UPMC
SORBONNE UNIVERSITÉS

HABILITATION A DIRIGER LES RECHERCHES
DE l'UNIVERSITÉ PIERRE ET MARIE CURIE

Spécialité

Mathématiques

Présentée par

Silvère BONNABEL

Une approche géométrique pour certains problèmes de filtrage non-linéaires.

soutenue le 8 septembre 2014 devant le jury composé de :

Mme. Brigitte d'Andéa-Novel	Examinateur
M. Marc Arnaudon	Rapporteur
M. Yacine Chitour	Rapporteur
M. Robert Mahony	Rapporteur
M. Frank Nielsen	Examinateur
M. Emmanuel Trélat	Président

Remerciements

Je souhaiterais remercier tous les membres de Jury, et en particulier les rapporteurs, pour le temps qu'ils ont consacré à ce travail d'HDR. Je suis honoré de leur présence dans ce jury.

Je souhaiterais également remercier tous les collaborateurs que j'ai pu avoir à ce jour: j'ai eu beaucoup de plaisir à travailler avec tous. Je pense en particulier aux étudiants qui m'ont accordé leur confiance.

J'aimerais remercier le centre de Robotique, et plus largement mon institution Mines ParisTech pour me permettre de travailler dans de si bonnes conditions et pour me laisser si libre dans mes choix.

Finalement, j'aimerais remercier mes proches, et tout particulièrement mon épouse, pour leur soutien et leur compréhension de ma passion pour la science.

Résumé

Ce mémoire présente les travaux que j'ai effectués à Mines ParisTech au centre Automatique et Systèmes pendant ma thèse de 2004 à 2007, puis à l'Université de Liège en 2008, et à Mines ParisTech depuis 2009 au centre de Robotique.

Le problème du filtrage consiste à estimer l'état interne d'un système à partir d'une part d'un modèle d'évolution du système, et d'autre part de mesures bruitées ou incomplètes de l'état de ce système mais accessibles expérimentalement. Le filtrage non-linéaire est une partie des processus stochastiques qui a suscité de nombreux travaux théoriques, mais la synthèse de filtres utilisables en pratique rencontre encore d'importantes difficultés. Les ingénieurs utilisent généralement le filtre de Kalman étendu, pour lequel on a peu de garanties théoriques et qui est sujet à divers problèmes. Le fil conducteur de nos travaux sur ce sujet consiste à concevoir des filtres non-linéaires pour certains problèmes pratiques de filtrage, en utilisant des méthodes issues de la géométrie différentielle. Ces travaux sont regroupés en deux parties. La première traite de la synthèse de filtres pour estimer l'état interne d'un système dynamique, lorsque ce dernier admet des symétries. La seconde partie traite essentiellement du filtrage de mesures bruitées sur la variété des matrices semi-définies positives de rang fixé (et éventuellement faible). Pour ce problème on cherchera à munir cette variété de métriques pertinentes, et l'on formulera les problèmes de débruitage comme des problèmes d'optimisation sur les variétés.

Il est intéressant de remarquer que chaque partie s'appuie sur l'un des deux principaux courants de géométrie non-euclidienne généralisant les travaux de Gauss sur les surfaces. Le premier est due à Klein avec le programme d'Erlangen, et les groupes de Lie. L'idée sous-jacente est qu'une géométrie est caractérisée par ses automorphismes, codés par des actions de groupe. Les géométries obtenues possèdent une structure très "homogène" du fait des symétries. Aussi, les systèmes qui admettent des symétries font preuve d'une certaine rigidité qui leur confère de fortes propriétés géométriques dont on tirera partie dans notre travail. Le deuxième courant a été proposée par Riemann dans son "Habilitationsschrift". Il s'agit d' une approche très flexible de la notion d'espace (variété différentiable), dans laquelle on spécifie en chaque point un élément de longueur qui permet de mesurer des distances entre points proches (une métrique riemannienne). Dans les problèmes de filtrage non-linéaires considérés dans la seconde partie, on s'efforcera d'interpréter certaines non-linéarités du problème comme l'appartenance de l'état à une variété riemannienne, et on cherchera à munir cette variété de métriques afin de pouvoir définir des opérations élémentaires de filtrage (calcul de moyenne, interpolation) intrinsèques.

Ce document est organisé de la façon suivante. Dans la première partie, nous allons nous intéresser au filtrage invariant, c'est-à-dire à la conception de filtres pour des systèmes dynamiques qui possèdent des symétries (invariances). L'exemple phare qui motive ces développements relève du domaine de la navigation inertielle, où un porteur (véhicule aérien, terrestre, sous-marin) cherche à estimer sa position et son orientation dans l'espace, à partir de données inertielles (gyromètres, accéléromètres), ainsi qu'à partir de mesures de divers capteurs (GPS, baromètre, vitesse de l'air, caméras, télémètre laser). Les symétries découlent naturellement des invariances galiléennes. L'explosion récente de capteurs bas-coût équipant les mini-drônes, et les besoins en algorithmes simples engendrés, ont valu à cet exemple d'être l'objet de nombreux travaux récents.

Dans le premier chapitre, on s'intéressera à la synthèse d'observateurs invariants. Un observateur est un estimateur qui fournit en temps réel une estimation de l'état interne d'un système dynamique à partir de mesures incomplètes et indirectes de l'état, dans un cadre déterministe. Lorsque le système possède des propriétés d'invariance, on peut imposer que l'observateur préserve ces symétries. Les observateurs invariants ne sont apparus il n'y a qu'une dizaine d'années [3]. Cela peut paraître étonnant car en statistiques, la notion d'estimateur invariant est connue depuis longtemps, et est utilisée comme une façon de restreindre la classe dans laquelle on cherche un estimateur optimal. De plus, l'invariance implique que l'erreur minimale que l'on peut espérer commettre avec un estimateur invariant ne dépend que d'un nombre réduit de composantes de l'état (i.e. elle est constante le long des orbites). Ainsi, nous verrons que les observateurs invariants apparaissent comme des analogues des estimateurs invariants pour un système dynamique déterministe, puisqu'ici aussi, il s'agit de restreindre la classe des estimateurs considérés, et ici aussi l'erreur d'estimation possède certains propriétés d'indépendance par rapport au point de fonctionnement.

Dans le deuxième chapitre, on introduit une approche probabiliste du filtrage invariant qui tient compte explicitement des bruits de mesure. On considérera en effet un processus stochastique admettant des propriétés d'invariance, et on supposera que des mesures incomplètes, indirectes et bruitées de l'état sont disponibles. On cherchera alors à donner la meilleure approximation possible de l'état conditionnellement aux mesures. En pratique, le filtre le plus populaire est le filtre de Kalman étendu, dont on proposera une version invariante. Dans le cas particulier où l'état évolue sur un groupe de Lie, on proposera une classe de filtres invariants pour lesquels on peut montrer que la distribution de probabilité de l'erreur d'estimation oublie sa condition initiale et converge vers une distribution fixe.

Finalement le troisième chapitre est consacré à certains résultats théoriques et également expérimentaux autour des filtres invariants obtenus ces dernières années, ainsi qu'à des ouvertures vers d'autres problèmes pouvant être reliés au filtrage invariant.

Dans la deuxième partie, nous étudierons la géométrie de l'espace des matrices semi-définies positives, pour pouvoir définir des opérations élémentaires de filtrage sur cet espace. En effet, les matrices définies positives sont devenues des outils fondamentaux de calcul matriciel dans plusieurs domaines de l'ingénierie et des mathématiques appliquées. Elle apparaissent comme matrices de covariance en statistiques, comme des éléments de l'espace de recherche en optimisation convexe et programmation semi-

définie, comme des noyaux en apprentissage automatique, comme tenseurs de diffusion en imagerie médicale. Elles jouent également un rôle fondamental en théorie du contrôle à travers leur utilisation dans la théorie des systèmes linéaires : filtrage optimal (équations de Kalman et Riccati), étude de stabilité et construction de modèles d'états équilibrés (équation de Lyapunov), commande optimale LQ (équation de Riccati), reformulation sous la forme d'inégalités matricielles linéaires (LMIs). On peut finalement citer le domaine de l'information quantique où les matrices définies positives hermitiennes apparaissent comme matrices de densité. Bon nombre d'algorithmes dans ces différentes disciplines font usage d'une géométrie riemannienne naturelle sur le cône des matrices positives. Mais l'application des algorithmes à des problèmes de grande taille est une motivation centrale dans les mathématiques appliquées d'aujourd'hui à travers le traitement de bases de données dont la taille ne cesse d'augmenter. Or, dans un algorithme itératif, les opérations matricielles courantes associées à une matrice carrée de taille $n \times n$ telles que le calcul de valeurs propres ou l'inversion ont un coût numérique d'ordre $O(n^3)$. Mais pour les problèmes de très grande taille cette complexité doit être au plus linéaire $O(n)$. Une solution consiste à approximer la matrice $n \times n$ par une matrice de rang faible $p \ll n$ qui peut être manipulée par des opérations de complexité $O(np^2)$.

Ici on s'intéressera essentiellement à la variété des matrices semi-définies positives de taille $n \times n$ et de rang p fixé. On cherchera à munir cette variété de métriques raisonnables. Une multitude de métriques étant possible, on restreindra notre recherche à des métriques possédant certaines propriétés d'invariance. Ensuite, on considérera des problèmes où l'on cherche à estimer/filtrer une matrice de cette variété à partir de mesures bruitées.

Plus précisément, on commencera par proposer dans le chapitre 4 une géométrie riemannienne pour les matrices semi-définies positives de rang fixé. Puis on proposera dans le chapitre 5 une notion de moyenne (géométrique) sur cette variété. En effet, l'opération d'interpolation joue un rôle central dans le filtrage d'un signal bruité, puisque lorsqu'on moyenne des variables aléatoires indépendantes on tend à réduire les fluctuations. Finalement, on proposera dans le chapitre 6 une méthode générale de descente de gradient stochastique sur les variétés riemanniennes. On obtiendra certains résultats généraux de convergence, et l'on appliquera la méthode à un problème d'estimation d'une grande matrice semi-définie positive de rang faible dont on mesure aléatoirement certaines coordonnées. Le problème s'apparente à de la complétion de matrice en ligne.

Contents

Résumé **5**

I Filtrage invariant 1

1 Estimateurs invariants, observateurs invariants **3**
1.1 Estimateurs invariants en statistiques 4
1.2 Observateurs invariants déterministes 6

2 Filtres invariants stochastiques **11**
2.1 Le problème du filtrage non-linéaire 12
2.2 Filtre de Kalman étendu invariant (IEKF) 12
2.3 Cas particulier d'une dynamique sur un groupe 15

3 Divers résultats autour des observateurs invariants **19**
3.1 Résultats en théorie du contrôle 19
3.2 Un observateur invariant en dimension infinie 20
3.3 Deux observateurs non-linéaires contractants 22
3.4 Applications en robotique mobile 24

II Filtrage sur les variétés riemanniennes et rang faible 27

4 Géométrie des matrices semi-définies positives **29**
4.1 Quelques variétés matricielles d'intérêt 31
4.2 Géométrie des matrices SDP de rang fixé 33
4.3 Une application: le filtre de Kalman de rang faible 36

5 Moyennes de matrices semi-définies positives **39**
5.1 Moyenne géométrique de matrices semi-définies positives 40
5.2 Tenseurs de diffusion et anisotropie 44

6 Gradient stochastique riemannien **47**
6.1 Méthode générale et résultats 47
6.2 Régression dans l'espace des matrices de rang faible 51

Conclusion et perspectives **55**

Part I

Filtrage invariant

Chapter 1

Estimateurs invariants, observateurs invariants

Un observateur fournit en temps réel une estimation de l'état interne d'un système dynamique, à partir de mesures incomplètes et indirectes de l'état accessibles expérimentalement, dans un cadre déterministe. Lorsque le système possède des propriétés d'invariance, on peut imposer que l'observateur préserve ces symétries. Les observateurs invariants ne sont apparus qu'assez tardivement [3]. Cela peut paraître assez étonnant car en statistiques, la notion d'estimateur invariant est bien établie (voir e.g. [78, 18]), et est utilisée comme une façon de limiter la recherche d'estimateur optimal à une classe restreinte d'estimateurs raisonnables. De plus, l'invariance implique que l'erreur minimale que l'on peut espérer commettre avec ces estimateurs ne dépend que d'un nombre réduit de composantes du point de fonctionnement autour duquel on l'étudie (i.e. elle est constante le long des orbites de l'action du groupe de symétries). Aussi, nous verrons que les observateurs invariants peuvent être vus comme l'analogue des estimateurs invariants pour un système dynamique déterministe, puisqu'ici aussi, il s'agit de restreindre la classe des estimateurs considérés, et ici aussi l'erreur d'estimation possède certains propriétés d'indépendance par rapport à la trajectoire autour de laquelle on l'étudie.

L'introduction par la notion d'estimateur invariant en statistiques, qui se réfère au cas bruité statique, nous semble posséder des vertus pédagogiques. Le parallèle entre estimateurs invariants en statistiques et observateurs invariants en théorie du contrôle permet de souligner les ressorts fondamentaux de l'invariance et ce qu'on est en droit d'en attendre pour l'estimation, à savoir: tout d'abord le fait que l'invariance peut fournir un a priori sur la forme de ce qu'on cherche (une solution d'équation différentielle, une fonction etc.), et restreindre énormément l'espace de recherche. Cette utilisation de l'invariance est récurrente en physique, comme le livre [58] en fournit une belle illustration par exemple, en retrouvant de façon quasiment systématique l'expression du lagrangien d'une particule isolée dans l'espace, en ayant postulé les invariances galiléennes, et le principe de moindre action. Ensuite, les invariances sont codées dans des actions de groupe, qui définissent des transformations du système le laissant inchangé, donc, le long d'une orbite du groupe (la classe des points qu'on peut relier par une telle transformation) le système paraît inchangé (ou encore "invariant"). Ainsi, on peut espérer, même s'il faut se méfier des conclusions hâtives en la matière, que certaines propriétés

obtenues en un point soient vérifiées en chaque point de son orbite.

1.1 Estimateurs invariants en statistiques

Supposons que $X \in \mathcal{X}$ est une variable aléatoire (vectorielle) dont la loi $P(x|\theta)$ dépend d'un paramètre a priori inconnu (l'état) appartenant à un espace Θ. Chaque mesure de X apporte une information sur θ et un estimateur $\hat{\theta}(x)$ est une fonction qui à une (ou plusieurs) réalisation associe une valeur possible du paramètre. Il est généralement impossible de déterminer θ à partir d'une seule mesure, mais on peut espérer avoir une estimation correcte à partir d'un échantillon suffisamment large. La fonction de perte $L(\hat{\theta}(x), \theta)$ est une mesure de l'écart entre état estimé et état vrai, qu'on supposera être partout positive. Le risque associé à un état θ et un estimateur $\hat{\theta}(x)$ est la perte moyennée sur un grand nombre de mesures. Il quantifie la précision de l'estimateur et s'écrit

$$R(\theta, \hat{\theta}) = \int L(\theta, \hat{\theta}(x)) P(x|\theta) dx$$

Le principe d'invariance en statistiques est un concept attractif pour la synthèse d'estimateurs, connu depuis longtemps (on parle parfois de statistiques équivariantes), et que l'on appréhende sans doute mieux par un exemple. Nous empruntons l'exemple suivant à [18].

Example 1. *On considère la loi exponentielle régissant une désintégration radioactive*

$$P(x|\theta) = \theta^{-1} \exp(\frac{-x}{\theta}) I_{\{x \geq 0\}}$$

On aimerait estimer θ sur la base d'une ou de plusieurs observations. On propose alors comme mesure d'erreur la perte relative $L(\theta, \hat{\theta}) = \log^2(\hat{\theta}/\theta)$. Si l'on mesure la décroissance en minutes plutôt qu'en secondes, le résultat d'observation sera $Y = X/60$, et la densité associée à cette nouvelle variable s'écrit

$$P(y|\theta) = \eta^{-1} \exp(\frac{-y}{\eta}) I_{\{x \geq 0\}}$$

où $\eta = \theta/60$ est le temps moyen de décroissance exprimé en minutes.

La structure du problème est invariante par changement d'échelle temporelle, et il paraît raisonnable de demander à l'estimateur de ne pas dépendre d'un choix arbitraire d'unité de temps. Mais cette contrainte introduit une grande rigidité, car l'estimateur doit se comporter de la même façon quelque soit le choix des unités, ce choix n'étant pas connu à l'avance. Ainsi, si la mesure est faite en minutes, l'estimateur donne une valeur $\hat{\theta}(y) = \hat{\theta}(x/60)$ qui doit correspondre à sa valeur si la mesure avait été faite en secondes, mais convertie en minutes, i.e. $\hat{\theta}(x)/60$. Cela implique $\hat{\theta}(x/60) = \hat{\theta}(x)/60$. Pour que la relation soit vraie quelque soit le choix des unités l'estimateur doit être une fonction homogène de degré 1 de la mesure i.e.

$$\hat{\theta}(x) = x\hat{\theta}(1) = \lambda x$$

ce qui restreint la recherche d'estimateur à la classe des fonctions linéaires de l'observation. La formalisation de l'exemple ci-dessus implique la notion de groupe de transformations.

Definition 1. *Soit G un groupe de Lie. On dit que G est un groupe de transformations d'un espace M si il existe une famille de difféomorphimes $\phi_g(x)$ de M dans lui-même, paramétré par les éléments $g \in G$, et tel que*

- $\phi_{g_2}\big(\phi_{g_1}(x)\big) = \phi_{g_2 g_1}(x)$ *pour tout* g_1, g_2, x.
- $\phi_e(x) = x$ *pour tout* x *où* e *est l'élément neutre de* G.

Dans l'exemple ci-dessus de changement d'unité métrique, G est le groupe des homotéties de rapport positif, i.e. $\mathbb{R}_+^*$ muni de la multiplication. Pour tout $g > 0$ l'action du groupe s'écrit $\phi_g(x) = x/g$.

Definition 2. *Une famille de densités $\big(P(x|\theta)\big)_{\theta\in\Theta}$ avec $x \in \mathcal{X}$, est dite invariante sous l'action d'un groupe G de transformations de $\mathcal{X}$ si pour tout $g \in G$, et $\theta \in \Theta$, il existe un difféomorphisme $\rho_g : \Theta \to \Theta$ tel que la variable $Y = \phi_g(X)$ a pour densité $P(y|\rho_g(\theta))$. La fonction de perte est dite invariante si $L(\theta_1, \theta_2) = L(\rho_g(\theta_1), \rho_g(\theta_2))$ pour tout g. Finalement un estimateur est invariant si c'est une fonction dite équivariante de la mesure, i.e. pour tout $g \in G$ on a $\hat{\theta}(\phi_g(x)) = \rho_g(\hat{\theta}(x))$.*

Le résultat immédiat et important que l'on est en droit d'attendre, est que la précision d'un estimateur invariant est indépendante de certains choix arbitraires faits par l'utilisateur lors de la mise en équation, comme le prouve le théorème suivant (voir e.g. [18]):

Theorem 1. *Si la famille de densités et la fonction de perte sont invariantes, le risque associé à un estimateur invariant $\hat{\theta}$ est constant le long des orbites du groupe, où l'on définit l'orbite d'un élément θ comme l'ensemble $\{\rho_g(\theta)\ ;\ g \in G\}$.*

Mathématiquement, le résultat s'écrit

$$\forall\theta \in \Theta\ \forall g \in G \quad R(\rho_g(\theta), \hat{\theta}) = R(\theta, \hat{\theta})$$

L'intérêt pratique est le suivant: si un estimateur invariant est optimal (i.e. minimise le risque) associé à un paramètre θ, on a la garantie qu'il est optimal pour l'ensemble des paramètres de l'orbite de θ. Dans le cas particulier d'une action transitive, i.e. si il n'existe qu'une seule orbite, un estimateur invariant optimal en un point le sera en tout point.

L'invariance contraint le choix des estimateurs, en liant rigidement leur comportement autour de différents points de fonctionnement. La conséquence est que les garanties obtenues autour d'un point particulier s'étendent immédiatement à un ensemble de points, éventuellement très large. C'est précisément ce type de propriétés qui fait la force des observateurs invariants en filtrage non-linéaire où des garanties sont dures à obtenir. Ils représentent donc une contrepartie des estimateurs invariants, mais pour des systèmes dynamiques.

Notons que les observateurs invariants [4, 21] ne sont apparus qu'assez longtemps après les articles fondateurs de Kalman et Luenberger [56, 65] sur les observateurs. Mais les observateurs invariants sur les groupes de Lie pour l'estimation d'orientation de drône ont été l'objet de divers travaux ces dernières années, e.g. [66, 88]. Remarquons finalement que les idées d'invariance ci-dessus ne se limitent pas au domaine de l'estimation, et peuvent aussi être déclinées à certains problèmes d'optimisation (voir e.g. [8] et notre travail [69]).

1.2 Observateurs invariants déterministes

Dans cette section nous faisons quelques rappels sur la théorie des observateurs invariants telle que construite dans [28]. Considérons un système dynamique sous forme d'état

$$\frac{d}{dt}x = f(x,u) \tag{1.1}$$

$$y = h(x) \tag{1.2}$$

où $x(t) \in \mathcal{X} \subset \mathbb{R}^n$ est un vecteur qui représente l'état du système a priori inconnu, $u \in \mathcal{U} \subset \mathbb{R}^m$ est l'entrée, et $y \in \mathcal{Y} \subset \mathbb{R}^p$ la sortie. La première équation est une équation différentielle ordinaire qui modélise l' évolution du système dans le temps. $y(t)$ est la sortie du système, c'est-à-dire une certaine fonction h de l'état $x(t)$ à l'instant t, disponible à l'utilisateur à travers les mesures effectuées par des capteurs. Un observateur est un système dynamique dont l'évolution est régie par une équation différentielle de la forme

$$\frac{d}{dt}\hat{x} = F(\hat{x},u,y), \tag{1.3}$$

implémentable en temps réel par l'utilisateur et dont l'état $\hat{x}$ est censé fournir une estimation de l'état vrai x. On parle d'observateur asymptotique lorsque $E(\hat{x}(t),x(t)) \to 0$ quand $t \to \infty$ pour une certaine fonction d'écart E entre l'état et son estimée.

Théorie Nous commençons avec quelques définitions.

Definition 3. *Considérons un groupe de transformations $(\phi_g)_{g\in G}$ de M. On dit que c'est un groupe de symétries d'un système d'équations différentielles sur M si chacune des transformations ϕ_g envoie chaque fonction solution du système d'équations dans une autre solution de ce système. On dit alors que le système est invariant.*

Definition 4. *Un champ de vecteurs w sur M est dit invariant si le système $\frac{d}{dt}z = w(z)$ est invariant.*

Definition 5. *Un invariant scalaire est une fonction $I_s : M \to \mathbb{R}$ telle que pour tout $g \in G$ on a $I_s(\phi_g(x)) = I_s(x)$, c'est-à-dire une fonction constante le long des orbites. Un ensemble complet d'invariants scalaires est un vecteur dont chaque composante est un invariant scalaire, et tel que tout invariant scalaire est une fonction de ses composantes.*

Considérons un groupe de transformations de $\mathcal{X} \times \mathcal{U}$ défini pour $x \in \mathcal{X}, u \in \mathcal{U}, g \in$ par

$$\phi_g(x,u) = \big(\varphi_g(x), \psi_g(u)\big), \tag{1.4}$$

où $(\varphi_g)_{g\in G}$ et $(\psi_g)_{g\in G}$ sont des groupes de transformations de resp. $\mathcal{X}$ et $\mathcal{U}$.

Definition 6. *Le système $\frac{d}{dt}x = f(x,u)$ est dit invariant s'il est invariant sous l'action de groupe* (1.4).

Le groupe envoie les solutions de l'équation différentielle dans des solutions si l'on a $\frac{d}{dt}X = f(X,U)$, où $(X,U) = (\varphi_g(x), \psi_g(u))$ pour tout $g \in G$. Par conséquent u représente les entrées habituelles, mais peut aussi représenter n'importe quelle donnée de l'environnement devant être transformée pour rendre le système invariant (par exemple des vitesses dont le système dépendrait doivent être transformées en conséquence quand on change d'unité métrique).

Definition 7. *Un observateur déterministe est invariant si le système le définissant* (1.3) *est invariant sous l'action du groupe* $(\hat{x}, x, u, y) \mapsto \big(\varphi_g(x), \varphi_g(\hat{x}), \psi_g(u), h(\varphi_g(x), \psi_g(u))\big)$.

Dans ce cas la structure de l'observateur mime la structure non-linéaire du système.

Definition 8. *La sortie est équivariante si il existe un groupe de transformations* $(\rho_g)_{g\in G}$ *tel que* $h(\varphi_g(x), \psi_g(u)) = \rho_g(h(x,u))$ *pour tous* g, x, u.

Si la fonctions f et h sont C^∞, que la sortie est équivariante, et que l'action du groupe est libre et régulière, on sait construire localement (c'est-à-dire pour g dans un voisinage de e) la forme de l'ensemble des observateurs invariants, voir [28]. La construction repose sur la méthode du repère mobile de Cartan [44] (voir aussi [73] pour une introduction pédagogique à cette dernière notion). Les propriétés remarquables de ces observateurs apparaissent grâce à l'introduction de ce qu'on appelle une erreur d'état invariante.

Definition 9. *L'application régulière* $(\hat{x}, x) \mapsto \eta(\hat{x}, x) \in \mathbb{R}^n$ *est appelée erreur d'état invariante si l'application* $x \mapsto \eta(\hat{x}, x)$ *est inversible pour tout* $\hat{x}$, *si* $\eta(x,x) = 0$ [erreur] *et si* $\eta(\varphi_g(\hat{x}), \varphi_g(x)) = \eta(\hat{x}, x)$ [invariance].

Finalement nous avons le résultat important suivant [28]

Theorem 2. *Si* $\hat{x}$ *est l'état d'un observateur invariant, la dynamique de l'erreur* η *ne dépend que de* η *et des invariants scalaires de* x, u *i.e.*

$$\frac{d}{dt}\eta(\hat{x}, x) = \Gamma\big(\eta(\hat{x}, x), I(x,u)\big)$$

où Γ *est une fonction régulière et* I *est un ensemble complet d'invariants de l'action de* G *sur* $\mathcal{X} \times \mathcal{U}$.

Corollary 1. *Sous les mêmes conditions l'équation d'erreur linéarisée, c'est-à-dire l'équation vérifiée par* $\delta\eta$ *qui est une approximation au premier ordre de la différence* $\eta - 0$, *s'écrit*

$$\frac{d}{dt}\delta\eta = \Lambda(\delta\eta, I(u,x))$$

L'intérêt est le suivant: si un observateur invariant est asymptotiquement convergent (i.e. $\eta(\hat{x}(t), x(t)) \to 0$ quand $t \to \infty$) autour d'une trajectoire donnée, il sera convergent autour d'un ensemble de trajectoires: considérons l'ensemble des trajectoires $x(t), u(t)$ sur lesquelles I est indépendant du temps i.e.,

$$\forall t > 0 \quad I(x(t), u(t)) \equiv I_0$$

où I_0 est un vecteur fixé. De telles trajectoires sont appelées "permanentes" [29] et ont souvent un sens physique bien établi. Autour de ces trajectoires, l'équation d'erreur linéarisée est autonome. Par conséquent, il est souvent aisé de construire un observateur dépendant explicitement de I (on peut utiliser $I(\hat{x}, u)$ dans l'implémentation puisque l'erreur est de toute façon linéarisée autour de la vraie trajectoire) tel que l'erreur linéarisée tende vers 0 asymptotiquement autour de *toute* trajectoire associée à I_0.

Cas particulier d'une dynamique sur un groupe de Lie Considérons maintenant une dynamique invariante à gauche sur un groupe de Lie G. Pour simplifier l'exposé, on va supposer une dynamique un peu moins générale que dans l'article [29] et de la forme

$$\frac{d}{dt}g = DL_g\omega, \quad y = h(g)$$

où $L_g(x) = gx$ est la multiplication à gauche sur le groupe par g, et DL_g l'application tangente induite. La dynamique est invariante sous l'action de la multiplication à gauche, i.e., $\phi_g = L_g$. L'application de sortie est $h : G \to \mathbb{R}^p$ et on suppose qu'elle est équivariante. On parlera de sortie équivariante à gauche si pour tous $g_1, g_2 \in G$ on a $h(g_1 g_2) = \rho_{g_1} h(g_2)$, et de sortie équivariante à droite si pour tous $g_1, g_2 \in G$ on a $h(g_1 g_2) = \rho_{g_2} h(g_1)$.

Example 2. [fondamental] *Soit $G = SO(3)$ le groupe des rotations de l'espace euclidien ordinaire, i.e. l'espace de configuration d'un solide ayant un point fixe. Le mouvement du solide est alors décrit par une courbe sur G et constitue un prototype de dynamique invariante à gauche sur un groupe*

$$\frac{d}{dt}R = R(\omega)_\times$$

où $\omega \in \mathbb{R}^3$ est le vecteur vitesse angulaire exprimée dans le repère du solide, et où l'on note $(b)_\times \in \mathbb{R}^{3\times 3}$ la matrice antisymétrique associée à l'opérateur produit vectoriel avec b, i.e. pour $a, b \in \mathbb{R}^3$ on a $(b)_\times a = b \times a \in \mathbb{R}^3$. Ici, on peut voir R comme la rotation qui transforme un vecteur exprimé dans le repère du solide en ce même vecteur mais exprimé dans un repère fixe attaché au sol. Si l'on considère par exemple la mesure par des magnétomètres attachés au solide du champ magnétique terrestre (vecteur B supposé constant dans le repère fixe attaché au sol)

$$y = h(R) = R^{-1}B$$

on obtient une sortie équivariante à droite.

Nous allons considérer pour ce problème une classe d'observateurs invariants à droites de la forme suivante

$$\frac{d}{dt}\hat{g} = DL_{\hat{g}}\omega + DR_{\hat{g}}k(\rho_{\hat{g}^{-1}}y) \tag{1.5}$$

où $\hat{g}$ représente l'état estimé, $DR_{\hat{g}}$ l'application tangente à l'opération de multiplication à droite par $\hat{g}$ sur le groupe, et où k est une fonction de l'espace de sortie vers l'espace tangent à l'identité: l'algèbre de Lie. L'invariance à droite de cet observateur ne paraît pas évidente mais peut être justifiée en laissant le groupe agir sur l'entrée ω. On a le résultat important suivant [29]

Theorem 3. *Considérons un système invariant à gauche sur un groupe de Lie, avec une sortie équivariante à droite. Considérons un observateur de la forme* (1.5)*. Alors l'erreur d'état invariante à droite $\eta = \hat{g}g^{-1}$ vérifie une équation autonome*

$$\frac{d}{dt}\eta = \Gamma(\eta)$$

totalement indépendante de $\omega(t)$.

L'intérêt est immédiat. Tout d'abord, un observateur qui stabilise localement asymptotiquement l'erreur à $\eta(e)$ sera convergent autour de *toute* trajectoire. De plus, on peut espérer construire plus facilement des observateurs globalement convergents, puisqu'il "suffit" d'obtenir une fonction de Lyapunov pour η. En pratique il reste tout de même rare d'obtenir une convergence globale. La propriété mise en lumière par le théorème a été utilisée dans [57] où les auteurs ont construit des observateurs invariants gradient destinés à faire décroître l'erreur de sortie, et ont obtenu des résultats généraux sur la convergence de la sortie estimée vers la sortie mesurée.

Chapter 2

Filtres invariants stochastiques

Le filtrage non-linéaire est une sous-partie de la théorie des processus stochastiques. Il a fait l'objet de nombreux travaux de recherche essentiellement des années soixante à la fin du vingtième siècle, et de nombreux résultats mathématiques sont disponibles [74]. En revanche, la conception d'algorithmes de filtrage efficaces, implémentables en pratique, et possédant de bonnes garanties théoriques a connu et connaît encore de sérieuses difficultés. Le filtre de Kalman étendu a été l'algorithme dominant pendant de nombreuses années, car il produit de bons résultats en pratique, mais il ne possède pas de garanties de convergence, et il peut être mis en défaut sur des exemples très simples (même sur un système de très petite dimension plusieurs linéarisations autour de la trajectoire estimée peuvent conduire à des résultats aberrants si celle-ci est éloignée de la vraie trajectoire). Depuis plus d'une décennie, les approches numériques de type filtrage particulaire ont permis d'énormes progrès dans la qualité des estimations, et ont connu un succès croissant dans les applications. Cela dit, elles requièrent généralement des moyens de calculs importants qui souvent ne sont pas disponibles à bord, lorsqu'il s'agit de robotique mobile et de navigation, et un certain savoir-faire que l'utilisateur ne possède pas forcément.

Dans ce chapitre, nous nous intéressons au problème du filtrage non-linéaire lorsque le système admet certaines propriétés d'invariance. On présente d'abord une méthode d'intérêt pratique, le filtre de Kalman invariant, introduit pour la première fois dans [20] et développé et appliqué depuis dans plusieurs travaux. Il s'agit simplement d'"invariantiser" les équations du filtre de Kalman lorsque le système considéré admet des symétries. Dans un deuxième temps, nous évoquons des travaux récents effectués dans le cadre de la thèse d'Axel Barrau. On considère un système bruité invariant à gauche sur un groupe de Lie, on propose une discrétisation exacte des équations, puis on étudie une classe de filtres invariants discrets ayant la propriété remarquable que quelque soit la distribution initiale de l'erreur d'état, cette dernière converge vers une distribution fixe.

2.1 Le problème du filtrage non-linéaire

Mathématiquement le problème du filtrage se pose ainsi. On considère un système dynamique dont la trajectoire est perturbée par des termes fluctuants

$$\frac{d}{dt}x = f(x,u) + L(t)$$

où $x \in \mathcal{X} \subset \mathbb{R}^n$, $u \in \mathcal{U} \subset \mathbb{R}^m$, et $L(t)$ est un terme aléatoire de type bruit blanc (dérivée au sens des distributions d'un processus de Wiener) ou force de Langevin [59] qui représente certaines perturbations ou imperfections du modèle. Cette équation peut être ré-écrite dans le formalisme moderne de la théorie des processus stochastiques

$$dx_t = f(x_t,u_t)dt + g(x_t)dw_t \tag{2.1}$$

où w_t est un processus de Wiener n-dimensionnel standard et l'équation est à interpréter au sens d'Itô ou de Stratonovich. Les capteurs mesurent un signal de sortie corrompu par un bruit généralement supposé blanc (des parasites du signal)

$$y_t = h(x_t) + R(t)$$

Afin d'éviter de faire appel à la théorie des distributions on formule généralement le problème du filtrage en utilisant de façon équivalente comme sortie le processus

$$y_t = \int_0^t h(x_s)ds + v_t$$

où v_t est un processus de Wiener (que l'on suppose standard pour simplifier mais qui est généralement prémultiplié par une matrice). Le problème du filtrage consiste à calculer à chaque instant la loi conditionnelle de x_t sachant $\mathcal{Y}_t = \sigma\{y_s; 0 \leq s \leq t\}$, cette dernière tribu représentant l'information que l'on a observée jusqu'à l'instant t. Dans le cas où f est l'addition d'une fonction linéaire de x et d'une fonction linéaire de u, g est une matrice constante, h une fonction linéaire, et la distribution initiale de l'état x_0 une gaussienne, il est bien connu que la solution du problème de filtrage $\hat{x}(t)$ est la solution d'une équation différentielle stochastique (EDS): l'équation du filtre de Kalman-Bucy [56] (détaillée plus loin au chapitre 4). Cette équation peut être discrétisée et facilement implémentée en pratique (filtre de Kalman). Mais dans le cas général, rien ne garantit qu'il existe un filtre de dimension finie implémentable qui résout le problème du filtrage.

2.2 Filtre de Kalman étendu invariant (IEKF)

Lorsque le système considéré possède des invariances, les équations du filtre de Kalman étendu peuvent êtres rendues invariantes assez aisément. On obtient alors le filtre de Kalman étendu invariant (dont l'abréviation en anglais est "IEKF") introduit dans [20, 31]. Nous allons présenter très schématiquement le principe. On reprend les équations

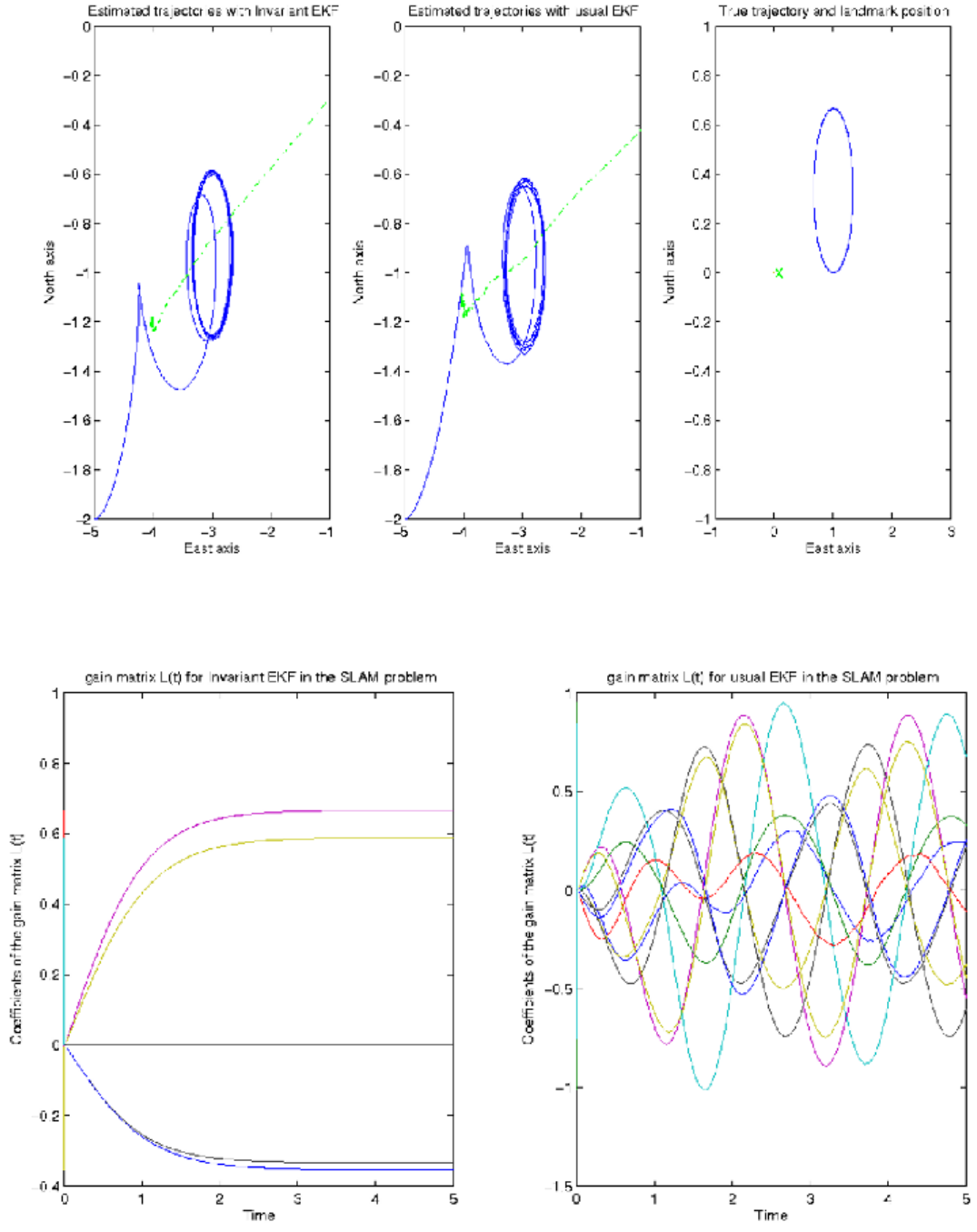


Figure 2.1: Figures extraites de l'article [23], où un robot décrit une trajectoire circulaire en observant avec du bruit la position dans son propre repère d'un point fixe dans le repère attaché au sol, et cherche à retrouver sa trajectoire et la position du point (à une rotation-translation près). Les résultats sont commentés dans le corps du texte.

ci-dessus mais où les bruits sont vus comme des perturbations déterministes qui entrent dans le système. On obtient le système

$$\frac{d}{dt}x = f(x,u) + L(x,t), \quad y = h(x) + R(x,t)$$

Si ce système est invariant sous l'action d'un groupe au sens de la définition 6, on peut construire un observateur invariant de la forme $\frac{d}{dt}\hat{x} = F(\hat{x},u,y)$ où la fonction F n'utilise pas le terme de perturbation L supposé inconnu. On obtient alors une équation d'erreur invariante qui peut être linéarisée (on ne retient que les termes de premier ordre en $\hat{x}-x$). On obtient alors un système d'erreur linéaire contenant certains termes de bruit. En supposant alors que ces termes qui entrent dans le système d'erreur linéarisé sont des bruits gaussiens, on peut construire un filtre de Kalman-Bucy pour le système linéarisé. En implémentant le gain obtenu sur le modèle non-linéaire, comme classiquement dans la méthode du filtre de Kalman étendu, on obtient un filtre de Kalman étendu *invariant*.

Ce filtre hérite des propriétés des observateurs invariants: autour de toute trajectoire permanente l'équation d'erreur est autonome, et par conséquent les coefficients qui apparaissent dans l'équation de Riccati permettant de calculer le gain du filtre sont indépendants du temps. Sous certaines hypothèses d'observabilité et de détectabilité, cela entraîne la convergence asymptotique du gain vers une matrice fixée, ainsi que de la matrice de covariance estimée de l'erreur. Cela confère donc au filtre des garanties théoriques de convergence autour d'une série de trajectoires qu'il ne possède pas en général.

Quelques remarques sur la méthode Tout d'abord, le filtre de Kalman invariant n'a pas été construit dans le cadre du formalisme moderne des processus stochastiques. Considérer les bruits comme des entrées déterministes, transformer les équations via une série de calculs, puis calculer des moyennes et variances sur les formules finales en supposant les bruits aléatoires revient à utiliser la technique élaborée par Langevin [59] devenue standard en physique [87]. Elle suggère qu'on ait implicitement opté pour une interprétation de l'équation d'évolution (2.1) au sens de Stratonovich, selon le résultat de Wong et Zakai [90], ou Sussman [85], qui montre que lorsqu'on résout l'équation différentielle non-stochastique associée pour chaque trajectoire particulière $w_t(\omega)$ d'une approximation différentiable du processus de Wiener w_t, la solution obtenue est une approximation de la solution au sens de Stratonovich. Ensuite, il faut dire que lorsqu'on linéarise le système, on néglige en moyenne certains termes où le bruit entre de façon multiplicative $\delta\eta L(t,x)$ considérés comme des termes de "second ordre", c'est-à-dire petits devant l'erreur $\delta\eta$ ou devant les termes de type $L(t,x)$. Cela repose sur une hypothèse de bruits faibles, somme toute très classique en pratique, et explicitée dans l'article [31]. Même si on peut argumenter que les approches de type filtre de Kalman étendu sont de toute façon basées sur une série d'approximations, ces linéarisations mériteraient d'être revues pour prendre ces termes en compte de façon plus satisfaisante.

Applications de l'IEKF Le filtre de Kalman étendu a été appliqué à plusieurs reprises. Ses performances se sont avérées comparables ou légèrement meilleures que

le filtre de Kalman étendu conventionnel, mais il possède d'intéressantes propriétés. D'une part les équation du filtres sont telles que certaines contraintes d'état sont automatiquement préservées par les équations du filtre (une matrice de rotation reste sur $SO(3)$, une concentration reste positive). D'autre part, les résultats sont plus facilement interprétables, car sur toute une série de trajectoires usuelles en pratique, la matrice censée refléter la covariance de l'erreur converge vers une valeur fixée, alors qu' elle oscille dans le cas du filtre de Kalman étendu conventionnel. Ce phénomène est illustré sur la figure 2.1 extraite de l'article [23], où en simulation un robot décrit une trajectoire circulaire en observant la position d'un point fixe de l'environnement (amer) relativement à sa propre position, et cherche à retrouver sa trajectoire et la position de l'amer (à une rotation-translation près). Les trois graphiques du haut montrent la trajectoire du robot (en bleu) et la position de l'amer (en vert) de gauche à droite 1) estimées par un IEKF 2) estimées par un EKF conventionnel 3) réelles. On observe la convergence de la trajectoire du robot vers une trajectoire circulaire et de l'amer vers une position fixe. Sur les graphiques du bas sont tracés les coefficients de la matrice de gain de Kalman pour les deux filtres. Les gains convergent dans le cas de l'IEKF, et notamment le gain correspondant à la partie inobservable du système reste nul.

De manière plus générale, nous avons testé expérimentalement le filtre de Kalman étendu invariant sur un problème d'estimation d'orientation d'une centrale inertielle dans [31], ainsi que dans diverses applications en robotique [23, 53] détaillées dans la suite.

2.3 Cas particulier d'une dynamique sur un groupe

On a vu que dans le cas déterministe, pour le cas particulier d'une dynamique invariante à gauche sur un groupe de Lie avec mesure équivariante à droite, on peut obtenir une équation d'erreur autonome. Cette propriété remarquable incite à poser le problème du filtrage non-linéaire sur un groupe de Lie comme on l'a fait récemment dans [15]. Tout d'abord on va choisir de mener les calculs sur le groupe de Lie $SO(3)$ pour des raisons de simplicité, et parce que la problématique du filtrage sur $SO(3)$ admet des applications immédiates à l'estimation d'attitude. On considère le problème suivant

$$\frac{d}{dt}R_t = R_t(\omega_t + w_t)_\times \tag{2.2}$$

$$Y_t = R_t L_t \tag{2.3}$$

où R_t représente l'orientation d'un solide dans l'espace, et en reprenant les notations de l'exemple 2, ω_t est le vecteur rotation instantané, w_t est un bruit blanc de $\mathbb{R}^3$, et L_t est un bruit blanc dans $SO(3)$ de sorte que $R_t L_t$ est une mesure bruitée de R_t (il faudra donner un sens à cette dernière formule). La première équation représente l'évolution de R_t où la vitesse angulaire est mesurée par des gyroscopes bruités. La seconde équation représente une mesure bruitée de tout l'état (situation a priori plus simple qu'une mesure partielle de l'état). Vu qu'il n'existe pas de capteur de rotation, ce cas est essentiellement académique. Cela dit on peut supposer dans le cas d'un satellite par exemple, qu'un algorithme de traitement d'image nous fournit en permanence à

travers par exemple l'observation d'étoiles lointaines une estimation bruitée de notre orientation (le problème statique associé est connu sous le nom de problème de Wahba, posé en 1965).

Afin de ré-écrire les équations du système dans le formalisme moderne de la théorie des processus stochastiques, il faut d'abord donner une interprétation à l'équation (2.2). Il faut évidemment considérer cette équation au sens de Stratonovich car 1- au sens d'Itô la dynamique ne préserve pas la relation $R_t R_t^T = I$ alors que la compatibilité de l'intégrale de Stratonovich avec les règles de calcul habituelles nous donne $d(R_t R_t^T) = 0$, c'est donc cette interprétation qui est privilégiée quand on considère des équations différentielles stochastiques (EDS) sur les groupes [64] 2- si les bruits des gyroscopes ne sont pas tout à fait blancs la solution de l'équation (2.2) est un processus bien défini, et bien approximé par la solution de l'EDS au sens de Stratonovich d'après le résultat de Wong et Zakai [90]. Il faut ensuite donner un sens à l'équation (2.3) censée refléter une mesure bruitée de la rotation. Afin d'éviter de faire appel à la théorie des distributions on formule généralement le problème du filtrage à partir de l'intégrale temporelle des mesures bruitées comme expliqué plus haut. Mais ici cela implique de définir un processus qui soit la somme d'éléments d'un groupe de Lie. Cela nous éloigne du filtrage.

Une façon simple de contourner le problème d'une définition rigoureuse dans le formalisme des processus stochastiques a été proposée dans [15] et consiste à considérer des mesures discrètes. Dans la pratique les mesures sont toujours discrètes, et dans certains cas elles peuvent même être délivrées à relativement basse fréquence (les bips GPS sont disponibles à quelques Hertz). Si l'on suppose que les bruits sont identiques sur les trois axes du gyroscope, alors le bruit w_t admet une matrice de covariance de la forme σI, et une discrétisation exacte peut être écrite. Elle nécessite de faire appel au processus brownien de rotation, qui est le mouvement décrit par une particule subissant un très grand nombre de chocs provoquant des petites rotations, i.e. la solution de l'équation au sens de stratonovich $dR_t = R_t \circ (dw_t)_\times$. Ce mouvement a été étudié pour la première fois par Francis Perrin dans sa thèse [76]. La distribution associée au temps t est donnée sous forme d'une série dont le terme général s'appuie sur des fonctions trigonométriques. Considérons une discrétisation $t_0, t_1, \cdots$ de l'axe temporel. Nous avons le résultat suivant:

Proposition 1. *Soit $R_n = R_{t_n}$. Alors le système discret (R_n, Y_n) satisfait les équations suivantes:*

$$\begin{aligned} R_{n+1} &= W_n R_n A_n \\ Y_n &= V_n R_n \end{aligned} \tag{2.4}$$

où W_n est la valeur au temps $t_{n+1} - t_n$ d'un mouvement brownien de rotation initialisé à l'identité et ayant une matrice de diffusion σI_d et A_n est la valeur au temps t_{n+1} de la solution dans $SO(3)$ de l'équation $\frac{d}{dt} Q_s = Q_s (\omega_s)_\times$ avec valeur initiale $Q_{t_n} = I$.

En mimant la forme du filtre de Kalman linéaire en temps discret, mais en remplaçant les additions par des multiplications sur le groupe (l'addition est commutative mais ici la multiplication ne l'est pas, conduisant à certains choix motivés par le résultat suivant

sur l'équation d'erreur), on propose la classe de filtres invariants

$$\hat{R}_{n+1} = K_n(Y_n\hat{R}_n^T)\hat{R}_n A_n \tag{2.5}$$

où K_n est une fonction de son argument à choisir par l'utilisateur. En définissant l' erreur invariante par multiplication à droite

$$\eta_n = R_n\hat{R}_n^T \tag{2.6}$$

nous avons le résultat important suivant

Proposition 2. *Les erreurs d'état invariantes sont des chaînes de Markov définies récursivement par des formules indépendantes de A_n:*

$$\eta_{n+1} = W_n R_n A_n A_n^T \hat{R}_n^T [K(V_n R_n \hat{R}_n^T)]^T \tag{2.7}$$
$$= W_n [K_n(V_n\eta_n)]^T \eta_n \tag{2.8}$$

On constate que l'évolution de l'erreur est indépendente de ω_t. Nous obtenons donc un filtre analogue au filtre de Kalman discret, défini intrinsèquement, et dont l'équation d'erreur se comporte comme si le processus d'état était constant.

Conséquences importantes Les conséquences d'une équation d'erreur autonome dans le cadre stochastique discret sont importantes pour la conception de filtres efficaces et facilement implémentables numériquement sur les groupes de Lie. En effet

- la suite de gains $(K_n)_{n\geq 1}$ peut être optimisée hors ligne, puisque l'évolution de l'erreur ne dépend pas de la trajectoire. Les gains sont ensuite stockés et le coût de calcul associé au filtre (2.5) est extrêmement faible.

- On peut régler les gains K_n à travers une linéarisation de l'erreur autour de l'identité, et utiliser les équations du filtre de Kalman discret pour le modèle linéaire. On construit alors un IEKF discret et on obtient la convergence asymptotique vers un gain K fixé.

- On peut ainsi également prendre en compte des modèles de bruits très généraux (i.e. non-gaussiens). Cela permet par exemple de tenir explicitement compte de la présence de mesures aberrantes ("outliers"). On peut par exemple alors optimiser K en simulation (simulations de Monte-Carlo) dans la classe des fonctions affines par morceaux, de sorte que le gain associé à une large erreur soit plus faible que celui associé à une petite erreur afin d'être robuste aux points aberrants.

De plus, la propriété d'autonomie de l'équation d'erreur permet d'obtenir des résultats de stabilité du filtre, i.e. d'oubli de sa condition initiale. En effet, lorsque les fonctions de gain $K_n(\cdot) \equiv K(\cdot)$ ne dépendent pas de n, la formule d'évolution de l'erreur d'état (2.8) définit une chaîne de Markov homogène. Sous des hypothèses techniques ad hoc, cette chaîne de Markov admet une distribution stationnaire unique et la suite des erreurs η_n converge en distribution vers cette distribution stationnaire. Les conditions appropriées reposent sur la machinerie des chaînes de Harris. Le résultat suivant en fournit un exemple

Theorem 4. *Considérons le filtre discret défini par* (2.5) *pour le système* (2.4). *Choisissons* $K_n(\cdot) \equiv K(\cdot) = \exp(-k \log(\cdot))$ *avec* $0 < k < 1$. *Alors la fonction densité de la variable d'état* (η_n) *converge en norme* L^1 *vers une distribution stationnaire independante de la distribution initiale de* η_0 *lorsque* $n \to \infty$.

Dans l'article [14] on présente des conditions de convergence bien plus générales sur un groupe de Lie quelconque. Un tel résultat permet d'optimiser numériquement la fonction de gain $K(\cdot)$ dans une certaine famille de gains (pour l'exemple ci-dessus le paramètre k dans $[0, 1]$) en fonction du niveau de bruit présent dans le système. En effet, si l'on connaît les lois qui sous-tendent les bruits, on peut chercher numériquement le gain qui minimise la dispersion de l'erreur asymptotique. Cette méthode a permis d'obtenir des filtres performants pour un exemple d'horizon artificiel où la mesure contient des points aberrants (phases de fortes accélérations) dans l'article [14].

Chapter 3

Divers résultats autour des observateurs invariants

Nous présentons ici brièvement plusieurs autres articles qui traitent du sujet de la synthèse d'observateurs non-linéaires tenant compte de certaines symétries du système, ou qui ne traitent pas de ce sujet mais qui s'inspirent des idées développées plus haut.

3.1 Résultats en théorie du contrôle

Principe de séparation Dans cette section nous présentons brièvement les résultats obtenus avec Erwan Salaun, Philippe Martin et Pierre Rouchon [30]. Considérons le système linéaire temps-invariant suivant

$$\frac{d}{dt}x = Ax + Bu, \quad y = Cx$$

où $x \in \mathbb{R}^n$ est l'état, $u \in \mathbb{R}^m$ le contrôle, et $y \in \mathbb{R}^p$ la sortie. Supposons que l'on dispose d'une loi de contôle $u-u_r = -K(x-x_r)$ qui stabilise exponentiellement le système autour d'une trajectoire de référence (x_r, u_r, y_r) lorsque l'état entier est connu. Si la matrice $A-BK$ est Hurwitz, le système est bien stabilisé puisque l'écart à la trajectoire désirée vérifie l'équation $\frac{d}{dt}(x-x_r) = (A-BK)(x-x_r)$. Supposons que l'on dispose également d'un observateur asymptotique de la forme $\frac{d}{dt}\hat{x} = A\hat{x} - L(C\hat{x}-y)$ avec $A-LC$ Hurwitz. En pratique seul l'état estimé $\hat{x}$ est disponible pour le controleur. On est donc en droit de se demander si la loi de contrôle $u-u_r = -K(\hat{x}-x_r)$ stabilise le système autour de la trajectoire. Le principe de séparation garantit que si l'on combine un observateur et un contrôleur stables le système bouclé résultant sera stable. On le montre simplement de la façon suivante. On a supposé $A-LC$ Hurwitz, donc $\tilde{x} = \hat{x} - x \to 0$ exponentiellement. De plus, en utilisant la loi de contrôle $u - u_r = -K(\hat{x} - x_r)$ l'équation suivante est vérifiée

$$\frac{d}{dt}(x - x_r) = A(x - x_r) - BK(\hat{x} - x_r) = (A - BK)(x - x_r) - BK\tilde{x}$$

En écrivant le système d'erreur complet pour les variables $(x - x_r, \tilde{x})$, on obtient un système linéaire triangulaire convergent et donc la convergence vers 0 de $x - x_r$.

On voit bien que la forme triangulaire du système d'erreur est au coeur du principe de séparation. Aussi l'on peut espérer obtenir un principe de séparation sur les groupes de Lie (a minima) autour des trajectoires permanentes lorsqu'un observateur invariant est combiné avec un contrôleur invariant, puisque l'équation d'erreur de l'observateur est autonome autour de ces trajectoires. C'est en effet bien le cas sous certaines hypothèses légères comme nous l'avons montré dans [30].

Coordination sur un groupe de Lie Dans cette section nous évoquons les résultats obtenus avec Alain Sarlette et Rodolphe Sepulchre [81]. Au cours de la dernière décennie, une grande attention dans la communauté du contrôle a été portée au problème de la coordination d'agents autonomes: chaque agent possède une loi de contrôle ne devant dépendre que de sa position relative à certains de ses voisins. Le but du problème de coordination est de définir des lois pour chaque agent aboutissant à une formation ou mouvement organisé, comme par exemple le vol en formation des oiseaux, ou le mouvement global des bancs de poissons, chaque poisson ne voyant que ses proches voisins. Mathématiquement cela peut être défini comme un mouvement où l'ensemble des agents évolue comme un objet rigide, ou encore on peut imposer que les positions relatives convergent vers une valeur fixe. Lorsque les agents évoluent sur un groupe de Lie, par exemple si l'on cherche à synchroniser des oscillateurs sur le cercle i.e. $G = S^1$, ou des robots mobiles i.e. $G = SE(2)$ représentant l'espace des vitesses de robots, les symétries du problème de contrôle peuvent êtres exploitées de façon systématique afin de produire des lois de contrôle pour la coordination. La notion d'erreur invariante telle que définie pour les observateurs sur les groupes de Lie, peut être utilisée pour définir la position relative entre deux agents. Il s'agit alors de faire converger cette erreur, et l'on se trouve dans un cadre où les idées développées pour la synthèse d'observateurs invariants s'avèrent utiles.

3.2 Un observateur invariant en dimension infinie

Dans cette section nous présentons brièvement des résultats obtenus avec Didier Auroux [12] dans le cadre du projet ANR intitulé LEFE-BFN (les enveloppes fluides et l'environnement - Back and Forth Nudging) et piloté par Jacques Blum.

On s'est intéressé dans ce travail à un problème classique d'océanographie [55], où l'on cherche à estimer les courants marins dominants à partir de données sur la hauteur de l'eau mesurée par des satellites, et corrompues par des bruits liés au passage des signaux à travers l'atmosphère, ainsi qu'à partir d'un modèle non-linéaire simplifié du mouvement de l'eau: l'océan est décomposé en plusieurs couches verticales dont seule la première est supposée mobile et le mouvement de l'eau y est décrit par un modèle de type eau peu profonde. L'article [12] dans un premier temps explore le sujet pour le cas simplifié d'un modèle de Saint-Venant sur un domaine rectangulaire Ω, et les méthodes sont ensuite appliquées au modèle complet en simulation et conduisent à des résultats très satisfaisants. Intéressons-nous donc aus équations de Saint-Venant sur un domaine

rectangulaire muni d'un repère orthonormal horizontal (x, y)

$$\frac{\partial}{\partial t}h = -\nabla(hv), \qquad \frac{\partial}{\partial t}v = -v\nabla v - g\nabla h$$

où $hv = h(v_x\mathbf{i} + v_y\mathbf{j})$ est le transport horizontal, avec $\mathbf{i}, \mathbf{j}$ les vecteurs unitaires pointant vers l'est et le nord. On suppose que la hauteur de l'eau $h(x, y, t)$ est mesurée pour tous x, y, t. Le but est alors de débruiter les mesures $h(x, y, t)$ et d'estimer les courants marins $v(x, y, t)$ en tout point (x, y) du domaine.

Il est clair que les équations de la mécanique des fluides ne dépendent pas d'un choix particulier d'axes orthonormaux (x, y): elles sont invariantes par translation et rotation (il faut bien sûr transformer le domaine en conséquence). On veut construire un observateur invariant de la forme "copie de la dynamique + terme de correction". Le moyen le plus simple pour construire des termes de corrections *intégraux*[1] invariants par rotation et translation consiste à former un produit de convolution entre un noyau isotrope et l'erreur de sortie habituelle $h(x, y, t) - \hat{h}(x, y, t)$. Cela conduit à l'observateur invariant suivant

$$\begin{aligned}
\frac{\partial}{\partial t}\hat{h} &= -\nabla(\hat{h}\hat{v}) + \int\int \phi_h(\xi^2 + \zeta^2)(h - \hat{h})_{(x-\xi, y-\zeta, t)}\ d\xi d\zeta \\
&= -\nabla(\hat{h}\hat{v}) + \varphi_h * (h - \hat{h}) \\
\frac{\partial}{\partial t}\hat{v} &= -\hat{v}\nabla\hat{v} - g\nabla\hat{h} + \int\int \phi_v(\xi^2 + \zeta^2)\nabla(h - \hat{h})_{(x-\xi, y-\zeta, t)}\ d\xi d\zeta \\
&= -\hat{v}\nabla\hat{v} - g\nabla\hat{h} + \varphi_v * \nabla(h - \hat{h})
\end{aligned}$$

L'originalité de cet observateur réside en ce que les valeurs estimées de la hauteur $\hat{h}$ et de la vitesse $\hat{v}$ en un point x, y dépendent des valeurs mesurées non seulement en ce point mais aussi dans un voisinage de ce point. Et, la façon dont on utilise les valeurs mesurées dans un voisinage ne privilégie aucune direction, car a priori il n'y a pas de raison que les mesures au nord du point x, y soient de meilleure qualité que celles à l'est du point x, y, les erreurs de mesures venant essentiellement des déformations du signal à travers les couches d'atmosphère traversées. Cet observateur est particulièrement robuste à un bruit de mesure centré car les termes de correction sont eux-mêmes des filtres gaussiens, ces derniers étant des outils classiques pour débruiter les images (voir e.g. [6]). On mélange donc ici un filtre "statique" (filtre gaussien isotrope) et un filtre "dynamique" (un observateur).

En considérant des vitesses faibles $\delta v = v - 0 \ll \sqrt{g\bar{h}}$ et des hauteurs proches de la hauteur d'équilibre $\bar{h}$ dans le sens où $\delta h = h - \bar{h} \ll \bar{h}$ on peut linéariser l'équation d'erreur autour de la position d'équilibre $(h, v) = (\bar{h}, 0)$ où $\bar{h}$ est une fonction constante sur le domaine. Les erreurs d'estimation linéarisées $\tilde{h} = \delta\hat{h} - \delta h$ et $\tilde{v} = \delta\hat{v} - \delta v$, sont alors solutions de l'équation

$$\frac{\partial}{\partial t}\tilde{h} = -\bar{h}\nabla\tilde{v} - \varphi_h * \tilde{h}, \qquad \frac{\partial}{\partial t}\tilde{v} = -g\nabla\tilde{h} - \varphi_v * \nabla\tilde{h}.$$

[1] on veut intégrer le bruit et surtout ne pas le dériver.

En éliminant $\tilde{v}$ on obtient une équation des ondes amorties avec frottement visqueux externe "modifiée" (on retrouve l'équation des ondes si les noyaux de convolution sont des distributions de Dirac)

$$\frac{\partial^2}{\partial t^2}\tilde{h} = g\bar{h}\Delta\tilde{h} + \varphi_v * \Delta\tilde{h} - \varphi_h * \frac{\partial}{\partial t}\tilde{h} \tag{3.1}$$

Nous avons alors le résultat de convergence suivant

Theorem 5. *Si φ_v et φ_h sont définies par*

$$\varphi_v(\xi,\zeta) = \phi_v(\xi^2+\zeta^2) = \beta_v \exp(-\alpha_v(\xi^2+\zeta^2)), \quad \varphi_h(\xi,\zeta) = \phi_h(\xi^2+\zeta^2) = \beta_h \exp(-\alpha_h(\xi^2+\zeta^2))(3.$$

où $\beta_v, \beta_h, \alpha_v, \alpha_h > 0$, alors l'approximation (3.1) *du système d'erreur sur le domaine rectangulaire Ω est fortement asymptotiquement convergente. Si l'on considère en effet l'espace de Hilbert et la norme suivants*

$$\mathcal{H} = H^1(\Omega) \times L^2(\Omega), \quad \|(u,w)\|_{\mathcal{H}} = \left(\int_\Omega \|\nabla u\|^2 + |w|^2\right)^{1/2}, \tag{3.3}$$

alors, pour toute solution de l'équation d'erreur $\tilde{h}$,

$$\lim_{t\to\infty} \left\| \left(\tilde{h}(t), \frac{\partial\tilde{h}}{\partial t}(t)\right) \right\|_{\mathcal{H}} = 0\,. \tag{3.4}$$

La preuve du théorème est relativement simple et s'appuie sur la décomposition du système d'erreur en série de Fourier, le domaine rectangulaire s'y prêtant parfaitement, et les produits de convolution devenant de simples multiplications dans le domaine fréquentiel.

3.3 Deux observateurs non-linéaires contractants

Un observateur intrinsèque pour une classe de systèmes mécaniques Dans cette section nous présentons brièvement les résultats obtenus dans l'article [22]. On considère un système mécanique lagrangien, dont on mesure à tout instant la position q (généralisée dans l'espace de configuration), et on souhaiterait estimer sa vitesse. Il est bien entendu exclu de dériver la mesure de position par rapport au temps, car la mesure est supposée bruitée. Dans l'article [5], les auteurs ont proposé pour ce problème un observateur localement exponentiellement convergent autour de toute trajectoire du système. L'observateur est, comme les équations d'Euler-Lagrange, totalement intrinsèque, c'est-à-dire invariant sous l'action du groupe infini des changements de coordonnées dans l'espace de configuration. En effet, il s'appuie sur la structure riemannienne définie sur l'espace de configuration par l'énergie cinétique, cette dernière quantité ne dépendant naturellement pas d'un choix particulier de coordonnées.

Dans [22] nous avons considéré un cas particulier, pour lequel on a obtenu une convergence plus forte. On suppose en effet que la variable position décrit dans l'espace

de configuration une géodésique au sens d'une certaine métrique riemannienne. C'est le cas pour un système lagrangien soumis à une force dérivant d'un potentiel $U(q)$, puisqu'il décrit une géodésique de l'espace de configuration muni de la métrique de Jacobi, selon le principe de Maupertuis [9]. On propose l'observateur réduit intrinsèque suivant

$$\frac{d}{dt}\hat{\xi} = -\frac{1}{2\lambda}\mathrm{grad}_{\hat{\xi}} D^2(\hat{\xi}, q), \quad \lambda > 0 \tag{3.5}$$

où $\hat{\xi}$ est une estimée du point $\xi(t)$ défini plus bas, q est la mesure de position, et la dynamique est définie par le gradient riemannien au point $\hat{\xi}$ de la distance géodésique au carré $D^2(\hat{\xi}, q)$ entre $\hat{\xi}$ et q.

Theorem 6. *Soit $\mathcal{M}$ une variété riemannienne lisse. Soit $T < \infty$. Supposons que $t \mapsto q(t) \in \mathcal{M}$ satisfait $\nabla_{\dot{q}}\dot{q} = 0$ pour $t \in [0, T]$. Soit $\xi(t) = \exp_{q(t)}(-\lambda\dot{q})$. Considérons l'observateur* (3.5) *et l'estimation de vitesse définie par*

$$\hat{v} = \mathcal{T}_{//\hat{\xi}\to q}\frac{d}{dt}\hat{\xi} \tag{3.6}$$

Soit (3.7) *l'inégalité suivante*

$$D(\hat{\xi}(t), \xi(t)) \leq e^{-\frac{1}{\lambda}t} D(\hat{\xi}(0), \xi(0)) \quad \forall t \in [0, T] \tag{3.7}$$

- *Supposons que la courbure sectionnelle de la variété est partout négative. Si pour tout $t \in [0, T]$, $D(\hat{\xi}(t), q(t))$ est borné par le rayon d'injectivité $I(t)$ en $q(t)$,* (3.7) *est vraie pour tout $\lambda > 0$. En particulier, sur une variété de Cartan-Hadamard, l'égalité est toujours vraie et $\hat{\xi}(0)$ peut être choisi arbitrairement. De plus pour tout $t \in [0, T]$*

$$\lambda\|\hat{v}(t) - \dot{q}(t)\| \leq D(\hat{\xi}(t), \xi(t)) \leq e^{-\frac{1}{\lambda}t} D(\hat{\xi}(0), \xi(0)) \tag{3.8}$$

- *Supposons la courbure sectionnelle majorée par $A > 0$.* (3.7) *est vraie tant que la distance $D(\hat{\xi}(t), q(t))$ reste inférieure à $\max(\frac{\pi}{4\sqrt{A}}, I(t))$ pour tout $t \in [0, T]$. Si la variété est simplement connexe et $\lambda > \frac{\pi}{4\|\dot{q}\|_g\sqrt{A}}$,* (3.7) *est vraie dès que $D(\hat{\xi}(0), q(0)) < \frac{\pi}{4\sqrt{A}}$. De plus, on a convergence dans ce cas de l'observateur dans le sens où pour $t \in [0, T]$:*

$$\lambda|\ \|\hat{v}(t)\| - \|\dot{q}\|\ | \leq D(\hat{\xi}(t), \xi(t)), \qquad 0 \leq \sin(\alpha_A(t)) \leq \frac{\sqrt{A}}{\sin(\sqrt{A}\ \lambda\|\dot{q}\|)} D(\hat{\xi}(t), \xi(t)) \tag{3.9}$$

où l'angle $\alpha(t)$ entre $\hat{v}(t)$ et $\dot{q}(t)$ satisfait $0 \leq \alpha(t) \leq \alpha_A(t) \leq \pi$.

Le résultat se comprend aisément lorsque la courbure sectionnelle est partout négative. Dans ce cas, on sait que les géodésiques émanant d'un point (par exemple q) s'écartent (divergent) exponentiellement les unes des autres. Par conséquent notre observateur réduit qui "remonte" les géodésiques vers le point q définit une dynamique contractant. Le résultat est assez contre-intuitif car les cas de courbure négative correspondent à des dynamiques sensibles aux conditions initiales, voire chaotiques, justement car les trajectoires s'écartent exponentiellement les unes des autres (voir les travaux d'Arnold sur l'hydrodynamique des fluides parfaits [10] et les conclusions qu'il en tire sur la difficulté à prévoir le temps).

Cas des observateurs positifs Dans cette section nous présentons brièvement les résultats obtenus dans [25, 47] avec Rodolphe Sepulchre et Alessandro Astolfi, et Dinh N. Thach dont je co-supervise la thèse avec Frédéric Mazenc. Motivées par des applications en biologie et en écologie, où les variables d'état représentent des concentrations ou des populations, plusieurs études ont été consacrées à la synthèse d'observateurs positifs. En effet, certains systèmes en biologie ou en écologie sont régis par des équations linéaires (éventuellement non-autonomes) de la forme habituelle

$$\frac{d}{dt}x = A(t)x + B(t)u, \quad y = C(t)x$$

où x est un vecteur dont toutes les composantes sont positives. On omettra désormais les dépendances en temps. Le but est alors de construire un observateur dont les estimations sont positives en permanence : il est impossible dans ce contexte de donner une interprétation physique à une coordonnée non-positive.

Le système linéaire $\frac{d}{dt}x = Ax,\ y = Cx$ admet les homotéties $G = \mathbb{R}^*$ comme groupe de symétrie via l'action $\phi_g(x) = gx$. Un ensemble complet d'invariants scalaires est donné par $I(x) = x/\|x\|$. $I(x)$ et $\|x\|$ constituent des coordonnées alternatives, les coordonnées polaires généralisées. Une propriété des systèmes invariants est que la dynamique de $I(x)$ est autonome: ici elle ne dépend évidemment pas de $\|x\|$ et c'est un fait général. Si la dynamique autonome de $I(x)$ est naturellement stable ou contractante, on doit pouvoir sous certaines hypothèses assurer la convergence, puisqu'il suffit de stabiliser l'erreur sur une sous-partie de l'état en bijection avec le groupe. C'est essentiellement ce que l'on a utilisé dans l'exemple de réacteur chimique de l'article [28] et dans le présent exemple.

En effet, si la dynamique induite sur la sphère est une contraction, alors une simple copie du système

$$\frac{d}{dt}\hat{x} = A\hat{x} + Bu$$

sera telle que $\hat{x}/\|\hat{x}\| - x/\|x\|$ convergera vers 0. On pourra alors estimer la norme du vecteur x comme le rapport de n'importe quelle coordonnée de y avec la coordonnée correspondante de $C\hat{x}/\|\hat{x}\|$. Les conditions sous lesquelles la dynamique induite sur la sphère est une contraction sont dérivées du théorème de Birkhoff (1957) complété par les travaux de Bushell [42], qui fournit une généralisation de la théorie de Perron-Frobenius sur des cônes. De plus, dans [47], nous avons proposé l'ajout de certains termes de correction destinés à accélérer la convergence de l'observateur dans l'espace projectif, i.e. de $I(x) - I(\hat{x})$ vers 0.

3.4 Applications en robotique mobile

EKF SLAM Pour un robot, le SLAM ("simultaneous localisation and mapping") consiste à construire une carte de son environnement tout en se localisant dans cet environnement. C'est un sous-domaine de la robotique très étudié car il constitue la clé de la navigation autonome en environnement inconnu. Il a ainsi souvent été qualifié de "Graal" de la robotique mobile. Dans l'article [23], nous avons considéré un robot

mobile équipé d'odomètres, et d'un télémètre laser rotatif, qui se déplace dans un plan. On suppose que grâce à un traitement embarqué des données laser, il connaît à chaque instant, dans son propre repère, la position bruitée de certains points d'intérêt fixes de l'environnement, appelés amers. Le but est d'estimer, à une rotation translation près, la carte des amers, ainsi que la position du robot, et d'attribuer à ces estimations un certain degré de confiance à travers une matrice de covariance. La technique usuelle pour traiter ce sous-problème est le EKF-SLAM. Elle consiste à construire un grand vecteur d'état constitué de l'orientation et la position du véhicule, ainsi que de la position de tous les amers. Un filtre de Kalman étendu est alors utilisé pour estimer le vecteur d'état à partir des mesures de position des amers dans le repère véhicule.

Le système admet des symétries par rotation translation - les invariances galiléennes. Le nombre d'amer pouvant être très important, l'espace d'état est bien plus grand que le groupe de symétries associé à la dynamique. Cela dit, puisque les amers sont immobiles, il a été montré dans [23] que la dynamique de l'ensemble véhicule et carte peut être vue comme une dynamique invariante à gauche sur un groupe de Lie. On peut donc construire pour ce problème un filtre de Kalman étendu invariant, et on obtient la convergence des gains du filtre, comme on l'a déjà montré sur la figure 2.1 du chapitre précédent. Les gains peuvent donc être calculés hors-ligne et remplacés par leur valeur asymptotique. Le résultat est digne d'intérêt, puisque de nombreux articles ont été consacrés à diminuer le coût numérique des méthodes d'EKF-SLAM, au moins quadratique avec le nombre d'amers, qui peut être très élevé.

3D SLAM Dans cette section nous présentons brièvement les résultats obtenus avec Thibaut Hervier lors de son stage de fin d'étude à Polytechnique, co-encadré avec François Goulette [53]. L'idée dans cet article est de construire une carte dense (sans points d'intérêt) de l'environnement à partir d'une image de profondeur générée par une caméra kinect, et de se localiser dans cette carte. Le robot mobile est également équipé d'odomètres et de gyroscopes. A chaque pas de temps, les gyroscopes et les odomètres donnent une prédiction du déplacement effectué par le robot, et un algorithme de traitement d'image donne une estimation de déplacement basée sur l'image de profondeur. Un filtre de Kalman étendu invariant permet alors d'obtenir une estimation de la position et de l'orientation du robot, avec certaines propriétés de convergence. Des résultats expérimentaux ont confirmé l'intérêt de la méthode. Certaines expériences avec des robots mobiles, menées par le post-doctorant Martin Barczyk co-encadré avec François Goulette, sont encore en cours à ce jour et les résultats s'avèrent encourageants.

Application à l'estimation de trajectoire d'un véhicule terrestre Une autre application à un problème pratique a été réalisée en 2010 avec l'aide notamment de Jean-Emmanuel Deschaud. Il s'est agi pour un véhicule terrestre (voiture), à partir de mesures de position d'un GPS, de vitesse angulaire d'un gyroscope bas-coût, et de vitesse d'une roue par un odomètre, de reconstruire l'orientation du véhicule et sa position. Un observateur simple a été proposé. Il estime en permanence également les biais du gyroscope, et la taille des roues, ces deux quantités dérivant lentement avec le temps. Grâce à la connaissance précise de ces quantités, en cas de perte temporaire du

GPS (dans une rue étroite les signaux GPS sont très dégradés en raison du manque de visibilité des satellites et des réflexions multiplies sur les façades) les estimations boucle ouverte restent correctes pendant une période non-négligeable. De plus, des garanties de convergence presque globales ont pu être obtenues grâce aux propriétés d'invariance de l'observateur. L'observateur est proposé et étudié dans [33], et les expériences y sont également décrites. Les résultats s'appliquent à un mouvement plan. Une extension au cas où l'on cherche à reconstruire l'orientation et la position du véhicule dans l'espace tridimensionnel est proposée dans [27].

Autre application et pistes Une question que l'on pourrait se poser est celle de la synthèse d'observateurs invariants sur les espaces homogènes. En effet, on pourrait reprendre la théorie dans le cas où le groupe de symétries a une dimension supérieure à celle de l'espace d'état. Nous n'avons jamais creusé ce sujet, et il pourrait être une piste de recherches futures. Cela dit, nous avons appliqué la théorie des observateurs invariants à un problème d'estimation de paramètres pour un système quantique à deux états dans l'article [32] où la dimension du groupe de symétries est un peu supérieure à celle de l'état. On construit un observateur invariant qui possède certaines propriétés de convergence sans rencontrer de difficulté particulière.

Part II

Filtrage sur les variétés riemanniennes et rang faible

Chapter 4

Géométrie des matrices semi-définies positives

Dans ce chapitre nous nous intéressons à l'espace des matrices semi-définies positives de rang fixé (et éventuellement faible). On présente essentiellement une géométrie riemannienne sur cet espace que l'on a proposée dans [34], ainsi qu'une application au filtrage de Kalman de rang faible évoquée dans [36].

Les matrices positives définies sont devenues des outils fondamentaux de calcul matriciel dans plusieurs domaines de l'ingénierie et des mathématiques appliquées. Elle apparaissent en effet comme matrices de covariance en statistiques, comme des éléments de l'espace de recherche en optimisation convexe et programmation semi-définie, comme des noyaux en apprentissage automatique, comme tenseurs de diffusion en imagerie médicale. Elles jouent également un rôle fondamental en théorie du contrôle à travers leur utilisation dans la théorie des systèmes linéaires : filtrage optimal (équations de Kalman et Riccati), étude de stabilité et construction de modèles d'états équilibrés (équation de Lyapunov), commande optimale LQ (équation de Riccati). Plus généralement, de nombreux problèmes de contrôle peuvent être reformulés sous la forme d'inégalités matricielles linéaires (LMIs) dont les inconnues sont des matrices définies positives. On peut finalement citer le domaine de l'information quantique où les matrices positives définies hermitiennes apparaissent comme matrices de densité.

Plusieurs algorithmes dans ces différentes disciplines font usage d'une géométrie riemannienne naturelle sur le cône des matrices positives, à travers le recours à la géométrie de l'information en statistiques, traitement du signal (e.g. [84]), et théorie de l'information quantique (e.g. [13]), à l'affine invariance en traitement d'images tensorielles (e.g. [49, 75]) et aux fonctions barrières pour les méthodes de point intérieur en optimisation (e.g. [70]). Deux raisons principales motivent la généralisation de ce cadre algorithmique aux matrices positives de rang faible: la taille des problèmes et les contraintes de rang.

L'application des algorithmes à des problèmes de grande taille est une motivation centrale dans les mathématiques appliquées d'aujourd'hui à travers le traitement de bases de données dont la taille ne cesse d'augmenter (big data). Or dans un algorithme itératif, les opérations matricielles courantes associées à une matrice carrée de taille $n \times n$ telles que le calcul de valeurs propres ou l'inversion ont un coût numérique d'ordre

$O(n^3)$. Mais pour les problèmes de très grande taille cette complexité doit être au plus linéaire $O(n)$. Une solution consiste à approximer la matrice $n \times n$ par une matrice de rang faible $p \ll n$ qui peut être manipulée par des opérations de complexité $O(np^2)$ par exemple. Une telle approche a déjà été préconisée en océanographie (filtrage de Kalman de rang faible [77], voir aussi [36, 63]), dans un grand nombre d'articles récents en apprentissage ou en traitement du signal (e.g. [43]), ou en réduction de modèles (e.g. [80]).

Sous espaces et réductions de données Nous allons commencer par quelques rappels sur la variété de Grassmann $Gr(p, n)$, qui désigne l'ensemble des sous-espaces linéaires de dimension p dans $\mathbb{R}^n$. Là aussi, les motivations proviennent essentiellement du rang faible $p \ll n$. Une géométrie digne d'intérêt est celle où un sous-espace $\mathcal{Y}$ est représenté par une matrice $Y \in \mathbb{R}^{n\times p}$ de rang plein qui sous-tend $\mathcal{Y}$. La représentation n'est pas unique, puisque la matrice YM sous-tend le même sous espace pour n'importe quelle matrice $M \in Gl(p)$. Un sous-espace est donc représenté par une classe d'équivalence de matrices dans un espace de matrices linéaire. La géométrie *quotient* qui en résulte est particulièrement adaptée au calcul sur les sous-espaces parce qu'elle peut tirer profit de la simplicité du calcul dans un espace linéaire ambiant de dimension comparable à celle de l'espace d'intérêt.

Les travaux présentés ici s'inscrivent dans la même logique. Ils s'appuient sur deux géométries quotient pour les matrices semi-définies positives de rang faible. Ces géométries trouvent leur source dans des factorisations classiques de matrices positives X sous la forme $X = YY^T$, où la matrice $Y \in \mathbb{R}^{n\times p}$ est de rang plein, et sous la forme $X = UR^2U^T$, où la matrice $U \in \mathbb{R}^{n\times p}$ satisfait la contrainte d'orthogonalité $U^TU = I_p$ et où la matrice R^2 est une matrice définie positive de dimension $p \times p$. Comme dans le cas de la variété des sous-espaces, ces géométries permettent une paramétrisation peu coûteuse de l'ensemble $S^+(p, n)$ tout en tirant profit de la simplicité du calcul dans l'espace ambiant. Dans notre volonté d'étendre la géométrie riemannienne dite naturelle du cône des matrices définies positives au cas où le rang n'est pas plein, nous nous sommes concentrés sur la deuxième géométrie qui s'appuie sur la factorisation $X = UR^2U^T$. En effet, la géométrie naturelle du cône fournit une métrique naturelle pour le facteur central R^2. Il reste à trouver alors une métrique pour le facteur U (on pourra utiliser la métrique naturelle des sous-espace de Grassman) et à combiner les deux proprement.

Le calcul sur les sous-espaces est motivé depuis longtemps par un grand nombre d'applications qui gravitent autour de la réduction de données et en particulier du calcul de sous-espaces propres dominants de grandes matrices. De la même manière, un large champ d'applications motive le calcul sur les matrices positives de rang faible. L'intérêt est plus récent que pour le calcul de sous-espaces, mais il a connu une accélération dans les dernières années avec l'intérêt croissant pour les relaxations convexes de problèmes d'optimisation durs et leur application à des problèmes de très grande dimension. L'approche considérée ici est non-convexe mais elle permet de traiter des problèmes de très grande taille qui ne peuvent pas (encore) être traités dans le cadre de l'optimisation convexe. Si l'espace des matrices semi-définies positives est un cône solide dans un espace euclidien,

et possède des propriétés et une géométrie qu'on appréhende facilement, le sous-ensemble des faces du cône définies par une contrainte de rang n'est pas convexe et s'avère même être une espace assez compliqué [80].

Le présent chapitre est organisé comme suit: on commence par quelques rappels sur les variétés de Stiefel, de Grassman et du cône des matrices définies positives. Puis on introduit une géométrie riemannienne sur l'ensemble des matrices semi-définies positives de rang fixé admettant plusieurs propriétés d'invariance, et on termine en soulignant des liens entre cette géométrie et d'éventuelles propriétés de contraction du filtre de Kalman de rang faible. Le lecteur intéressé sera dirigé vers d'autres applications dans l'article [83].

4.1 Quelques variétés matricielles d'intérêt

Variété de Stiefel $St(p,n) = O(n)/O(n-p)$ Elle est définie comme l'ensemble des matrices $n \times p$ avec la contrainte d'orthogonalité $U^T U = I_p$ où I_p représente la matrice identité de $\mathbb{R}^{p\times p}$. La présentation sous la forme d'une variété quotient découle de la propriété de $St(p,n)$ d'être un espace homogène. En effet, le groupe $O(n)$ agit sur l'espace $St(p,n)$ des $p-$repères orthonormaux transitivement, i.e. permet de transformer tout p-repère en un autre p-repère arbitraire. Le groupe d'isotropie, i.e. le groupe qui stabilise un élément U_0 donné de $St(p,n)$ est composé des rotations de $O(n-p)$ qui agissent sur l'orthocomplément de l'espace engendré par les colonnes de U_0.

Variété de Grassman $Gr(p,n)$ Il s'agit de l'ensemble des sous-espaces linéaires de dimension p dans $\mathbb{R}^n$. On peut la représenter comme l'ensemble des classes d'équivalences $St(p,n)/O(p)$ puisque pour $O \in O(p)$ et $U \in St(p,n)$, les colonnes des matrices U et UO engendrent le même sous-espace linéaire de $\mathbb{R}^n$, pouvant être identifié au projecteur UU^T.

Cône des matrices définies positives La géométrie du cône des matrices définies positives symétriques $P(n)$, espace auquel on se référera comme "le cône" dans la suite, a été bien étudiée. Le groupe $Gl(n)$ des matrices de $\mathbb{R}^{n\times n}$ a une action transitive sur $P(n)$ par congruence

$$A \mapsto LAL^T, \tag{4.1}$$

où $L \in Gl(n)$. En effet toute matrice A peut être transformée en l'identité en choisissant $L = A^{-1/2}$ où $A^{1/2}$ est la racine carrée de A que l'on définit aisément en diagonalisant A. $P(n)$ peut donc être vu comme un espace homogène $Gl(n)/O(n)$ (puisque l'action est transitive) et réductif[1] [48]. La conséquence principale est l'existence d'une métrique

[1]Toute matrice de l'algèbre de Lie $\mathfrak{gl}(n) = T_I\mathrm{Gl(n)} = \mathbb{R}^{n\times n}$ peut être exprimée comme la somme de sa partie symétrique et de sa partie anti-symétrique, donc $\mathfrak{gl}(n) = \mathfrak{m} + \mathfrak{so}(n)$ est une somme directe de l'ensemble des matrices symétriques $\mathfrak{m}$ et de l'algèbre de Lie $\mathfrak{so}(n) = T_I\mathrm{O(n)}$. De plus pour tout $S \in \mathfrak{m}$ et $O \in \mathrm{O(n)}$ on a $Ad_O(S) = OSO^{-1} = OSO^T \in \mathfrak{m}$, et donc $Ad_{\mathrm{O(n)}}(\mathfrak{m}) \subset \mathfrak{m}$. L'existence d'un tel espace $\mathfrak{m}$ prouve que $\mathrm{Gl(n)/O(n)}$ est un espace homogène *réductif.*

$Gl(n)$-invariante sur la variété $Gl(n)/O(n)$. On l'appelle métrique naturelle du cône symétrique. A un facteur près elle correspond aussi à la métrique affine-invariante en traitement d'images tensorielles [75], la métrique de Siegel en géométrie symplectique, et également à la métrique de l'information de Fisher associée à la famille des lois normales de moyenne nulle en dimension n. De façon similaire, tout vecteur tangent $X \in T_I P(n)$ à l'identité $I \in P(n)$ peut être transformé en un vecteur tangent $A^{1/2}XA^{1/2} \in T_A P(n)$. Tout d'abord, remarquons que l'espace tangent à l'identité est l'ensemble des matrices symétriques: $T_I P(n) = \mathrm{Sym(n)}$. A l'identité I, la métrique $Gl(n)$-invariante est définie par le produit scalaire associé à la norme de Frobenius $g_I^{P_n}(X_1, X_2) = \mathrm{Tr}\left(X_1 X_2^T\right) = \mathrm{Tr}\left(X_1 X_2\right)$. L'invariance de la métrique implique alors $g_A^{P_n}(A^{1/2}X_1A^{1/2}, A^{1/2}X_2A^{1/2}) = \mathrm{Tr}\left(X_1 X_2\right)$, et donc pour tout élément arbitraire $A \in P(n)$ la métrique naturelle s'écrit

$$g_A^{P(n)}(D_1, D_2) = \mathrm{Tr}\left(D_1 A^{-1} D_2 A^{-1}\right) \tag{4.2}$$

L'invariance de la métrique (4.2) a des implications directes sur l'expression des géodésiques. L'application exponentielle en I est l'exponentielle matricielle usuelle

$$\exp_I^{P(n)} X = \exp X = \sum_{k=0}^{\infty} (1/k!) X^k.$$

La norme de Froebenius $\|X\|_F$ est la distance géodésique $d(\exp X, I) = \|X\|_F$, d'où la formule à l'identité

$$d(A, I) = \|\log A\|_F.$$

L'invariance de la métrique permet d'étendre la caractérisation des géodésiques pour tout point arbitraire $A \in P(n)$:

$$\exp_A^{P_n}(tX) = A^{1/2} \exp(tA^{-1/2}XA^{-1/2})A^{1/2}, \quad t > 0,$$

et la distance géodésique correspondante

$$d_{P_n}(A, B) = d(A^{-1/2}BA^{-1/2}, I) = \|\log(A^{-1/2}BA^{-1/2})\|_F = \sqrt{\sum_k \log^2(\lambda_k)}, \tag{4.3}$$

où λ_k sont les racines de l'équation $\det(AB^{-1} - \lambda I)$. La distance est également invariante par inversion matricielle $(A, B) \mapsto (A^{-1}, B^{-1})$ car $\log^2(\lambda_k)$ est invariant par inversion $\lambda_k \mapsto \lambda_k^{-1}$. De plus, la distance géodésique permet de calculer une expression de la moyenne riemannienne de deux matrices $A, B \in P(n)$ définie comme le point milieu de la géodésique les reliant. En effet la géodésique $A(t)$ reliant A et B est

$$A(t) = \exp_A^{P_n}(tX) = A^{1/2} \exp(t \log(A^{-1/2}BA^{-1/2}))A^{1/2},$$

avec $A^{-1/2}XA^{-1/2} = \log(A^{-1/2}BA^{-1/2}) \in \mathrm{Sym(n)}$. Le point milieu est obtenu pour $t = 1/2$, conduisant à la moyenne:

$$A \circ B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}. \tag{4.4}$$

D'un point de vue géométrique, la métrique naturelle a de multiples raisons d'être préférée à la métrique "plate" associée à la distance $\|A - B\|_F$ sur le cône des matrices définies positives. En effet, cet espace n'est pas un espace vectoriel. Aussi, le traiter comme un espace vectoriel dans les applications peut conduire à des performances algorithmiques dégradées [84]. Un autre désavantage de la métrique plate est que $P(n)$ n'est pas un espace géodésiquement complet puisque la géodésique $A + t(B - A)$ n'est pas un élément du cône pour tout t. En revanche l'espace muni de la métrique naturelle est géodésiquement complet, les bords du cône étant à une distance infinie de n'importe quel point. Une conséquence est que la métrique naturelle est adaptée aux méthodes de point intérieur dans $P(n)$. Il est en effet intéressant de noter qu'elle coïncide avec la métrique définie par les barrières logarithmiquement homogènes et auto-concordantes en optimisation sur l'ensemble convexe $P(n)$ [40]. De plus l'invariance sous l'action de $Gl(n)$ est désirable dans de nombreuses applications (e.g., [75, 84]). En particulier, lorsque A apparaît comme matrice de covariance $A = \mathbb{E}(zz^T)$, l'action (4.1) correspond à un changement de base $z \mapsto Lz$. La conséquence est que cette métrique est bien adaptée à l'estimation intrinsèque de matrices de covariances (voir [84] et l'article didactique rédigé dans le cadre de la thèse en cours d'Axel Barrau [16]). Dans [75, 49], les auteurs considèrent que l'invariance par inversion est bien adaptée à la physique de l'imagerie médicale par tenseur de diffusion. Finalement, la moyenne riemannienne associée à la métrique naturelle, possède diverses propriétés qui permettent de la qualifier de moyenne "géométrique" par opposition à la moyenne arithmétique. Le prochain chapitre reviendra plus en détail sur les moyennes de matrices semi-définies positives.

4.2 Géométrie des matrices SDP de rang fixé

La métrique naturelle (4.2) de $P(n)$ n'est bien définie qu'à l'intérieur du cône en raison des inversions matricielles impliquées dans sa définition. Comme évoqué plus haut, deux raisons principales motivent pourtant la généralisation de ce cadre algorithmique aux matrices positives de rang faible: la taille des problèmes et les contraintes de rang. Notre généralisation de la géométrie naturelle du cône à l'ensemble $S^+(p, n)$ des matrices de dimension $n \times n$ semi-définies positives de rang p s'appuie sur la métrique naturelle de l'ensemble $Gr(p, n)$ des sous-espaces de dimension $p < n$ et sur la métrique naturelle du cône à travers la décomposition suivante de toute matrice $A \in S^+(p, n)$:

$$A = UR^2U^T$$

où $R^2 \in P(p)$ et $U \in St(p, n)$. En considérant l'action du groupe orthogonal $O(p)$:

$$\begin{aligned} R &\mapsto O^T RO \quad \in P(p), \\ U &\mapsto UO \quad \in St(p, n). \end{aligned} \tag{4.5}$$

on constate que la représentation $A = UR^2U^T$ avec $(U, R^2) \in St(p, n) \times P(p)$ est univoque à la relation d'équivalence près $(U, R^2) \equiv (UO, O^T R^2 O)$ pour tout $O \in O(p)$. L'ensemble $S^+(p, n)$ peut donc être vu comme le quotient

$$S^+(p, n) \cong (St(p, n) \times P(p))/O(p).$$

Cette représentation admet l'interprétation géométrique suivante: les matrices semi-définies positives peuvent être vues comme des ellipsoïdes plats, dont les axes principaux sont donnés par les vecteurs propres de la matrice. Le facteur U apparaît alors comme une base du sous-espace dans lequel vit l'ellipsoïde et le facteur R^2 définit la forme de l'ellipsoïde dans cette base de dimension réduite.

Afin de définir une métrique riemannienne sur cet ensemble, on peut remarquer que le sous-ensemble des matrices dont le facteur R^2 est l'identité, i.e. de la forme UU^T, est en bijection avec la variété de Grassman $Gr(p,n)$ à travers sa représentation quotient $St(p,n)/O(p)$. Puisque la variété de Grassman apparaît donc comme un sous-ensemble de $S_+(p,n)$ pour lequel on dispose d'une distance naturelle, on propose d'imposer que la distance infinitésimale entre deux matrices de $S^+(p,n)$ proches, de la forme UU^T et VV^T, coïncide avec la distance naturelle de Grassman (voir e.g. [1]), et que la distance infinitésimale entre deux matrices de $S^+(p,n)$ proches, de la forme UR^2U^T et UQ^2U^T, coïncide avec la distance naturelle entre les facteurs R^2 et Q^2 sur le cône de dimension réduite $P(p)$. En effet si $(U,R^2) \in S^+(p,n) \times P(p)$ représente $A \in S^+(p,n)$, il est tentant de représenter les vecteurs tangents $T_A S^+(p,n)$ par la variation infinitésimale (Δ, D), où

$$\begin{aligned} \Delta &= U_\perp B, \qquad B \in \mathbb{R}^{(n-p)\times p}, \\ D &= RD_0R \end{aligned} \tag{4.6}$$

et où les colonnes de $U_\perp \in St(n-p,n)$ forment une base du sous-espace vectoriel orthogonal à celui sous-tendu par U, c'est-à-dire $U^TU_\perp = 0$, et $D_0 \in \mathrm{Sym(p)} = T_I P(p)$ est un vecteur tangent à l'identité I_p transporté dans l'espace tangent à R^2 par l'action naturelle de $Gl(p)$ sur $P(p)$. La métrique riemannienne retenue sur $S^+(p,n)$ est une simple somme pondérée de distances infinitésimales dans respectivement $Gr(p,n)$ et dans $P(p)$:

$$g_{(U,R^2)}((\Delta_1,D_1),(\Delta_2,D_2)) = \mathrm{Tr}\left(\Delta_1^T\Delta_2\right) + k\ \mathrm{Tr}\left(R^{-1}D_1R^{-2}D_2R^{-1}\right), \quad k>0, \tag{4.7}$$

généralisant (4.2) de façon assez naturelle. La métrique ci-dessus est définie pour un couple de vecteurs tangents associés à un choix particulier de représentant (U,R^2) parmi l'ensemble des couples de $St(p,n)\times P(p)$ décrivant un même élément de $S^+(p,n)$ vu comme un espace quotient. Il s'agit donc de vérifier que l'on dispose d'une métrique bien définie sur la variété, i.e. indépendante du choix de représentants. Pour cela on décompose l'espace tangent à l'espace produit $St(p,n)\times P(p)$ en un espace tangent aux orbites de l'action (4.5) de $O(p)$ sur $St(p,n)\times P(p)$, appelé espace vertical, et un espace complémentaire transverse aux orbites appelé espace horizontal. La figure 4.1 illustre la notion de variété quotient. Pour plus de détails sur les variétés quotients le lecteur pourra se référer à l'ouvrage [1]. Le résultat suivant montre que nos définitions conduisent bel et bien à une structure de variété riemannienne sur $S^+(p,n)$.

Theorem 7. *L'espace* $S^+(p,n) \cong (St(p,n)\times P(p))/O(p)$ *muni de la métrique* (4.7) *est une variété riemannienne quotient avec espace horizontal*

$$\mathcal{H}_{(U,R^2)} = \{(\Delta,D) : \Delta = U_\perp B,\ B\in\mathbb{R}^{(n-p)\times p},\ D = RD_0R,\ D_0 \in \mathrm{Sym(p)}\}.$$

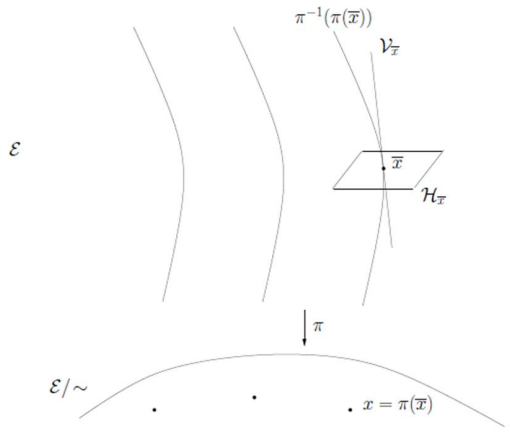


Figure 4.1: Représentation schématique empruntée à [1] d'une variété quotient $\mathcal{E}/\sim$ avec projection canonique π. Une classe d'équivalence $\pi^{-1}(\pi(\bar{x}))$ est représentée comme sous-espace de l'espace total $\mathcal{E}$ et correspond à un point unique $\pi(\bar{x})$ sur la variété quotient. En $\bar{x}$, l'espace tangent à $\mathcal{E}$ se décompose en un espace vertical $\mathcal{V}_{\bar{x}}$ tangent à la classe d'équivalence dont $\bar{x}$ est un représentant, et un espace horizontal choisi comme son complémentaire. Dans le cas présent l'espace total est $\mathcal{E} = St(p,n) \times P(p)$, un représentant s'écrit $\bar{x} = (U, R^2)$, la relation d'équivalence est définie par $(U, R^2) \sim (UO, OR^2O^T)$ pour $O \in O(p)$, et π est l'opération abstraite de projection sur l'espace quotient $(\mathcal{E}/\sim) = S^+(p,n)$.

La métrique n'hérite pas de l'invariance de $P(n)$ sous n'importe quelle transformation de $Gl(n)$, mais conserve l'invariance par rotation et par changement d'échelle. Ce sont deux invariances d'intérêt car la première correspond à l'invariance par choix arbitraire d'orientation des axes, et la seconde peut être reliée à l'invariance par choix d'échelle (par exemple un choix arbitraire d'unités où l'on choisit des km plutôt que des miles). De plus, si l'on substitue à l'opération d'inversion celle de pseudo-inversion, adaptée aux matrices de rang déficient, on conserve la propriété d'invariance par inversion:

Theorem 8. *La métrique* (4.7) *est invariante sous l'action par congruence des transformations orthogonales, homotéties, et par pseudo-inversion.*

Puisque l'on connaît explicitement les géodésiques de $St(p,n)$ et de $P(p)$, on pourrait penser que les géodésiques de $S^+(p,n)$ s'expriment simplement comme produit de géodésiques de ces deux espaces. Ce raisonnement se justifie lorsque la projection π de l'espace total $St(p,n) \times P(p)$ est une submersion riemannienne. Ici ce n'est pas le cas car l'espace horizontal et l'espace vertical ne sont pas orthogonaux au sens de la métrique produit de l'espace total $St(p,n) \times P(p)$. On conçoit aisément que, par exemple lorsque le facteur k apparaissant dans la métrique (4.7) est très grand, le chemin le plus court entre deux ellipsoïdes plats vivant dans le même sous-espace n'est pas nécessairement entièrement contenu dans ce sous-espace. On peut en effet par un jeu de rotations successives dans certains cas construire un chemin où un ellipsoïde est amené sur un ellipsoïde final qui est ce même ellipsoïde mais orienté différemment dans le sous-espace dans lequel il vit, et cela sans jamais avoir modifié la forme de l'ellipsoïde.

Dans [67] on a proposé une métrique légèrement différente qui permettait de rendre l'espace horizontal orthogonal à l'espace vertical, en retirant simplement sa composante parallèle aux fibres. Mais la métrique qui en découle perd les propriétés d'invariances décrites ci-dessus qui sont une motivation centrale de la présente approche destinée à étendre la métrique naturelle du cône au cas où le rang chute.

4.3 Une application: le filtre de Kalman de rang faible

Cette section reprend l'article [35, 36]. Ce dernier article ne contient pas de vrai résultat, et se contente d'évoquer des liens entre la géométrie introduite plus haut et le filtre de Kalman de rang faible. Nous prévoyons une étude plus poussée dans le futur, comme on l'évoque dans les perspectives à la fin de ce mémoire.

Le filtre de Kalman étendu est couramment utilisé en assimilation de données. Le problème d'estimation de courants marins traité à la Section 3.2 fournit en effet un exemple qui se prêterait à cette approche. Pour de tels systèmes régis par des équations aux dérivées partielles, l'implémentation numérique à chaque pas de temps de l'équation de Riccati qui sous-tend le calcul du gain du filtre de Kalman, discrétisée sur le domaine spatial, peut s'avérer extrêmement lourde en complexité de calcul et en stockage des données, puisque la matrice de covariance peut contenir jusqu'à 10 millions d'entrées. De plus, pour des pas de temps espacés il est souhaitable d'utiliser l'équation de Riccati discrète qui requiert des inversions matricielles qui peuvent être particulièrement coûteuses.

Une solution, pour diminuer la complexité numérique, consiste à projeter la covariance du vecteur d'état dans un espace de petite dimension. Cette idée est apparue il y a plusieurs dizaines d'années (voir par exemple [46]). Cette opération réduit le stockage et la multiplication de matrices de taille $n \times n$ à des matrices de taille $r \times n$ permettant de passer d'une complexité quadratique en la dimension de l'état n à une complexité linéaire en la taille du problème. Le sous-espace de faible dimension est censé capturer les valeurs propres les plus grandes de la matrice de covariance, où l'erreur doit être le plus atténuée. Une méthode basée sur cette idée et employée en océanographie est le filtre SEEK [41].

Mathématiquement, les équations du filtre de Kalman-Bucy s'énoncent ainsi. Considérons le système linéaire temps variant suivant

$$\begin{array}{rcl} \frac{d}{dt}x(t) & = & A(t)x(t) + G(t)w(t) \\ y(t) & = & C(t)x(t) + H(t)\eta(t), \end{array} \tag{4.8}$$

où $x(t) \in \mathbb{R}^n$, $y(t) \in \mathbb{R}^p$ et où $w(t) \in \mathbb{R}^m$, $\eta(t) \in \mathbb{R}^p$ sont des bruits blancs gaussiens. Les matrices $A(t)$, $G(t)$, $C(t)$ et $H(t)$ sont supposées avoir des dimensions appropriées.

Les équations du filtre de Kalman-Bucy définissent une équation différentielle régissant l'évolution de la variable

$$\hat{x}(t) = \mathbb{E}[x(t)|\{y(s)\}_{0 \leq s \leq t}]$$

qui est la meilleure prédiction de l'état courant x à partir des mesures courantes et passées, à travers l'évolution de la matrice de covariance de l'erreur $P(t)$

$$\frac{d}{dt}\hat{x} = (A - PC^T(HH^T)^{-1}C)\hat{x} + PC'(HH^T)^{-1}y \tag{4.9}$$

$$\frac{d}{dt}P = \Phi_t(P) = AP + PA^T + GG^T - PC^T(HH^T)^{-1}CP \tag{4.10}$$

La fonction Φ_t définit une équation matricielle de Riccati en temps continu, et pour tout $P \in P(n)$, $\Phi_t(P)$ est un vecteur tangent au cône, i.e. $\Phi_t(P) \in T_P P(n)$.

Lorsque l'on cherche à transposer l'équation de Riccati sur la variété des matrices semi-définies positives de rang (faible) fixé, on est confronté au problème suivant: $\Phi_t(P) \notin T_P S^+(p,n)$ pour $P \in S^+(p,n)$. Une première solution, que l'on a adoptée, consiste à dériver les hypothèse adéquates sur le bruit d'état afin que l'équation de Riccati soit bien définie sur $S^+(p,n)$. Cette méthode a été reprise et étendue dans [62] où le vecteur tangent $\Phi_t(P)$ est projeté orthogonalement sur la variété $S^+(p,n)$. Cette dernière modification permet de prendre en compte des bruits d'états plus généraux, mais les propriétés de contraction associées ne sont pas étudiées.

Propriétés de contraction dans le cas rang plein En temps continu, sous certaines hypothèses d'observabilité et de détectabilité, P. Bougerol a montré des propriétés de contraction (au sens de la métrique naturelle du cône) de l'équation de Riccati en temps discret [39], le résultat étant aussi valable en temps continu. En termes vagues, un système dynamique contractant possède la propriété selon laquelle la distance entre deux trajectoires infinitésimalement proches tend à décroitre. Lorsqu'un système dynamique est (strictement) contractant toutes les trajectoires tendent les unes vers les autres. Les propriétés de contraction de l'équation de Riccati en temps continu peuvent être appréhendées à travers un calcul analogue au calcul effectué en temps discret dans [39] (voir [36] pour plus de détails).

La propriété de contraction de l'équation de Riccati est très désirable, car elle implique l'oubli de la matrice de covariance initiale du filtre $P(0)$. Ainsi, après une phase transitoire, les performances ne dépendent pas d'une mauvaise initialisation de P, souvent mal connue. L'oubli de la condition initiale confère de plus au filtre une grande robustesse vis-à-vis d'erreurs imprévues (non-modélisées) et importantes. En effet l'enjeu est la capacité du filtre à se remettre d'une perte prolongée d'observations ou d'observabilité (on peut penser par exemple à un robot se localisant avec un télémètre laser et passant devant un miroir).

Propriétés de contraction dans le cas rang faible Si l'on suppose le bruit d'état isotrope, i.e. $G = \mu I$, l'équation de Riccati sur $S^+(p,n)$ est bien définie et s'écrit alors

$$\frac{d}{dt}URU^T = AURU^T + URU^TA^T + \mu^2 UU^T - URU^TC^T(HH^T)^{-1}CURU^T \tag{4.11}$$

où (U, R) paramétrise la matrice de rang faible. Elle correspond au flot suivant dans l'espace tangent horizontal

$$\frac{d}{dt}U = (I - UU^T)AU \tag{4.12}$$

$$\frac{d}{dt}R = A_U R + RA_U^T + \mu^2 I - RC_U^T(HH^T)^{-1}C_U R \tag{4.13}$$

où $A_U = U^T AU$ et $C_U = CU$ sont les matrices projetées dans le sous-espace engendré par les colonnes de U. L'intérêt de cette décomposition réside en ce que 1- le système est triangulaire et 2- la deuxième équation est une équation de Riccati définie sur l'intérieur du cône de faible dimension $P(p)$ et elle hérite des propriétés de contraction de l'équation de Riccati dans le cas rang plein. Bien sûr pour des raisons topologiques la première équation (4.12) ne peut pas être une contraction globale (penser au cas $p = 1$ où U est un vecteur de la sphère unitaire tournant à vitesse constante). Cependant, elle est connue sous le nom de flot d'Oja, et possède certaines propriétés de convergence.

En effet, lorsque la matrice A est symétrique et présente un écart strictement positif entre sa p-ième valeur propre et la suivante (valeurs propres classées par ordre décroissant) la matrice U converge presque globalement vers le sous-espace dominant U_∞ de A [89]. On a alors le fait que 1- lorsque U est suffisamment proche de U_∞ le flot (4.12) est contractant au sens de la métrique naturelle de Grassman et 2- si l'on fixe les matrices $A_U \equiv A_{U_\infty}$ et $C_U \equiv C_{U_\infty}$ dans (4.13) on obtient une équation contractante au sens de la métrique naturelle du cône de petite dimension. On peut donc s'attendre à un comportement asymptotiquement contractant pour l'équation de Riccati rang faible au sens de notre métrique (4.7).

Chapter 5

Moyennes de matrices semi-définies positives

La notion d'interpolation ou de moyenne pondérée joue un rôle central dans le débruitage des mesures. En effet, bon nombre d'algorithmes de filtrage destinés à débruiter des signaux reposent sur l'idée qu'en ajoutant (en moyennant) des mesures indépendantes, les fluctuations liées au bruit tendent à se compenser et donc à se réduire. De plus, une pondération peut apparaître dans la moyenne en fonction du bruit de chaque capteur. Un exemple simple est le suivant: si l'on dispose d'une série de deux mesures $y_i = a + w_i$ avec $i = \{1, 2\}$ où $a \in \mathbb{R}$ est un paramètre que l'on veut estimer et w_i est une gaussienne de moyenne nulle et de variance $\sigma_i > 0$, l'estimateur par maximum de vraisemblance s'écrit $\hat{a} = (y_1/\sigma_1 + y_2/\sigma_2)/(1/\sigma_1 + 1/\sigma_2)$ et apparaît comme une moyenne des signaux, pondérée par la confiance que l'on peut accorder à chaque signal.

Lorsque la variété est munie d'une structure riemannienne, on peut procéder classiquement pour définir des moyennes intrinsèquement comme suit: étant donné deux points A, B, de la variété, la géodésique liant ces deux points fournit une courbe naturelle permettant d'interpoler entre les deux matrices. Une moyenne pondérée entre deux matrices s'exprime alors comme un point de la géodésique et la pondération joue sur la distance aux extrémités. Quant à la moyenne pondérée de N matrices $A_1, \cdots, A_N$, on peut avoir recours à la moyenne de Karcher qui consiste à définir la moyenne comme un point X de la variété qui minimise la somme des distances géodésiques au carré $\sum_1^N w_i d^2(X, A_i)$ où les w_i sont des poids de somme égale à 1, et qui ainsi fournit une généralisation de la notion de barycentre euclidien usuelle. Bien sûr l'existence et l'unicité de la moyenne de Karcher reposent sur des conditions, essentiellement liées à la courbure de la variété. Notamment, quand la courbure sectionnelle de la variété est partout négative la moyenne de Karcher existe et est unique, et possède de remarquables propriétés de stabilité (elle est peu sensible à une petite variation des matrices $A_1, \cdots, A_N$) qui la rendent très attractive pour des opérations de filtrage, comme l'ont mis en évidence certains travaux [75].

Dans ce chapitre, nous allons considérer deux variétés particulières pour lesquelles nous avons introduit de nouvelles métriques riemanniennes, courbes interpolantes, et notions de moyenne. Tout d'abord nous considérerons la variété des matrices semi-définies positives de rang fixé $S^+(p, n)$. Nous nous poserons la question de la possibilité

de définir une moyenne "géométrique" sur cette variété. Notre but est d'étendre la moyenne géométrique de $P(n)$, i.e., la moyenne de Karcher au sens de la métrique naturelle du cône au cas dégénéré. Une telle moyenne peut s'avérer utile lorsqu'on travaille avec des approximations de rang faible de matrices de covariance intervenant dans de très grands systèmes par exemple. La question est mathématiquement digne d'intérêt car si la moyenne géométrique sur le cône s'étend facilement par densité aux bords du cône, elle ne préserve pas le rang.

Nous mentionnerons également un travail sur le cône des matrices définies positives de dimension trois $P(3)$. Cette variété intervient dans les applications de traitement d'image tensorielles en imagerie médicale. Si la moyenne de Karcher au sens de la métrique naturelle du cône a permis de filtrer efficacement ces images tensorielles (voir [49, 75]), elle possède certains défauts auxquels on peut remédier en introduisant une métrique riemannienne et des courbes interpolantes qui s'inspirent de notre travail sur $S^+(p,n)$.

5.1 Moyenne géométrique de matrices semi-définies positives

Ce paragraphe reprend le travail effectué avec Anne Collard et Rodolphe Sepulchre [26]. Le but de notre approche est de généraliser la moyenne de Karcher au sens de la métrique naturelle du cône au cas dégénéré des matrices semi-définies positives de rang fixé p, en une moyenne qui préserve le rang. Il existe plusieurs moyennes sur $P(n)$ pouvant être appelées moyennes géométriques. L'approche axiomatique d'Ando, Li et Mathias [7] qualifie en effet de "géométrique" une moyenne de matrices positives semi-définies qui possèderait les propriétés suivantes:

(P1) Cohérence avec les scalaires. Si A et B commutent, alors $M(A,B) = (AB)^{1/2}$.

(P2) Homogénéité jointe. Pour $\alpha, \beta > 0$

$$M(\alpha A, \beta B) = (\alpha\beta)^{1/2} M(A,B).$$

(P3) Invariance par permutation. $M(A,B) = M(B,A)$.

(P4) Monotonicité. Si $A \leq A_0$ (i.e. $(A_0 - A)$ est une matrice positive) et $B \leq B_0$, les moyennes vérifient $M(A,B) \leq M(A_0,B_0)$.

(P5) Continuité supérieure. Si $\{A_n\}$ et $\{B_n\}$ sont des suites décroissantes convergeant vers A, B nous avons $\lim M(A_n,B_n) = M(A,B)$.

(P6) Invariance par congruence. Pour tout $G \in \mathrm{Gl(n)}$, on a $M(GAG^T, GBG^T) = GM(A,B)G^T$.

(P7) Auto-dualité. $M(A,B)^{-1} = M(A^{-1},B^{-1})$.

Sur $P(n)$, la moyenne de Karcher au sens de la métrique naturelle du cône existe et est unique, et elle vérifie cet ensemble de propriétés. Nous cherchons désormais à étendre cette moyenne aux bords du cône, et plus exactement à la variété $S^+(p,n)$. L'approche classique de la littérature consiste à étendre la moyenne géométrique par un argument de continuité. Par conséquent, le sujet n'a reçu que peu d'attention. Pourtant cette approche a de sérieuses limitations puisque la moyenne induite ne préserve pas le rang. On peut même montrer que la moyenne de deux matrices de rang $p < n/2$ prises aléatoirement dans $S^+(p,n)$ est presque toujours nulle.

Nous proposons de voir les matrices de $S^+(p,n)$ comme des matrices de $P(n)$ projetées dans un sous-espace de dimension p. Aussi, la notion de sous-espace de dimension p nous semble jouer un rôle central, et nous imposons que la moyenne de deux matrices de $S^+(p,n)$ vive dans le sous-espace moyen, au sens du point milieu de la géodésique dans $Gr(p,n)$ reliant les deux sous-espaces qui sous tendent les matrices en question. Le prix à payer pour imposer que le rang soit conservé est une modification de certaines de propriétés listées ci-dessus. En particulier, c'est l'invariance par congruence sous l'action des matrices inversibles qui entraîne une chute du rang de la moyenne de deux matrices de $S^+(p,n)$. Il faut donc remplacer cette propriété par une invariance plus réduite. Nous proposons de remplacer trois propriétés comme suit:

(P1') Cohérence avec les scalaires. Si A, B commutent et sont sous-tendues par le même sous-espace $A \circ B = (AB)^{1/2}$.

(P6') Invariance par homotétie et rotation. Pour $(\mu, P) \in \mathbb{R}_+^* \times \mathrm{O(n)}$ on a

$$(\mu P^T A \mu P) \circ (\mu P^T B \mu P) = \mu P^T (A \circ B) \mu P$$

(P7') Auto-dualité. $(A \circ B)^\dagger = A^\dagger \circ B^\dagger$, où $\dagger$ représente la pseudo-inversion.

Lorsqu'on cherche à définir une moyenne géométrique d'un nombre arbitraire de matrices semi-définies positives, on peut imposer qu'elle respecte l'extension des propriétés listées ci-dessus à un nombre arbitraire de matrices. Cette extension vient naturellement et il est inutile de la détailler ici. En revanche, l'extension de la moyenne (4.4) à trois matrices ou plus a fait l'objet d'un certain nombre d'articles de mathématiques récents [19, 54, 60].

Moyenne proposée La moyenne proposée part du constat que la variété de Grassman $Gr(p,n)$ des sous-espaces de dimension p dans l'espace euclidien ambiant de dimension n peut être vue comme un sous-ensemble de $S^+(p,n)$. En effet, $Gr(p,n)$ peut être identifiée au projecteurs de $\mathbb{R}^n$ sur les sous-espaces de dimension p, qui forment un sous-ensemble de $S^+(p,n)$:

$$\{P \in \mathbb{R}^{n \times n} /\ P^T = P,\ P^2 = P,\ \mathrm{Tr}\,(P) = p\}, \tag{5.1}$$

à travers notamment sa géométrie quotient $St(p,n)/O(p)$ mentionnée au chapitre 4. Il convient donc de savoir *a minima* définir une moyenne qui préserve le rang sur ce sous-espace. Or, cette sous-variété possède une structure riemannienne naturelle. La

distance au carré entre deux sous-espaces est simplement la somme des carrés des angles principaux [50] entre les deux sous-espaces considérés. La moyenne de Karcher de N sous-espaces $S_1, \cdots S_N$ de $Gr(p,n)$ est alors définie comme la solution du problème aux moindres carrés

$$\text{argmin}_{X \in Gr(p,n)} \sum_{1}^{N} d_{\text{Gr(p,n)}}(X, S_i)^2$$

Cette dernière fonction à minimiser s'écrit $\sum_{i=1}^{N} \sum_{j=1}^{p} \theta_{ij}^2$ où θ_{ij} est le j-ième angle principal entre X et S_i. Pour $N > 2$, on ne connait pas de solution explicite au problème, mais cette moyenne peut être facilement approximée [82]. De plus, on peut montrer que la moyenne de Karcher existe et est unique si les sous-espaces sont contenus dans une boule géodésique de rayon inférieur à $\pi/(4\sqrt{2})$ [2].

L'extension depuis l'espace des projecteurs de rang p à l'ensemble $S^+(p,n)$ repose sur la décomposition

$$A = UR^2U^T,$$

de toute matrice $A \in S^+(p,n)$, où U est une matrice de $St(p,n)$ et R^2 est une matrice positive définie de $P(p)$. Cette décomposition souligne l'interprétation des éléments de $S^+(p,n)$ comme des *ellipsoïdes plats* dans $\mathbb{R}^n$. U définit une base de l'espace qui sous-tend l'ellipsoïde, et R^2 définit la forme de l'ellipsoïde dans le cône de petite dimension $P(p)$. L'idée de la moyenne géométrique de N matrices $A_1, \cdots, A_N$ de $S^+(p,n)$ consiste à choisir le sous-espace moyen (moyenne des projecteurs $U_iU_i^T$) pour sous-tendre la moyenne des matrices A_i. Ainsi, l'ellipsoïde plat moyen vit dans ce sous-espace et sa forme est définie par la une moyenne géométrique (disons la moyenne de Karcher au sens de la métrique naturelle) des facteurs R_i^2 sur le cône de petite dimension $P(p)$.

Cependant, lorsqu'on considère N matrices $A_1, \cdots, A_N$ la décomposition matricielle $A_i = U_iR_i^2U_i^T$ pour $1 \leq i \leq N$ est définie à une transformation orthogonale près

$$U_i \mapsto U_iO_i, \qquad R_i^2 \mapsto O_i^T R_i^2 O_i\,, \tag{5.2}$$

avec $O_i \in O(p)$ car

$$A_i = U_iR_i^2U_i^T = U_iO_i(O_i^T R_i^2 O_i)O_i^T U_i.$$

Ces transformations orthogonales n'affectent pas la moyenne des sous-espaces mais affectent la moyenne géométrique des facteurs de rang faible puisque généralement $R_1^2 \circ R_2^2 \neq R_1^2 \circ (O^T R_2^2 O)$ pour une rotation arbitraire $O \in O(p)$, et où $\circ$ désigne la moyenne de Karcher (4.4) au sens de la métrique naturelle du cône.

Ce problème peut être résolu en choisissant des représentants (U_i, R_i) liés rigidement entre eux de la façon suivante. Concentrons-nous d'abord sur le cas où l'on souhaite définir la moyenne de deux matrices A_1, A_2 de $S^+(p,n)$. On peut décomposer les matrices de la façon suivante $A_1 = U_1R_1^2U_1^T$ et $A_2 = U_2R_2^2U_2^T$. Les représentants $(U_i, R_i^2), i = 1, 2$, sont définis à des transformations orthogonales près

$$U_i \mapsto U_iO, \qquad R_i^2 \mapsto O^T R_i^2 O.$$

Toutes les bases $U_i\text{O(p)}$ correspondent au même projecteur $U_iU_i^T$. Nous proposons de choisir dans $St(p,n)$ les représentants $Y_1 = U_1Q_1$ et $Y_2 = U_2Q_2$ où Q_1, Q_2 sont définis

par

$$(Q_1, Q_2) = \text{argmin}_{(O_1,O_2)\in O(p)\times O(p)} d_{St(p,n)}(U_1O_1, U_2O_2)\,. \tag{5.3}$$

Proposition 3. *La décomposition en valeurs singulières de $U_1^TU_2$ s'écrit*

$$U_1^TU_2 = O_1(\cos\Sigma)O_2^T, \tag{5.4}$$

où Σ est une matrice diagonale dont les entrées sont les angles principaux entre les sous-espaces engendrés par U_1 et U_2 [50]. Si la paire (O_1, O_2) est définie par (5.4)*, c'est une solution de* (5.3)*.*

Ce choix de représentants admet également l'interprétation géométrique suivante:

Proposition 4. *Soit $Y_1 = U_1Q_1$ et $Y_2 = U_2Q_2$ avec (Q_1, Q_2) une solution de (5.4). Alors la rotation $\Pi \in SO(n)$ qui envoie la base Y_1 sur la base $Y_2 = \Pi Y_1$ est une rotation qui minimise $d_{SO(n)}(\Pi, I)$ parmi toutes les rotations qui envoient Y_1 sur le sous-espace engendré par les colonnes de U_2.*

On peut définir la moyenne de ces deux bases bien choisies Y_1, Y_2 comme point milieu W de la géodésique les reliant dans $St(p, n)$. En raison du choix de ces bases particulières, ce point milieu sous-tend la moyenne des sous-espaces au sens de Grassman. En écrivant

$$A_1 = Y_1Q_1^2Y_1^T, \qquad A_2 = Y_2Q_2^2Y_2^T. \tag{5.5}$$

la moyenne proposée est alors donnée par

$$A_1 \# A_2 = W(Q_1^2 \circ Q_2^2)W^T$$

où $Q_1^2 \circ Q_2^2$ est la moyenne géométrique (4.4) de Q_1^2 et Q_2^2 (qui coïncide avec la moyenne de Karcher sur le cône).

La proposition 4 fournit une interprétation géométrique simple de cette notion de moyenne. Il y a plusieurs façons de transporter les ellipsoïdes A_1 et A_2 dans le sous-espace moyen. Différentes rotations entraîneront différentes positions respectives entre les ellipsoïdes finaux. Ce choix peut être rendu univoque et sensé en choisissant une rotation minimale.

La moyenne de N matrices $A_1, \cdots, A_N$ de $S^+(p, n)$ est définie de façon similaire. On calcule le sous-espace moyen au sens de Grassman entre les différents sous-espaces engendrés par les matrices. Puis on amène tous les ellipsoïdes dans ce sous-espace par des rotations minimales. Ensuite, on calcule une moyenne de Karcher au sens de la métrique naturelle du cône dans ce sous-espace de petite dimension. Cette moyenne admet les propriétés suivantes:

Theorem 9. *Sur le sous-espace des projecteurs, la moyenne introduite coïncide avec la moyenne de Karcher au sens de la métrique naturelle de $Gr(p, n)$. Par ailleurs, lorsque toutes les matrices engendrent le même sous-espace, la moyenne introduite coïncide avec la moyenne de Karcher sur le cône de petite dimension $P(p)$.*

Theorem 10. *La moyenne introduite possède les propriétés (PP1'), (PP2-PP5), et (PP6'-PP7') et peut donc être qualifiée de moyenne "géométrique".*

Theorem 11. *Soit N projecteurs $U_1U_1^T, \cdots, U_NU_N^T \in S^+(p,n)$. La moyenne introduite est la moyenne de Karcher des matrices $U_1U_1^T, \cdots, U_NU_N^T$ au sens de la métrique* (4.7) *sur $S^+(p,n)$ proposée au chapitre 4.*

5.2 Tenseurs de diffusion et anisotropie

Ce paragraphe évoque brièvement le travail effectué avec Anne Collard et Rodolphe Sepulchre dans l'article [45]. On introduit une nouvelle notion de moyenne sur $P(3)$ différente de la moyenne associée à la métrique naturelle du cône, mais qui s'inspire de la moyenne développée ci-dessus dans le cas rang faible à la section 5.1, en ce qu'elle repose sur un calcul de moyenne de "forme" et un calcul de moyenne d'"orientation". Ici l'approche est beaucoup plus appliquée que ce qu'elle ne l'est dans la section précédente, et sans détailler le cadre mathématique, nous ne présenterons que les motivations applicatives et une illustration des résultats.

L'application concernée est le traitement d'images de tenseurs de diffusion. La Diffusion Tensor Imaging (DTI) en anglais est une technique non-invasive qui permet de fournir une image de la diffusion des molécules d'eau dans les tissus biologiques. Dans les tissus organisés en fibres admettant des directions privilégiées, le mouvement résultant des molécules d'eau est très anisotrope puisque les molécules parcourent une plus grande distance le long des fibres. Les mesures de diffusion des molécules d'eau (qui utilisent les mêmes outils que l'imagerie par résonance magnétique IRM) permettent de caractériser cette anisotropie. Le représentation la plus répandue de cette diffusion directionnelle consiste à utiliser des tenseurs de diffusion, selon un formalisme introduit en 1994 [17].

Un tenseur de diffusion peut être codé par une matrice définie positive symétrique en chaque voxel. Tout tenseur de diffusion peut être ainsi graphiquement représenté par un ellipsoïde. Cet ellipsoïde admet comme axes principaux les vecteurs propres de la matrice associée (représentant les trois principales directions de diffusion). Selon [61], les principales informations pertinentes pour le diagnostic que l'on peut espérer extraire de ces images sont 1- la diffusivité moyenne, qui doit être un invariant par rotation, car c'est une caractéristique du milieu, et qui peut par exemple être exprimée comme la trace du tenseur de diffusion $(\lambda_1 + \lambda_2 + \lambda_3)/3$, 2- l'indice d'anisotropie de la diffusion qui doit être complètement défini par la forme de l'ellipsoïde indépendamment de ses dimensions, par exemple l'anisotropie relative $\sqrt{\sum_i \lambda_i - \langle\lambda_i\rangle^2}/\langle\lambda\rangle$ ou l'anisotropie dite "de Hilbert" que nous avons introduite dans notre travail $HA = \log(\frac{\lambda_{\max}}{\lambda_{\min}})$ où $\lambda_{\max}$ et $\lambda_{\min}$ représentent respectivement la plus grande et la plus petite valeur propre du tenseur de diffusion, en référence à la métrique projective de Hilbert dans l'orthant positif $\mathbb{R}^3_+$ [42], et 3- l'orientation des fibres, selon l'hypothèse usuelle de la colinéarité du vecteur associé à la plus grande valeur propre $\lambda_{\max}$ avec la direction des fibres.

Les données étant extrêmement bruitées, les images brutes doivent être filtrées afin d'obtenir une image interprétable biologiquement. Les traitements d'image classiques ayant été d'abord développés pour des champs scalaires, les premiers traitements d'images tensorielles ont consisté à filtrer séparément les différentes informations

scalaires contenues dans les indices de diffusivité moyenne et d'anisotropie. Cela dit, la nature tensorielle des images DTI a rapidement motivé le développement de méthodes de traitement d'images adaptées. En particulier, des méthodes basées sur la géométrie naturelle du cône $P(3)$ ont rapidement émergé (voir par exemple les premiers travaux [75, 49]) puisque le cadre riemannien permet de généraliser plusieurs méthodes de calcul dans les espaces vectoriels au contexte non-linéaire.

Notre démarche dans l'article [45] a consisté à adopter également un cadre géométrique, mais l'emphase a été placée sur la synthèse de courbes d'interpolation entre deux tenseurs. L'interpolation est une opération mathématique centrale dans le débruitage des données, basée sur des moyennes pondérées. L'interpolation proposée dans l'espace non-linéaire $P(3)$ peut apparaître comme un compromis entre des considérations d'ordre pratique (temps de calcul associé au traitement) et d'ordre théorique (propriétés d'invariance sous l'action de certains groupes de transformation). Son principal avantage est de pouvoir interpoler entre deux tenseurs anisotropes sans que le tenseur médian subisse une perte d'anisotropie (contrairement au cas où on utilise les géodésiques naturelles de $P(3)$). C'est dans l'importance qu'on donne à l'orientation du tenseur qu'on s'inspire de la métrique développée sur $S^+(p,n)$ (la chute de rang pouvant être vue comme cas limite d'une anisotropie infinie). La figure 5.1 présente des résultats d'expériences numériques. Pour plus de détails on renvoie le lecteur à l'article [45].

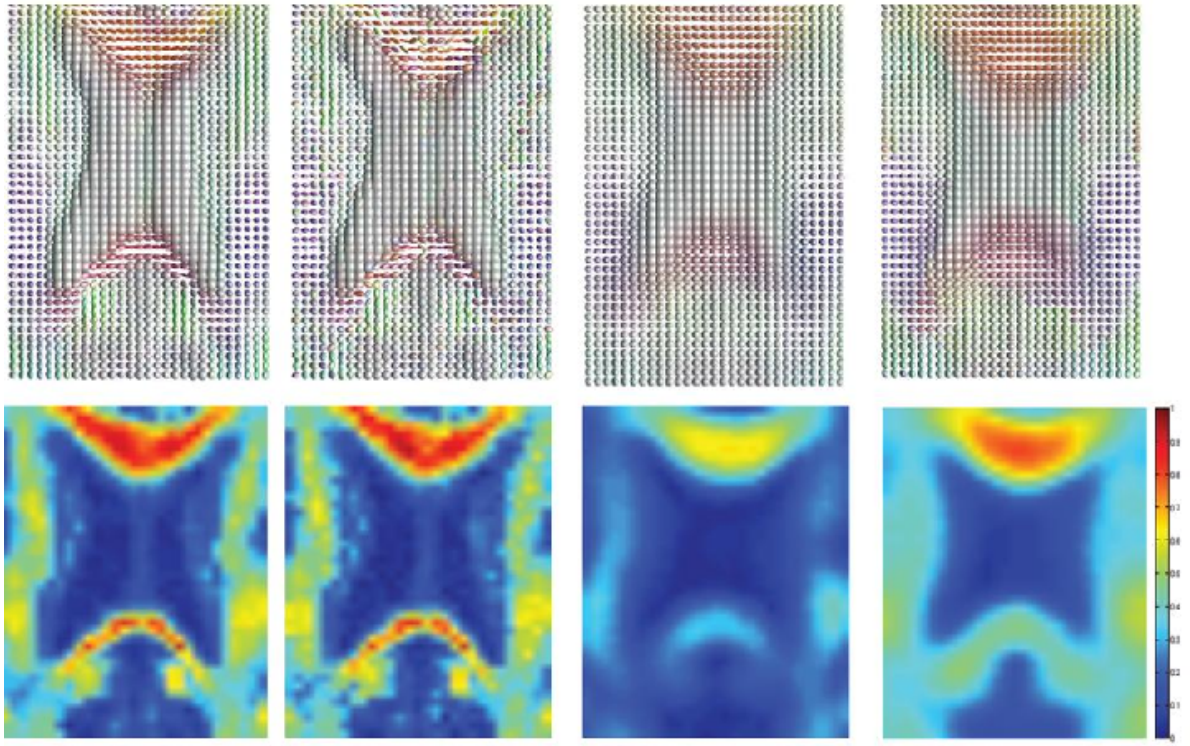


Figure 5.1: Filtrage gaussien d'une image de tenseurs de diffusion obtenue au centre de recherche du cyclotron de l'université de Liège. En haut: tranches 2D de l'image. En bas: anisotropie des tenseurs de l'image. De gauche à droite: tranche de l'image construite à partir des données réelles, image avec ajout de bruit, résultat d'un filtrage gaussien avec la méthode log-euclidienne [11] qui constitue l'état de l'art, filtrage gaussien basé sur l'interpolation proposée dans [45]. Les deux méthodes conduisent naturellement à un floutage des formes. On peut observer néanmoins que les zones de forte anisotropie dans l'image originale restent des zones d'anisotropie importante lorsqu'on utilise notre cadre [45]. L'utilisation de la métrique log-euclidienne conduit donc à une perte supérieure d'information d'anisotropie dans le filtrage.

Chapter 6

Gradient stochastique riemannien

Dans l'article [68], (et plus généralement dans [67]), on cherche à filtrer une matrice de rang faible dont on mesure aléatoirement (et éventuellement avec un bruit) certaines coordonnées. On a pour cela construit un algorithme de gradient stochastique sur la variété $S^+(p,n)$. J'ai apporté par la suite des preuves de convergence générales de la méthode [24]. Ce chapitre est divisé en deux parties. D'abord on présente l'algorithme de descente de gradient stochastique sur les variétés et ses propriétés de convergence presque sûre. Ensuite, on applique les résultats pour montrer la convergence d'un algorithme de [68].

6.1 Méthode générale et résultats

Considérons le problème d'optimisation suivant

$$\text{minimiser} \quad C(w) \tag{6.1}$$

où $w \in \mathcal{W} \subset \mathbb{R}^n$ est un paramètre de minimisation et C une fonction de coût supposée positive et trois fois continûment différentiable (a priori non convexe). La méthode sans doute la plus simple pour attaquer ce problème numériquement est appelée descente de gradient et consiste en la procédure récursive définie par la formule

$$w_{t+1} = w_t - \gamma_t H(w_t)$$

où w_t est le paramètre à l'étape $t \in \mathbb{N}$, et où $H(w_t) = \nabla C(w_t)$ est le gradient de la fonction de coût au point w_t, et où le pas $\gamma_t \geq 0$ est un gain à ajuster à chaque étape. Sous certaines hypothèses, on peut espérer que pour n'importe quelle initialisation $w_0 \in \mathcal{W}$ la suite $(w_t)_{t\geq 0}$ converge vers l'argument minimum de la fonction de coût, qui est un zéro du gradient H. Le résultat suivant par exemple est bien connu:

Proposition 5. *Si H est continue à croissance linéaire, admet un unique zéro w^* et vérifie la condition*

$$\forall w \neq w^* \quad \langle H(w); w - w^* \rangle > 0$$

et la suite des pas la condition

$$\sum \gamma_t^2 < \infty \quad \textit{and} \quad \sum \gamma_t = +\infty \tag{6.2}$$

alors $w_t \to w^*$ *quand* $t \to \infty$.

La preuve consiste à étudier les variations de la suite de terme général $\|w_t - w^*\|$ qui joue le rôle de fonction de Lyapunov, à montrer qu'elle tend vers une limite, grâce à la deuxième condition sur le pas, ce qui implique que $(w_t)_{t\geq 0}$ est bornée. La première condition sur le pas implique que $\liminf_t \langle H(w); w - w^* \rangle = 0$. On peut extraire une sous-suite de $(w_t)_{t\geq 0}$ qui ne peut donc que converger vers w^* par continuité de H. On en déduit $\|w_t - w^*\| \to 0$.

On peut supposer maintenant que l 'on n'a pas accès aux valeurs numériques du gradient $H(w)$ mais à une évaluation bruitée $H(Z, w)$ où Z est un vecteur aléatoire à valeurs dans un espace mesurable $\mathcal{Z}$, de loi dP, et où le gradient de la fonction de coût d'intérêt admet une représentation intégrale de la forme

$$\nabla C(w) = \mathbb{E}_z H(z, w) = \int H(z, w) dP(z)$$

ce qui revient à une hypothèse de bruit centré. On peut aussi formuler le problème en supposant que la fonction de coût admet la représentation intégrale $C(w) = \mathbb{E}_z Q(z, w) = \int Q(z, w) dP(z)$, où $Q(z, w)$ est la fonction de coût évaluée sous "l'évènement" z, et où $H(z, w)$ apparaît comme le gradient de la fonction de coût évaluée sous z. La méthode de gradient stochastique est une procédure récursive aléatoire définie par

$$w_{t+1} = w_t - \gamma_t H(z_t, w_t)$$

où $z_1, z_2, \cdots$ est une suite de réalisations indépendantes de la variable aléatoire Z de loi dP. On est en droit d'espérer que sous certaines hypothèses cette procédure converge vers un zéro de la fonction vectorielle $\nabla C(w)$.

On parle pour ce type de problème d'approximation stochastique, car à chaque étape on approxime une somme par un de ces éléments. La formulation et la procédure de gradient stochastique qui s'ensuit sont adaptées au moins à deux problèmes d'intérêt pratique: 1- le problème de minimisation *en ligne*, où l'utilisateur a accès à une suite d'évènements aléatoires $z_1, z_2, \cdots$ tirés selon la loi dP inconnue, et 2- si $H(z, w)$ est facile à calculer et que la loi dP est facile à simuler, on peut rendre la recherche de minima locaux aléatoire pour des questions d'efficacité numérique. En effet le calcul de l'intégrale dans sa totalité à chaque itération débouche sur une procédure déterministe mais peut être lourd. Une espérance étant une moyenne pondérée, si il y a beaucoup de redondance dans les données la somme peut être bien approximée par quelques un de ses éléments. De plus, lorsque $C(w)$ est elle-même une approximation empirique d'une vraie fonction de coût que l'on cherche à minimiser et à laquelle on n'a pas accès, mettre en place des méthodes lourdes pour trouver avec une grande précision son argument minimum peut s'avérer inutile. C'est entre autres en raison de cette idée, introduite et détaillée dans [38], que les méthodes de gradient stochastiques ont connu un regain d'intérêt récent dans le domaine de l'apprentissage statistique sur de grandes bases de données.

Méthode de gradient stochastique dans un contexte riemannien Ici on s'intéresse au problème (6.1) où $C(w) = \mathbb{E}_z Q(z, w)$ est une fonction non négative trois fois différentiable et où w est un paramètre de minimisation appartenant à une variété riemannienne lisse $\mathcal{M}$. Sur $\mathcal{M}$, on propose de remplacer la procédure récursive habituelle par

$$w_{t+1} = \exp_{w_t}(-\gamma_t H(z_t, w_t)) \tag{6.3}$$

où $\exp_w$ est l'application exponentielle (riemannienne) en w, et $H(z, w)$ peut être vu comme le gradient riemannien de la fonction de coût approximé, i.e. $H(z, w)$ est un vecteur aléatoire de l'espace tangent à w qui vérifie $\mathbb{E}_z H(z, w) = \int H(z, w) dP(z) = \nabla C(w)$ où $\nabla C(w)$ représente le gradient riemannien de C au point $w \in \mathcal{M}$. La mise à jour proposée (6.3) est une généralisation directe de l'algorithme de gradient stochastique dans le cas euclidien. En effet, le gradient riemannien repère la direction dans laquelle une fonction s'accroît le plus dans une petite boule (au sens de la métrique riemannienne choisie) autour d'un point. Ainsi, à chaque étape, le paramètre se déplace selon la géodésique qui émane de w_t, dans la direction définie par $H(z_t, w_t)$ avec une intensité $\|H(z_t, w_t\|$. Si la variété est $\mathbb{R}^n$ muni du produit scalaire euclidien habituel, les géodésiques sont des droites, l'exponentielle une addition, et les définitions coïncident. On peut remarquer que la procédure définie plus haut est totalement intrinsèque, c'est-à-dire qu'elle ne dépend pas d'un choix arbitraire de coordonnées locales sur la variété.

Dans de nombreux exemples, on ne dispose pas de formule explicite pour l'application exponentielle, et son calcul peut être numériquement coûteux. Etant donné que dans les méthodes de gradient stochastique le pas γ_t tend vers 0 quand t tend vers l'infini, on pourra avoir recours à une rétraction, qui est une approximation de l'exponentielle au premier ordre. En effet, on définit une rétraction comme une application $R_w(v) : T_w\mathcal{M} \mapsto \mathcal{M}$ qui envoie l'espace tangent à w vers la variété, et qui vérifie $d(R_w(tv), \exp_w(tv)) = O(t^2)$. Cela conduit à la procédure alternative

$$w_{t+1} = R_{w_t}(-\gamma_t H(z_t, w_t)) \tag{6.4}$$

Résultats de convergence Dans les applications sur les variétés, les fonctions de coût sont rarement convexes. Aussi, l'on ne peut espérer montrer qu'une convergence vers un point critique de la fonction de coût. C'est l'étude de ces points critiques qui permet alors de conclure. Sous certaines hypothèses on peut montrer la convergence presque sûre vers un point critique de la fonction de coût pour l'algorithme euclidien décrit plus haut. Ce type de convergence nous intéresse tout particulièrement, car en pratique elle assure que quelle que soit la séquence de bruits aléatoire à laquelle on est confrontée la convergence est garantie. L'étude de la vitesse de convergence est en revanche beaucoup plus difficile. Dans l'article [24], nous avons obtenu les résultats de convergence suivants comme généralisation du cas euclidien.

Theorem 12. *Considérons l'algorithme* (6.3) *sur une variété riemannienne lisse et connexe* $\mathcal{M}$ *ayant un rayon d'injectivité uniformément minoré par* $I > 0$. *Supposons que la suite des pas* $(\gamma_t)_{t\geq 0}$ *satisfait les conditions standard* (6.2). *Supposons qu'il existe un compact* K *tel que* $w_t \in K$ *pour tout* $t \geq 0$. *Supposons également que le gradient est borné sur* K, *i.e.* $\|H(z, w)\| \leq A$ *pour tout* $w \in K$ *avec* $A > 0$. *Alors* $C(w_t)$ *converge presque sûrement (p.s.) et* $\nabla C(w_t) \to 0$ *p.s.*

Theorem 13. *Considérons l'algorithme* (6.4) *avec R_w une rétraction deux fois continûment différentiable. Sous les hypothèses du théorème 12, $C(w_t)$ converge presque sûrement (p.s.) et $\nabla C(w_t) \to 0$ p.s.*

La preuve du théorème 12 est assez proche de celle développée dans le cadre euclidien dans l'article [37], et de façon générale elle repose sur des arguments classiques en approximation stochastique (voir e.g. [79]) transposés dans le contexte riemannien. Comme on l'a déjà évoqué plus haut, la preuve du théorème 13 est basée sur le fait que le pas tend vers 0, ce qui implique qu'une approximation au premier ordre de l'application exponentielle est suffisante pour garantir la convergence asymptotique.

On peut remarquer que la condition de confinement du paramètre dans un compact qui apparaît ci-dessus est évidemment vérifiée lorsque l'on travaille sur une variété compacte. Cette hypothèse peut être relaxée au prix d'une série d'hypothèses supplémentaires dans le cas euclidien. L'article [24] contient un troisième théorème où l'on donne des conditions sous lesquelles on peut se passer de cette hypothèse de confinement sur des variétés de Hadamard, dont le cône des matrices définies positives muni de sa métrique naturelle est un exemple. Ces variétés ont une courbure négative dont les effets peuvent conduire à déstabiliser l'algorithme de descente de gradient. Une modification du pas permet alors de lutter contre ces effets. Ici, nous ne rappelons pas ce théorème car en pratique les hypothèses sont difficiles à vérifier. Cela dit, il peut être appliqué à l'espace euclidien de dimension n et apparaît comme un peu plus général que le théorème de [37] qui l'a inspiré, comme illustré par la section applicative 6.2 ci-dessous.

Un exemple illustratif d'application Une méthode d'analyse des données bien connue est l'analyse en composantes principales (ACP). A partir d'un ensemble de N données décrites par une série de vecteurs $z_1, \cdots, z_N$ de $\mathbb{R}^n$ (par exemple la répartition des voix attribuées à n candidats d'une élection présidentielle dans chacun des N départements), l'ACP consiste à, après avoir centré les données, former la matrice de covariance empirique de taille $n \times n$ des données, à la diagonaliser, et à retenir les vecteurs principaux associées aux plus grandes valeurs propres comme variables explicatives des disparités dans les données. Cette méthode est principalement utilisée pour analyser les données (par exemple dans le cas de l'élection le premier axe va capturer les différences entre sensibilités de gauche ou de droite, le second la tendance sécuritaire etc.), et/ou pour réduire les données, car en pratique il est fréquent qu'un nombre réduit d'axes permette de capturer les disparités dans les données, les axes associées aux faibles valeurs propres apportant généralement peu d'information.

Le problème de l'ACP pour la réduction de données en ligne se pose ainsi: étant donné une suite de vecteurs aléatoires $z_1, \cdots z_t, \cdots$ de dimension n indépendants et identiquement distribués, de matrice de covariance $E(z_t z_t^T) = A$, on cherche une procédure récursive d'estimation du sous-espace engendré par les p plus grandes valeurs propres de A. Le problème revient à identifier un élément de la variété de Grassman $Gr(p, n)$ introduit au chapitre 4. On peut représenter les éléments de $Gr(p, n)$ par les projecteurs WW^T où W appartient à la variété de Stiefel $St(p, n)$. On définit la fonction de coût

suivante

$$C(W) = -\frac{1}{2}\mathbb{E}_z[z^T W^T W z] = -\frac{1}{2}\mathrm{Tr}\left(W^T A W\right)$$

qui est minimale lors que W est une base du sous-espace dominant de A. La représentation est invariante sous l'action $W \mapsto WO,\ O \in \mathcal{O}(p)$, et l'espace d'état est donc l'ensemble des classes d'équivalences $[W] = \{WO\},\ O \in \mathcal{O}(p)$ qui forment une représentation quotient $St(p,n)/O(p)$ de la variété de Grassman $\mathrm{Gr}(p,n)$. Le gradient riemannien sous l'évènement z s'écrit $H(z,W) = (I - WW^T)zz^T W$. Nous avons le résultat suivant

Proposition 6. *Supposons que $z_1, z_2, \cdots$ sont uniformément bornés. Considérons l'algorithme de gradient stochastique*

$$W_{t+1} = W_t V_t \cos(\gamma_t \Theta_t) V_t^T + U_t \sin(\gamma_t \Theta_t) V_t^T \tag{6.5}$$

où $U_t \Theta_t V_t$ est la décomposition en valeurs singulières compacte de la matrice $(I - W_t W_t^T) z_t z_t^T W_t$. Alors W_t converge p.s. vers un sous-espace invariant de A.

La preuve est une application directe du théorème 12. D'après le résultat, W_t converge p.s. vers un point vérifiant $\nabla C(W) = 0$, i.e. $AW = WW^T AW$. Pour un tel point il existe M tel que $AW = WM$, ce qui prouve que W est un sous-espace invariant de A. Une analyse locale montre que seul le sous-espace dominant est stable sous certaines hypothèses basiques [72]. Le théorème 13 implique directement:

Proposition 7. *Soit R_W une rétraction deux fois différentiable. L'algorithme*

$$W_{t+1} = R_{W_t}\left(W_t + \gamma_t (I - W_t W_t^T) z_t z_t^T W_t\right) \tag{6.6}$$

converge p.s. vers un sous-espace invariant de A.

Considérons la rétraction définie par $R_W(\gamma H)$=qf$(W+\gamma H)$ où qf() extrait le facteur orthogonal dans la décomposition QR de son argument. Pour des petits pas γ_t, cette rétraction consiste à retrancher le gradient dans l'espace euclidien ambiant $\mathbb{R}^{n\times p}$, puis à orthonormaliser les colonnes de la matrice obtenue. C'est une rétraction lisse [1]. L'algorithme (6.6) avec cette rétraction est appelé algorithme d'Oja et sa convergence a été démontrée dans [71]. Ici il découle directement du théorème 13.

Cet exemple permet d'illustrer les bénéfices apportés par la rétraction en terme d'efficacité numérique. En effet, la mise à jour géodésique (6.5) requiert $O(nr^2) + O(r^3)$ opérations à chaque étape, quand la mise à jour (6.6) requiert $O(nr^2)$ opérations, ce qui peut être avantageux lorsque r est grand.

6.2 Régression dans l'espace des matrices de rang faible

Nous présentons ici les résultats de l'article [68] écrit dans le cadre du co-encadrement avec Rodolphe Sepulchre de la thèse de Gilles Meyer. Motivé par des applications

en apprentissage statistique, on s'intéresse au problème de régression suivant. Soit le modèle

$$y = \mathrm{Tr}\,(AX) + \epsilon$$

où ϵ est un bruit centré, $X = xx^T$ (avec x vecteur aléatoire à valeurs dans $\mathbb{R}^n$) est une matrice aléatoire semi-définie positive de rang 1, et où A est une matrice semi-définie positive inconnue de rang fixé p, i.e. $A \in S^+(p,n)$. On cherche à résoudre un problème de régression dans l'espace des matrices de rang faible, donc à inférer le paramètre inconnu A à partir de réalisations du couple (X,y) tiré selon une loi de probabilité inconnue. Pour ce faire on adopte une approche basée sur les moindres carrés, en cherchant à minimiser la fonction de coût

$$C(W) = \int (\mathrm{Tr}\,(WX) - y)^2 dP(x,y) = \int (x^T W x - y)^2 dP(x,y) \tag{6.7}$$

où dP représente la probabilité inconnue selon laquelle le couple (xx^T, y) est tiré. Sur le cône (convexe) des matrices semi-définies positives, la fonction de coût est convexe, et le problème admet un unique minimum. En revanche pour $W \in S^+(p,n)$, la contrainte de rang rend le problème non-convexe.

Ce problème est motivé par une méthode de classification de données. Supposons qu'un certain type de donnée (une image) est décrite par un vecteur $z \in \mathbb{R}^n$ de grande dimension (descripteurs). On peut évaluer la similarité de deux données z_i, z_j au sens d'une certaine métrique sur $\mathbb{R}^n$ appelée distance de Mahalanobis

$$d^2_W(z_i, z_j) = (z_i - z_j)^T W (z_i - z_j)$$

où W est une matrice semi-définie positive. Si $d^2_W(z_i, z_j)$ dépasse un certain seuil, on considère que z_i et z_j appartiennent à des classes différentes (par exemple une image de chien et une image de chat). Le but de l'apprentissage de distance de Mahalanobis est d'identifier à partir d'un ensemble d'entraînement (une série de paires auxquelles un expert assigne une distance qui lui semble raisonnable) un paramètre W tel que la distance intra-classe soit faible et la distance inter-classe importante, afin de correctement prédire la classe d'appartenance d'une nouvelle donnée. De plus, pour des raisons de stockage, et de coût de calcul à chaque itération, lorsque n est très grand, on peut imposer que W soit de rang faible $p \ll n$. Cette approximation ne dégrade pas nécessairement le pouvoir de prédiction de W, puisqu'en imposant le rang faible on cherche à ne retenir que les valeurs propres les plus importantes de W, et donc les plus discriminantes. On choisit ici de traiter le problème de l'apprentissage en ligne. En posant $x = z_i - z_j$, et y la distance assignée par l'expert au couple z_i, z_j l'idée est de chercher le paramètre $W \in S^+(p,n)$ comme minimum de la fonction de coût (6.7), à partir d'une suite d'exemples $(x_t, y_t)_{t\geq 0}$ tirés selon la loi dP.

Calcul des gradients riemanniens On peut considérer deux géométries d'intérêt sur $S^+(p,n)$. La première est une géométrie naïve où l'on écrit toute matrice $W \in S^+(p,n)$ sous la forme $W = GG^T$ avec $G \in \mathbb{R}^{n\times p}_*$ i.e. G est une matrice de dimension $n \times p$ et de rang plein. Cette décomposition induit une géométrie quotient $S^+(p,n) \simeq GL(n)/O(p)$

qu'on peut munir d'une métrique riemannienne induite par la norme de Frobenius dans l'espace total $\mathbb{R}_*^{n\times p}$. Le gradient riemannien de la fonction de coût (6.7) évaluée sous l'évènement $z_t = (x_t, y_t)$ s'écrit dans cette géométrie

$$H(z_t, G_t) = 2(\hat{y}_t - y_t)x_t {x_t}^T G_t$$

où la distance au carré prédite par le paramètre courant est $\hat{y}_t = \mathrm{Tr}\left(W_t x_t x_t^T\right) = \|G_t^T x_t\|^2$. L'algorithme de gradient riemannien stochastique (6.3) s'écrit alors

$$G_{t+1} = G_t - \gamma_t(\|G_t^T x_t\|^2 - y_t)x_t x_t^T G_t \tag{6.8}$$

D'autre part, en considérant la géométrie polaire décrite au chapitre 4 et basée sur la décomposition $W = UR^2U^T$, munie de la métrique (4.7) où l'on écrit le facteur $k = \lambda/(1-\lambda)$, on obtient le gradient riemannien

$$H(z_t, U_t, R_t) = \left(2\lambda(\hat{y}_t - y_t)(I - U_tU_t^T)X_tU_tR_t^2, (1-\lambda)(\hat{y}_t - y_t)R_tU_t^TX_tU_tR_t\right)$$

En utilisant la rétraction vue dans la section précédente on obtient l'algorithme en ligne

$$\begin{aligned} U_{t+1} &= \mathrm{qf}\left(U_t - 2\lambda\gamma_t(\hat{y}_t - y_t)(I - U_tU_t^T)\mathrm{Sym}(X_t)U_tR_t^2\right), \\ R_{t+1} &= R_t \exp\left(-\frac{1}{2}(1-\lambda)\gamma_t(\hat{y}_t - y_t)R_tU_t^T\mathrm{Sym}(X_t)U_tR_t\right), \end{aligned} \tag{6.9}$$

où le facteur λ apparaît comme un paramètre de réglage permettant de favoriser l'identification du du sous-espace ($\lambda \to 1$) ou de la matrice de petite dimension R ($\lambda \to 0$).

Résultats théoriques Pour montrer la convergence de l'algorithme (6.9) avec nos résultats il faut d'abord prouver le confinement du paramètre. Ce dernier point semble difficile, et l'on se contentera des performances satisfaisantes en pratique de l'algorithme (6.9) soulignées dans [68]. En revanche, nos résultats théoriques généraux sur le gradient stochastique riemannien permettent de prouver simplement la convergence de l'algorithme (6.8) pourvu que la séquence $(\gamma_t)_{t\geq 0}$ soit bien choisie:

Proposition 8. *Soit $(x_t)_{t\geq 0}$ vecteurs gaussiens de $\mathbb{R}^n$ indépendants et ayant pour matrice de covariance la matrice identité. Supposons que $y_t = x_t^T V x_t$ est généré par un paramètre inconnu $V \in S^+(p,n)$. L'algorithme de descente de gradient riemannien* (6.8) *où le pas vérifie $\gamma_t = s_t/f(G_t)$ avec $f(G_t) = \max(1, \|G_t\|^6)$ et où la séquence $(s_t)_{t\geq 0}$ vérifie la condition classique* (6.2) *est telle que $G_tG_t^T = W_t \to W_\infty$ et $\nabla C(W_t) \to 0$ p.s. De plus, si W_∞ a pour rang $p = n$ nécessairement $W_\infty = V$ p.s. Si $p < n$, alors W converge presque sûrement vers un sous-espace invariant de V de dimension r. Si il s'agit du sous-espace dominant de V, alors $W_\infty = V$.*

Résultats numériques et extensions Certaines géométries ont aussi été développées sur l'ensemble des matrices de taille $n \times d$ et de rang p fixé. Elles permettent d'attaquer le problème de la complétion de matrice, qui a reçu énormément d'attention au cours des dernières années. De manière générale, les méthodes développées sont compétitives avec l'état de l'art, comme le montrent les articles [68, 67, 69].

Conclusion et perspectives

Nous avons présenté deux types d'approches de certains problèmes de filtrage. Dans une première partie, on introduit une approche assez systématique où un système dynamique admet des symétries et l'on cherche des procédures de filtrage qui admettent les mêmes symétries. Les résultats de convergence qui en découlent rappellent fortement les résultats déjà connus sur les estimateurs invariants en statistiques (cas statique). Dans une deuxième partie, on considère des problèmes de filtrage statique ou débruitage lorsque l'état appartient à une variété riemannienne, qu'on cherche alors à munir de métriques pertinentes. Cette deuxième approche est moins systématique, car de nombreuses géométries sont possibles, et il faut faire des choix. De plus, il faut composer avec des considérations sur la complexité calculatoire des méthodes, car si l'approche du filtrage par les variétés riemanniennes permet de maintenir numériquement les contraintes sur l'état estimé, son coût numérique peut être élevé.

Attardons-nous maintenant sur quelques pistes de continuation de ces travaux pour les années à venir. Commençons avec le sujet du filtrage appliqué à la localisation. La capacité d'un robot à se localiser est essentielle pour la navigation (le contrôleur aérien doit savoir précisément où sont les avions) et pour l'autonomie en robotique (de l'aspirateur autonome à la voiture sans conducteur). L'explosion récente de capteurs aidant à la localisation (GPS, caméras vidéo, télémètres lasers fournissant des images de profondeur), certains de très bas-coût aux performances dégradées, mais qu'on peut multiplier sur une même plateforme, ainsi que l'apparition incessante de nouveaux capteurs (surtout dans le domaine des images de profondeur), rendent le sujet de la fusion de données, et de la robotique probabiliste (qui consiste à tenir explicitement compte des bruits de mesure et incertitudes de chaque capteur), plus actuels que jamais. On peut dénombrer plusieurs pistes de recherche liées à ce domaine.

L'approche stochastique du filtrage invariant, à laquelle nous nous sommes intéressés récemment [15, 14] pour pouvoir attribuer des intervalles de confiance à nos estimations (une information précieuse pour un contrôleur aérien par exemple), paraît prometteuse en résultats théoriques. Les travaux en cours montrent qu'il est possible d'obtenir des garanties de convergence de la loi de l'erreur d'estimation vers une distribution stationnaire lorsqu'on considère une dynamique sur un groupe de Lie, l'espace de configuration d'un véhicule étant SE(3) en navigation. On peut penser en particulier à l'extension suivante. Lorsqu'on cherche à estimer de surcroît des paramètres constants (les nombreux biais des capteurs), la théorie actuelle est insuffisante pour obtenir la convergence. Mais on doit pouvoir tirer parti, avec un algorithme Espérance-Maximisation (EM), ou

en s'ouvrant sur le vaste domaine du filtrage particulaire de type Rao-Blackwell par exemple (où l'on conditionne l'estimation par le biais), du fait qu'à biais connu la loi de l'erreur d'estimation peut être bien estimée. Pouvoir étendre les résultats à ce cas serait peut-être la clé de son véritable succès pour des applications industrielles de robotique mobile et surtout de navigation inertielle, où l'estimation des biais des capteurs (qu'ils soient de coût faible ou élevé) est inévitable.

Toujours sur ce thème, on souhaite aborder le champ extrêmement vaste de la fusion avec capteurs visuels (images classiques ou de profondeur). Pour le problème très étudié ces dernières années du SLAM (localisation et cartographie simultanée), l'approche par filtrage invariant semble ouvrir une avenue de recherche, comme l'indiquent les idées introduites dans [23] (voir aussi le paragraphe 3.4 du présent mémoire). La robotique probabiliste [86], qui constitue l'état de l'art sur le sujet est très axée sur des méthodes numériques extensives et bénéficierait de méthodes simples apportant des garanties mathématiques. D'un point de vue pratique, le centre de robotique de Mines ParisTech dispose d'un véhicule routier expérimental équipé de divers capteurs, dont certains ne sont apparus que récemment (on pense au capteur Velodyne qui fournit des images de profondeur à partir de télémètres laser multi-couches). L'adaptation des méthodes invariantes au problème particulier de la localisation de ce véhicule est riche en pistes de recherche.

Un autre domaine d'intérêt concerne l'estimation de l'état interne de systèmes dynamiques de grande dimension et les matrices de rang faible. Le filtre de Kalman de rang faible est une modification du filtre de Kalman étendu où la matrice de covariance intervenant dans l'équation de Riccati est approximée par une matrice de rang faible. Son étude théorique est encore quasi-inexistante et reste à faire (en temps continu et en temps discret). Nous souhaiterions ensuite l'appliquer à de l'assimilation de données en océanographie où la dynamique est décrite par des équations aux dérivées partielles qui, discrétisées, conduisent à un état de grande dimension. Le but serait alors d'obtenir des garanties pour des systèmes décrits par les équations de Saint-Venant, et de Burgers multi-couches. Nous avons déposé un projet de collaboration franco-indienne sur le sujet avec Maëlle Nodet et Didier Auroux. Notons également qu'il est possible que notre travail antérieur sur l'interpolation de matrices de rang faible s'avère utile lorsque les observations sont éparses dans le temps ou dans l'espace.

Les problèmes d'estimation et d'inférence où l'on dispose d'une très large quantité de données à traiter prennent une place croissante dans les mathématiques appliquées d'aujourd'hui (le SLAM cité ci-dessus est un autre exemple), et l'on pourra aussi aborder certains problèmes d'apprentissage comme on l'a déjà fait. Notamment, on aimerait consacrer encore un peu de temps à l'étude de l'approximation stochastique riemannienne, qui peut s'avérer utile dans un contexte d'apprentissage sur de grandes bases de données. De nombreux résultats existent dans le cas euclidien, et on peut s'interroger sur l'existence d'une contrepartie riemannienne. Notamment, on peut accélérer la convergence en moyennant les itérations [79]. On aimerait appliquer ce genre de technique sur les variétés, mais cela implique un calcul de moyenne sur la variété à chaque étape, et débouche sur de nouveaux problèmes intéressants, puisque les preuves de convergence

actuelles risquent de ne pas pouvoir être facilement transposables.

Un axe de recherche d'un autre ordre concerne la planification de mouvement. Nous pensons ici au travail de thèse de Jean Grégoire, que je co-encadre avec Arnaud de la Fortelle, sur la coordination et la planification de trajectoires de véhicules aux carrefours [52, 51], qui n'a pas été détaillé dans ce mémoire par souci de cohérence. Nous aimerions dans le futur introduire de l'incertitude dans la mesure et la commande, et chercher à tenir compte explicitement de ces incertitudes dans nos méthodes actuelles.

Mentionnons finalement depuis 2011, une collaboration avec MANITOWOC, leader mondial des grues de chantier, sur le domaine du contrôle de grues. Même si le travail est confidentiel, on espère pouvoir extraire des problématiques rencontrées certaines questions théoriques à creuser dans le futur.

Bibliography

[1] P-A. Absil, R. Mahony, and R. Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2009.

[2] B. Afsari. Riemannian Lp center of mass: existence, uniqueness, and convexity. *Proceedings of the American Mathematical Society*, 139:655–673, 2011.

[3] N. Aghannan and P. Rouchon. On invariant asymptotic observers. In *41st IEEE Conference on Decision and Control*, pages 1479–1484, 2002.

[4] N. Aghannan and P. Rouchon. On invariant asymptotic observers. In *Decision and Control, 2002, Proceedings of the 41st IEEE Conference on*, volume 2, pages 1479–1484. IEEE, 2002.

[5] N. Aghannan and P. Rouchon. An intrinsic observer for a class of Lagrangian systems. *Automatic Control, IEEE Transactions on*, 48(6):936–945, 2003.

[6] L. Alvarez, F. Guichard, P.L. Lions, and J.M. Morel. Axioms and fundamental equations of image processing. *Archive for rational mechanics and analysis*, 123(3):199–257, 1993.

[7] T. Ando, C.K. Li, and R. Mathias. Geometric means. *Linear algebra and its applications*, 385:305–334, 2004.

[8] L. Arnold, A. Auger, N. Hansen, and Y. Ollivier. Information-geometric optimization algorithms: a unifying picture via invariance principles. *arXiv preprint arXiv:1106.3708*, 2011.

[9] V. Arnold. *Mathematical Methods of Classical Mechanics*. Mir Moscou, 1976.

[10] V.I. Arnol'd. Sur la géométrie différentielle des groupes de lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. *Ann. Inst. Fourier*, 16:319–361, 1966.

[11] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Log-euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic resonance in medicine*, 56(2):411–421, 2006.

[12] D. Auroux and S. Bonnabel. Symmetry-based observers for some water-tank problems. *Automatic Control, IEEE Transactions on*, 56(5):1046–1058, 2011.

[13] R. Balian, Y. Alhassid, and H. Reinhardt. Dissipation in many-body systems: a geometric approach based on information theory. *Phys. Reports*, 131:1–146, 1986.

[14] A. Barrau and S. Bonnabel. Intrinsic filtering on Lie groups with discrete-time observations. In *In preparation*, 2013.

[15] A. Barrau and S. Bonnabel. Intrinsic filtering on SO(3) with discrete-time observations. In *Decision and Control, 2013, Proceedings of the 52nd IEEE Conference on*. IEEE, 2013.

[16] A. Barrau and S. Bonnabel. A note on the intrisic Cramer-Rao bound. In *Geometric Science of Information*. Springer LCNS, 2013.

[17] P.J. Basser, J. Mattiello, and D. LeBihan. Mr diffusion tensor spectroscopy and imaging. *Biophysical journal*, 66(1):259–267, 1994.

[18] J. O. Berger. *Statistical decision theory and Bayesian analysis*. Springer Verlag, 1985.

[19] R. Bhatia and R. Karandikar. Monotonicity of the geometric mean. *Math. Ann.*, 353(4):1453–1467, 2012.

[20] S. Bonnabel. Left-invariant extended Kalman filter and attitude estimation. In *Decision and Control, 2007 46th IEEE Conference on*, pages 1027–1032. IEEE, 2007.

[21] S. Bonnabel. *Observateurs asymptotiques invariants: théorie et exemples*. PhD thesis, Ecole des Mines de Paris, 2007.

[22] S. Bonnabel. A simple intrinsic reduced-observer for geodesic flow. *Automatic Control, IEEE Transactions on*, 55(9):2186–2191, 2010.

[23] S. Bonnabel. Symmetries in observer design: review of some recent results and applications to EKF-based SLAM. In *Robot Motion and Control 2011*, pages 3–15. Springer, 2012.

[24] S. Bonnabel. Stochastic gradient descent on riemannian manifolds. *IEEE Trans. on Automatic Control*, 58(9), 2013.

[25] S. Bonnabel, A. Astolfi, and R. Sepulchre. Contraction and observer design on cones. In *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pages 7147–7151. IEEE, 2011.

[26] S. Bonnabel, A. Collard, and R. Sepulchre. Rank-preserving geometric means of positive semi-definite matrices. *Linear Algebra and Its Applications*, 438(8):3202–3216, 2013.

[27] S. Bonnabel, J-E. Deschaud, and E. Salaun. A simple nonlinear filter for low-cost ground vehicle localization system. In *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pages 3270–3275. IEEE, 2011.

[28] S. Bonnabel, P. Martin, and P. Rouchon. Symmetry-preserving observers. *Automatic Control, IEEE Transactions on*, 53(11):2514–2526, 2008.

[29] S. Bonnabel, P. Martin, and P. Rouchon. Non-linear symmetry-preserving observers on Lie groups. *Automatic Control, IEEE Transactions on*, 54(7):1709–1713, 2009.

[30] S. Bonnabel, P. Martin, P. Rouchon, and E. Salaun. A separation principle on lie groups. In *18th IFAC world congress 2011*, pages 8004–8009, Milano, Italy, 2011.

[31] S. Bonnabel, P. Martin, and E. Salaun. Invariant extended Kalman filter: theory and application to a velocity-aided attitude estimation problem. In *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*, pages 1297–1304. IEEE, 2009.

[32] S. Bonnabel, M. Mirrahimi, and P. Rouchon. Observer-based hamiltonian identification for quantum systems. *Automatica*, 45(11):1144–1155, 2009.

[33] S. Bonnabel and E. Salaün. Design and prototyping of a low-cost vehicle localization system with guaranteed convergence properties. *Control Engineering Practice*, 19(6):591–601, 2011.

[34] S. Bonnabel and R. Sepulchre. Riemannian metric and geometric mean for positive semidefinite matrices of fixed rank. *SIAM J. matrix anal. appl.*, 31, 2009.

[35] S. Bonnabel and R. Sepulchre. The geometry of low-rank kalman filters. In *Arxiv preprint*. 2012.

[36] S. Bonnabel and R. Sepulchre. The geometry of low-rank kalman filters. In *Matrix Information Geometry*, pages 53–68. Springer, 2013.

[37] L. Bottou. *Online Algorithms and Stochastic Approximations*. Online Learning and Neural Networks, Edited by David Saad, Cambridge University Press, 1998.

[38] L. Bottou and O. Bousquet. *The trade-offs of large scale learning*. J.C. Platt, D. Koller, Y. Singer, and S. Roweis, editors. Advances in Neural Information Processing Systems, volume 20, MIT Press, 2008.

[39] P. Bougerol. Kalman filtering with random coefficients and contractions. *Siam J. Control and Optimization*, 31:942–959, 1993.

[40] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[41] P. Brasseur and J. Verron. The seek filter method for data assimilation in oceanography: a synthesis. *Ocean Dynamics*, 56(5-6):650–661, 2006.

[42] P.J. Bushell. Hilbert's metric and positive contraction mappings in a banacb space. *Archive for Rational Mechanics and Analysis*, 1973.

[43] E.J. Candès and Y. Plan. Tight oracle bounds for low-rank matrix recovery from a minimal number of random measurements. *IEEE Transactions on Information Theory*, 57(4):2342–2359, 2009.

[44] E. Cartan. La méthode du repère mobile, la théorie des groupes continus, et les espaces généralisés. *Exposés de Géométrie, Hermann, Paris*, 1935.

[45] A. Collard, S. Bonnabel, C. Phillips, and R. Sepulchre. An anisotropy preserving metric for dti processing. *arXiv:1210.2826, accepted in Journal of Computer Vision*, 2012.

[46] Dick P Dee. Simplification of the kalman filter for meteorological data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 117(498):365–384, 1991.

[47] N.T. Dinh, S. Bonnabel, , and R. Sepulchre. Contraction and observer design on cones. In *Decision and Control and European Control Conference (CDC-ECC), 2013 52nd IEEE Conference on*, 2013.

[48] J. Faraut and A. Korányi. *Analysis on symmetric cones*. Clarendon Press Oxford, 1994.

[49] P.D. Fletcher and S. Joshi. Riemannian geometry for the statistical analysis of diffusion tensor data.Signal Processing, 2007.

[50] G.H. Golub and C. Van Loan. *Matrix Computations*. John Hopkins University Press, 1983.

[51] J. Gregoire, S. Bonnabel, and A. de La Fortelle. Optimal cooperative motion planning for vehicles at intersections. In *In IEEE Intelligent Vehicles Symposium*, 2012.

[52] J. Gregoire, S. Bonnabel, and A. de La Fortelle. Priority-based coordination of robots. *arXiv preprint arXiv:1306.0785*, 2013.

[53] T. Hervier, S. Bonnabel, and F. Goulette. Accurate 3D maps from depth images and motion sensors via nonlinear Kalman filtering. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pages 5291–5297. IEEE, 2012.

[54] J. Holbrook. No dice: a deterministic approach to the cartan centroid. *J. Ramanujan Math. Soc.*, 27(4):509–521, 2011.

[55] S. Jiang and M. Ghil. Tracking nonlinear solutions with simulated altimetric data in a shallow water model. *Journal of physical oceanography*, 27:72–95, 1997.

[56] R. E. Kalman and R. S. Bucy. New results in linear filtering and prediction theory. *Journal of basic Engineering*, 83(3):95–108, 1961.

[57] C. Lageman, J. Trumpf, and R. Mahony. Gradient-like observers for invariant dynamics on a lie group. *Automatic Control, IEEE Transactions on*, 55(2):367–377, 2010.

[58] L. Landau and E. Lifshitz. *Mécanique*. Mir, Moscou, 4th edition, 1982.

[59] P. Langevin. Sur la théorie du mouvement brownien. *CR Acad. Sci. Paris*, 146(530-533), 1908.

[60] J. Lawson and Y. Lim. Monotonic properties of the least squares mean. *Math. Ann.*, 351(2):267–279, 2011.

[61] D. Le Bihan, J.F. Mangin, C. Poupon, C. Clark, S. Pappata, N. Molko, and H. Chabriat. Diffusion tensor imaging: Concepts and applications. *Journal of Magnetic Resonance Imaging*, 13:534–546, 2001.

[62] C. Le Bris and P. Rouchon. Low rank approximation for the numerical simulation of high dimensional lindblad and riccati equations. *arXiv preprint arXiv:1207.4580*, 2012.

[63] C Le Bris and P Rouchon. Low-rank numerical approximations for high-dimensional lindblad equations. *Physical Review A*, 87(2):022125, 2013.

[64] M. Liao. *Lévy processes in Lie groups*, volume 162. Cambridge university press, 2004.

[65] D. Luenberger. An introduction to observers. *IEEE Trans. on Automatic Control*, 16(6):596–602, 1971.

[66] R. Mahony, T. Hamel, and J-M Pflimlin. Nonlinear complementary filters on the special orthogonal group. *IEEE-Trans. on Automatic Control*, 53(5):1203–1218, 2008.

[67] G. Meyer, S. Bonnabel, and R. Sepulchre. Linear regression under fixed-rank constraints: a riemannian approach. In *Proc. of the 28th International Conference on Machine Learning (ICML)*, 2011.

[68] G. Meyer, S. Bonnabel, and R. Sepulchre. Regression on fixed-rank positive semidefinite matrices: a riemannian approach. *Journal of Machine Learning Reasearch (JMLR)*, 12:593–625, 2011.

[69] B. Mishra, G. Meyer, S. Bonnabel, and R. Sepulchre. Fixed-rank matrix factorizations and riemannian low-rank optimization. *arXiv preprint, accepted in Computational Statistics*, 2012.

[70] Y.E. Nesterov and M.U. Todd. On the riemannian geometry defined for self-concordant barriers and interior point methods. *Found. Comput. Math.*, 2:333–361, 2002.

[71] E. Oja. *Subspace methods of pattern recognition.* Research Studies Press, 1983.

[72] E. Oja. Principal components, minor components, and linear neural networks. Neural Networks, 1992.

[73] P.J. Olver. Moving frames. *Journal of Symbolic Computation*, 36(3):501–512, 2003.

[74] E. Pardoux. Filtrage non linéaire et équations aux dérivées partielles stochastiques associées. Springer, 1991.

[75] X. Pennec, P. Fillard, and N. Ayache. A riemannian framework for tensor computing. *International Journal of Computer Vision*, 66:41–66, 2006.

[76] F. Perrin. *Étude mathématique du mouvement brownien de rotation.* PhD thesis, Université de Paris, 1928.

[77] D.T. Pham, J. Verron, and M.C. Roubaud. A singular evolutive extended Kalman filter for data assimilation in oceanography. *Journal of Marine Systems*, 16:323–340, 1998.

[78] E.J.G. Pitman. The estimation of the location and scale parameters of a continuous population of any given form. *Biometrika*, 30(3/4):391–421, 1939.

[79] B.T. Polyak and A.B. Juditsky. Acceleration of stochastic approximation by averaging. *SIAM Journal on Control and Optimization*, 30(4):838–855, 1992.

[80] B. Recht, M. Fazel, and P.A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Review*, 52(3):471–501, 2010.

[81] A. Sarlette, S. Bonnabel, and R. Sepulchre. Coordinated motion design on lie groups. *Automatic Control, IEEE Transactions on*, 55(5):1047–1058, 2010.

[82] A. Sarlette and R. Sepulchre. Consensus optimization on manifolds. *SIAM Journal on Control and Optimization*, 48(1):56–76, 2009.

[83] R. Sepulchre, P.A. Absil, and S. Bonnabel. Géomètrie des matrices semi-définies positives de rang fixé: un peu de théorie et beaucoup d'applications. In *CIFA*, 2010.

[84] S.T. Smith. Covariance, subspace, and intrinsic crame´ r-rao bounds. *Signal Processing, IEEE Transactions on*, 53(5):1610–1630, 2005.

[85] H.J. Sussmann. On the gap between deterministic and stochastic ordinary differential equations. *The Annals of Probability*, 6(1):19–41, 1978.

[86] A. Thurn, W. Burgard, and D. Fox. *Probabilistic Robotics*. MIT Press, Cambridge, Massachussets, 2006.

[87] N.G. Van Kampen. *Stochastic processes in physics and chemistry*, volume 1. North holland, 1992.

[88] J.F. Vasconcelos, R. Cunha, C. Silvestre, and P. Oliveira. Landmark based nonlinear observer for rigid body attitude and position estimation. In *Proccedings of the 2007 IEEE CDC*, pages 1033 – 1038, 2007.

[89] Y. Wei-Yong, U. Helmke, and J.B. Moore. Global analysis of oja's flow for neural networks. *IEEE Trans. on Neural Networks*, 5:674–683, 1994.

[90] E. Wong and M. Zakai. On the convergence of ordinary integrals to stochastic integrals. *The Annals of Mathematical Statistics*, 36(5):1560–1564, 1965.