faut qu'un faible recalage).

Manolis Perrot **16.3 Un cas typique pour comprendre la dynamique des landmarks.**

Nous allons nous intéresser à la dynamique du mouvement de landmarks dans un cas simplifié. Pour autant, il est porteur d'une grande richesse, car il permet de comprendre qualitativement les trajectoires avec lesquelles nous aurons affaire dans les sections suivantes. Considérons un jeu de landmarks  $q$  constitué trois groupes éloignés de points :

- un point isolé  $q_1$  ;
- trois points en triangle d'arrête  $d, q_2, q_3, q_4$  ;
- deux points proches à une distance  $d, q_5, q_6$  .

Si on pose  $a = k(d)$ , et que l'on considère que les points éloignés n'interagissent pas entre eux, la cométrie est alors :

$$K_q = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & a & a & 0 & 0 \\ 0 & a & 1 & a & 0 & 0 \\ 0 & a & a & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & a \\ 0 & 0 & 0 & 0 & a & 1 \end{pmatrix}. \tag{16.17}$$

**Analyse spectrale de la métrique** Afin de regarder quelles sont les vitesses privilégiées, il faut calculer l'inverse de cette matrice, ce qui revient à calculer l'inverse de chaque bloc. Le bloc suivant correspond à un regroupement de  $n$  points équidistants :

$$M = \begin{pmatrix} 1 & & & & \\ & 1 & & (a) & \\ & & 1 & & \\ & (a) & & 1 & \\ & & & & 1 \end{pmatrix}. \tag{16.18}$$

Un vecteur propre de  $M$  est  $(1) = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ , associé à la valeur propre  $1 + (n - 1)a$ .  $M$  étant diagonalisable, car symétrique, on peut l'écrire :

$$M = \sum \lambda_i e_i e_i^T, \tag{16.19}$$

avec  $(e_i)$  une famille orthonormée de vecteurs propres. Posons  $e = \frac{(1)}{\sqrt{n}}$ , on a alors :

$$M - (1 + (n - 1)a)ee^T = I_n(1 - a) + a(1)(1)^T - (1 + (n - 1)a)ee^T = (1 - a)(I_n - ee^T). \tag{16.20}$$

En vérifiant que  $1 - a$  est valeur propre, la décomposition de  $M$  est :

$$M = (1 + (n - 1)a) ee^T + (1 - a)(I_n - ee^T) \text{ et } g = M^{-1} = \frac{1}{1 + (n - 1)a} ee^T + \frac{1}{1 - a}(I_n - ee^T). \tag{16.21}$$

Ainsi, pour  $v$  un champ de vitesses sur nos landmarks, il est aisé de calculer son coût quadratique  $v^T M v$  en l'écrivant

$$v = e(e^T v) + (v - e(e^T v)) \tag{16.22}$$

$$= v_{\text{col}} + v_{\text{ncol}}, \tag{16.23}$$

où  $v_{\text{col}}$  est la composante *colinéaire* de  $v$ , et où le résiduel orthogonal  $v_{\text{ncol}}$  est de moyenne nulle. On trouve alors simplement

**Résultats limites** D'une part, lorsque les points sont très éloignés ( $d$  grand donc  $a \sim 0$ ), on a : et on retrouve que tous les points sont indépendants, il n'y a pas d'influence. En revanche, s'ils sont proches ( $d$  petit,  $a \sim 1$ ) : où la dernière écriture est un abus, et doit être compris par bloc. Pour être plus précis, il faut voir que le défaut de colinéarité sera très fortement pénalisé, ce qui incitera des points proches à bouger dans la même direction : à la limite, un vecteur vitesse de norme finie s'écrira nécessairement

$$v = (w_1, w_2, w_2, w_2, w_3, w_3)^T, \quad (16.24)$$

et sera pénalisé en

$$v^T g_q v = w_1^2 + \frac{1}{3}(w_2^2 + w_2^2 + w_2^2) + \frac{1}{2}(w_3^2 + w_3^2) \quad (16.25)$$

$$= w_1^2 + w_2^2 + w_3^2. \quad (16.26)$$

Un groupe de  $n$  points très proches se déplaceront donc comme un seul ! On retrouve là notre interprétation "mécanique des fluides" des métriques de déformations à noyaux, où le coût d'une déformation est assimilé à l'effort nécessaire à un opérateur extérieur pour "déformer la nappe élastique" de l'espace ambiant, sans considération pour ce qui est dessiné ici.

Les outils que nous avons développés dans la première section ont un aspect pratique indéniable, et apparaissent adéquats à nos problèmes d'appariement. Cependant, la pénalisation introduite induit des comportements indésirés lors du transport des points, que nous allons maintenant décrire.

Manolis Perrot

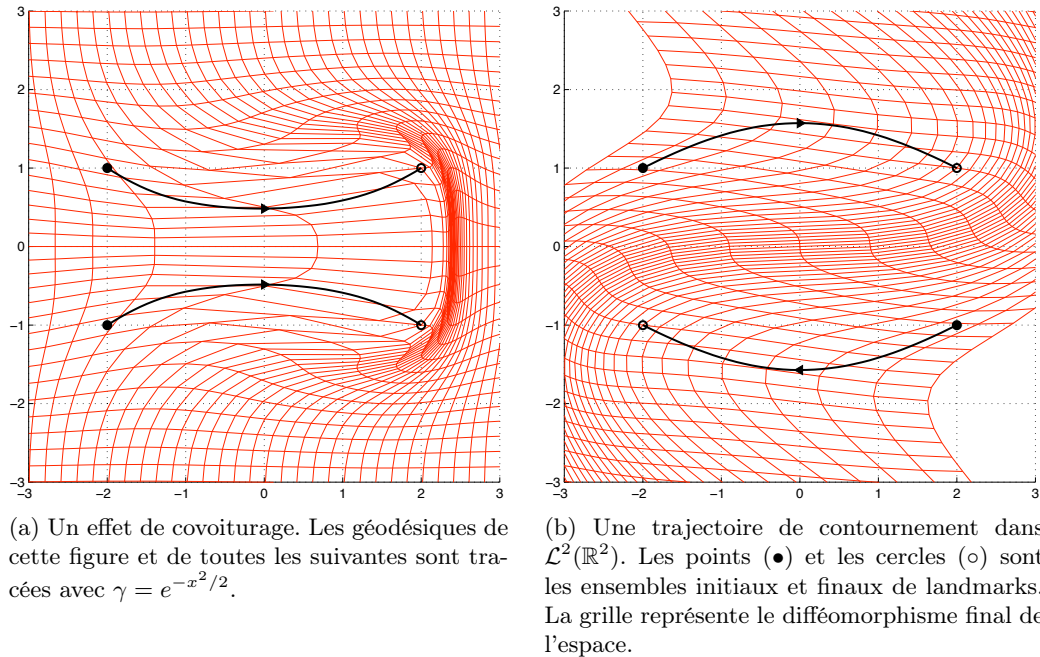


FIGURE 16.1 – Au sens des métriques à noyaux, les trajectoire optimales ne sont pas toujours celles auxquelles on s'attendrait de prime abord. C'est qu'en plus de pénaliser les déchirures, ces métriques induisent un effet "de covoiturage", de mutualisation du coût de transport entre particules proches.

## 16.4 Intuitions sur les géodésiques et comportements problématiques.

Nous allons voir dans cette section une conséquence directe du fait que les trajectoires de moindre coût sont celles qui ont tendances à faire bouger les points proches en bloc.

### 16.4.1 Compression des géodésiques, effets de "covoiturage".

Nous venons de voir que le déplacement de  $n$  points proches dans une même direction coûtait autant que le déplacement d'un seul point. Ainsi les points assez proches qui désirent aller dans une direction similaire vont avoir tendance à se rapprocher pour abaisser le coût du trajet : c'est le principe du covoiturage. La figure 16.1a montre cet effet. Out-of-the-box, les métriques à noyaux ne génèrent donc pas de *translations*, et auront tendance à comprimer puis dilater les objets pour les transporter sur de longues distances.

### 16.4.2 Phénomène de contournement.

La figure 16.1b nous montre comment deux points qui désirent se déplacer en sens opposé doivent s'éviter afin de parcourir un chemin de moindre coût (en effet, deux vitesses effectives opposées pour des points corrélés, i.e. proches, sont fortement pénalisées dans notre modèle). Encore une fois, ce n'est pas vraiment souhaitable.

Les deux exemples ci-dessus nous montrent que dans certain cas, le modèle exclut les simples translations de points, car elles ne sont pas optimales au sens de notre modèle.

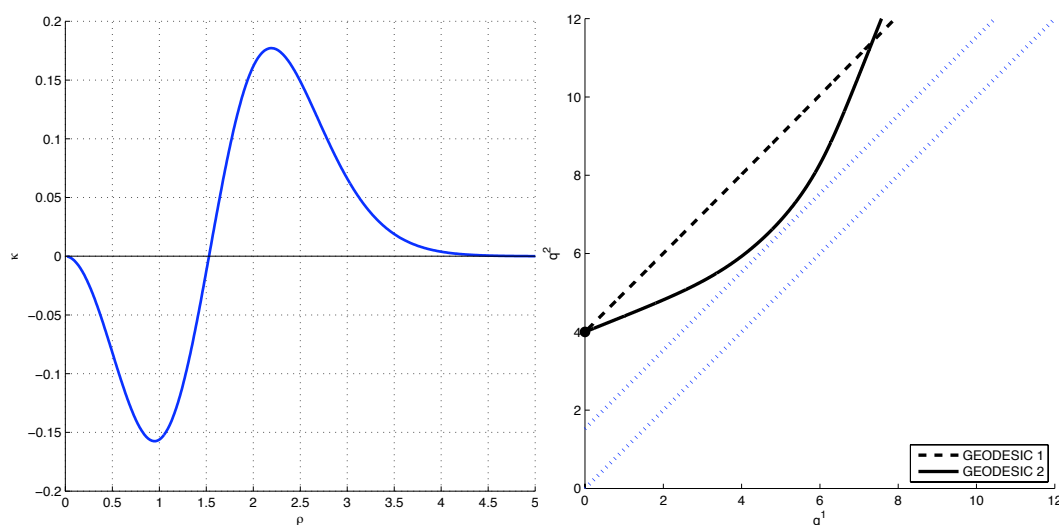


FIGURE 16.2 – A gauche : courbure sectionnelle  $\kappa$  pour  $\mathcal{L}^2(\mathbb{R}^1)$ . A droite : deux trajectoires dans  $\mathcal{L}^2(\mathbb{R}^1)$  dans le plan  $(q^1, q^2)$ . Les deux géodésiques ont la même origine, et restent dans la région où  $\kappa > 0$  (la ligne supérieure en pointillés bleus indique  $\kappa = 0$ ) ; elles ont comme moments initiaux  $(p_1, p_2) = (1, 1)$  et  $(p_1, p_2) = (1, 0.4)$ .

## 16.5 Courbure positive et points conjugués.

### 16.5.1 Étude du cas $\mathcal{L}^2(\mathbb{R}^1)$ .

Dans le cas de deux points sur une droite, on peut calculer explicitement la courbure sectionnelle, qui ne dépend que de la distance  $\rho = |x - y|$  entre les points :

$$\kappa = \frac{\gamma_0 - \gamma(\rho)}{\gamma_0 + \gamma(\rho)} \gamma''(\rho) - \frac{2\gamma_0 - \gamma(\rho)}{(\gamma_0 + \gamma(\rho))^2} (\gamma(\rho))^2. \quad (16.27)$$

*Remarque 16.4.* Voilà comment on peut interpréter le signe de la courbure. Une courbure négative provoque une divergence des géodésiques issues d'un même point, alors qu'une courbure positive se traduit par un croisement ultérieur, et donc l'apparition de *points conjugués* (i.e. des points qui peuvent être reliés par plus d'une géodésique).

Sur la figure 16.2, on remarque que la courbure est négative lorsque les points se rapprochent, et est positive pour des plus grandes valeurs de  $\rho$ , puis que la variété "s'aplatit" quand  $\rho \rightarrow \infty$  (les points ne sont plus corrélés). La courbure positive induit notamment l'existence de points conjugués.

Essayons de comprendre les deux géodésiques présentées sur la figure 16.2. Supposons que les deux points veuillent se déplacer d'une distance fixée vers la droite. S'ils sont assez éloignés, ils peuvent se déplacer plus ou moins indépendamment (c'est la géodésique 1). Mais si (i) le point de derrière accélère tandis que celui de devant ralentit, alors (ii) une fois que les points sont proches, ils se déplacent en tandem en utilisant moins d'énergie car ils sont proches et finalement (iii) le point de derrière ralentit et celui de devant accélère lorsqu'ils sont vraiment très proches (géodésique 2).

Manolis Perrot

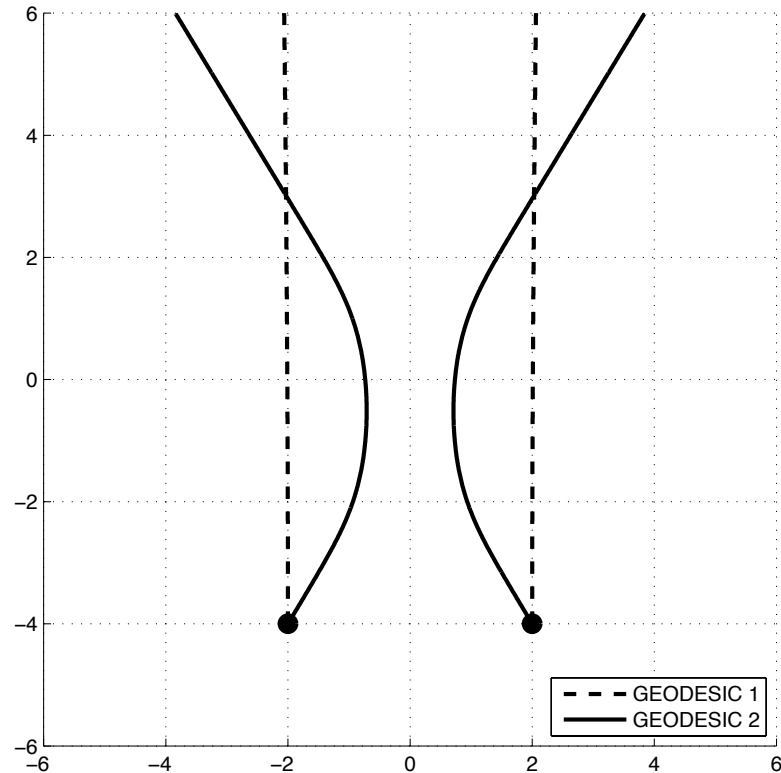


FIGURE 16.3 – Points conjugués dans  $\mathcal{L}^2(\mathbb{R}^2)$ .

*Remarque 16.5.* Cette figure nous montre l'existence de *points conjugués*. L'exemple le plus simple de tels points se voit sur la sphère; en partant d'un pôle, on peut rejoindre l'autre par n'importe quel grand cercle. Ce phénomène pose un problème pour l'analyse des données. En effet, rappelons qu'on procède à cette analyse en regardant les  $w_i$ , vecteurs de tirs initiaux. Ainsi, deux vecteurs différents (par exemple opposés dans le cas de la sphère), et donc traités comme tels, aboutiront à la même forme. On voit ici l'importance de faire uniquement de faibles déformations, donc de "petits tirs", pour éviter ce genre d'interprétations erronées.

### 16.5.2 Points conjugués dans le cas $\mathcal{L}^2(\mathbb{R}^2)$ .

La figure 16.3 met également en évidence l'existence de points conjugués dans  $\mathbb{R}^2$ . Ces deux exemples permettent d'extrapoler l'existence de comportements similaires pour un nombre de points et des dimensions supérieures.

## 16.6 Courbure négative et comportements chaotiques.

Le calcul explicite de la courbure sectionnelle à partir de la cométrie  $\mathbb{K}_q$  montre l'existence de points où elle est négative. Ceci entraîne l'existence de comportements chaotiques à plus ou moins long terme.

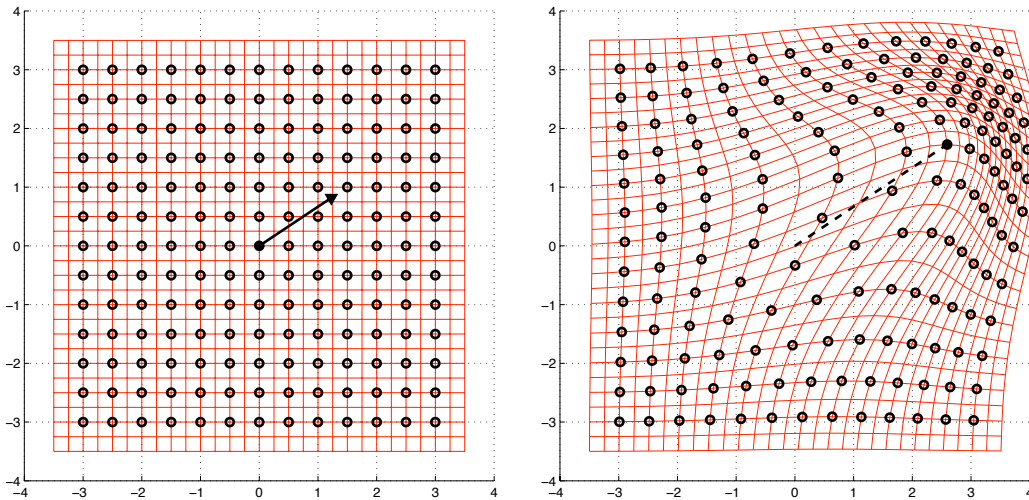


FIGURE 16.4 – Effet de charriage avec un noyau gaussien, pour un seul porteur de moment non nul  $q^1$ . A gauche : configuration initiale. A droite : configuration à  $t = 1$ . La courbure négative s’interprète aisément à partir de cette figure. Considérons en effet le triangle  $(AOB)$ , où  $A$  est la configuration de droite,  $O$  (pour origine) celle de gauche, et où  $B$  est simplement obtenue comme  $A$  à partir de  $O$ , mais avec une direction de tir différente. Les côtés  $[OA]$  et  $[OB]$  du triangle sont donc les tirs géodésiques décrits ci-dessus, avec effet de charriage. Maintenant, quelle est la géodésique  $[AB]$ , le troisième côté du triangle ? Pour passer d’une grille déformée à une autre, les points “comprimés” doivent prendre des trajectoires différentes les unes des autres. Pour éviter de payer le coût de “déchirure”, la trajectoire optimale aura donc tout intérêt à détendre la grille, à espacer les points comprimés en revenant de  $A$  sur  $O$ , avant de les redéformer sur  $B$ . Autrement dit, le troisième côté  $[AB]$  est très proche de la trajectoire  $[AO] + [OB]$ , le triangle est *fin, pincé* : c’est bien ce que nous avons appris à associer à une courbure négative.

### 16.6.1 Déplacement d’un seul point, effet de charriage.

Imaginons que l’on désire déplacer qu’un seul point  $q_1$ . C’est un cas où la courbure est toujours négative. Le point  $q^1$  qui porte le moment se déplace en ligne droite à vitesse constante, tandis que les points  $q^a$  ( $a > 1$ ) sont portés par le mouvement général que le déplacement de  $q^1$  cause et se déplacent à la vitesse  $\dot{q}^a = k(q^a, q^1)p_1$ , qui est parallèle à  $p_1$  (mais pas constante). On voit bien que plus les points sont proches, plus ils sont charriés fortement par  $q^1$ . La courbure négative implique la divergence rapide des géodésiques, et donc un comportement chaotique. En effet, si l’on imagine que l’on modifie légèrement la direction de  $p_1$ , la configuration finale des landmarks (à  $t=1$ ) sera très différente de celle générée par le moment de référence  $p_1$ .

*Remarque 16.6.* Il faut garder en tête que nos méthodes sont faites pour être calculées numériquement, donc avec une précision finie (et même faible, par exemple on prend pour un schéma d’Euler  $dt = 0.1$ ). On désire donc qu’une faible erreur sur la connaissance des conditions initiales donne tout de même les bonnes formes. Cette sensibilité aux conditions initiales de tir n’est pas souhaitable, car cela signifie aussi que deux vecteurs  $w_i$  proches seront interprétés comme très corrélés, alors qu’ils donnent en réalité des déformations très différentes.

En modifiant légèrement les rapports entre certaines composantes des moments et l’énergie totales, on peut même exhiber des comportements totalement désordonnés, comme l’illustre la figure 16.5.

Manolis Perrot

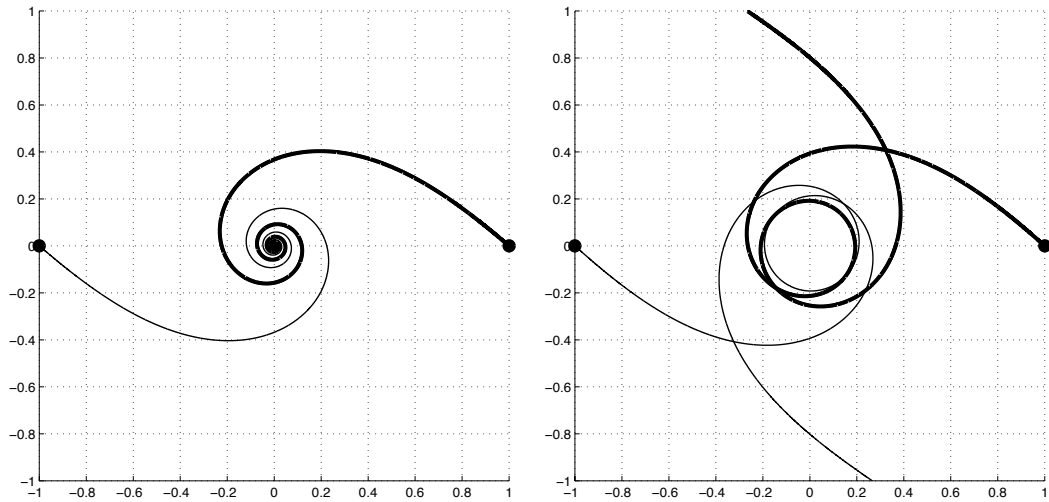


FIGURE 16.5 – Trajectoires convergentes et divergentes de deux landmarks. Remarquons que les tracés sont effectués pour  $(q^1(0), q^2(0)) = ((0, -1), (0, 1))$ ,  $(p_1(0), p_2(0)) = ((-10, 9), (10, 9))$  à gauche, et  $(p_1(0), p_2(0)) = ((-10, 8.6), (10, 8.6))$  à droite.

### 16.6.2 Conclusion

On l’a vu, le cadre LDDMM + métriques à noyaux est, au sens de la complexité algorithmique, l’une des plus simples structures riemannienne “sans déchirure” dont on peut munir un espace de landmarks : on sait depuis Hamilton qu’un tir géodésique efficace reposera sur la cométrie  $K_q$ , et quelle manière plus efficace d’associer une matrice à un nuage de point que la formule totalement parallèle

$$K_{q,ij} = k(x_i, x_j) = \exp(-\|x_i - x_j\|^2 / 2\sigma^2) \quad ? \quad (16.28)$$

Cette simplicité algorithmique est la grande raison de la popularité des métrique “à noyaux”. Il ne faut néanmoins pas nous leurrer sur leur pertinence : si les résultats obtenus sont crédibles pour des déformations d’une taille inférieure au diamètre des formes, la séance d’aujourd’hui aura eu le mérite de montrer le manque de fiabilité/pertinence de ces métriques pour des trajectoires de tir plus longues.

Un sujet de recherche actuel serait donc de chercher à dépasser la théorie présentée lors de la séance 9, pour combiner les propriétés désirables des métriques de Wasserstein (i.e. transport optimal, qui se comportent très bien vis à vis des translations, mais n’imposent aucune corrélation aux particules voisines) et des métriques à noyaux (qui permettent justement d’imposer – à un coût raisonnable – des a priori de régularité spatiale sur le plan de déformation).



# Chapitre 17

## Remise en question du cadre d'étude riemannien

*Séance 13*

*[LP13], [Ash07]*

*Nicolas Masson*

### 17.1 Introduction

Au cours des séances précédentes, nous nous sommes penchés sur des techniques de recalage d'image LDDMM (Large Deformation Diffeomorphic Metric Mapping) utilisant la géométrie Riemannienne, dans le but de construire un atlas permettant une analyse statistique sur des données biologiques. Cependant, malgré la pertinence et le bien-fondé de ces méthodes, celles-ci sont très coûteuses, voire à la limite du praticable dans certains contextes. De plus, si l'on s'est jusqu'ici concentrés sur une application en particulier, l'intérêt du *matching* ne se limite pas à l'étude de population. Le problème du recalage concerne aussi bien l'encodage d'une vidéo que le suivi individuel d'un patient ; il faut donc bien avoir à l'esprit que les exigences varient considérablement selon l'utilisation que l'on cherche à en faire. En ce sens, il est légitime de s'interroger sur la nécessité des méthodes LDDMM, et de se demander si d'autres approches ne pourraient pas nous fournir un meilleur rapport efficacité/coût que l'approche Riemannienne dans certains cas. L'objet de cette séance sera donc d'étudier une idée "concurrente" : le *matching* SVF (pour Stationary Velocity Field), et de le comparer aux procédés LDDMM. On commencera par une brève introduction aux groupes de Lie, qui constituent le cadre mathématique adapté pour la méthode SVF, avant d'exposer le principe de ladite méthode, pour enfin donner quelques éléments de comparaison avec le point de vue Riemannien.

### 17.2 Éléments de théorie des groupes de Lie

Un groupe de Lie  $\mathbb{G}$  est une variété lisse munie d'un élément neutre  $id$ , d'une loi de composition lisse  $(g, h) \in \mathbb{G} \times \mathbb{G} \rightarrow gh$  et d'une loi d'inversion également lisse  $g \rightarrow g^{-1}$ , toutes deux compatibles avec la structure de variété différentielle. On appellera champ de vecteurs toute application lisse  $X$  sur  $\mathbb{G}$  telle que  $\forall g \in \mathbb{G}, X(g) \in T_g \mathbb{G}$ , que l'on peut identifier à la dérivée directionnelle d'une fonction scalaire  $\varphi$  le long du champ :  $\partial_X \varphi(g) = \frac{\partial \varphi(g+tX(g))}{\partial t}$ . On définit le crochet de Lie entre deux champs  $X$  et  $Y$  comme suit :

$$[X, Y]\varphi = \partial_X \partial_Y \varphi - \partial_Y \partial_X \varphi$$

Notons que celui-ci qui agit comme une multiplication interne sur l'algèbre des champs de

Nicolas Masson vecteurs. Si  $a \in \mathbb{G}$ , on définit la translation à gauche  $L_a : g \rightarrow ag$ . On dit qu'un champ de vecteurs est invariant à gauche s'il reste inchangé sous l'action de la différentielle de la translation, au sens où :  $dL_a(g).X(g) = X(ag)$ , c'est-à-dire  $(L_a)_*X = X$  en reprenant les notations du cours. L'ensemble des champs de vecteurs invariants à gauche, noté  $\mathcal{G}$ , forme une sous-algèbre de Lie de l'espace des champs de vecteurs. Dans la mesure où un champ invariant à gauche est entièrement déterminé par sa valeur en  $id$ , on pourra identifier  $\mathcal{G}$  à  $T_{id}\mathbb{G}$  grâce à la carte  $x \in T_{id}\mathbb{G} \rightarrow X \in \mathcal{G}$  défini par  $X(g) = dL_g(id).x$ . Par conséquent, on notera indifféremment  $X$  un champ de vecteur invariant à gauche ou un élément de  $T_{id}\mathbb{G}$ . On remarque que les éléments de  $\mathcal{G}$  sont complets, au sens où leur flot est défini en tout temps, et que celui-ci vérifie  $\varphi_X^t(g) = g\varphi_X^t(id)$  par invariance à gauche. Ces propriétés nous permettent de définir l'exponentielle de Lie comme suit :

$$\begin{aligned} \exp & : \mathcal{G} & \rightarrow & \mathbb{G} \\ & X & \rightarrow & \varphi_X^1(id) \end{aligned}$$

En particulier, l'exponentielle permet de définir le sous-groupe à un paramètre associé à  $X$  comme le morphisme :

$$\begin{aligned} \mathbb{R} & \rightarrow & \mathbb{G} \\ t & \rightarrow & \exp(tX) \end{aligned}$$

qui vérifie  $\varphi_X^t = \exp(tX)$ . Ainsi, pour tout vecteur tangent  $X \in T_{id}\mathbb{G}$ , le sous-groupe à un paramètre qui lui est associé est une courbe qui part de l'identité avec  $X$  pour vecteur tangent initial.

Il faut bien le reconnaître, la théorie est au premier abord un peu indigeste pour le lecteur profane, mais l'utilisation que nous en feront est à la fois limitée et illustrative.

### 17.3 La méthode SVF

Le principe de la méthode est assez aisé à comprendre compte tenu des séances précédentes, dans la mesure où elle apparaît (du moins au début) comme une simplification des méthodes LDDMM. En effet, les difféomorphismes que l'on avait utilisés pour le *matching* étaient obtenus en intégrant une équation différentielle à champ de vitesse variable. Ici, l'idée est simplement d'utiliser des champs ne dépendant pas du temps.

Avant de continuer, il convient tout d'abord de questionner la pertinence de ce choix. Il n'est pas aisé de le justifier a priori, dans la mesure où le procédé SVF semble au premier abord clairement moins adapté que le LDDMM pour l'imagerie médicale. Prenons l'exemple du suivi longitudinal d'un patient : on réalise de manière régulière et assez espacée dans le temps plusieurs images du cerveau (par exemple) d'un patient, et on cherche à avoir un indicateur révélant le développement d'une pathologie. Pour ce faire, on effectue un *matching*, et on regarde le champ de vitesses qui a engendré ce recalage. Ainsi, les zones où les champs de vecteurs sont non-négligeables nous renseignent sur ce qui a bougé et nous disent donc si une pathologie s'est développée. Cependant, la contrainte de constance dans le procédé SVF empêche le champ de vecteurs d'être localisé sur ce qui bouge, d'où une perte d'information, contrairement à la méthode LDDMM qui va permettre au champ de "suivre" les particules mouvantes. En fait, le problème apparaît lorsque les déformations sont trop grandes, et seule une expérience nous permettra de conclure sur ce point.

Même si l'on accepte pour un temps la pertinence de l'approche pour les petites déformations, le procédé apparaît à première vue comme un appauvrissement du LDDMM. Cependant, le changement n'est pas des moindres, car il permet d'utiliser la théorie des groupes de Lie, le flot d'une ODE (équation différentielle autonome) étant en fait un sous-groupe à un paramètre d'un certain groupe de Lie, ce qui n'est pas du tout le cas des flots dans le cadre Riemannien. Ce qui fait tout l'intérêt du cadre Lie est que le *matching* est considérablement plus rapide à la fois au niveau de la recherche du meilleur champ de vitesses que du calcul de la déformation.

### 17.3.1 Calcul de l'exponentielle : le *scaling and squaring algorithm*

Si le cadre Lie permet autant de rapidité pour l'intégration du flot, c'est parce que la théorie remet la composition à l'honneur au sens suivant : si  $X$  est un vecteur tangent, on a, pour  $n$  suffisamment grand :

$$\exp(X) = \left( \exp\left(\frac{X}{2^n}\right) \right)^{2^n} \simeq \left( id + \frac{X}{2^n} \right)^{2^n} \quad (17.1)$$

Ainsi, au lieu d'intégrer un flot à proprement parler, ce qui en terme de complexité est forcément linéaire en le temps, on compose simplement 8 fois un procédé peu coûteux (on prend typiquement  $n = 256$ ) !

### 17.3.2 Recherche du champ de vitesses optimal et formule BCH

Dans les séances précédentes, on a vu comment trouver le champ de vitesses optimal : on tirait une première fois avec un moment  $p_0$  pour arriver en  $q_0$ . Le gradient de l'attache aux données nous donnait un "désir de mouvement"  $\delta q_0$  vers notre cible, que l'on transformait en "désir de mouvement"  $\delta p_0$  sur le moment, avant de tirer à nouveau avec un moment  $p_0 + \delta p_0$  et de réitérer le processus. Cependant, cette transformation d'un "désir" en un autre est loin d'être évidente, et consiste dans ses grandes lignes en une intégration inverse du flot, avec un coût conséquent.

Ici, le principe est le même, mais la théorie des groupes de Lie nous aide beaucoup en nous donnant la formule Baker-Campbell-Hausdorff (BCH) :

$$\log(\exp(X) \circ \exp(\delta X)) = X + \delta X + \frac{1}{2}[X, \delta X] + \frac{1}{12}[X, [X, \delta X]] \dots \quad (17.2)$$

Il est assez clair que calculer une somme est nettement plus rapide que d'intégrer un flot, d'où une efficacité nettement accrue en comparaison au LDDMM.

Maintenant qu'il est établi que la méthode SVF dépassait largement sa concurrente en terme de rapidité, il reste à déterminer au moyen d'expériences dans quels contexte l'approche Lie trouve sa pertinence.

Nicolas Masson

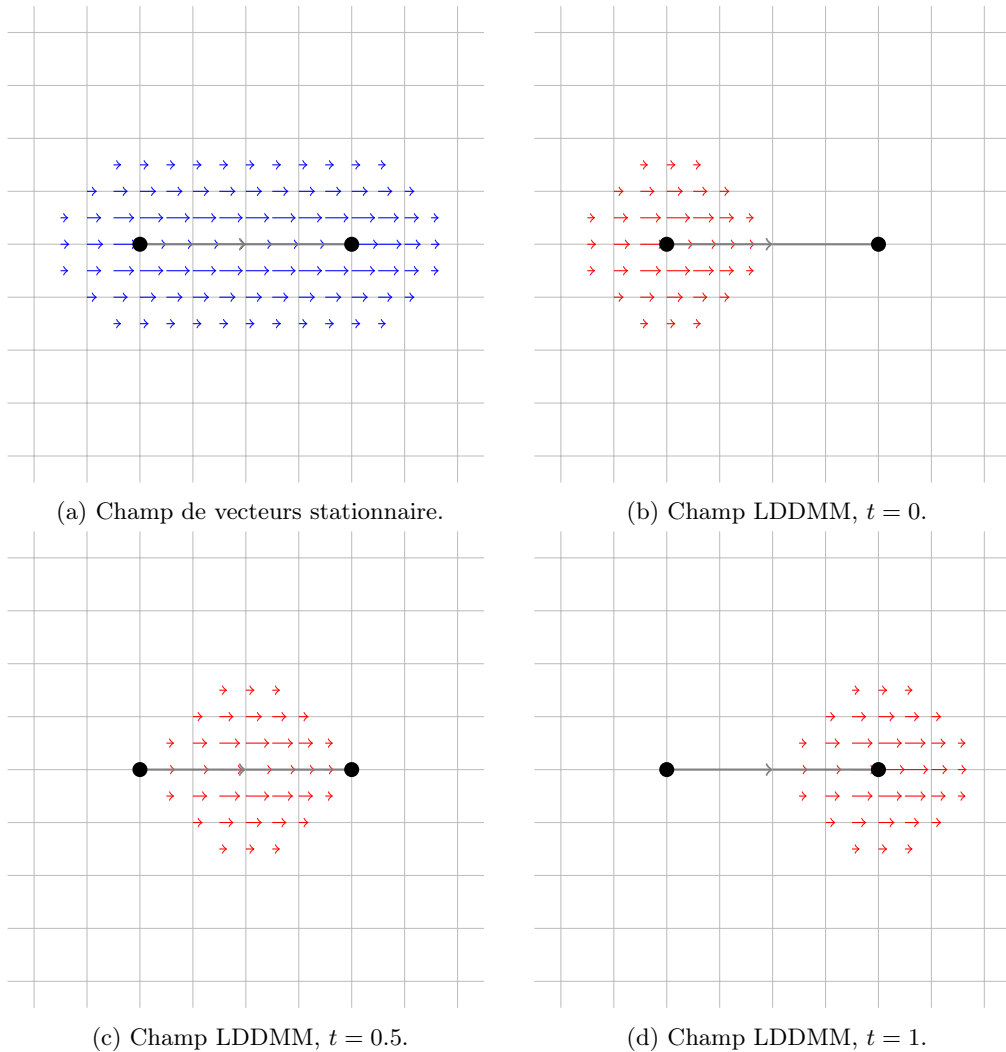


FIGURE 17.1 – Le cadre “Champs de Vecteurs Stationnaires” et la théorie LDDMM n’ont pas la même manière d’interpréter le matching du point de gauche sur le point de droite. Là où le premier repose sur la théorie des groupes de Lie, et construit la déformation par intégration d’un champ de vecteurs stationnaire au cours du temps, le second s’appuie sur une notion de “transport à coût minimal” qui sera concentré autour des particules en mouvement (principe de *réduction*). Dans les deux cas, un a priori de régularité gaussien est utilisé, que ce soit sous la forme d’une simple pénalisation ou d’une métrique RKHS.

Si la théorie des champs de vecteurs stationnaires permet d’obtenir des transports difféomorphiques à moindre coût, elle souffre de sévères limitations pour l’analyse : l’extrapolation (i.e. laisser filer les trajectoires jusqu’à  $t = 1.5$ ,  $t = 2$ ) sera généralement médiocre, et l’information peu localisée. Là où le champ LDDMM pris en  $t = 0$  (peu ou prou le moment de tir initial) permet de bien voir que c’est la particule de droite qui est en mouvement (et pas le reste de l’image), cette trajectoire a complètement “bavé” dans l’espace avec la théorie “groupe de Lie”. Sur une image complexe, cette diffusion de l’information le long des trajectoires peut grandement compliquer l’analyse, qui pourrait se résumer à prendre la norme  $L^2$  du moment de tir initial dans le cas Riemannien.

Cette figure est l’exact analogue de la Figure 17.3.

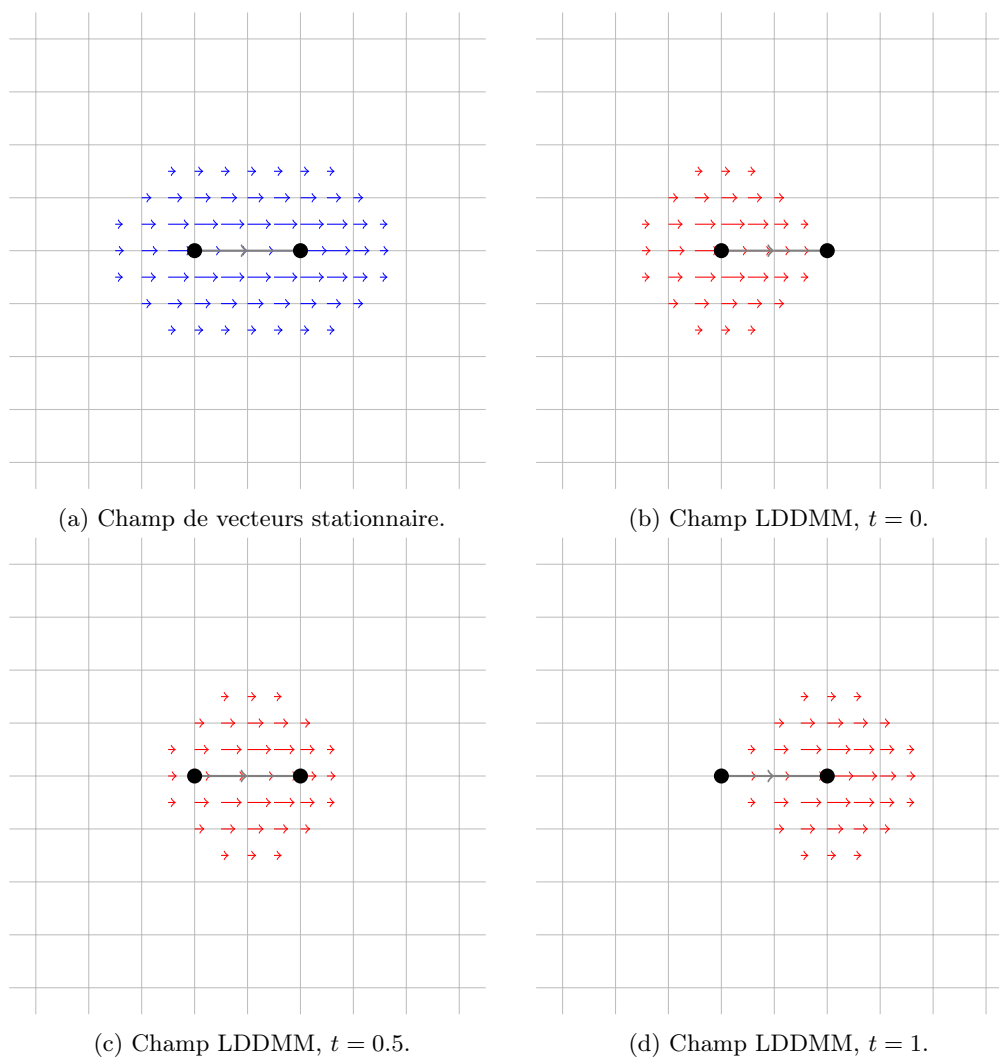


FIGURE 17.2 – Pour un déplacement plus petit par rapport à l'échelle de régularisation, ce problème de dilution de l'information est moins gênant. Avec un coût numérique bien inférieur, on aura donc tendance à privilégier les méthodes "Lie" pour du matching intra-sujet où aucune extrapolation sur le long terme n'est nécessaire. Pensons par exemple à une application de suivi longitudinal, avec comparaison de scans IRM d'un même sujet âgé/dément pris à un an d'intervalle : une forte norme du champ de déformation autour de l'hippocampe est un indicateur anatomique pertinent (souvent corrélé à la maladie d'Alzheimer), que l'on peut analyser automatiquement et faire remonter au radiologue.

Cette figure est l'exacte analogue de la Figure 17.4.

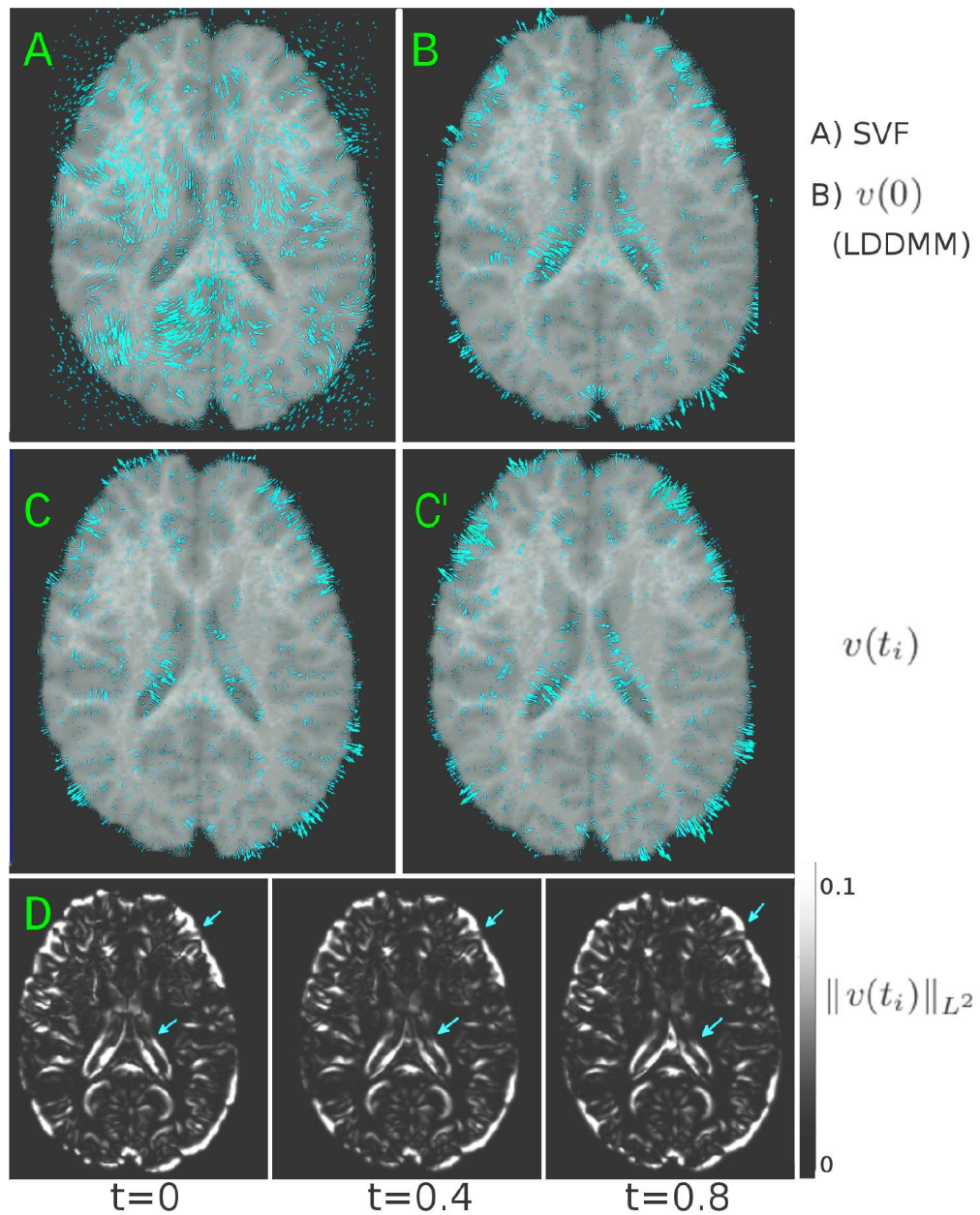


FIGURE 17.3 – Comparaison inter-sujet. En bleu : champ de vitesses optimal pour le matching avec les différentes méthodes. En A), SVF, en B) C) C'), LDDMM respectivement aux temps 0, 0.4, et 0.8. En D), magnitude associée mesurée avec la norme  $L^2$ .



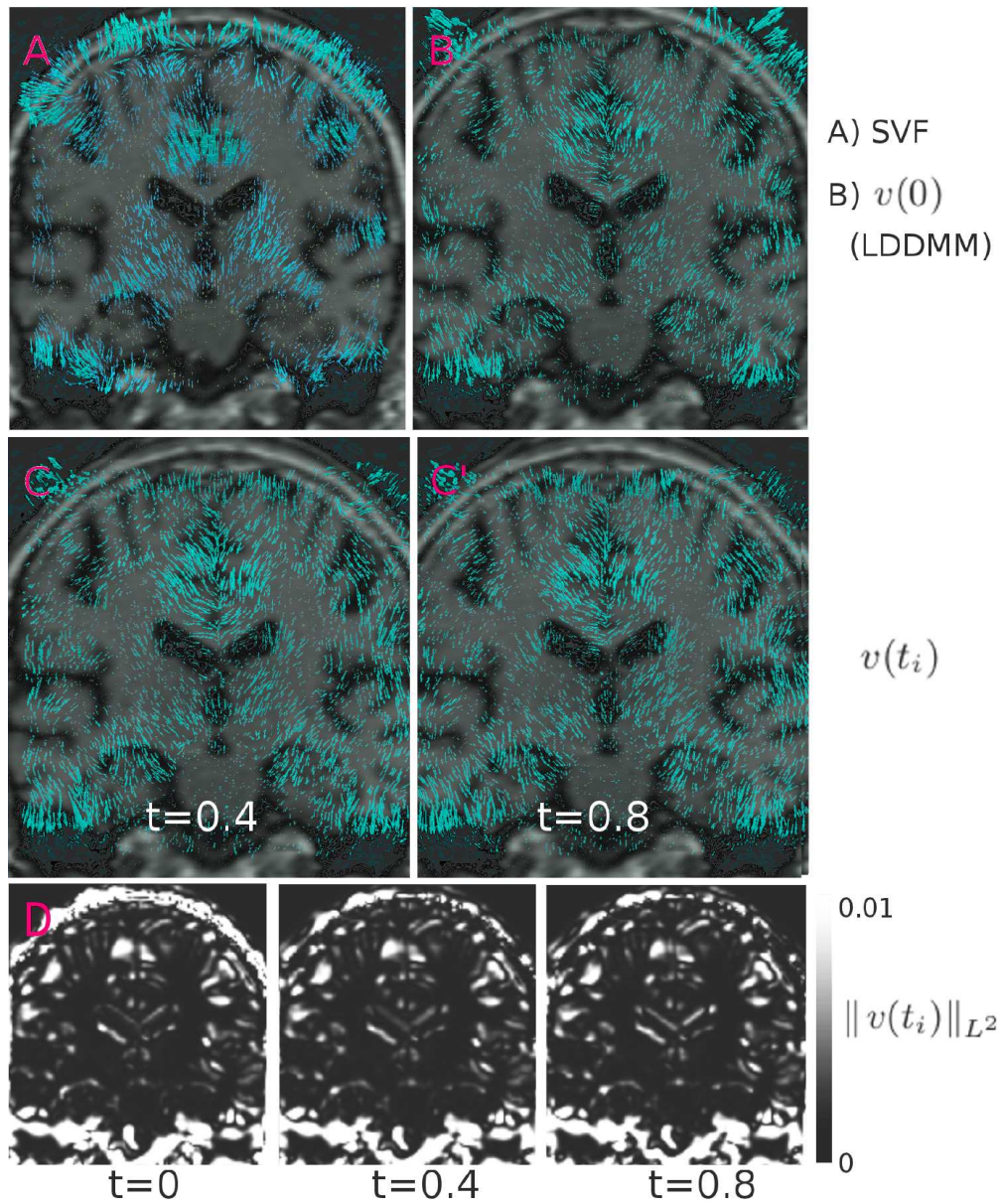


FIGURE 17.4 – Suivi longitudinal intra-sujet. En bleu : champ de vitesses optimal pour le matching avec les différentes méthodes. En A), SVF, en B) C) C'), LDDMM respectivement aux temps 0, 0.4, et 0.8. En D), magnitude associée mesurée avec la norme  $L^2$ .

Nicolas Masson **17.4 Récapitulatif des notions vues dans les théories Riemanniennes et Lie**

Notion	Riemann-LDDMM	Lie-SVF
Espace étudié Qui agit sur :	$\text{Diff}(\mathbb{R}^d)$ des formes $q$ – des landmarks.	$\text{Diff}(\mathbb{R}^d)$ des images $I$ .
Outil élémentaire :	Métrique RKHS sur un espace $V \subset (\mathbb{R}^d \rightarrow \mathbb{R}^d)$ , à <b>choisir de manière biologiquement pertinente</b> .	Opération de composition $\circ$ , fixée par les mathématiques.
Trajectoires recherchées :	Courbes géodésiques, qui minimisent la longueur.	Sous-groupes à un paramètre, morphisms de $(\mathbb{R}, +)$ dans $(\text{Diff}, \circ)$ .
Résultat théorique fondamental :	Description Hamiltonienne des géodésiques, qui sont uniquement déterminées par les conditions de Cauchy position-moment initiaux.	Les sous-groupes à un paramètre de $\text{Diff}$ sont les $t \mapsto \exp(tX)$ , au sens de l'intégration des flots.
Principe de réduction :	La déformation optimale se concentre autour des particules en mouvement.	Le champ optimal est concentré autour des trajectoires.
Représentant :	Moment de tir initial $p_0$ , porté par les particules.	Flot stationnaire $X$ , sur l'espace ambiant.
Critère à minimiser :	$\frac{\gamma_{\text{reg}}}{2} p_0^T K_{q_0} p_0 + \frac{\gamma_{\text{att}}}{2} \text{Att}(q_1, \tilde{q})$	$\frac{\gamma_{\text{reg}}}{2} \ X\ _{\text{reg}}^2 + \frac{\gamma_{\text{att}}}{2} \text{Att}(\varphi \cdot I_0, I_1)$
Routine forward, exp :	Shoot Hamiltonien, $p_0 \mapsto q_1$ ,	Exponentiation rapide, $X \mapsto \varphi$ .
Routine backward, ( $d_{p_0/X} \exp$ ) <sup>*</sup> :	Flot rétrograde des équations adjointes, $\partial_{q_1} C \mapsto \partial_{p_0} C$ .	Basée sur la formule BCH, $\partial_\varphi C \mapsto \partial_X C$ .
Complexité d'une itération :	Linéaire en la précision temporelle $1/dt$ .	<b>Logarithmique</b> en la précision temporelle $1/dt$ .

Le cadre le plus "crédible" biologiquement est celui de la géométrie Riemannienne, qui permet d'encoder dans la *métrique* un a priori anatomique. En délaissant ce degré de liberté "géométrique" pour gagner en structure *algébrique* comme le fait la méthode SVF, le gain algorithmique est considérable : on passe du linéaire au logarithmique en la précision temporelle pour le matching d'images pleines – en opposition au recalage de nuages de points épars, de surface de codimension 1 pour lesquelles stocker un flot sur l'espace ambiant est un gros gachis.



## 17.5 Comparaison expérimentale

Les expériences qui feront l'objet de notre étude ont été réalisées dans deux contextes très différents : la comparaison inter-sujet et le suivi longitudinal intra-sujet. Le premier cas correspond à ce que l'on a étudié lors des dernières séances du cours : on étudie plusieurs sujets que l'on cherche à comparer (par exemple afin de construire un atlas), impliquant des déformations relativement importantes. Dans le second, l'objectif est celui qui a été un peu plus haut : étudier un même individu au cours du temps afin d'observer son évolution, ce qui nécessite des déformations beaucoup plus faibles.

### 17.5.1 Suivi longitudinal intra-sujet

Pour le suivi longitudinal intra-sujet, les résultats expérimentaux sont présentés Figure 17.4.

On constate que dans ce cas, les résultats sont assez similaires avec les méthodes SVF et LDDMM : le champ de vitesses est localisé à peu près aux mêmes endroits, nous procurant les mêmes informations. De plus, on constate également que le champ de vitesses dans le cas LDDMM ne présente pas de variations pertinentes. Les résultats sont donc très similaires, mais avec des coûts incomparables !

### 17.5.2 Comparaison inter-sujet

Pour la comparaison inter-sujet, les résultats expérimentaux sont présentés Figure 17.3.

La situation est ici radicalement différente : les champs de vitesse ne sont pas du tout localisés au même endroit, synonyme d'une grande perte d'information et d'un manque de pertinence pour la méthode SVF. Ce n'est si étonnant : comme on l'a remarqué précédemment, la pertinence de la méthode SVF repose sur de petites déformations, et il y a bien évidemment une différence beaucoup plus importante entre deux sujets qu'entre un sujet à deux instants différents.

## 17.6 Conclusion

Le champ d'applications du recalage d'image est vaste, et les méthodes utilisées en sont donc d'autant plus variées. Les exigences différant considérablement entre deux applications, il est nécessaire de bien savoir dans quel contexte on se trouve pour évaluer l'adéquation d'une méthode. Ainsi, l'approche Riemannienne, bien que très séduisante et relativement efficace dans certains domaines (notamment la comparaison inter-sujet), peut être largement dépassée en terme de rapport précision/coût par des méthodes Lie comme il a été vu ici. Notons tout de même afin de nuancer le propos que la présentation des résultats a volontairement été simplifiée et que certains paramètres sous-jacents ayant leur importance n'ont pas été présentés.

Nicolas Masson

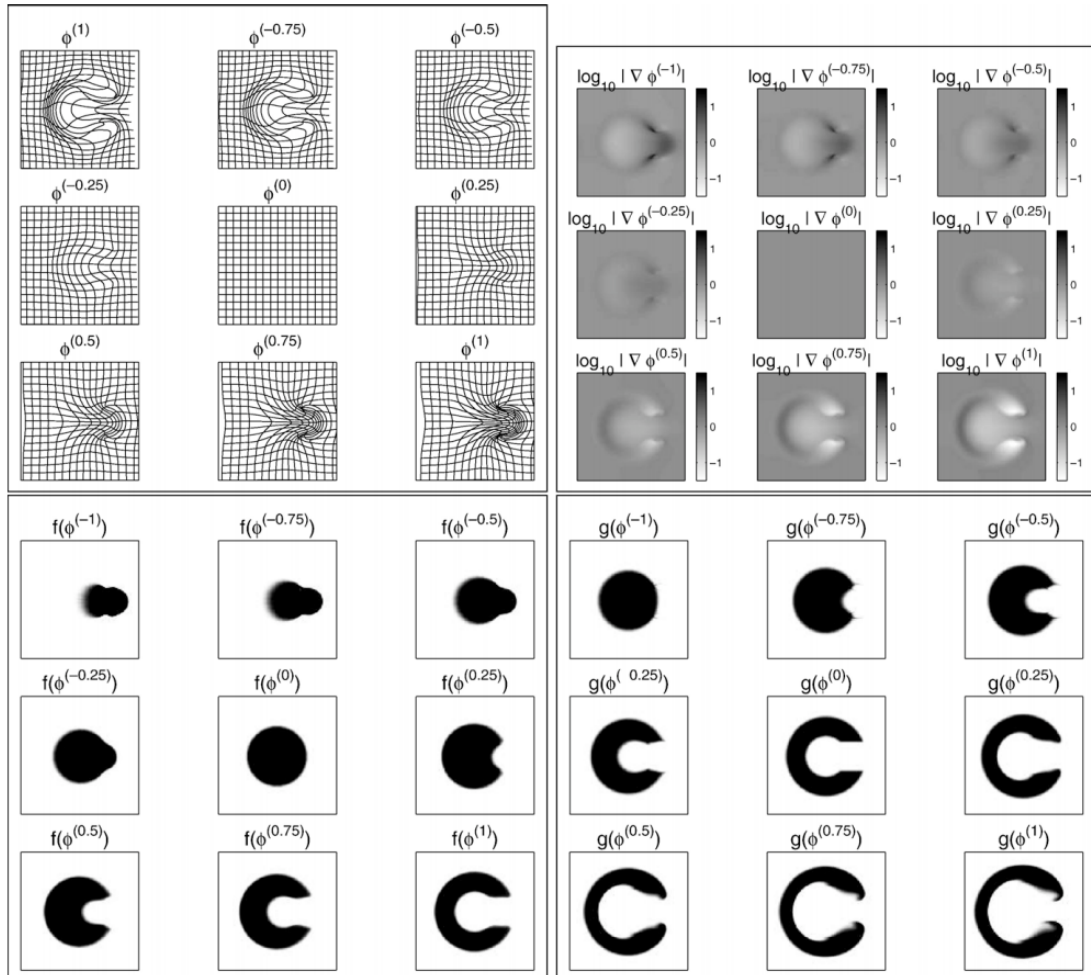


FIGURE 17.5 – Exemple de matching SVF  $\varphi$  obtenu entre une boule  $f$  et une image de “C”  $g$ . Dans cette image tirée d’un article “pratique” présentant une implémentation efficace des idées SVF, [Ash07] de John Ashburner, on voit que la continuité du difféo est bien assurée numériquement.

# Chapitre 18

## Autres utilisations de la géométrie Riemannienne en imagerie médicale

Séance 14

[Pen09]

Pascal Millet

### 18.1 Synthèse sur la géométrie Riemannienne

#### 18.1.1 Concepts théoriques principaux

**Rappels de la première partie du cours** Pour utiliser les outils de la géométrie riemannienne, nous avons eu besoin d'introduire la notion de variété différentielle, qui permet de munir l'espace topologique d'un espace vectoriel tangent en chaque point. On peut ainsi donner un sens à la notion de vecteur tangent à une courbe ainsi qu'à la différentielle d'une fonction. Cependant, cela ne suffit pas pour définir la dérivée d'un champ de vecteur car l'espace tangent global (comme union des espaces tangent en chaque point) n'est pas muni d'une structure d'espace vectoriel. On a donc besoin de la notion de connexion pour savoir comment déplacer les vecteurs d'un espace tangent à l'autre. Cela permet en outre de définir la notion de géodésique comme des courbes de vecteur vitesse constant. Enfin, nous avons vu que l'ajout de la métrique Riemannienne détermine de façon unique la connexion (connexion de Levi-Civita associée à la métrique).

**La carte exponentielle** Etant donné un espace de Riemann et un point  $P$ , on peut définir une application lisse entre l'espace tangent et une partie de la variété en associant à chaque vecteur  $\vec{v}$  le point de la géodésique de condition initiale  $\vec{v}$  au temps 1 (c'est le flot géodésique au temps 1 comme défini en cours de calcul différentiel). Cependant, cette application n'est injective qu'au voisinage de 0 car, car certains points peuvent être atteints au temps 1 par plusieurs géodésiques (exemple du pôle sud de la sphère si on considère la carte exponentielle associée au pôle nord). On peut définir le logarithme d'un point  $Q$  de la variété comme le vecteur de norme minimale dont l'image par l'exponentielle est  $Q$ . Le logarithme est ainsi défini en tout point où un tel vecteur existe et est unique. On appelle cut locus vectoriel, le bord de l'image du logarithme et cut locus l'ensemble des exponentielles des point du cut locus vectoriel. Ainsi sur la sphère au pôle nord, le cut locus vectoriel est le cercle de rayon  $\pi r$ .

**Structure vectorielle transportée depuis le log-domaine** Interprétation : On peut interpréter l'exponentielle comme une généralisation de la somme d'un point et d'un vecteur, ainsi  $\exp_P(v)$  est l'analogie dans un espace affine de  $P + \vec{v}$ . De même  $\log_P(Q)$  est l'analogie de  $\vec{PQ}$ . Garder ces analogies à l'esprit permet de généraliser facilement les algorithmes applicables dans

Pascal Millet des espaces vectoriels au cas riemannien. Il existe des algorithmes efficaces pour calculer l'exponentielle et le logarithme.

On peut également remarquer que la métrique riemannienne induit une mesure sur la variété. Cela permet de définir l'intégrale d'une fonction de la variété dans un espace vectoriel de dimension finie.

### 18.1.2 Extension des notions de moyenne et de covariance

Une fois ce cadre défini rigoureusement, il est possible de généraliser les notions de moyenne et de covariance. Tout d'abord la moyenne de Fréchet qui est le minimiseur de la quantité

$$\sigma^2(P) = \int_M \text{dist}(P, t)^2 dt = \int_M \|\log_P(t)\|^2 dt \quad (18.1)$$

soit pour une distribution discrète de  $n$  points sur la variété :

$$\sigma^2(P) = \sum_{k=1}^n \|\log_P(Q_k)\|^2. \quad (18.2)$$

On a ensuite la définition barycentrique de Karcher qui consiste à trouver un point  $P$  tel que  $\int_M \log_P(t) dt = 0$  et dans le cas d'une distribution discrète

$$\sum_{k=1}^n \log_P(Q_k) = 0. \quad (18.3)$$

Lors du calcul de la moyenne de Fréchet, on est amené à calculer le gradient de  $\sigma^2$ .

**Proposition 18.1.** *Soit  $p \in M$ , si le cut-locus de  $p$  est de mesure nulle, alors  $\sigma^2$  est différentiable en  $p$  et on a :*

$$\vec{\nabla} \sigma^2(p) = -2 \int \log_p(q) d\mu(q) \quad (18.4)$$

ou (dans le cas fini qui nous intéresse en pratique) :

$$\vec{\nabla} \sigma^2(p) = -2 \sum_{i=1}^n \log_p(q_i). \quad (18.5)$$

*Démonstration.* Nous allons démontrer la proposition dans le cas fini. Pour cela, il suffit de calculer le gradient en  $p$  de  $d(p, q)^2$ . Rappelons le lemme de Gauss, vu à la fin de la première partie du cours :

**Lemme 18.1.** *Soit  $U$  une boule géodésique centrée en  $p \in M$  sur laquelle l'exponentielle est injective (pour que le champ radial soit bien défini), le champ de vecteurs radial  $\frac{\partial}{\partial r}$  est orthogonal aux sphères géodésiques de  $U$ .*

On aura besoin du corollaire suivant :

**Proposition 18.2.** *La fonction distance radiale à  $q$  est différentiable sur  $U - \{q\}$  et*

$$\vec{\nabla} r(p) = \frac{\partial}{\partial r}(p) \quad (18.6)$$

*Preuve du corollaire.* Pour démontrer cette proposition, il suffit de remarquer que la carte exponentielle au point  $q$  fournit une carte locale de  $U$  donc la différentiabilité de  $r$  est équivalente (par définition d'une fonction différentiable sur une variété) à la différentiabilité de  $r \circ \exp_q$  et on peut calculer la différentielle dans cette carte (qui est pour l'instant une forme linéaire sur  $T_{\log_q(p)}(T_q M)$ ) : le produit scalaire contre  $\frac{\partial}{\partial r}$  qu'il faut ramener à une forme linéaire sur  $T_p M$  au moyen du difféomorphisme  $d \exp_q(p)$  induit par la carte entre  $T_{\log_q(p)}(T_q M)$  et  $T_p M$ . On obtient

$$d(r \circ \exp_q)(\log_q(p)) = dr(p) \circ d \exp_q(\log_q(p)) \quad (18.7)$$

Quand on revient à la définition des espaces tangents (pour  $v$  vecteur de  $T_{\log_q(p)}(T_q M)$ ), l'exponentielle envoie la classe de courbes passant par  $\log_q(p)$  en 0 et de vecteur dérivé en 0 égal à  $v$  sur une classe de courbe correspondant à un vecteur de  $T_p M$  ce qui définit l'application  $d \exp_q(p)$ , puisque que les courbes de niveau de  $r \circ \exp_q$  sont envoyées sur les courbes de niveau de  $r$ ,  $d \exp_q(p)$  induit un isomorphisme entre les espaces tangents à la sphère géodésique au point  $p$ . Donc, d'après l'équation 18.7 le noyau de  $dr$  est exactement l'hyperplan tangent à la sphère au point  $p$ . Enfin, par définition du champ radial (sur la variété)  $\frac{\partial}{\partial r}$  est tel que

$$\frac{\partial}{\partial r}(p) = d \exp_q(p)\left(\frac{\partial}{\partial r}\right) \quad (18.8)$$

Cela signifie en évaluant 18.7 en  $\frac{\partial}{\partial r}$  que

$$dr(p)\left(\frac{\partial}{\partial r}(p)\right) = 1 \quad (18.9)$$

Par le lemme de Gauss,  $\frac{\partial}{\partial r}(p)$  est orthogonal à l'espace tangent à la sphère donc  $dr(p)$  est simplement le produit scalaire contre  $\frac{\partial}{\partial r}(p)$ . C'est pourquoi on trouve bien le résultat annoncé.  $\square$

**Retour au calcul du gradient de la variance sur une variété Riemannienne** On peut maintenant démontrer la proposition. Dans les variétés riemanniennes que l'on considère ici, la distance radiale est égale à la distance euclidienne. Soit  $q \in M$  qui n'appartient pas au cut-locus. On définit le champ de vecteur radial par rapport au point  $q$ . La proposition précédente donne, par composition avec la fonction carrée réelle (différentiable) :

$$\text{grad}_p d(p, q)^2 = 2d(p, q) \frac{\partial}{\partial r}(p) \quad (18.10)$$

Or, par définition,

$$\frac{\partial}{\partial r}(p) = d \exp_q(\log_q(p))\left(\frac{\partial}{\partial r}\right) \quad (18.11)$$

En revenant à la définition en terme de dérivée directionnelle, on a

$$\frac{\partial}{\partial r}(p) = \frac{d}{dt}(\exp_q(\log_q(p) + t \frac{\partial}{\partial r}))|_0 \quad (18.12)$$

$$= \frac{d}{dt}(\exp_q(\log_q(p) + t \frac{\log_q(p)}{d(p, q)}))|_0 \quad (18.13)$$

. Par le lemme de changement d'échelle (rescaling lemma vu dans la séance 5),

$$\exp_q(\log_q(p) + t \frac{\log_q(p)}{d(p, q)}) = \gamma(1 + \frac{t}{\log_q(p)}) \quad (18.14)$$

Pascal Millet où  $\gamma$  est l'unique géodésique minimale telle que  $\gamma(0) = q$  et  $\gamma(1) = p$ . donc on obtient finalement :

$$\frac{\partial}{\partial r}(p) = \frac{\gamma'(1)}{d(p, q)} \quad (18.15)$$

On se place maintenant dans la carte exponentielle au point  $p$ . Par définition, si on note  $\tilde{\gamma}$  l'unique géodésique minimale telle que  $\tilde{\gamma}(0) = p$  et  $\tilde{\gamma}(1) = q$ , on a :

$$\log_p(q) = \tilde{\gamma}'(0) \quad (18.16)$$

Par unicité, on a  $\gamma(t) = \tilde{\gamma}(1-t)$  et donc :

$$\gamma'(1) = -\tilde{\gamma}'(0) = -\log_p(q) \quad (18.17)$$

Finalement en rassemblant 18.10, 18.15 et 18.17, on obtient

$$\text{grad}_p d(p, q)^2 = -2 \log_p(q) \quad (18.18)$$

On en déduit (par linéarité du gradient) la formule du gradient de la variance dans le cas fini (le cas continu ne se déduit pas directement de ce qui précède car la dérivation sous l'intégrale ne s'obtient pas simplement) :

$$\vec{\nabla} \sigma^2(p) = -2 \sum_{i=1}^n \log_p(q_i) \quad (18.19)$$

dans le cas fini. □

**Bilan : un algorithme de minimisation de la variance** On peut voir avec cette formule que la moyenne de Karcher consiste à définir une moyenne comme un point critique de la variance. La moyenne de Fréchet est donc une moyenne de Karcher ou un point en lequel la variance n'est pas différentiable (donc un point dont le cut-locus n'est pas de mesure nulle). L'algorithme de descente du gradient géodésique permet bien de trouver le minimiseur d'une énergie sur une variété riemannienne. On va ainsi généraliser très facilement la notion de moyenne pondérée. Si on se donne un poids  $w_i$  pour chaque point de l'échantillon, on peut considérer l'énergie

$$E(p) = \sum_i w_i d(f_i, p)^2 \quad (18.20)$$

et on constate, dans le cas euclidien que le minimiseur de cette énergie est bien la moyenne pondérée des  $f_i$ .

**Formules pratiques** Une fois que l'on a trouvé la moyenne, on peut s'intéresser à la variance d'un ensemble continu ou discret de données. Si on note  $\bar{p}$  la moyenne des données, on peut se placer dans le plan tangent (muni d'une base orthonormée) et calculer la quantité :

$$\Sigma = \int \log_{\bar{p}}(q) \log_{\bar{p}}(q)^T dP(q) \quad (18.21)$$

ou dans le cas discret :

$$\Sigma = \frac{1}{n} \sum_{i=1}^n \log_{\bar{p}}(q_i) \log_{\bar{p}}(q_i)^T \quad (18.22)$$

Ces formules sont obtenues par généralisation des formules vectorielles pour le calcul de la matrice de covariance. Une telle matrice sert à effectuer une analyse statistiques des données (par exemple par une analyse en composantes principales).

## 18.2 Quelques méthodes vectorielles utilisées en analyse d'image et leurs analogues Riemanniennes

Le traitement d'images à valeurs *réelles* est un problème bien connu, pour lequel de nombreuses théories ont été développées. Avant de tenter de les généraliser à des images à valeurs dans une *variété riemannienne* quelconque, on rappelle les problèmes et solutions les plus classiques dans le cas vectoriel.

### 18.2.1 Interpolation

L'interpolation est le fait de retrouver la valeur d'une fonction à partir de l'image de quelques points. On a donc pour une interpolation linéaire entre deux points d'antécédent 0 et 1.

$$f(x) = (1 - x)f(0) + xf(1) \tag{18.23}$$

Sous cette forme, la valeur apparait comme une moyenne pondérée des données. On peut généraliser cette notion en dimension  $n$  en exprimant  $x$  comme barycentre de  $n+1$  points  $x = \sum_{i=1}^{n+1} \lambda_i x_i$ . On obtient donc une formule pour calculer l'image de  $x$  :

$$f(x) = \sum_{i=1}^{n+1} \lambda_i f(x_i) \tag{18.24}$$

Donc si  $f$  est à valeur dans une variété riemannienne, on comprend l'intérêt d'avoir généralisé la notion de moyenne pondérée. On verra par la suite un exemple dans lequel l'interpolation linéaire permet de reconstituer un champ de tenseurs à partir de quelques points.

### 18.2.2 Filtrage/lissage d'images

Les images que nous considérons ici peuvent être représentées par des fonctions  $L^2(\llbracket 0, n \rrbracket^2, \mathbb{R})$  où  $n$  est le nombre de pixels par ligne et par colonne. Le principe lorsque l'on traite une image (par exemple pour la débruiter) est de trouver, étant donnée une image donnée  $\tilde{f}$ , une image qui minimise la quantité :

$$\text{Reg}(f) + \text{Similarite}(f, \tilde{f}) \tag{18.25}$$

Le terme  $\text{Reg}(f)$  (critère de régularisation) est un terme qui permet de placer un a priori sur l'image que l'on veut obtenir. Par exemple, on peut choisir de pénaliser les variations de  $f$  comme nous le verrons plus loin mais on peut également vouloir imposer un certain spectre fréquentiel et pénaliser l'écart par rapport à celui-ci.  $\text{Similarite}(f, \tilde{f})$  est un terme qui pénalise l'éloignement par rapport à l'image de départ. En général, sur des images denses, on le prend égal à  $\|f - \tilde{f}\|_2^2$ .

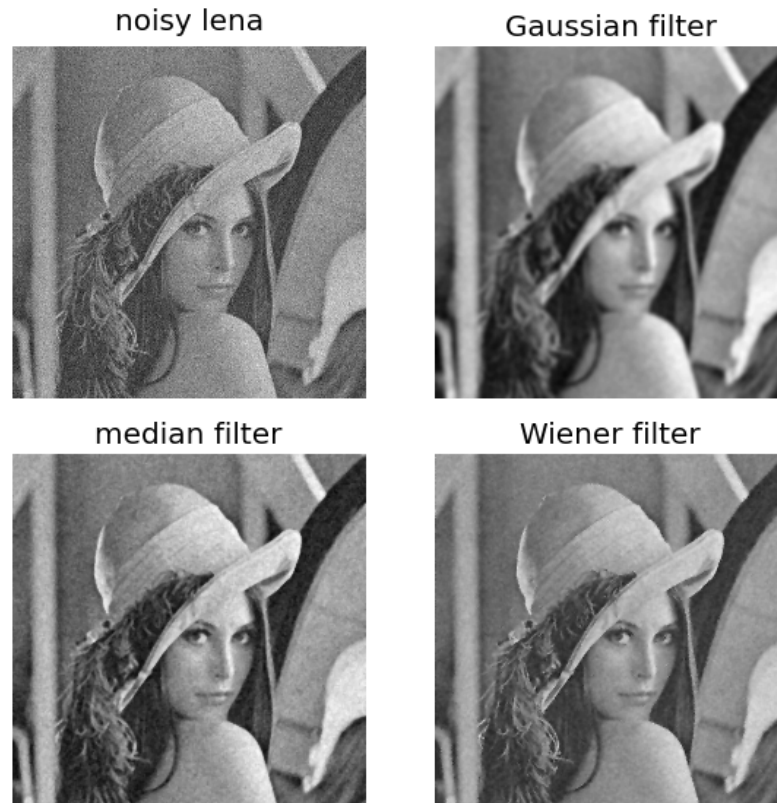
On peut bien sûr envisager une descente de gradient pour trouver  $f$  mais pour certains critères, il est possible d'obtenir une expression explicite de  $f$  (par des calculs d'analyse de Fourier qui seront vus en cours de "Traitement du signal") comme une moyenne pondérée des images de  $\tilde{f}$

$$f(x) = \sum_i \omega_i(x) \tilde{f}(i). \tag{18.26}$$

On peut donc ranger dans ce cadre le simple floutage gaussien, ou convolution avec une gaussienne

$$\omega_i(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\|i-x\|^2}{2}}. \tag{18.27}$$

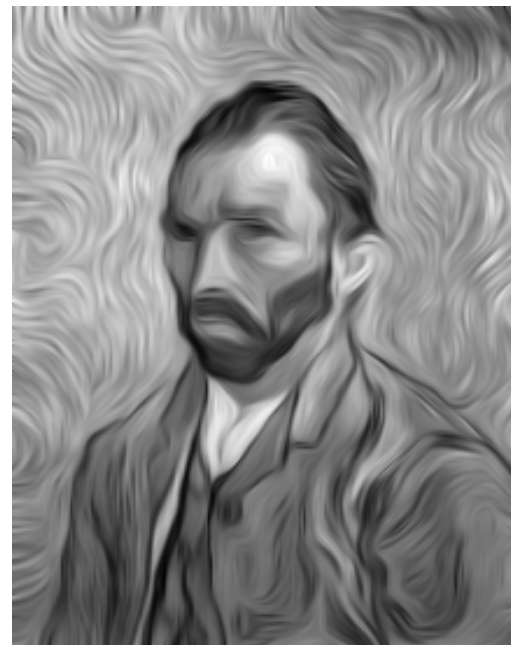
Pascal Millet



(a) Quelques filtrages linéaires d'une image "Lena" bruitée.



(b) Un tableau de Van Gogh...



(c) Après application d'un algo de diffusion anisotrope. Remarquons comme les structures "en ligne", les arêtes sont préservées.

FIGURE 18.1 – Exemples d'algorithmes de lissage d'images à valeurs réelles. Images tirées de [pybonacci.github.io/scipy-lecture-notes-ES/intro/scipy.html](https://pybonacci.github.io/scipy-lecture-notes-ES/intro/scipy.html) et *Anisotropic Diffusion in Image Processing*, de Joachim Weickert.



On peut interpréter le calcul en disant que  $f(x)$  est la moyenne des  $\tilde{f}(i)$  pondérés selon une gaussienne centrée en  $x$ . On comprend donc que la fonction  $\tilde{f}$  est lissée par ce processus. Lorsque la fonction est à valeur dans une variété riemannienne (nous allons en voir un exemple par la suite), nous allons donc utiliser la notion de moyenne pondérée définie précédemment.

Il existe aussi des manière non-linéaires de filtrer le bruit sur un image. La diffusion harmonique consiste à minimiser une quantité qui pénalise les fortes variations de la fonction :

$$\text{Reg}(p) = \frac{1}{2} \int_{\Omega} \|\nabla p\|_{p(x)}^2 dx \quad (18.28)$$

Par un calcul d'analyse fonctionnelle (équation d'Euler-Lagrange), on peut calculer le gradient de cette quantité (qui est donc une fonction) :

$$\nabla \text{Reg}(p)(x) = -\Delta p(x) \quad (18.29)$$

Suivre ce gradient pendant un temps  $\delta t$  revient à suivre l'évolution spontanée du système pendant  $\delta t$  si on considère que la fonction donne la température en chaque point. On comprend bien que ce procédé a tendance à lisser l'image. Il faut suffisamment d'itérations pour obtenir une image débruitée mais pas trop pour ne pas obtenir une répartition de "température" homogène. En pratique, on ajoute donc à Reg, un terme qui pénalise l'éloignement par rapport aux données initiales. Par exemple dans le cas d'un bruit gaussien, on peut ajouter le terme :  $\int_{\Omega} \text{dist}^2(p(x), q(x)) dx$  où  $q$  est la donnée initiale. Il peut également être intéressant de filtrer les régions homogènes de l'image sans atténuer les contours de ces régions (on considère en quelque sorte que l'image est constituée d'aplats). On utilise pour cela la diffusion anisotrope dans laquelle on minimise

$$\text{Reg}_{\varphi}(p) = \int_{\Omega} \varphi(\|\nabla p(x)\|_{p(x)}) dx \quad (18.30)$$

Cela permet, en choisissant bien la fonction  $\varphi$  d'obtenir une diffusion anisotrope.

### 18.2.3 Analyse en composantes principales

Enfin, l'analyse en composante principale (CPA) est une méthode statistique permettant de trouver les axes de variabilité maximaux dans un ensemble de données. Pour cela, on calcule la covariance empirique sous la forme

$$C = \frac{1}{k} M^T M \quad (18.31)$$

où  $M$  est la matrice des observations empiriques centrées sur leur moyenne (chaque ligne correspond à une mesure conjointe des  $n$  d'une valeur dans les  $n$  direction de l'espace [pour certaines applications, chaque direction représente une variable aléatoire] et on note  $k$  le nombre de mesures, ainsi  $M$  est une matrice  $k$  par  $n$ ). La matrice  $C$  est symétrique positive réelle et pour une combinaison linéaire des  $n$  vecteurs de base  $\sum_{i=1}^n \lambda_i e_i$ , on note  $X$  le vecteur coordonnées. Ainsi  $X^T C X$  est la variabilité de l'échantillon selon la direction  $X$ . D'après le théorème spectral  $C$  est diagonalisable dans une base orthonormée et cela permet de trouver les axes de variabilités propres et en particulier l'axe de variabilité maximal. On peut utiliser cette technique pour analyser des données mais également pour les débruiter (on ne retient que les axes de variabilité principaux pour décrire les données, les axes de plus faible variabilité correspondent au bruit).

Il existe plusieurs moyens de généraliser l'analyse en composante principale. L'analyse en géodésique principale consiste à trouver une sous-variété de faible dimension engendrée par un ensemble de géodésiques qui expliquent la variabilité des données. On peut également envisager une analyse en composantes principales dans l'espace tangent à la moyenne de l'échantillon.

Pascal Millet **18.2.4 Questions pratiques : Robustesse**

En statistique, il peut être intéressant que la moyenne ne soit pas trop sensible aux valeurs extrêmes. Cette propriété s'appelle la robustesse. En géométrie Riemannienne, en prenant une distance qui approxime localement la distance Riemannienne mais qui sous-estime cette distance pour des points très éloignés, on peut obtenir une moyenne plus robuste. Une manière de définir une telle distance est de considérer une fonction  $\varphi$  croissante concave telle que  $\varphi(0) = 0$  et  $\varphi'(1) = 1$  et de poser  $d'(x, y) = \varphi(d(x, y))$ . En général, si l'espace est plongé dans  $\mathbb{R}^n$ , la distance euclidienne de  $\mathbb{R}^n$  fournit automatiquement une distance robuste et simple à calculer. Cependant, pour cette distance, la de moyenne de Fréchet, minimiseur de

$$E(p) = \int_{\Omega} d'(p, q)^2 dq \tag{18.32}$$

n'est plus un barycentre de la distribution (i.e. un point tel que

$$\int_{\Omega} \frac{d'(p, q)}{d(p, q)} \log_p(q) dq \tag{18.33}$$

car

$$\nabla E(p) = -2 \int_{\Omega} \varphi'(d(p, q)) \frac{d'(p, q)}{d(p, q)} \log_p(q) dq \tag{18.34}$$

Il faut donc choisir la définition qui convient le mieux. De manière générale, on utilise la métrique intrinsèque pour évaluer la performance et le domaine de validité d'une métrique approchée (ce domaine de validité est un domaine de variance des données).

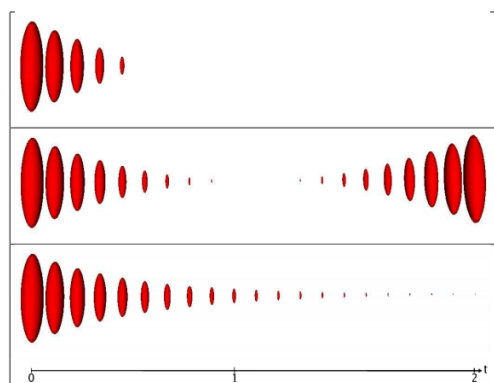
## 18.3 Application de la géométrie Riemannienne à l'analyse d'images de tenseurs

Miracle de la science moderne : les IRM permettent aujourd'hui d'obtenir des images à valeurs tensorielles, où le produit scalaire en chaque point quantifie la diffusion de l'eau dans le corps du patient. Avec des applications simples (diagnostic d'hémorragies vs ischémie rapide, ce qui peut se révéler crucial en cas d'AVC) ou complexes (suivi de fibres neuronales, ...), cette nouvelle modalité est appelée à jouer un rôle toujours plus important et stimule grandement la recherche en traitement d'images.

### 18.3.1 L'espace des matrices symétriques définies positives

**Rappels sur les produits scalaires** L'espace  $S_n^{++}$  des matrices symétriques définies positives (on appellera aussi une telle matrice un tenseur) est un cône dans l'espace des matrices symétriques (ie si  $A \in S_n^{++}$ ,  $\forall \lambda \in \mathbb{R}_+$ ,  $\lambda A \in S_n^{++}$  et A est convexe). A priori on peut donc effectuer des moyennes entre éléments de  $S_n^{++}$  sans sortir de l'espace. Cependant, ces moyennes ne sont en général pas satisfaisantes et l'implémentation d'algorithme telle que la descente de gradient pose des problème. On risque en effet de sortir de l'espace à cause de la discrétisation du problème! (voir figure 18.2a). Pour pallier à cela, on va munir l'espace d'une structure riemannienne et envoyer les bords à l'infini.

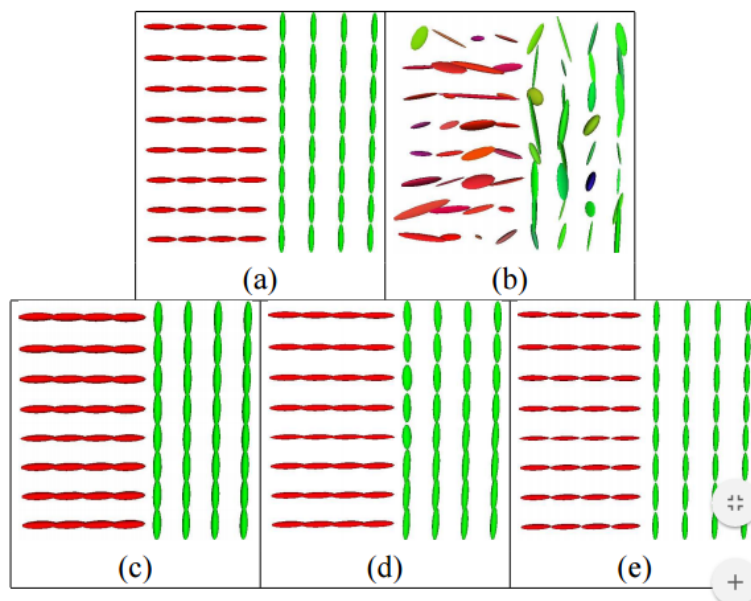
**Métrique affine-invariante** Pour choisir une métrique canonique, on va imposer que la métrique soit invariante par changement de coordonnées affine. Ici les matrices que l'on considère



(a) Simulation d'un shoot géodésique avec différentes métriques. En haut, un shoot dans  $S_n^{++}$  vu comme une sous-variété riemannienne de  $S_n$ , munie de la métrique euclidienne des matrices : les géodésiques sortent du cône. Au milieu, un shoot pour la métrique de Cholesky. En bas, un shoot pour la métrique affine-invariante. La dernière ligne est bien la plus satisfaisante, avec des bords repoussés à l'infini : un algorithme de descente de gradient (comme les diffusions) ne risque donc pas de sortir du cône  $S_n^{++}$ .



(b) Interpolation géodésique entre deux tenseurs. La première ligne est une simple interpolation linéaire “coordonnée à coordonnée”, la seconde est le shoot géodésique dans  $S_n^{++}$  muni de la métrique affine-invariante, et la troisième correspond au cadre log-euclidien. À droite, on trace le volume de l'ellipsoïde en fonction du temps : l'effet de gonflement des moyennes, indésirable dans l'interpolation naïve, est bien évité.



(c) Comparaison entre les différentes méthodes. (a) : image initiale. (b) : image bruitée. (c) : régularisation euclidienne. (d) : régularisation de Cholesky. (e) : régularisation log-euclidienne

FIGURE 18.2 – Intérêt de la métrique affine-invariante et de son approximation algébrique “log-euclidienne” pour le traitement d'images à valeurs tensorielles.

Pascal Millet sont des matrices de diffusion donc l'action de  $GL_n$  sur  $S_n^{++}$  est naturellement donnée par :

$$A * \Sigma = A^T \Sigma A \quad (18.35)$$

En effet, ainsi, si  $\Sigma$  est la matrice de diffusion dans la base  $\mathcal{B}$ , et si  $A = \text{Mat}_{\mathcal{B}', \mathcal{B}}(id)$ ,  $A * \Sigma$  est la matrice de diffusion dans la nouvelle base  $\mathcal{B}'$ . Ainsi, intuitivement, on voudrait une distance entre les diffusions qui ne doit donc pas dépendre du choix de la base utilisée pour les représenter. Il faut maintenant préciser formellement ce que signifie une métrique invariante sous l'action. Il faut notamment comprendre comment l'action transporte les vecteurs tangents. L'espace tangent en tout point est  $S_n$ . L'action de  $A \in GL_n$  induit un difféomorphisme de la variété. On utilise la différentielle pour transporter les vecteurs tangent d'un point à un autre. La condition d'invariance s'écrit donc :

$$\forall V, W \in S_n, \langle V, W \rangle_{\Sigma} = \langle A^T V A, A^T W A \rangle_{A^T \Sigma A} \quad (18.36)$$

L'action est transitive donc une métrique vérifiant la condition d'invariance ci-dessus est entièrement déterminée par sa données en l'identité. De plus, puisque les transformations orthogonales laissent l'identité invariante, on a, en notant  $O$  une matrice orthogonale :

$$\forall V, W \in S_n, \langle V, W \rangle_{Id} = \langle O^T V O, O^T W O \rangle_{Id} \quad (18.37)$$

On peut montrer que les métrique vérifiant cette condition sont de la forme

$$\langle V, W \rangle_{Id} = \text{Tr}(VW) + \beta \text{Tr}(V) \text{Tr}(W) \text{ avec } \beta > -\frac{1}{n} \quad (18.38)$$

Pour connaître la métrique en un autre point  $\Sigma$ , on utilise la transitivité et l'invariance. On pose  $A = \Sigma^{\frac{1}{2}} \in GL_n$  et on a alors ( $A$  est en fait symétrique définie positive) :  $A * Id = \Sigma$ . Cela permet très facilement de réécrire la métrique en  $\Sigma$  :

$$\forall V, W \in S_n, \langle V, W \rangle_{\Sigma} = \langle A^{-1} V A^{-1}, A^{-1} W A^{-1} \rangle_{Id} \quad (18.39)$$

$$= \text{Tr}(V \Sigma^{-1} W \Sigma^{-1}) + \beta \text{Tr}(V \Sigma^{-1}) \text{Tr}(W \Sigma^{-1}) \quad (18.40)$$

Il est intéressant de remarquer que toutes ces métriques sont associées à la même connexion de Levi-Civita. En particulier, elles partagent toutes les même géodésiques, cependant la distance le long d'une géodésique n'est pas la même d'une métrique à l'autre. On dispose également d'expressions explicites de l'exponentielle et du logarithme.

$$\text{Exp}_{\Sigma}(W) = \Sigma^{\frac{1}{2}} \exp(\Sigma^{\frac{1}{2}} W \Sigma^{\frac{1}{2}}) \Sigma^{\frac{1}{2}} \quad (18.41)$$

$$\text{Log}_{\Sigma}(V) = \Sigma^{\frac{1}{2}} \log(\Sigma^{\frac{1}{2}} V \Sigma^{\frac{1}{2}}) \Sigma^{\frac{1}{2}} \quad (18.42)$$

Avec la formule du logarithme, on peut calculer la distance et voir que les matrices qui ont une valeur propre nulle (c'est à dire les bords du cône) ont bien été envoyées à l'infini !

**Métrique Log-euclidienne** Dans la métrique précédente, la moyenne est obtenue par une descente du gradient. Il serait plus efficace d'avoir une formule explicite. La métrique log-euclidienne permet d'obtenir une telle formule tout en restant satisfaisante en terme de précision. Ainsi, elle se situe entre la métrique affine-invariante (très bien fondée mathématiquement) et la simple moyenne dans  $S_n$  (très rapide à calculer mais présentant de nombreux défauts). Pour la définir, il faut tout d'abord remarquer les faits suivants :

**Proposition 18.3.** *L'exponentielle des matrices est une bijection de  $S_n$  dans  $S_n^{++}$*

La preuve se fait simplement en utilisant le théorème spectral qui permet de diagonaliser la matrice symétrique dans une base orthonormée. Cela permet également d'obtenir une formule explicite pour le calcul de l'exponentielle et du logarithme.

L'intérêt d'une telle bijection est qu'elle permet de transporter la structure d'espace euclidien de l'ensemble des matrices symétriques vers l'ensemble des matrices symétriques définies positives. La métrique Log-euclidienne est alors la métrique ainsi transportée. Ainsi, on a donc un moyen simple de traiter des données : On envoie les matrices de  $S_n^{++}$  dans  $S_n$  grâce au logarithme puis on effectue les calculs (par exemple la moyenne) puis on prend l'exponentielle du résultat. Ainsi on peut obtenir une formule explicite pour la moyenne et la calculer efficacement et appliquer l'algorithme de descente du gradient sans craindre de rencontrer les bords du cône. Par ailleurs, l'espace obtenu présente de nombreuses similarités avec l'espace muni de la métrique affine-invariante. Par exemple, les géodésiques passant par l'identité sont les mêmes pour les deux métriques. Par ailleurs, les tests sur des données réelles montrent que les résultats obtenus lors des calculs avec ces deux métriques sont très similaires (on observe typiquement un écart de 1%) alors que les algorithmes implémentant la métrique Log-euclidienne tournent 4 à 10 fois plus vite!

### 18.3.2 IRM de diffusion (DTI : diffusion tensor imaging)

On s'intéresse ici au débruitage d'images de tenseurs de diffusion obtenues par IRM. Chaque pixel de l'image est maintenant une matrice symétrique définie positive donc une image peut être vue comme une application à valeurs dans cette variété. On va donc pouvoir utiliser les outils de géométrie riemannienne à deux niveaux : pendant la construction de l'image à partir des données brutes fournies par l'IRM et pendant le traitement de l'image proprement dit.

Nous allons d'abord voir comment la géométrie riemannienne permet de mieux prendre en compte la spécificité du bruit lors de la construction de l'image. En effet, le processus d'acquisition est très bruité. Il consiste à observer la diffusion du signal pour différentes directions du gradient du champ électrique. Si on note  $S_i$  le signal (valeur scalaire) et  $g_i$  le gradient, on a la relation suivante entre ces valeurs (fournie par l'analyse physique du phénomène) :

$$S_i = S_0 \exp(-bg_i^T Dg_i) \quad (18.43)$$

Comme  $D$  est symétrique il suffit en principe de 6 équations scalaires pour retrouver tous les coefficients. Cependant, étant donné le niveau de bruit, il peut être intéressant de surdéterminer le système (ie prendre plus de mesures que nécessaire) et trouver une solution au sens des moindres carrés. Il s'agit donc de minimiser la quantité suivante :

$$\text{Sim}_G(D) = \sum (\hat{S}_i - S_i(D))^2 \quad (18.44)$$

où  $\hat{S}_i$  désigne la valeur mesurée et  $S_i(D)$  la valeur calculée en prenant la matrice de diffusion égale à  $D$ . Cependant, si l'on résout cela de manière naïve dans  $S_n \mathbb{R}$  en cherchant quand le gradient (de la fonction vue comme allant de  $S_n$  dans  $\mathbb{R}$ ) est nul, le résultat peut ne pas être une matrice de  $S_n^{++}$  à cause de l'imprécision des mesures. Un meilleur choix est donc de considérer cette fonction comme allant de la variété riemannienne  $S_n^{++}$  munie de l'une des métriques définies plus haut dans  $\mathbb{R}$  puis d'effectuer une descente du gradient. Par exemple pour la métrique Log-euclidienne, cela correspond simplement à minimiser dans  $S_n$  la fonction suivante :

$$\text{Sim}_G(L) = \sum (\hat{S}_i - S_i(\exp(L)))^2 \quad (18.45)$$

car l'exponentielle est une isométrie entre  $S_n$  et  $S_n^{++}$  par définition de la métrique. L'avantage de procéder ainsi est double : les valeurs obtenues sont toujours dans  $S_n^{++}$  et les résultats obtenus sont plus satisfaisants. Le critère de régularisation est ici un critère des moindres carrés ce qui

Pascal Millet correspond à un bruit gaussien. En pratique, le bruit est d'un autre type (bruit ricien) et le critère de régularisation est différent mais le formalisme riemannien s'adapte bien à ce cas et fourni un algorithme satisfaisant.

La seconde étape du débruitage consiste à traiter l'image de tenseurs obtenue précédemment. Il s'agit donc simplement d'appliquer la généralisation des algorithmes de filtrage/lissage vus précédemment. En particulier un filtrage linéaire ou un algorithme de diffusion anisotrope permettent de débruiter correctement l'image. L'avantage du formalisme riemannien est que les quantités pertinentes telles que le volume initial du tenseur sont bien mieux retrouvées lors du débruitage qu'avec l'utilisation d'algorithme plus simples pour lesquels on observe un net épaissement des tenseurs. Les fibres obtenues sont bien plus nette et le nombre de fibres qui peuvent être correctement reconstitué est aussi plus important. Cependant, une évaluation complète et quantitative des avantages de la méthode reste à effectuer.

### 18.3.3 Analyse de la variabilité des lignes sulcales

Une autre utilisation ces métriques sur l'espace des tenseurs peut être faite dans le cadre de l'étude de la variabilité des lignes sulcales. Après acquisition d'un échantillon d'images sur des individus et un recalage rapide qui permet d'obtenir des correspondances de points le long des lignes. On peut calculer pour chaque point la matrice de covariance qui est alors un tenseur. On dispose ainsi également d'une fonction à valeur dans  $S_n^{++}$ . On peut alors chercher à réduire la quantité d'information nécessaire pour encoder la variabilité. Pour cela on cherche un sous ensemble optimal de matrices telles que une distance entre le champ de matrice empirique et le champ interpolé linéairement à partir des matrices sélectionnées soit inférieure à une valeur prescrite. On utilise ainsi la généralisation riemannienne de l'algorithme d'interpolation linéaire. Il est intéressant de constater que seule 366 tenseurs suffisent à encoder la variabilité sans perte d'information significative.

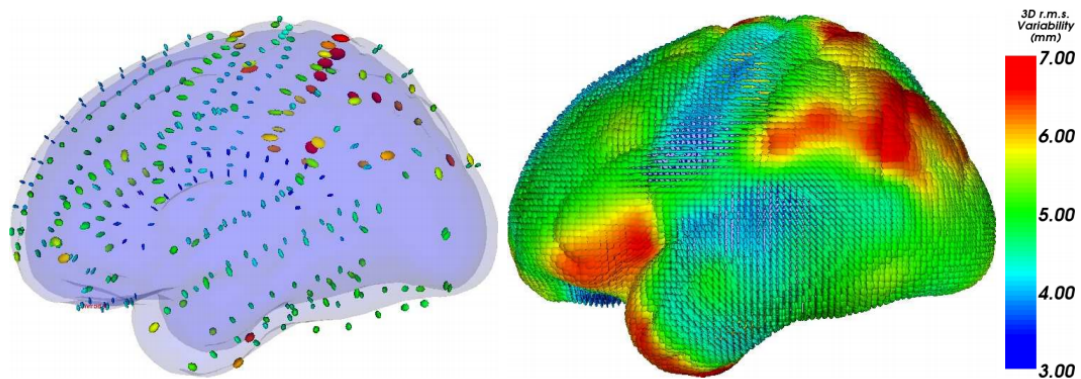
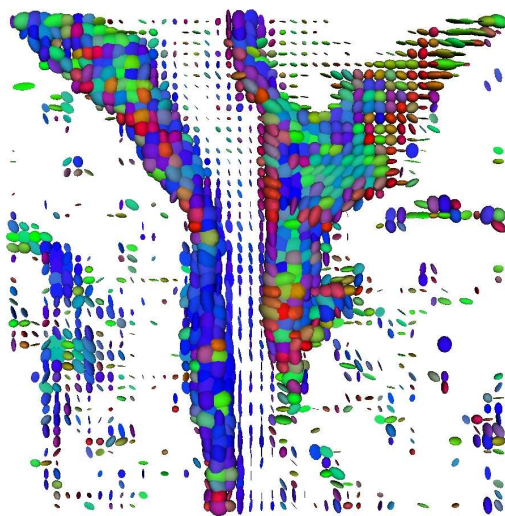
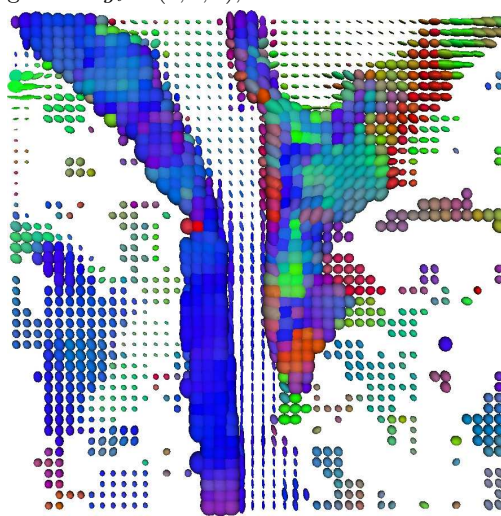


FIGURE 18.3 – Carte de variabilité des lignes sulcales obtenue par le protocole précédent. À gauche, les tenseurs retenus pour expliquer la variabilité et à gauche l'interpolation linéaire entre eux pour reconstituer la carte complète.

(a) Champ au repos  $S_0$ .(b) Exemple d'image mesurée  $S_i$  – ici pour un gradient  $g_i = (1, 0, 1)$ , bruitée.

(c) En répétant l'opération pour au moins 6 gradients, on peut estimer point-par-point un champ de tenseurs bruité.



(d) Un algorithme de débruitage adapté permet de diminuer le rapport signal/bruit, et d'augmenter la fiabilité des algorithmes de traitement ultérieurs – suivi de fibre, par exemple.

FIGURE 18.4 – Exemple pratique de traitement d'image DTI, sur le haut de la moelle épinière - rachis d'un patient.

## Pascal Millet 18.4 Pour aller plus loin : conclusion et perspectives

Quel bilan retirer de ces dix-huit séances ? Nous y avons développé la notion de *métrique* locale, qui permet d'étendre le vocabulaire et les outils classiques de géométrie à de toutes nouvelles classes de problème. Nous attarder sur l'anatomie computationnelle, un domaine de recherche très actif, nous a ensuite permis de mettre ces idées en application, de faire le lien entre mathématiques "fondamentales" classiques et questions concrètes posées par les médecins, chercheurs ou non.

En difféomorphométrie comme en traitement d'images à valeurs tensorielles, on a pu mettre en évidence la présence d'au moins trois méthodes :

- Une méthode "naïve" : flot optique ou moyennage coordonnées à coordonnées ; extrêmement rapide à calculer, adapté aux situations où l'on dispose de beaucoup de données peu dégradées.
- Une méthode Riemannienne : LDDMM et métrique Affine-Invariante ; dotée de garanties théoriques fortes, qui mettent en avant les a priori bien compris que l'on souhaite mettre dans nos algos.
- Un compromis "tangential" : SVF et cadre log-euclidien ; en renonçant à la richesse et l'homogénéité du cadre Riemannien, on accède à une structure algébrique qui accélère grandement les algorithmes.

C'est bien l'interaction, la comparaison entre ces différentes méthodes qui intéressera le spécialiste : chaque point de vue a ses points forts et points faibles, qu'il importe de bien comprendre pour arriver à des performances optimales dans une application voulue.

**Perspectives de carrière** J'espère vous avoir donné un bon aperçu de ce que peuvent être les mathématiques appliquées au traitement d'image. Vous montrer que « problème concret » ne rime pas toujours avec « idées simplistes » ! Aujourd'hui, les questions *mathématiquement intéressantes* liées à des applications réelles ne manquent pas ; elles ouvrent de plus des perspectives de carrière très intéressantes.

Dans le monde académique, d'abord : on en est rarement conscient en entrant à l'ENS, mais le nombre de postes ouverts aux mathématiciens "appliqués" est bien supérieur à celui proposé à leurs homologues algébristes, géomètres ou logiciens – voir à ce sujet la Figure 18.5. Soyez conscients que se lancer dans les domaines les plus fondamentaux, c'est rentrer dans un milieu extrêmement compétitif. Le confort matériel offert aux élèves normaliens ne doit pas vous leurrer : même pour vous qui avez "réussi les concours" à vingt ans (...), il sera difficile de percer et d'obtenir un emploi.

A contrario, se spécialiser en mathématiques appliquées, c'est s'ouvrir les portes du monde des start-ups, de l'industrie. Les entreprises de haute-technologie ont grand besoin d'*experts* qui puissent leur apporter des algorithmes à la pointe, une vraie plus-value sur des produits extraordinairement complexes. Pour développer un scanner IRM par exemple, des dizaines de compétences pointues sont mises à contribution. Un profil unique d'ingénieur généraliste ne peut donc pas suffire, et, avec votre formation très poussée en mathématiques, vous disposez d'un capital précieux. Les entreprises en sont bien conscientes, et sont prêtes à vous accueillir à bras ouverts : je peux en témoigner comme bon nombre de camarades.

Je vous encourage donc très fortement à rester ouverts et curieux : au moment de décider de votre orientation, de choisir votre domaine de spécialisation, prenez bien tous ces éléments en considération. Soyez sûrs de faire un choix que vous ne regretterez pas.



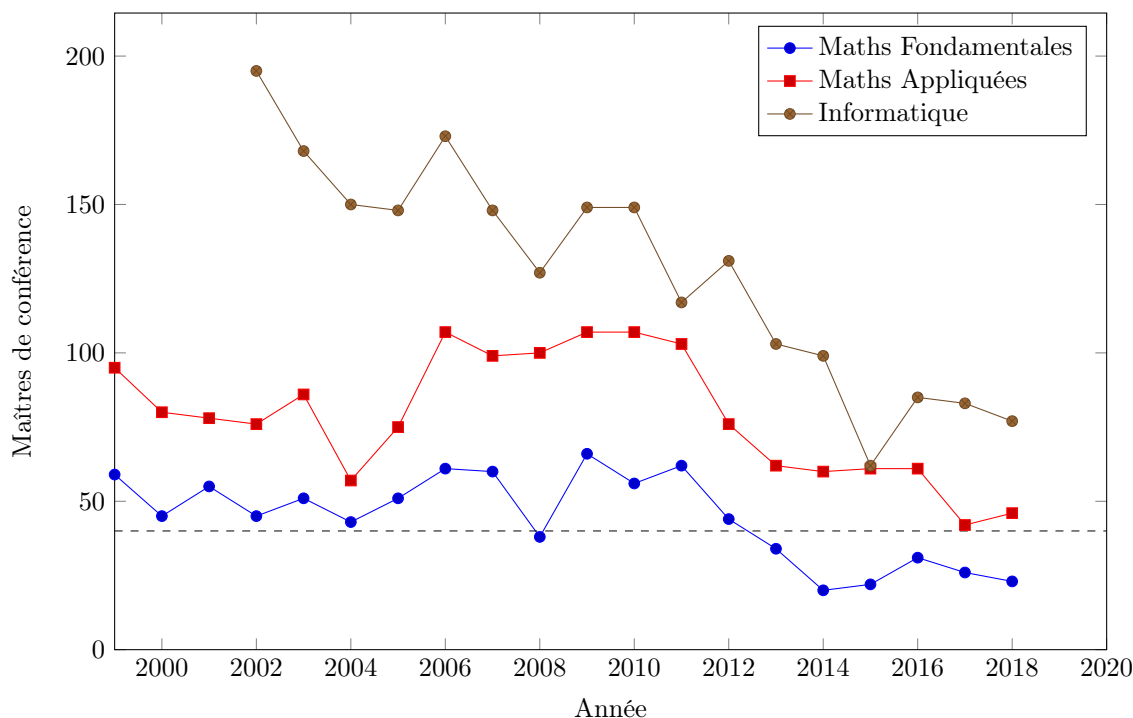


FIGURE 18.5 – Postes de « maître de conférence » ouverts en France, selon les filières. On peut ajouter à ceux-ci les postes de « chargé de recherches » CNRS, 8-10 par an au total pour les filières mathématiques.

Données tirées du site de référence, [postes.smai.emath.fr/2017/OUTILS/bilans.php](http://postes.smai.emath.fr/2017/OUTILS/bilans.php).

**Offre de cours** Pour clore ce polycopié comme il se doit, voici donc quelques suggestions de cours et références dans la continuité de notre groupe de lecture. De nombreuses questions restent à élucider, tant du point de vue théorique que pratique, et, pour ceux qui souhaiteraient s’y lancer, la route est encore pleine de surprises...

- Le cours de MM. Alain Trounev et Joan Glaunès, *Géométrie et espaces de formes*, au M2 MVA – à l’ENS Cachan, de janvier à mars au deuxième semestre. C’est le cours qui soutient les articles des séances 9, 10 et 11 ; la suite naturelle de ce groupe de lecture : le cours demande du travail, mais est tout à fait accessible à des élèves maîtrisant les bases de géométrie riemannienne présentées ici. C’est un bon exemple de cours de “traitement d’images” solide sur le plan mathématique.
- Autre cours du M2 MVA : *Advanced Medical Imaging*, par MM. Hervé Delingette et Xavier Pennec – semestres 1 et 2. Plus touche à tout, proche du milieu médical, ce cours est aussi l’occasion de découvrir la variété des applications, la pluralité des approches coexistant en anatomie computationnelle. Si le point de vue privilégié dans notre cours a été celui de la géométrie Riemannienne, avec tracé de géodésiques dans un espace de difféomorphismes de l’espace ambiant, il ne faut pas croire qu’il s’agit là de la seule manière de travailler : on l’a bien vu au cours des deux dernières séances, portant sur des articles de M. Pennec !

*Pascal Millet* Enfin, si l'esprit "maths appliquées à l'imagerie" de ce cours vous a plu, sachez que vous pourrez tout simplement le retrouver ici, au 45 rue d'Ulm : MM. Stéphane Mallat (au DI) et Bertrand Maury ou Gabriel Peyré (au DMA) sont d'excellents enseignants, qui sauront parfaitement vous conseiller et répondre à vos interrogations. Je ne saurais d'ailleurs trop vous recommander les cours de *Modélisation* (1A) et de *Data Science* (2A) : on s'y intéresse aux foules, aux photos, et à la musique plutôt qu'aux cerveaux et aux cœurs... mais le plaisir de découvrir et expérimenter reste le même !