



HAL
open science

Transport numérique de quantités géométriques

Guilhem Lepoutier

► **To cite this version:**

Guilhem Lepoutier. Transport numérique de quantités géométriques. Analyse numérique [math.NA]. Université Paris Sud - Paris XI, 2014. Français. NNT : 2014PA112202 . tel-01131612

HAL Id: tel-01131612

<https://theses.hal.science/tel-01131612>

Submitted on 14 Mar 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N° d'ordre:

THÈSE

Présentée pour obtenir

LE GRADE DE DOCTEUR EN SCIENCES
DE L'UNIVERSITÉ PARIS-SUD

Spécialité: Mathématiques

École doctorale de Mathématiques de la région Paris-Sud

par

Guilhem LEPOULTIER

Transport numérique de quantités géométriques

Soutenue le 25 septembre 2014 devant la Commission d'examen:

M.	Thierry COLIN	(rapporteur)
M.	Pascal FREY	(co-directeur de thèse)
M.	Jean-Frederic GERBEAU	(rapporteur)
M.	Frédéric LAGOUTIÈRE	(examineur)
M.	Bertrand MAURY	(Directeur de thèse)
M.	Jean-Marie MIREBEAU	(examineur)

Table des matières

1	Introduction	9
2	Présentation générale de la méthode	13
2.1	Introduction	13
2.1.1	Rappels sur l'équation de transport conservative linéaire	13
2.1.1.1	Modélisation	13
2.1.1.2	Approches eulérienne et lagrangienne	15
2.1.1.3	Forme et construction des solutions de l'équation de transport conservative linéaire	15
2.1.2	Présentation de la méthode étudiée	16
2.2	Advection d'un scalaire et d'un tenseur	18
2.2.1	Formulation théorique	18
2.2.2	Mise en œuvre numérique	19
2.3	Application à l'équation de transport conservative	21
2.3.1	Formulation théorique	21
2.3.2	Traduction dans le contexte de la théorie de la mesure	24
2.4	Dépendance de l'erreur en V	25
2.4.1	Erreur L1	25
2.4.2	Illustration numérique	28

TABLE DES MATIÈRES

2.4.2.1	un exemple 1d	28
2.4.2.2	Influence du conditionnement sur un cas 2d	31
2.4.3	Distance de Wasserstein	32
2.4.4	Dépendance en fonction du champ	33
2.4.4.1	Contrôle de l'erreur du flot et de sa différentielle . . .	34
2.4.4.2	Erreur entre la solution exacte et son approximation obtenue par transport par le champ approché	36
3	Comparaison de distances	39
3.1	introduction	39
3.2	Notations	40
3.3	Distance de Wasserstein	41
3.3.1	Introduction au transport optimal	42
3.3.1.1	Introduction : le problème de Monge	42
3.3.1.2	Le problème de Kantorovich	43
3.3.1.3	Principaux résultats	44
3.3.1.4	Espace de Wasserstein	45
3.3.1.5	Temps et géodésiques	46
3.3.1.6	Le cas particulier des gaussiennes	47
3.3.2	L'équation de transport est bien posée pour la distance de Wasserstein	48
3.4	Distance de Fisher	51
3.4.1	Introduction	51
3.4.2	Matrice de Fisher	52
3.4.2.1	Cas général	52

TABLE DES MATIÈRES

3.4.2.2	Expression de la métrique de Fisher pour les gaussiennes 1d pour le paramétrage classique	53
3.4.3	α -connexion	56
3.4.4	Les gaussiennes multidimensionnelles	57
3.5	Distance de Wasserstein pondérée en 1d	60
3.5.1	Généralités sur les distances construites à partir d'une fonctionnelle sur un espace métrique	61
3.5.2	Application aux gaussiennes 1d transportées	62
3.5.2.1	L'ensemble des gaussiennes 1d	62
3.5.2.2	L'ensemble des gaussiennes transportées	63
3.6	conclusion du chapitre	66
4	Courbes	69
4.1	Rappel sur les courbes et notations	69
4.2	Introduction et remarques préalables	70
4.2.1	Introduction	70
4.2.2	Points singuliers de Γ_t	71
4.3	Utilisation de la normale et reconstruction de degré 2	71
4.3.1	Remarques préliminaires	71
4.3.2	Existence et unicité du polynôme d'interpolation	73
4.3.3	Transport numérique et discrétisation adaptative	76
4.3.3.1	Apparition de points d'inflexion	76
4.3.3.2	Une technique de raffinement	78
4.4	Utilisation des splines cubiques	80
4.4.1	Notations	80

4.4.2	Principe	81
4.4.3	Dépendance en fonction du paramétrage	81
4.4.4	Cas test	82
4.4.5	Erreur géométrique	83
5	Applications	87
5.1	Calcul numérique de la distance de Fisher pour les gaussiennes 2D . .	87
5.1.1	Présentation de la méthode	88
5.1.2	Détail du champ et matrice de Fisher	90
5.1.3	Résultats numériques	92
5.1.3.1	Validation du code	92
5.1.3.2	Test sur un cas purement 2d	92
5.2	Élaboration d'un programme de mise en oeuvre de la méthode pour l'équation de transport conservative	94
5.2.1	Considérations générales sur le programme	94
5.2.1.1	Introduction	94
5.2.1.2	Caractéristique du projet et choix du langage Python	96
5.2.1.3	Première présentation du code	97
5.2.2	La classe ChampSCCOOS	99
5.2.2.1	Les fichiers océanographiques et leur traitement . . .	99
5.2.2.2	champ à données ponctuelles	100
5.2.3	La classe ChampMaillage	101
5.3	Premier cas test	101
5.3.1	Première situation : V petit	102
5.3.2	Deuxième situation : V moins petit	103

TABLE DES MATIÈRES

5.4	Application à des fichiers océanographiques	104
5.4.1	Carte océanographique	104
5.4.2	Bateau à la dérive	104
5.5	Champ de constriction	105
5.6	problème de Stokes diphasique	106
5.6.1	Présentation du problème	106
5.6.2	Résolution numérique du problème	107
5.6.2.1	Généralités	107
5.6.2.2	Intérieur/extérieur	108
5.6.2.3	Adaptation	110
5.6.2.4	Résultats numériques	110
6	Annexes 1	119
6.1	Fonctions définies par une discrétisation ponctuelle	119
6.1.1	Notations, hypothèses et énoncé du théorème	119
6.1.2	Lemmes préliminaires	121
6.1.3	Démonstration du théorème	127
6.2	Le cas 2d	128
7	Annexes 2	131
7.1	Géométrie hyperbolique : le modèle du demi-plan de Poincaré	131
7.1.1	Construction de la métrique	131
7.1.2	Géométrie de l'espace	132
7.1.3	Références	133
7.2	Éléments aux statistiques	134

TABLE DES MATIÈRES

7.2.1	Introduction	134
7.2.2	Modèle statistique, estimation	134
7.2.3	Deux méthodes de constructions d'estimateurs : la méthode des moments et l'EMV	135
7.2.4	Information de Fisher 1d	136
7.3	Géométrie différentielle et riemannienne	139
7.3.1	Généralités	139
7.3.2	Connexions affines	141
7.3.2.1	Approche intuitive	141
7.3.2.2	Formalisation	142
7.3.2.3	Courbes et connexion	143
7.3.3	Rappel sur les méthodes particulières	144

Chapitre 1

Introduction

Une part importante de l'activité en calcul scientifique et analyse numérique est consacrée aux problèmes de transport d'une quantité par un champ donné (ou lui-même calculé numériquement). Les questions de conservations étant essentielles dans ce domaine, on formule en général le problème de façon eulérienne sous la forme d'un bilan au niveau de chaque cellule élémentaire du maillage, et l'on gère l'évolution en suivant les valeurs moyennes dans ces cellules au cours du temps. Une autre approche consiste à suivre les caractéristiques du champ et à transporter les valeurs ponctuelles le long de ces caractéristiques. Cette approche est délicate à mettre en oeuvre, n'assure pas en général une parfaite conservation de la matière transportée, mais peut permettre dans certaines situations de transporter des quantités non régulières avec une grande précision, et sur des temps très longs (sans conditions restrictives sur le pas de temps comme dans le cas des méthodes eulériennes).

Les travaux de thèse présentés ici partent de l'idée suivante : dans le cadre des méthodes utilisant un suivi de caractéristiques, transporter une quantité supplémentaire géométrique apportant plus d'informations sur le problème (on peut penser à un tenseur des contraintes dans le contexte de la mécanique des fluides, une métrique sous-jacente lors de l'adaptation de maillage, etc.).

Un premier pan du travail est la formulation théorique d'une méthode de transport de telles quantités. Elle repose sur le principe suivant : utiliser la différentielle du champ de transport pour calculer la différentielle du flot, nous donnant une information sur la déformation locale du domaine nous permettant de modifier nos quantités géométriques. Cette une approche a été explorée dans [12], dans le contexte des méthodes particulières. Des travaux sur ce sujet sont en cours, encore non publiés, par

exemple dans le domaine de la physique des plasmas (voir [11]).

Ces premiers travaux amènent à travailler sur des densités paramétrées par un couple point/tenseur, comme les gaussiennes par exemple, qui sont un contexte d'applications assez naturelles de la méthode. En effet, on peut par la formulation établie transporter le point et le tenseur. On s'attend alors, si la densité est concentrée au voisinage du point, à pouvoir reconstruire après transport une bonne approximation de la solution au moyen de notre point et notre tenseur transportés.

La question qui se pose alors et qui constitue le second axe de notre travail est celle du choix d'une distance sur des espaces de densités, permettant par exemple d'étudier l'erreur commise entre la densité transportée et son approximation en fonction de la « concentration » au voisinage du point. On verra que les distances L^p montrent des limites par rapport au phénomène que nous souhaitons étudier. Cette étude repose principalement sur deux outils, les distances de Wasserstein, tirées de la théorie du transport optimal ([37]), et la distance de Fisher ([2]), au carrefour des statistiques et de la géométrie différentielle.

Dans l'esprit assez exploratoire guidant ces travaux, différentes applications numériques ont été réalisées. Un code permettant d'effectuer le transport pour différents types de champs (donné de manière exacte, sur un maillage, à partir de fichiers de données) et d'effectuer parallèlement une résolution par volumes finis a été implémenté.

Le **Chapitre 2**, après quelques rappels sur l'équation de transport conservative linéaire, présente les bases de la méthode : le transport d'un scalaire et d'un tenseur, la résolution numérique du problème, puis l'application à l'équation de transport conservative avec une densité initiale concentrée au voisinage d'un point. On trouvera aussi quelques premières études d'erreur.

Le **Chapitre 3** s'intéresse à la question du choix de distances pour l'étude de densités concentrées. Il contient une introduction théorique au transport optimal ainsi qu'à la distance de Fisher.

Le **Chapitre 4** est une autre application de la méthode développée au chapitre 2. Il s'agit lors du phénomène de transport d'interface, d'obtenir une meilleure reconstruction en déplaçant des tangentes.

Le **Chapitre 5** regroupe les applications numériques effectuées. En plus du code déjà abordé, on trouvera la résolution d'un problème de Stokes diphasique au moyen

d'un suivi d'interface, et un outil de calcul de la distance de Fisher entre gaussiennes dans le cas 2d.

On trouvera deux annexes à la fin du manuscrit : **l'annexe 1** contient une méthode de reconstruction d'un champ donné sur une grille ponctuelle nous donnant une bonne précision sur le champ ainsi que sa différentielle, utile dans le développement du code mais présentée à part pour des questions de lisibilité.

l'annexe 2 contient quelques rappels ou éléments liés à différentes parties du manuscrit : le demi-plan de Poincaré, quelques rappels de Géométrie différentielle, notamment sur les connexions, quelques points de statistiques et un rappel sur les méthodes particulières.

Chapitre 2

Présentation générale de la méthode

2.1 Introduction

2.1.1 Rappels sur l'équation de transport conservative linéaire

2.1.1.1 Modélisation

Commençons par quelques rappels de modélisation liés à l'équation de transport conservative. On considère une substance transportée par un fluide, par exemple un polluant dans l'atmosphère ou l'océan. On suppose que le transport s'effectue dans tout l'espace et pour tout temps positif afin de s'affranchir des problèmes de bord. On représente le fluide par sa densité $\rho(x, t)$, de densité initiale ρ_0 , et le mouvement du fluide par un champ de vecteur

$$\begin{aligned} F: \mathbb{R}^d \times \mathbb{R} &\longrightarrow \mathbb{R}^d \\ (x, t) &\longmapsto F(x, t) \end{aligned}$$

dont on note $\phi: (x, t) \in \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}^d$ le flot défini par

$$\begin{cases} \partial_t \phi(x, t) = F(\phi(x, t), t) \\ \phi(x, 0) = x \end{cases} .$$

Un calcul formel que nous présentons plus bas nous montre que ce problème est modélisé par l'équation de transport conservative linéaire

$$\begin{cases} \partial_t \rho + \operatorname{div}(\rho F) &= 0 \\ \rho(x, 0) &= \rho_0(x) \end{cases} .$$

C'est lors de l'étude mathématique que nous donnerons un sens à cette équation suivant les hypothèses sur F et sur ρ_0 et que nous verrons si les hypothèses sur le fluide et la densité rendent pertinentes la modélisation. Donnons les grandes lignes du calcul conduisant à l'établissement de cette équation.

Soit Ω un ouvert régulier de \mathbb{R}^d . On note $\phi(x, t)$ le flot associé à F et on pose $\Omega_t = \phi(\Omega, t)$. Par hypothèse, la quantité de substance contenue dans Ω_t ne change pas au cours du temps, donc

$$\partial_t \left(\int_{\Omega_t} \rho(x, t) \, dx \right) = 0 \quad .$$

De là

$$\begin{aligned} 0 &= \partial_t \left(\int_{\Omega_t} \rho(x, t) \, dx \right) \\ &= \partial_t \int \rho(x, t) \mathbf{1}_{\phi(\Omega, t)}(x) \, dx \\ &= \partial_t \int \rho(x, t) \mathbf{1}_{\Omega}(\phi(x, -t)) \, dx \\ &= \int \partial_t \rho(x, t) \mathbf{1}_{\Omega}(\phi(x, -t)) + \rho(x, t) \partial_t \mathbf{1}_{\Omega}(\phi(x, -t)) \, dx \\ &= \int_{\phi(\Omega, t)} \partial_t \rho(x, t) \, dx + \int \rho(x, t) F(\phi(x, -t), -t) \cdot \nabla \mathbf{1}_{\Omega}(\phi(x, -t)) \, dx \\ &\quad \text{où } \nabla \mathbf{1}_{\Omega}(\phi(x, -t)) \, dx \text{ nous donne la mesure de surface du bord de } \phi(\Omega, t) \\ &= \int_{\phi(\Omega, t)} \partial_t \rho(x, t) \, dx + \int_{\phi(\Omega, t)} \operatorname{div} \rho F \, dx \quad \text{d'après la formule de Stokes .} \end{aligned}$$

Ω est quelconque, donc le résultat est vrai pour tout ouvert régulier de \mathbb{R}^d .

Avant de passer au traitement mathématique de cette équation, nous allons effectuer quelques rappels sur les approches lagrangienne et eulérienne des phénomènes de transport.

2.1 Introduction

2.1.1.2 Approches eulérienne et lagrangienne

La forme $\rho(x, t)$ est une description **eulérienne** de la solution. À x fixé, on regarde la substance (plus précisément sa densité) qui évolue en fonction du temps. L'approche **lagrangienne** va consister à suivre la densité avec le fluide, c'est à dire considérer $\rho(\phi(x, t), t)$.

Plus généralement, pour une fonction $\varphi(x, t)$, on appelle **dérivée lagrangienne** la quantité

$$D_t \varphi(x, t) := \partial_t (t \mapsto \varphi(\phi(x, t), t)) \quad .$$

On a

$$D_t \varphi(x, t) = \partial_t \varphi(x, t) + \nabla \varphi(x, t) \cdot F(x, t) \quad .$$

Remarquons que si $\varphi(x, t) = \mu(\phi(x, -t))$ avec $\mu: \mathbb{R}^d \rightarrow \mathbb{R}$, alors

$$\varphi(\phi(x, t), t) = \mu(x) \quad .$$

2.1.1.3 Forme et construction des solutions de l'équation de transport conservative linéaire

Nous allons maintenant revoir les éléments essentiels sur l'étude de l'équation de transport conservative. Suivant [15], nous mettons dès maintenant l'hypothèse suivante sur F , pour nous affranchir de la question des temps d'existence : il existe $k: \mathbb{R} \rightarrow \mathbb{R}^+$ continue tel que pour tout t , l'application $x \mapsto F(x, t)$ soit lipschitzienne de rapport $k(t)$.

Une description eulérienne de la solution est donnée par la proposition suivante

Proposition 1

Si F est \mathcal{C}^1 et ρ_0 est une distribution d'ordre 0, alors ρ_t défini par

$$\langle \rho_t, \varphi \rangle = \langle \rho_0, \varphi(\phi(\cdot, t)) \rangle \quad \forall \varphi \in \mathcal{C}^1(\mathbb{R}^d)$$

est solution de l'équation de transport conservative linéaire. Si ρ_0 est une fonction, on a en particulier

$$\rho_t(x) = \rho_0(\phi(x, -t)) |\det d\phi(\cdot, -t)_x|$$

Démonstration

F est \mathcal{C}^1 , donc ϕ est \mathcal{C}^1 . Donc ρ_t appartient à $\mathcal{D}'_1(\mathbb{R}^d)$ car pour tout t , $\nabla \phi(\cdot, t)$ est

borné sur tout compact.

Soit t fixé. Pour φ fixé dans \mathcal{C}_0^1 , on a

$$\lim_{h \rightarrow 0} \left\langle \frac{\rho_{t+h} - \rho_t}{h}, \varphi \right\rangle = \langle \rho_0, F(\phi(\cdot, t)) \cdot \nabla \varphi(\phi(\cdot, t)) \rangle = \langle \rho_t, F \cdot \nabla \varphi \rangle \quad .$$

On a donc $\partial_t \langle \rho_t, \varphi \rangle = \langle \rho_t, F \cdot \nabla \varphi \rangle$. On a alors

$$\begin{aligned} \partial_t \langle \rho_t, \varphi \rangle &= \langle \rho_t, F \cdot \nabla \varphi \rangle \\ &= \langle \rho_t, \sum F_i \partial_i \varphi \rangle \\ &= - \sum \langle \partial_i (F_i \rho_t), \varphi \rangle \\ &= - \langle \operatorname{div}(\rho_t F), \varphi \rangle \quad . \end{aligned}$$

$\operatorname{div}(\rho_t F) \in \mathcal{D}'_1$ donc on en déduit que $t \mapsto \rho_t$ est dérivable et que $\langle \partial_t \rho_t, \varphi \rangle = \partial_t \langle \rho_t, \varphi \rangle$ et de là $\partial_t \rho_t + \operatorname{div}(\rho_t F) = 0$.

Une description lagrangienne nous conduit à la proposition classique suivante

Proposition 2

Soit $\rho(x, t)$ la solution de l'équation de transport conservative. On a

$$\rho(\phi(x, t), t) = \rho_0(x) \exp \left(- \int_0^t \operatorname{div}(F(\phi(x, s), s)) \, ds \right)$$

2.1.2 Présentation de la méthode étudiée

On trouve essentiellement deux grandes classes de méthodes de résolutions numériques pour résoudre les équations de transport conservatives. La première englobe les méthode eulériennes, de manière naturelle les volumes finis ([25]) compte-tenu du caractère conservatif de l'équation, mais aussi les éléments finis. La seconde les méthodes de suivis de particules, qu'on peut schématiquement sous-diviser en deux sous-classes, suivant que ces dernières soient choisies de manière déterministe (méthodes particulières, [32], voir section 7.3.3, page 144) ou dans un cadre probabiliste (méthodes de Monte-Carlo, [20]).

Nous plaçant dans le domaine des méthodes de suivis de particules, nous allons formaliser et étudier une méthode visant à utiliser la différentielle du champ F pour

2.1 Introduction

transporter une information supplémentaire associée à chaque particule, permettant une meilleure reconstruction.

On considèrera que notre densité de départ ρ_0 est concentrée au voisinage d'un point. Plus précisément, on aura

$$\rho_0(x) = \frac{1}{\sqrt{|\det V|}} \eta \left({}^t(x - \mu) V^{-1} (x - \mu) \right)$$

où $\mu \in \mathbb{R}^d$ et V matrice symétrique définie positive. η est une fonction positive bornée intégrable sur \mathbb{R} telle que

$$\int_{\mathbb{R}^d} \eta(\|x\|^2) < +\infty \quad .$$

$\eta(x) = \exp(-\frac{x}{2})$ nous donnera par exemple la famille des gaussiennes.

Nous allons transporter μ et V , obtenant deux paramètres μ_t et V_t nous permettant de reconstruire une densité

$$\bar{\rho}_t(x) = \frac{1}{\sqrt{|\det V_t|}} \eta \left({}^t(x - \mu_t) V_t^{-1} (x - \mu_t) \right)$$

appartenant au même espace que la densité initiale. En toute généralité, la solution exacte (Proposition 1, page 15)

$$\rho_t(x) = \rho_0(\phi(x, -t)) |\det d\phi(\cdot, -t)_x|$$

n'appartiendra plus à cette famille de densités à deux paramètres. Néanmoins, on peut espérer si la densité est concentrée au voisinage de μ qu'elle reste proche de cet espace en un sens que nous quantifierons et donc obtenir une approximation convenable de cette solution.

Dans un premier temps, nous présenterons le formalisme théorique pour l'advection d'un scalaire et d'un tenseur, le moyen de résoudre numériquement ce problème et quelques illustrations.

Nous nous intéresserons ensuite à l'application à l'équation de transport conservative ainsi qu'au push-forward en théorie de la mesure.

Nous étudierons ensuite la qualité de l'approximation en fonction de V . Dans un premier temps, en distance L^1 . Nous regarderons ensuite brièvement en distance de Wasserstein, avant de nous intéresser plus longuement à cette distance au chapitre

suisant. On y trouvera (section 3.3, page 41) une introduction à cette théorie pour le lecteur qui ne serait pas familier avec ce sujet.

A partir de maintenant nous faisons les hypothèses suivantes sur F :

- F sera \mathcal{C}^2 pour garantir que le flot soit \mathcal{C}^2 ;
- on supposera qu'il existe $k: \mathbb{R} \rightarrow \mathbb{R}^+$ continue tel que pour tout t , l'application $x \mapsto F(x, t)$ soit lipschitzienne de rapport $k(t)$. Toujours pour nous affranchir des problèmes de temps d'existence.

2.2 Advection d'un scalaire et d'un tenseur

2.2.1 Formulation théorique

Définition 1

Soit Ψ un champ de vecteur. On définit son transporté au temps t par

$$\Psi_t(x) = d\phi(\cdot, -t)_x \Psi(\phi(x, -t))$$

Justifions cette définition. On se fixe un temps t . On se donne un arc $\gamma(s)$ qu'on suppose \mathcal{C}^1 . On définit la transportée de l'arc au temps t par

$$\gamma_t(s) = \phi(\gamma(s), t) \quad .$$

Le vecteur tangent en s à l'arc est donné par

$$\gamma'_t(s) = d\phi(\cdot, t)_{\gamma(s)} \gamma'(s) \quad .$$

Si l'on souhaite que le transport d'un vecteur conserve le caractère tangent, alors on en déduit la définition, traduite dans une approche eulérienne.

Définition 2

Soit Ψ un champ de formes bilinéaires symétriques définies positives sur \mathbb{R}^d . On définit son transporté au temps t par

$$\Psi_t(x) = {}^t d\phi(\cdot, -t)_x \Psi(\phi(x, -t)) d\phi(\cdot, -t)_x \quad .$$

Remarque 3

La définition découle naturellement de celle du transport d'un vecteur : on fait

2.2 Advection d'un scalaire et d'un tenseur

remonter les vecteurs auxquels on applique le tenseur ainsi que le « pied » de ces derniers par le flot.

Donnons une formulation lagrangienne de ce transport.

Proposition 4

On a

$$\Psi_t(\phi(x, t)) = {}^t d\phi(\cdot, t)_x^{-1} \Psi(x) d\phi(\cdot, t)_x^{-1} \quad .$$

Démonstration

On a

$$\Psi_t(\phi(x, t)) = {}^t d\phi(\cdot, -t)_{\phi(x, t)} \Psi(x) d\phi(\cdot, -t)_{\phi(x, t)} \quad .$$

Comme

$$\phi(\phi(x, t), -t) = x \quad ,$$

en différentiant, on obtient

$$d\phi(\cdot, -t)_{\phi(x, t)} \circ d\phi(\cdot, t)_x = \text{Id}$$

et donc

$$d\phi(\cdot, t)_x^{-1} = d\phi(\cdot, -t)_{\phi(x, t)} \quad .$$

Nous donnerons quelques illustration graphiques du phénomène dans la sous-section suivante.

2.2.2 Mise en œuvre numérique

Nous avons une formulation pour advecter des tenseurs, qui repose sur le calcul de la différentielle du flot. Dans une perspective numérique, il nous faut une manière de calculer cette dernière. La proposition suivante nous la donne, en nous ramenant à une EDO :

Proposition 5

On a

$$\partial_t d\phi(\cdot, t)_x = dF(\cdot, t)_{\phi(x, t)} \circ d\phi(\cdot, t)_x \quad .$$

Démonstration

Le flot étant \mathcal{C}^2 , on peut employer le théorème de Schwarz

$$\begin{aligned} \partial_t d\phi(\cdot, t)_x &= d(\partial_t \phi(\cdot, t))_x \\ &= d(F(\phi(\cdot, t)))_x \\ &= dF_{\phi(x,t)} \circ d\phi(\cdot, t)_x \quad . \end{aligned}$$

Le point crucial pour la mise en oeuvre de la méthode est la connaissance de la différentielle du champ d'advection. Il y a deux situations possibles :

- le champ nous est donné de manière analytique et nous pouvons calculer explicitement sa dérivée ;
- le champ est donné de manière discrète. Dans la suite, il s'agirait essentiellement de champs sur maillages et de champs donnés par des valeurs ponctuelles. Il s'agira alors :
 - soit d'intégrer dans le solveur nous fournissant le champ une méthode d'approximation de sa différentielle ,
 - soit de se donner une méthode d'approximation de la différentielle ainsi qu'une évaluation de sa précision à partir des données que nous avons. C'est par exemple ce que nous mettrons en oeuvre dans le cas des fichiers de données océanographiques (champ donné par des valeurs ponctuelle) en nous appuyant sur le Théorème 59 (page 129).

On trouvera sur les figures 2.1 à 2.4 quelques illustrations du phénomène dans \mathbb{R}^2 . Nous prenons comme champ de tenseurs initial le champ valant la matrice identité en tout point. Nous l'advectionnons de manière explicite en intégrant l'EDO. Nous représentons la solution sur $[-4, 4] \times [-4, 4]$ en prenant les points de coordonnées entiers et en y dessinant les boules unités. Nous représentons le champ exact, ainsi que celui calculé numériquement par un schéma d'Euler.

Remarque 6

Le schéma est exact dans le cas du champ $(\frac{1}{10}y^2, 0)$ et du cisaillement.

2.3 Application à l'équation de transport conservative

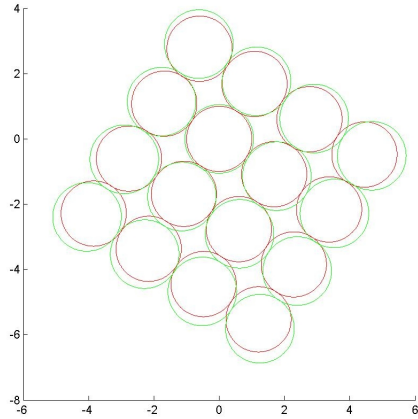


FIGURE 2.1 $F = (-y, x), T = 1$

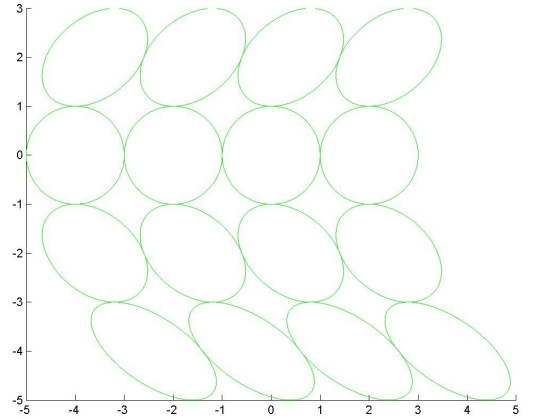


FIGURE 2.2 $F = (\frac{1}{10}y^2, 0), T = 1$

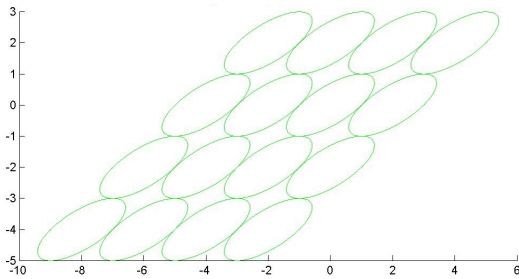


FIGURE 2.3 $F = (y, 0), T = 1$

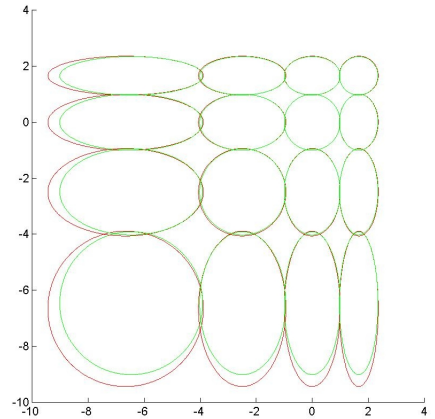


FIGURE 2.4 $F = -(x^2, y^2), T = 1$

2.3 Application à l'équation de transport conservative

2.3.1 Formulation théorique

On considère l'équation de transport conservative suivante, définie sur \mathbb{R}^d :

$$\begin{cases} \partial_t \rho + \operatorname{div}(\rho F) = 0 \\ \rho(x, 0) = \rho_0(x) \end{cases} \quad (2.1)$$

Avec les hypothèses définies précédemment sur F qu'on rappelle ici :

$$\begin{aligned} F &\in \mathcal{C}^2(\mathbb{R}^d \times \mathbb{R}, \mathbb{R}^d), \\ \exists k &\in \mathcal{C}^0(\mathbb{R}, \mathbb{R}^+), \text{ tq } x \mapsto F(x, t) \text{ lipschitzienne de rapport } k(t). \end{aligned} \quad (2.2)$$

La densité initiale est de la forme

$$\rho_0(x) = \frac{1}{\sqrt{|\det V|}} \eta \left({}^t(x - \mu) V^{-1} (x - \mu) \right) \quad . \quad (2.3)$$

avec η fonction positive bornée intégrable sur \mathbb{R} telle que $\int_{\mathbb{R}^d} \eta(\|x\|^2) < +\infty$, multipliée par une constante C_d rendant ρ_0 normalisé. μ appartient à \mathbb{R}^d et V est une matrice symétrique définie positive.

On rappelle (Proposition 1, page 15) que la solution exacte est donnée par

$$\begin{aligned} \rho(x, t) &= \rho_0(\phi(x, -t)) |\det d\phi(\cdot, -t)_x| \\ &= \frac{1}{\sqrt{|\det V|}} \eta \left({}^t(\phi(x, -t) - \mu) V^{-1} (\phi(x, -t) - \mu) \det d\phi(\cdot, -t)_x \right) \quad . \end{aligned} \quad (2.4)$$

Cette densité n'appartient plus, sans hypothèses supplémentaires, à l'espace à deux paramètres de la densité initiale. Nous allons définir un élément de cette espace correspondant naturellement à la solution exacte en se basant sur le transport que nous avons défini précédemment.

Dans un premier temps, suivant la Définition 2 (page 18), on définit les transportés de μ et V .

Définition 3

Les transportés au temps t de μ et V^{-1} sont définis par :

$$\begin{aligned} \mu_t &= \phi(\mu, t) \quad , \\ V_t^{-1} &= {}^t d\phi(\cdot, t)_\mu^{-1} \circ V^{-1} \circ d\phi(\cdot, t)_\mu^{-1} \quad . \end{aligned}$$

On définit maintenant l'approximation de la solution suivante, basée sur l'advection du point μ autour duquel la densité est concentrée ainsi que de l'advection du tenseur V :

2.3 Application à l'équation de transport conservative

Définition 4

On pose

$$\bar{\rho}(x, t) = \frac{1}{\sqrt{|\det V_t|}} \eta \left({}^t(x - \mu_t) V_t^{-1} (x - \mu_t) \right)$$

Nous allons voir à la section suivante que si $\|V\|$ tend vers zéro, la distance entre l'approximation et la solution exacte tend vers zéro en un sens que nous préciserons.

Exemple 7 (Champ radial 1d)

On considère une gaussienne $\rho_\varepsilon = \frac{1}{\sqrt{2\pi\varepsilon}} \exp\left(-\frac{(x - \mu)^2}{2\varepsilon^2}\right)$. On pose

$$F(x) = x$$

et on a donc

$$\phi(x, t) = e^t x$$

et

$$d\phi(\cdot, t)_x = \partial_1 \phi(x, t) = e^t$$

Dans ce cas, la solution exacte est donnée par

$$\begin{aligned} \rho_\varepsilon^t(x) &= \rho_\varepsilon(\phi(x, -t)) |\det [d\phi(\cdot, -t)_x]| \\ &= \frac{1}{\sqrt{2\pi\varepsilon}} \exp\left(-\frac{(e^{-t}x - \mu)^2}{2\varepsilon^2}\right) e^{-t} \\ &= \frac{1}{\sqrt{2\pi\varepsilon e^t}} \exp\left(-\frac{(x - e^t\mu)^2}{2(\varepsilon e^t)^2}\right) , \end{aligned}$$

l'espérance et la variance transportée sont données par :

- $\mu_t = e^t \mu$;
- $\varepsilon_t^2 = (e^t)^2 \varepsilon^2$.

L'approximation est donc

$$\bar{\rho}(x, t) = \frac{1}{\sqrt{2\pi \varepsilon e^t}} \exp\left(-\frac{(x - e^t\mu)^2}{2(\varepsilon e^t)^2}\right) = \rho_\varepsilon^t(x) .$$

Elle coïncide avec la solution exacte.

2.3.2 Traduction dans le contexte de la théorie de la mesure

Les formules pour la densité et son approximation restent valables si nous remplaçons le flot par un difféomorphisme ϕ ne dépendant pas du temps. On a ainsi une « solution exacte » donnée par

$$\rho(x) = \frac{1}{|\sqrt{\det V}|} \rho_0(\phi^{-1}(x)) |\det d\phi_x^{-1}|$$

et une « solution approchée »

$$\bar{\rho}(x) = \frac{1}{\sqrt{|\det V_\phi|}} \eta\left({}^t(x - \mu_\phi) V_\phi^{-1}(x - \mu_\phi)\right)$$

où l'on a $\mu_\phi = \phi(\mu)$ et $V_\phi^{-1} = {}^t d\phi(\mu)^{-1} \circ V^{-1} \circ d\phi(\mu)^{-1}$

Cela correspond au cadre formel du « push-forward » en théorie de la mesure. On se donne une mesure $d\rho_0$ absolument continue par rapport à la mesure de Lebesgue, de densité

$$\rho_0(x) = \frac{1}{\sqrt{|\det V|}} \eta\left({}^t(x - \mu) V^{-1}(x - \mu)\right) .$$

On définit la mesure $d\rho$ comme la mesure image de $d\rho_0$ par ϕ , c'est à dire

$$d\rho(f) = d\rho_0(f \circ \phi) .$$

C'est une mesure absolument continue par rapport à la mesure de Lebesgue, de densité

$$\rho(x) = \frac{1}{|\det \sqrt{V}|} \rho_0(\phi^{-1}(x)) |\det d\phi_x^{-1}| .$$

On peut ainsi définir une approximation de la mesure image en posant

$$\bar{\rho}(x) = \frac{1}{\sqrt{|\det V_\phi|}} \eta\left({}^t(x - \mu_\phi) V_\phi^{-1}(x - \mu_\phi)\right) .$$

Ce cadre nous servira dans certains exemples 2d. Il nous fournira aussi un cadre assez naturel pour utiliser la distance de Wasserstein

2.4 Dépendance de l'erreur en V

Plaçons nous un instant dans le cas du transport pour les gaussiennes 1d. Soit

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) .$$

Soit F un champ de vecteur de \mathbb{R} de flot $\phi(x, t)$. On considère la solution de l'équation de transport conservative 2.1 avec g comme donnée initiale

$$g_t(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\phi(x, -t) - \mu)^2}{2\sigma^2}\right) |\partial_x \phi(x, -t)| .$$

g_t n'est pas une gaussienne, mais si σ est petit, on s'attend à ce qu'elle en soit proche. Et donc que la distance avec l'approximation

$$\bar{g}_t(x) = \frac{1}{\sqrt{2\pi}\partial_x \phi(\mu, \sigma)} \exp\left(-\frac{(x - \phi(\mu, t))^2}{2\partial_x \phi(\mu, \sigma)^2}\right)$$

Soit petite.

Plus généralement, on s'attend à ce que l'écart entre l'approximation définie en Définition 4 (page 22) et la solution exacte de (2.1) décroisse avec V . Nous allons étudier ce phénomène dans un premier temps pour la distance L^1 , puis ensuite pour la distance de Wasserstein, dont on trouvera une introduction dans la section 3.3 (page 41).

2.4.1 Erreur L^1

On munit \mathbb{R}^d d'une norme quelconque et l'ensemble des endomorphismes de \mathbb{R}^d de la norme subordonnée. \sqrt{V} désignera l'unique matrice symétrique définie positive telle que $\sqrt{V}\sqrt{V} = V$.

La notation $C_{a,b,c}$ désignera une constante dépendant de a, b, c . On se permettra, pour simplifier les calculs, de garder la même notation de ligne en ligne pour une constante qui peut changer d'un facteur.

Pour alléger les notations dans les calculs, on se place ici dans le cas où le transport est défini directement par un difféomorphisme (section 2.3.2, page 24). Soit ϕ

un difféomorphisme. On pose $\psi = \phi^{-1}$. On considère

$$\rho(x) = \frac{1}{\sqrt{\det V}} \eta \left({}^t(\psi(x) - \mu) V^{-1} (\psi(x) - \mu) \right) \det d\psi_x$$

et

$$\bar{\rho}(x) = \frac{1}{\sqrt{\det V_t}} \eta \left({}^t(x - \mu_t) V_t^{-1} (x - \mu_t) \right)$$

où, suivant la Définition 3, on pose $\mu_t = \phi(\mu)$ et $V_t = {}^t d\phi_\mu^{-1} \circ V \circ d\phi_\mu^{-1}$. Nous allons établir un estimateur d'erreur L^1 entre ρ et son approximation $\bar{\rho}$.

Quitte à considérer $d\phi_\mu^{-1}(\phi(x) - \phi(\mu))$, on peut supposer que $\mu = 0$, $\phi(\mu) = 0$ et $d\phi_\mu = \text{Id}$.

Théorème 8

Si η est \mathcal{C}^2 à support compact, $\|\sqrt{V}\| \leq 1$, alors on a

$$\|\rho - \bar{\rho}\|_{L^1} \leq C \left(\text{Cond}(\sqrt{V}) \|\sqrt{V}\| + \text{Cond}(\sqrt{V})^2 \|\sqrt{V}\|^2 \right)$$

ou $\text{Cond}(\sqrt{V})$ désigne le conditionnement de \sqrt{V} .

Nous allons donner un corollaire immédiat.

Corollaire 9

Soit $(V_n)_n$ une suite de matrice symétrique définie positive tendant vers zéro dont le conditionnement est borné. Alors on a

$$\|\rho_n - \bar{\rho}_n\|_{L^1} \leq C \|\sqrt{V_n}\| \quad .$$

Donnons maintenant démonstration de la preuve.

Démonstration

On pose

$$e_1 = \int_{\mathbb{R}^d} |\rho(x) - \bar{\rho}(x)| dx \quad .$$

On a

$$e_1 = \int_{\mathbb{R}^d} \left| \frac{1}{\sqrt{\det V}} \eta \left({}^t\psi(x) V^{-1} \psi(x) \right) \det(d\psi_x) - \frac{1}{\sqrt{\det V}} \eta \left({}^t x V^{-1} x \right) \right| dx \quad .$$

On effectue le changement de variable $y = \sqrt{V^{-1}}x$ et on obtient

$$e_1 = \int_{\mathbb{R}^d} \left| \eta \left({}^t\psi(\sqrt{V}y) V^{-1} \psi(\sqrt{V}y) \right) \det(d\psi_{\sqrt{V}y}) - \eta(\|y\|^2) \right| dy \quad .$$

2.4 Dépendance de l'erreur en V

η étant à support compact et $\|\sqrt{V}\| \leq 1$, il existe une boule ouverte B , centrée en 0, de rayon r indépendant de V et dépendant de ψ , contenant le support de l'intégrande. On a de plus, si y appartient à B , par formule de Taylor

$$\psi(\sqrt{V}y) = \sqrt{V}y + R(\sqrt{V}y) \quad \text{où} \quad \|R(\sqrt{V}y)\| \leq \sup_B \|d^2\psi\| \|\sqrt{V}\|^2 \|y\|^2 \quad .$$

Donc on en déduit

$$e_1 = \int_B \left| \eta \left(\|y\|^2 + 2 {}^t R \sqrt{V}^{-1} y + {}^t R V^{-1} R \right) \det(d\psi_{\sqrt{V}y}) - \eta(\|y\|^2) \right| dy \quad .$$

On a

$$\det(d\psi_{\sqrt{V}y}) = 1 + O(\|\sqrt{V}\| \|y\|) \quad ,$$

donc

$$\begin{aligned} e_1 &\leq \int_B \left| \eta \left(\|y\|^2 + 2 {}^t R \sqrt{V}^{-1} y + {}^t R V^{-1} R \right) - \eta(\|y\|^2) \right| dy \\ &\quad + \sup_B \|d^2\psi\| \|\sqrt{V}\| \int_B \|y\| \eta \left({}^t \psi(\sqrt{V}y) V^{-1} \psi(\sqrt{V}y) \right) dy \quad . \end{aligned}$$

On note I_1 la première intégrale. Par formule de Taylor, on a

$$\begin{aligned} I_1 &\leq \sup |\eta'| \int_B \left| 2 {}^t R \sqrt{V}^{-1} y + {}^t R V^{-1} R \right| dy \\ &\leq C_{\eta,\psi} \int_B \left(\|\sqrt{V}\|^2 \|\sqrt{V}^{-1}\| \|y\|^3 + \|\sqrt{V}\|^4 \|\sqrt{V}^{-1}\|^2 \|y\|^4 \right) dy \\ &\leq C_{\eta,\psi} \left(\text{Cond}(\sqrt{V}) \|\sqrt{V}\| + \text{Cond}(\sqrt{V})^2 \|\sqrt{V}\|^2 \right) \quad . \end{aligned}$$

Comme la deuxième intégrale est majorée par $\sup \eta \int_B \|y\| dy$, on en déduit que

$$e_1 \leq C_{\eta,\psi} \left(\text{Cond}(\sqrt{V}) \|\sqrt{V}\| + \text{Cond}(\sqrt{V})^2 \|\sqrt{V}\|^2 \right) \quad .$$

Remarque 10

Pour des raisons de simplicité de l'exposé nous nous sommes limités au cas où la densité est à support compact. Le résultat reste valable pour des fonctions η tel que $\int |x|^k \eta(x) dx < \infty$. Il faut alors utiliser le théorème de Taylor avec reste intégral sur η . Et contrôler l'intégrabilité du reste au moyen de la propriété de décroissance rapide.

Remarque 11

\sqrt{V} étant symétrique définie positive, son conditionnement est égal au quotient

de la plus grande valeur propre par la plus petite.

Remarque 12

Hormis en dimension 1, nous n'avons pas la convergence pour les distances L^p , $p > 1$. Cela est naturel si on considère que la norme L^p des gaussiennes est de l'ordre de $\frac{1}{|\sqrt{V}|^{p-1}}$.

Donnons maintenant une traduction dans le cadre des équations de transport

Corollaire 13

Si $\rho(x, t)$ est une solution de (2.1) et $\bar{\rho}(x, t)$ son approximation suivant la Définition 4, alors

$$\|\rho(\cdot, t) - \bar{\rho}(\cdot, t)\|_{\mathbb{L}^1} \leq C_{\mu, t, F, \eta} \left(\text{Cond}(\sqrt{V}) \|\sqrt{V}\| + \text{Cond}(\sqrt{V})^2 \|\sqrt{V}\|^2 \right) .$$

2.4.2 Illustration numérique

2.4.2.1 un exemple 1d

On considère le champ $F(x) = \frac{1}{1+x^2}$. On considère l'équation différentielle

$$\begin{cases} x'(t) & = F(x(t)) \\ x(0) & = x_0 \end{cases}$$

On a

$$x'(t) = \frac{1}{x(t)^2 + 1} \quad \text{donc} \quad x'(t)x(t)^2 + x'(t) = 1 \quad .$$

En primitivant cette dernière expression, on obtient que

$$\frac{1}{3}x(t)^3 + x(t) = t + c$$

et donc

$$x(t)^3 + 3x(t) - (t + c) = 0 \quad .$$

On est donc ramené à calculer les racines d'un polynôme de degré 3. Les formules de Cardan, avec un peu de travail pour déterminer la constante pour retrouver la condition initiale, implique alors que

$$x(t) = \sqrt[3]{\frac{3t + x_0^3 + 3x_0 + \sqrt{(3t + x_0^2 + 3x_0)^2 + 4}}{2}} + \sqrt[3]{\frac{3t + x_0^3 + 3x_0 - \sqrt{(3t + x_0^2 + 3x_0)^2 + 4}}{2}} \quad .$$

2.4 Dépendance de l'erreur en V

Pour le transport de mesure, c'est le comportement de la solution en espace qui nous intéressera. On calcule à la main la différentielle du flot. Etant donné la lourdeur des calculs, nous ne les détaillons pas ici. On représente, sur la figure 2.5 pour un certain nombre de temps fixé la solution en espace, et pour les même temps, sur la figure 2.6, la différentielle du flot en espace

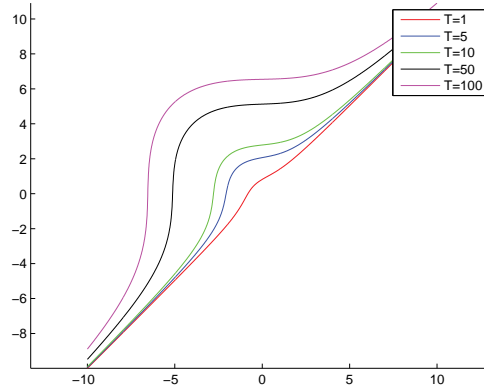


FIGURE 2.5 la solution en espace à temps fixé

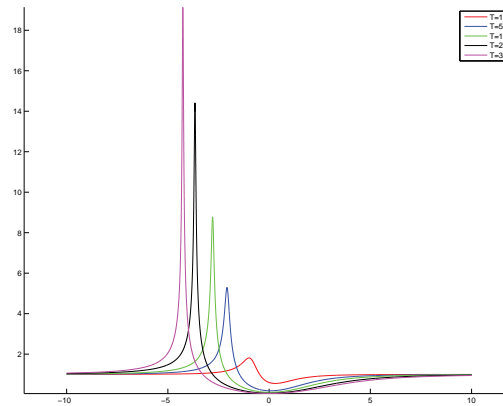


FIGURE 2.6 la différentielle du flot en espace à temps fixé

Remarque 14

On voit donc bien que la différentielle du flot nous garantit que la gaussienne transportée n'est pas une gaussienne, puisque le déterminant, à temps fixé, est non constant. Cependant, la perturbation reste relativement localisé.

On pose $\eta(x) = \exp(x) \times \zeta(x)$ où ζ est une fonction plateau \mathcal{C}^∞ valant 1 sur $[-10, 10]$ et 0 sur $[-11, 11]^C$. η est alors à support compact. On considère alors

$$\rho_\varepsilon = \frac{1}{\sqrt{2\pi\varepsilon}} \eta \left(-\frac{(x - \mu)^2}{2\varepsilon^2} \right) .$$

On note ρ_ε^t sa transportée au temps t . On a

$$\begin{aligned}\rho_\varepsilon^t(x) &= \rho_\varepsilon(\phi(x, -t)) |\partial_i \phi(x, -t)| \\ &= \rho_\varepsilon(\phi(x, -t)) \left| \frac{1}{\partial_i \phi(x, t)} \right| .\end{aligned}$$

L'approximation $\bar{\rho}_\varepsilon^t$ (Définition 4, page 22) est alors égale à :

$$\bar{\rho}_\varepsilon^t(x) = \frac{1}{\sqrt{2\pi\varepsilon} |\partial_i \phi(\mu, t)|} \eta \left(-\frac{(x - \phi(\mu, t))^2}{2(\varepsilon \partial_i \phi(\mu, t))^2} \right) .$$

On représente sur les figures 2.7 à 2.8 quelques résultats à $T = 5$ et $\mu = 1$ fixés. On

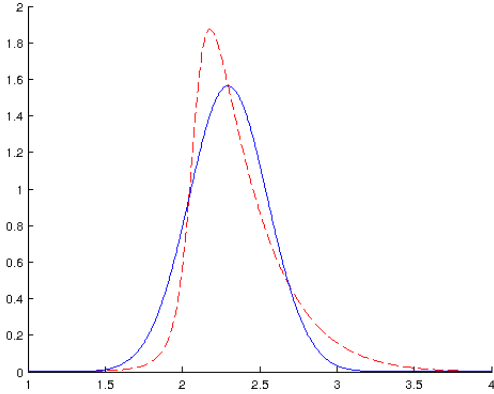


FIGURE 2.7 $\varepsilon = 0.8$, pointillés : solution exacte. traits pleins : solution approchée

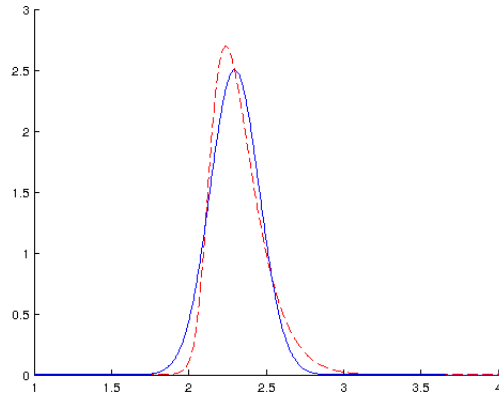


FIGURE 2.8 $\varepsilon = 0.5$

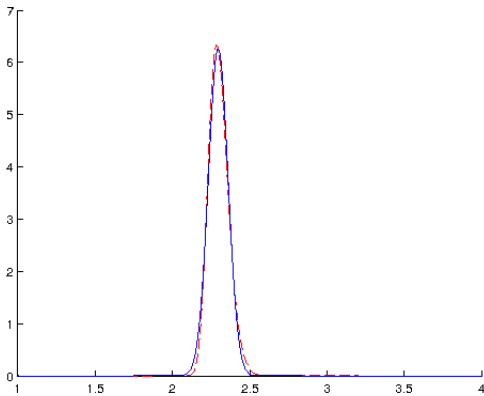


FIGURE 2.9 $\varepsilon = 0.2$

met en évidence sur la figure 2.10 la convergence d'ordre 1 pour la norme L^1

2.4 Dépendance de l'erreur en V

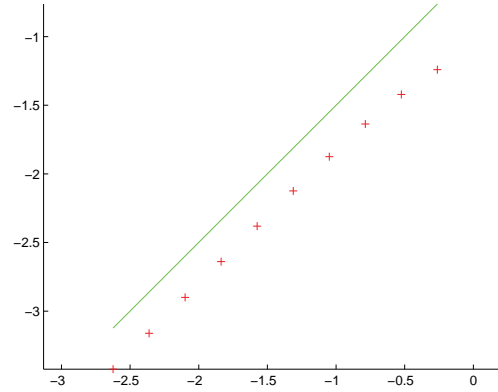


FIGURE 2.10 $T = 5, C = 1, \varepsilon$ variant de 1.3^{-1} à 1.3^{-10} . log de l'erreur en fonction de log d'epsilon et droite de référence de pente 1

2.4.2.2 Influence du conditionnement sur un cas 2d

On a vu dans le Théorème 8 (page 26) que l'erreur d'approximation dépend du conditionnement de \sqrt{V} . Nous allons montrer numériquement que l'explosion de ce conditionnement empêche la convergence L^1 .

On se donne ici directement un difféomorphisme, donné par

$$\phi(x, y) = (x(1 + y^2), y) \quad .$$

On prend $\mu = (0.3, 0.3)$, on se donne R matrice de rotation d'angle $\frac{\pi}{3}$ et on pose

$$V = {}^tR \circ \begin{pmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{pmatrix} \circ R \quad .$$

On prend $\eta(x) = \exp(x)\zeta(x)$.

Dans un premier temps, on prend $\varepsilon_2 = 2.5^{-k}$, $\varepsilon_1 = 2.5^{-k}$ avec k variant de 1 à 10. Le conditionnement est donc constant. On représente sur la 2.11 le log de l'erreur L^1 entre la densité de la mesure-image par ϕ et son approximation en fonction du log de la norme de \sqrt{V} et on retrouve l'ordre 1. Dans un deuxième temps on prend $\varepsilon_2 = 2.5^{-k}$, $\varepsilon_1 = \varepsilon_2^2$ avec k variant de 1 à 10. Le conditionnement est alors de l'ordre de 2.5^k , $|\sqrt{V}|$ de l'ordre de 2.5^{-k} . On représente sur la 2.12 le log de l'erreur L^1 entre la densité de la mesure-image par ϕ et son approximation en fonction du log de la norme de \sqrt{V} et on met en évidence qu'il n'y a pas convergence.

Intuitivement, nos densités concentrées sont proches de masses de Dirac, puis-

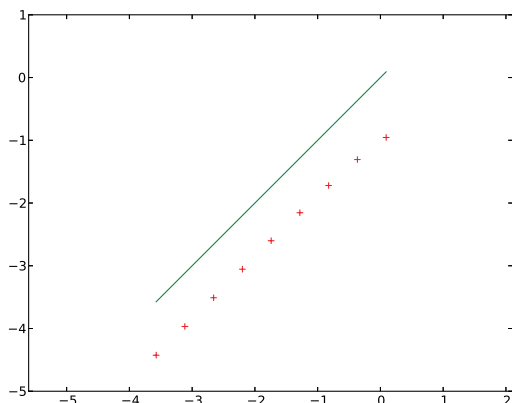


FIGURE 2.11
conditionnement constant

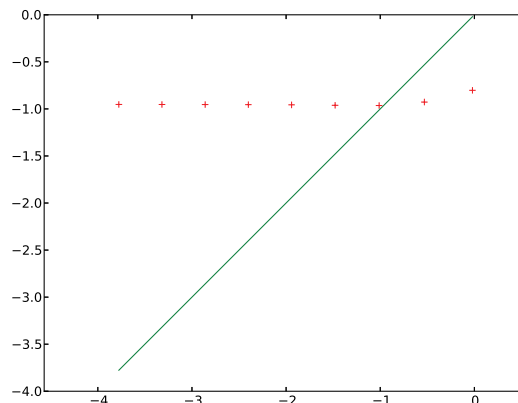


FIGURE 2.12 explosion du condi-
tionnement

qu'on peut les regarder comme des approximations de l'unité. L'espace L^1 ne comportant par les masses de Dirac, nos estimations ont lieu à proximité de sa « frontière », rendant les estimations sujettes à de fortes variations. C'est en particulier pour cette raison que le conditionnement intervient. Nous allons maintenant étudier l'erreur d'estimation au moyen de la distance de Wasserstein, définie pour les masses de Dirac et bien adapté à la théorie du transport.

2.4.3 Distance de Wasserstein

On rappelle qu'on trouvera un introduction détaillée à la théorie du transport optimal et aux distances de Wasserstein dans le chapitre 3, à la section 3.3. On se reportera à [37] pour une introduction complète.

Le théorème suivant nous fournit une estimation d'erreur entre l'approximation définie dans la Définition 4 et la solution de l'équation de transport conservative (2.1)

Théorème 15

Si ρ_0 a des moment d'ordre $2p$ finis, on a alors

$$W_p(\rho, \bar{\rho}) \leq K \|\sqrt{V}\|^2$$

Démonstration

On pose :

- $T_1(x) = \phi(x, t)$;

2.4 Dépendance de l'erreur en V

- $T_2(x) = d\phi(\cdot, t)_\mu(x - \mu) + \mu_t$.

T_1 transporte ρ_0 sur ρ et T_2 transporte ρ_0 sur $\bar{\rho}$. Ainsi, $T = T_1 \circ T_2^{-1}$ transporte $\bar{\rho}$ sur ρ . Majorons le coût total de ce transport pour une fonction coût $\|x - y\|^p$. On posera $L = d\phi(\cdot, t)_\mu(x - \mu)$.

$$\begin{aligned} C_T &= \int \|x - T(x)\|^p \bar{\rho}(x) dx \\ &= \int \|x - \phi(L^{-1}(x - \mu_t) + \mu, t)\|^p \rho(L^{-1}(x - \mu_t) + \mu) |\det L^{-1}| dx \\ &= \int \|\dots\|^p \frac{1}{|\det \sqrt{V}|} \rho(\sqrt{V}^{-1}(x - \mu_t) + \mu) |\det L^{-1}| dx \quad . \end{aligned}$$

On pose $y = \sqrt{V}^{-1}(L^{-1}(x - \mu_t)) + \mu$ et on obtient alors par changement de variable

$$C_T = \int \|L(\sqrt{V}(y - \mu)) + \mu_t - \phi(\sqrt{V}(y - \mu) + \mu, t)\|^p \rho(y) dy \quad .$$

Par formule de Taylor, on a alors

$$C_T \leq K \int \|\sqrt{V}\|_\infty^{2p} \|y - \mu\|^{2p} \rho(y) dy \quad .$$

Comme notre densité est au moins de moment d'ordre $2p$ fini, on retrouve la convergence d'ordre 2 en $\|\sqrt{V}\|$.

2.4.4 Dépendance en fonction du champ

Dans le contexte d'applications, nous n'aurons parfois qu'une connaissance approchée du champ, par exemple un champ P1 sur maillage. Nous allons ici établir un premier résultat nous permettant de contrôler l'erreur calculée avec la distance W_p en fonction de l'erreur sur le champ.

Soit $F^1(x, t)$ et $F^2(x, t)$ deux champs de vecteurs vérifiant les hypothèses (2.2) (page 22), avec de plus :

- $dF^i(\cdot, t)$ globalement lipschitzienne ;
- $\|F^1(\cdot, t) - F^2(\cdot, t)\|_\infty \leq \varepsilon_0$;
- $\|dF^1(\cdot, t)_x - dF^2(\cdot, t)_x\|_\infty \leq \varepsilon_1 \quad \forall x$.

Dans la suite, on posera $K(t)$ le max des quatres constantes de lipschitz au temps t , et on posera $C(t) = \int_0^t K(s) ds$. On commence par démontrer un lemme qui nous

servira par la suite.

2.4.4.1 Contrôle de l'erreur du flot et de sa différentielle

Lemme 16

il existe trois constantes \overline{C}_0, C_0, C_1 telles que

$$\begin{aligned} \|\phi_1(x, T) - \phi_2(x, T)\| &\leq \overline{C}_0(T) \varepsilon_0 \quad , \\ \|\mathrm{d}\phi_1(\cdot, T)_x - \mathrm{d}\phi_2(\cdot, T)_x\| &\leq C_0(T) \varepsilon_0 + C_1(T) \varepsilon_1 \quad . \end{aligned}$$

Démonstration

Commençons par le contrôle du flot. On a

$$\begin{aligned} \|\phi_1(x, T) - \phi_2(x, T)\| &\leq \int_0^T \left\| F^1(\phi_1(x, t), t) - F^2(\phi_2(x, t), t) \right\| dt \\ &\leq \int_0^T \left\| F^1(\phi_1(x, t), t) - F^1(\phi_2(x, t), t) \right\| dt \\ &\quad + \int_0^T \left\| F^1(\phi_2(x, t), t) - F^2(\phi_2(x, t), t) \right\| dt \\ &\leq \int_0^T K(t) \|\phi_1(x, t) - \phi_2(x, t)\| dt + \varepsilon_0(T - 0) \\ &\leq \varepsilon_0 T \exp\left(\int_0^T K(t) dt\right) \quad \text{par le lemme de Grönwall} \\ &\leq \varepsilon_0 T C(T) \quad . \end{aligned}$$

2.4 Dépendance de l'erreur en V

Passons maintenant au contrôle de la différentielle du flot

$$\begin{aligned}
\|d\phi_1(\cdot, T)_x - d\phi_2(\cdot, T)_x\| &\leq \int_0^T \left\| dF_{\phi_1(x,t)}^1(d\phi_1(\cdot, t)_x) - dF_{\phi_2(x,t)}^2(d\phi_2(\cdot, t)_x) \right\| dt \\
&\leq \int_0^T \left\| dF_{\phi_1(x,t)}^1(d\phi_1(\cdot, t)_x) - dF_{\phi_1(x,t)}^1(d\phi_2(\cdot, t)_x) \right\| dt \\
&\quad + \int_0^T \left\| dF_{\phi_1(x,t)}^1(d\phi_2(\cdot, t)_x) - dF_{\phi_2(x,t)}^2(d\phi_2(\cdot, t)_x) \right\| dt \\
&\leq \int_0^T \left\| dF_{\phi_1(x,t)}^1 \right\| \|d\phi_1(\cdot, t)_x - d\phi_2(\cdot, t)_x\| dt \\
&\quad + \int_0^T \left\| dF_{\phi_1(x,t)}^1 - dF_{\phi_2(x,t)}^1 \right\| \|d\phi_2(\cdot, t)_x\| dt \\
&\quad + \int_0^T \left\| dF_{\phi_2(x,t)}^1 - dF_{\phi_2(x,t)}^2 \right\| \|d\phi_2(\cdot, t)_x\| dt \quad .
\end{aligned}$$

$x \rightarrow dF^i(\cdot, t)_x$ est globalement lipschitzienne, donc

$$\begin{aligned}
\|d\phi_1(\cdot, T)_x - d\phi_2(\cdot, T)_x\| &\leq \int_0^T K(t) \|d\phi_1(\cdot, t)_x - d\phi_2(\cdot, t)_x\| dt \\
&\quad + \int_0^T K(t) \|\phi_1(x, t) - \phi_2(x, t)\| \|d\phi_2(\cdot, t)_x\| dt \\
&\quad + \int_0^T \varepsilon_1 \int_0^T \|d\phi_2(\cdot, t)_x\| dt \quad .
\end{aligned}$$

On a

$$d\phi_2(\cdot, t)_x = \text{Id} + \int_0^t dF^2(\cdot, s)_{\phi_2(x,s)} d\phi_2(\cdot, s)_x ds$$

et $\|dF^2(\cdot, t)_x\| \leq K(t)$. D'après le lemme de Grönwall, on en déduit que

$$\|d\phi_2(\cdot, t)_x\| \leq C(t)$$

et donc

$$\begin{aligned}
\|d\phi_1(\cdot, T)_x - d\phi_2(\cdot, T)_x\| &\leq \int_0^T K(t) \|d\phi_1(\cdot, t)_x - d\phi_2(\cdot, t)_x\| dt \\
&\quad + \int_0^T K(t) \varepsilon_0 TC(t) C(t) dt \\
&\quad + \varepsilon_1 \int_0^T C(t) dt
\end{aligned}$$

On utilise le lemme de Grönwall et on obtient

$$\begin{aligned}
 &\leq \left(\int_0^T \varepsilon_0 K(t) TC(t) C(t) dt + \varepsilon_1 \int_0^T C(t) dt \right) \exp \left(\int_0^T K(t) dt \right) \\
 &\leq \left(\varepsilon_0 TC(T)^2 \int_0^T K(t) dt + \varepsilon_1 \int_0^T C(t) dt \right) C(T) \\
 &\leq \left(\varepsilon_0 TC(T)^3 + \varepsilon_1 \int_0^T C(t) dt \right) C(T) \\
 &\leq \varepsilon_0 TC(T)^4 + \varepsilon_1 TC(T)^2 \quad .
 \end{aligned}$$

De là on en déduit :

$$\|\phi_1(x, T) - \phi_2(x, T)\| \leq \overline{C}_0(T) \varepsilon_0 \quad ,$$

$$\|d\phi_1(\cdot, T)_x - d\phi_2(\cdot, T)_x\| \leq C_0(T) \varepsilon_0 + C_1(T) \varepsilon_1 \quad .$$

2.4.4.2 Erreur entre la solution exacte et son approximation obtenue par transport par le champ approché

On considère ici un champ F vérifiant les hypothèses (2.2) (page 22), qu'on connaît approximativement par un champ \overline{F} vérifiant lui aussi les hypothèses précédentes, tel que :

- $\|F - \overline{F}\|_\infty \leq \varepsilon_0$;
- $\|dF - d\overline{F}\|_\infty \leq \varepsilon_1$.

On notera $\phi(x, t)$ et $\overline{\phi}(x, t)$ les flots associés. Soit

$$\rho_0(x) = \frac{1}{\sqrt{\det V}} \eta \left({}^t(x - \mu)V^{-1}(x - \mu) \right)$$

tel que ρ_0 densité de probabilités admet des moments d'ordre $2p$ finis. On considère la solution exacte donnée par

$$\rho(x, t) = \rho_0(\phi(x, -t)) |\det d\phi(\cdot, -t)_x| \quad .$$

On note $\overline{\rho}(x, t)$ l'approximation de

$$\rho_0(\overline{\phi}(x, -t)) |\det d\overline{\phi}(\cdot, -t)_x|$$

2.4 Dépendance de l'erreur en V

qui nous fournit une approximation de ρ où l'on a commis deux erreurs, une à cause de l'erreur sur le champ et l'autre à cause de l'approximation. Le théorème suivant nous fournit une estimation d'erreur.

Théorème 17

$$W_p(\rho, \bar{\rho}) \leq m_1 (C_0 \varepsilon_0 + C_1 \varepsilon_1) + M(t) m_2 \|\sqrt{V}\|^2$$

Démonstration

Reprenant le principe de la démonstration de Théorème 15, on a :

- $\phi(\cdot, t)$ transporte ρ_0 sur $\rho(\cdot, t)$;
- $d\bar{\phi}(\cdot, t)_\mu(x - \mu) + \mu$ transporte ρ_0 sur $\bar{\rho}$.

Ainsi

$$S(x) = d\bar{\phi}(\cdot, t)_\mu(\phi(x, -t) - \mu) + \mu$$

transporte ρ sur $\bar{\rho}$. Dans la suite, pour simplifier on supposera que $\mu = 0$.

On a alors

$$\begin{aligned} W_p(\rho, \bar{\rho})^p &\leq \int \|S(x) - x\|^p \rho(x, t) dx \\ &\leq \int \|d\bar{\phi}(\cdot, t)_0(\phi(x, -t)) + \bar{\phi}(0, t) - x\|^p \rho(x, t) dx \quad . \end{aligned}$$

On pose $y = \phi(x, -t)$ et on obtient

$$W_1(\rho, \bar{\rho}) \leq \int \|d\bar{\phi}(\cdot, t)_0(y) + \phi(0, t) - \phi(y, t)\|^p \rho_0(y) dy \quad .$$

Par une formule de Taylor sur $\phi(y + 0, t)$, on a alors, à constante de convexité près

$$W_1(\rho, \bar{\rho}) \leq \int \|d\bar{\phi}(\cdot, t)_0 - d\phi(\cdot, t)_0\|^p \|y\|^p \rho_0(y) dy + M(t) \int \|y\|^{2p} \rho_0(y) dy \quad .$$

On a

$$\int \|d\bar{\phi}(\cdot, t)_0 - d\phi(\cdot, t)_0\|^p \|y\|^p \rho_0(y) dy \leq (C_0(T) \varepsilon_0 + C_1(T) \varepsilon_1)^p \underbrace{\int \|y\|^{2p} \rho_0(y) dy}_{m_1^p} \quad .$$

Reste à maîtriser le deuxième terme. On a

$$\int \|y\|^{2p} \frac{C_d}{\sqrt{\det V}} \eta({}^t(y - \mu)V(y - \mu)) dy = \int \langle \sqrt{V^{-1}}x, \sqrt{V^{-1}}x \rangle^{2p} \eta(\|x\|^2) dx \quad .$$

Comme

$$\langle \sqrt{V^{-1}}x, \sqrt{V^{-1}}x \rangle^{2p} \leq \left\| \sqrt{V^{-1}} \right\|^{2p} \|x\|^{2p} \quad ,$$

on obtient

$$\int \|y\|^{2p} \frac{C_d}{\sqrt{\det V}} \eta \left({}^t(y - \mu)V(y - \mu) \right) dy \leq \left\| \sqrt{V} \right\|^{2p} \underbrace{\int \|x\|^{2p} \eta \left(\|x\|^2 \right) dx}_{m_2^p} \quad .$$

En combinant cette dernière inégalité avec celle portant sur la différentielle du flot, on a que :

$$\begin{aligned} W_p(\rho, \bar{\rho})^p &\leq m_1^p (C_0\varepsilon_0 + C_1\varepsilon_1)^p + m_2^p M(t) \left\| \sqrt{V} \right\|^{2p} \\ &\leq \left(m_1 (C_0\varepsilon_0 + C_1\varepsilon_1) + m_2 M(t) \left\| \sqrt{V} \right\|^2 \right)^p \end{aligned}$$

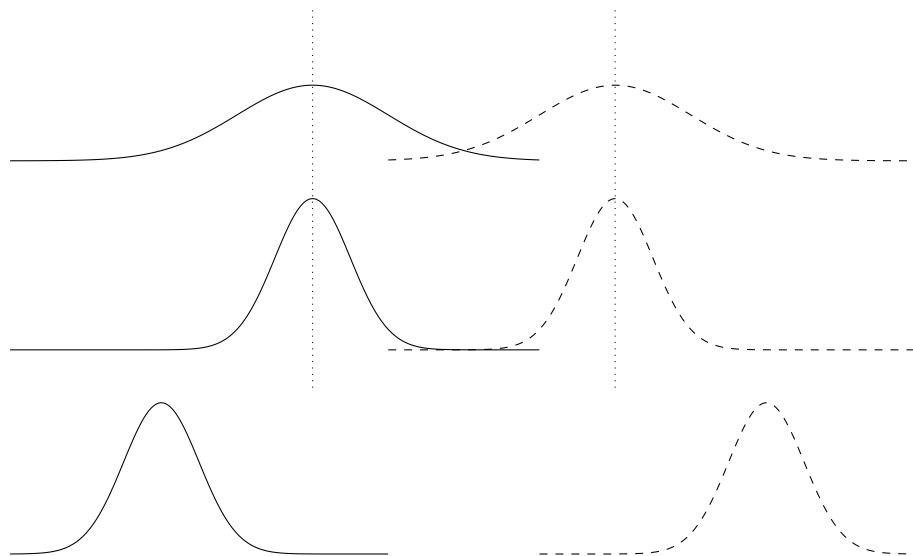
Nous avons introduit dans ce chapitre les éléments fondamentaux de la méthode. Elle amène naturellement à travailler avec des densités concentrées ((2.3), page 22 et (2.4), page 22). Nous avons vu, lors des estimateurs d'erreurs, (Théorème 8, page 26 et Théorème 15, page 32) que la concentration de la densité est ce qui gouverne la précision de l'estimation. Nous allons prolonger cette thématique au chapitre suivant.

Chapitre 3

Comparaison de distances pour l'étude de densités concentrées

3.1 introduction

Considérons les trois configurations suivantes, représentant des densités de probabilités



Pour la distance L^1 , les densités des deux dernières configurations sont à la même distance, et celles de la première à une distance plus petite. Dans une perspective d'applications, cela n'est pas toujours satisfaisant. Si l'on considère par exemple qu'une des densités représente un polluant et l'autre une densité de population, la

distance L^1 ne permet pas de mesurer le risque de manière très précise, la deuxième configuration étant plus problématique que la troisième

Pour la distance de Wasserstein, dont on parlera un peu plus bas les densités des deux premières configurations sont à la même distance, et celles de la dernière à une distance plus élevée : contrairement à la distance L^1 , elle perçoit bien « l'écartement », mais pas le « mélange » des densités.

Cette question du choix de la distance, qui n'est pas à proprement parler de la modélisation, mais y est liée dans l'esprit (choix d'un outil mathématique pour étudier un phénomène concret) est peu traitée en mathématiques appliquées hors du domaine de l'imagerie.

Nous allons étudier ici différentes distances dans l'idée de pouvoir les appliquer à la distance entre deux solutions de l'équation de transport conservative ((2.1), page 21) ou des densités définie par push-forward (section 2.3.2, page 24) avec pour données initiales des densités concentrées au voisinage d'une point comme nous les avons définies précédemment ((2.3), page 22).

Nous commencerons par nous intéresser aux distances de Wasserstein. Naturelles dans le contexte du transport, elles ne permettent cependant pas capter les phénomènes de concentration qui nous intéressent ici.

Nous étudierons ensuite la distance de Fisher, outil issu des statistiques. On verra sur le cas des gaussiennes qu'elle offre des propriétés intéressantes, mais reste d'un emploi malaisé.

Nous verrons alors qu'on peut moduler les distances de Wasserstein avec une fonctionnelle, permettant de retrouver certains comportements de la distance de Fisher, mais dans un formalisme théorique adapté à l'équation de transport.

3.2 Notations

Soit η est une fonction continue sur \mathbb{R} admettant des moments à tout ordre, C_d la constante tel que $x \in \mathbb{R}^d \mapsto C_d \eta(\|x\|_2^2)$ soit une densité de probabilité sur \mathbb{R}^d . On pose

$$\mathcal{X}_d = \left\{ x \mapsto \frac{C_d}{\sqrt{|\det V|}} \eta \left({}^t(x - \mu) V^{-1} (x - \mu) \right) \mid \mu \in \mathbb{R}^d, V \in S^{++} \right\} .$$

3.3 Distance de Wasserstein

En particulier, on a $\mathcal{X}_d \subset \mathcal{P}_p(\mathbb{R}^d)$ pour tout p . Dans le cas particulier des gaussiennes on posera

$$\mathcal{G}_d = \left\{ x \mapsto \frac{1}{(2\pi)^{\frac{d}{2}} \sqrt{\det V}} \exp\left(-\frac{1}{2} {}^t(x - \mu) V^{-1} (x - \mu)\right) \mid \mu \in \mathbb{R}^d, V \in S^{++} \right\} .$$

Si $d = 1$, on utilisera pour le paramétrage l'écart type noté σ , c'est à dire

$$\mathcal{G}_1 = \left\{ x \mapsto \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \mid \mu \in \mathbb{R}, \sigma \in \mathbb{R}^{+*} \right\} .$$

Soit ϕ un difféomorphisme de \mathbb{R}^d . Si ρ est une densité de probabilités, par abus de notation on notera $\phi_{\#}\rho$ la densité de la mesure image. On pose

$$\mathcal{X}_d^\phi = \{\phi_{\#}\rho \mid \rho \in \mathcal{X}_d\} \quad \text{et} \quad \mathcal{G}_d^\phi = \{\phi_{\#}g \mid g \in \mathcal{G}_d\} .$$

Si $\phi = \phi(\cdot, t)$ est le flot d'un champ de vecteur, \mathcal{X}_d^ϕ (respectivement \mathcal{G}_d^ϕ) est l'ensemble des solutions au temps t de l'équation de transport pour des données initiales dans \mathcal{X}_d (respectivement dans \mathcal{G}_d). Si $\rho \in \mathcal{X}_d$, on rappelle qu'on définit son approximation de la manière suivante (Définition 4, page 22)

$$\bar{\rho}(x, t) = \frac{1}{\sqrt{|\det V_t|}} \eta\left({}^t(x - \mu_t) V_t^{-1} (x - \mu_t)\right)$$

où $\mu_t = \phi(\mu)$ et $V_t^{-1} = {}^t d\phi_\mu^{-1} \circ V^{-1} \circ d\phi_\mu^{-1}$.

3.3 Distance de Wasserstein

Nous allons commencer par rappeler brièvement les éléments essentiels de la théorie du transport optimal. On pourra se référer à [37] pour une introduction approfondie du sujet.

3.3.1 Introduction au transport optimal

3.3.1.1 Introduction : le problème de Monge

Considérons un tas de sable, dont la répartition est représentée dans un espace X par une mesure μ : si $A \subset X$, $\mu(A)$ mesure la quantité de sable présente en A . On souhaite transporter ce tas dans un emplacement précis, représenté par une mesure ν fixée dans un espace Y : si $B \subset Y$, $\nu(B)$ mesure la quantité de sable devant se retrouver en B après transport. On doit donc avoir $\mu(X) = \nu(Y)$ et on peut en normalisant supposer que les mesures sont des mesures de probabilités.

On considère que transporter une unité de $x \in X$ sur $y \in Y$ coûte $c(x, y)$, où $c : X \times Y \rightarrow \mathbb{R} \cup \{+\infty\}$ est une fonction mesurable appelée **fonction de coût**. On cherche une fonction mesurable $T : X \rightarrow Y$ envoyant le tas de sable sur l'emplacement d'arrivée et minimisant le coût de transport (en un sens que nous allons préciser plus bas).

Dans un premier temps, il s'agit de déterminer les fonctions T acceptables, c'est à dire transportant effectivement le tas de sable sur l'emplacement d'arrivée. Soit $B \subset Y$. La matière présente en B provenant de $T^{-1}(B)$, on doit avoir $\nu(B) = \mu(T^{-1}(B))$, c'est à dire que ν doit être la mesure image de μ par T . On appellera **transport** de μ sur ν une telle fonction.

Intéressons nous maintenant à la question du coût de transport, c'est à dire de définir une fonction $\mathcal{C}(T)$ associant à chaque transport un coût. Transporter la matière de x en $T(x)$ coûte $c(x, T(x))$. En sommant les coûts de transports en les pondérant par la quantité de matière déplacée, on obtient

$$\mathcal{C}(T) = \int c(x, T(x)) d\mu(x) \quad .$$

Chercher à optimiser le coût de transport revient donc à résoudre le problème de minimisation suivant

$$\inf \mathcal{C}(T)$$

sur l'ensemble des fonctions T tel que $\nu = T_{\#}\mu$. On appelle ce problème le **problème de Monge**.

Exemple 18

Supposons que l'on souhaite envoyer tout le sable en $y_0 \in Y$, c'est à dire $\nu = \delta_{y_0}$.

3.3 Distance de Wasserstein

Il n'y a alors qu'un seul transport possible : $T(x) = y_0$ et l'on a alors

$$\inf \mathcal{C} = \int_X c(x, y_0) d\mu \quad .$$

Supposons à l'inverse que tout le sable est concentré en $x_0 \in X$ et qu'on souhaite l'envoyer sur $\nu = \frac{1}{2}(\delta_{y_1} + \delta_{y_2})$ avec $y_1 \neq y_2$. Alors il n'y a pas de solutions. En effet, on doit avoir $\nu(T(x_0)) = 1$, donc $\frac{1}{2}(\delta_{y_1}(T(x_0)) + \delta_{y_2}(T(x_0))) = 1$ et donc $T(x_0) = y_1$ ainsi que $T(x_0) = y_2$ ce qui est impossible.

Assez limpide dans sa formulation, le problème de Monge est cependant compliqué dans son traitement mathématique. En particulier, comme nous le montre la deuxième partie de l'exemple, à cause de l'impossibilité de scinder la masse présente en un point de l'espace de départ lors du transport. Une formulation alternative due à Kantorovich permet de traiter ce genre de configurations.

3.3.1.2 Le problème de Kantorovich

Dans la suite, $P(X \times Y)$ désignera l'ensemble des mesures de probabilité sur $X \times Y$.

Définition 5

Soit $\Pi(\mu, \nu) = \{\pi \in P(X \times Y) \mid \pi \text{ a pour lois marginales } \mu \text{ et } \nu\}$ On appelle **problème de Kantorovich** le problème de minimisation suivant

$$\inf_{\pi \in \Pi(\mu, \nu)} \int_{X \times Y} c(x, y) d\pi(x, y) \quad .$$

On appelle **plan de transport** un élément de $\Pi(\mu, \nu)$.

Remarque 19

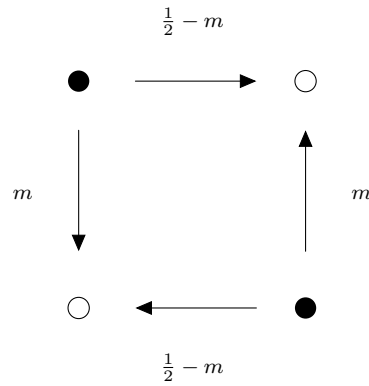
Ce problème est une version relaxée du problème de Monge. Si $A \subset X$ et $B \subset Y$, $\pi(A \times B)$ mesure la proportion de matière de A envoyée en B . Toute la matière présente en A devant être transporté dans Y , on a $\pi(A \times Y) = \mu(A)$. De même, $\pi(X \times B) = \nu(B)$. Ainsi, μ et ν sont les lois marginales de π .

Remarque 20

A l'inverse du problème de Monge, il existe toujours une façon d'effectuer le transport, fût-elle non optimale. Il suffit en effet de prendre $\pi = \mu \otimes \nu$. Pour ce qui concerne l'existence d'un plan de transport réalisant l'infimum (qu'on appellera plan de transport optimal), elle nécessite des hypothèses assez faible puisque l'on peut montrer qu'il suffit que le coût de transport soit semi-continu inférieurement.

Exemple 21

Nous allons illustrer ici le problème de Kantorovich. Dans le plan, on considère le cas où $\mu = \frac{1}{2}(\delta_{(1,0)} + \delta_{(0,1)})$ (points noirs sur la figure suivante) et $\nu = \frac{1}{2}(\delta_{(0,0)} + \delta_{(1,1)})$ (points blancs sur la figure suivante), pour le coût $c(x, y) = \|x - y\|$.



Il existe alors une infinité de solutions optimales : qu'on envoie toute la masse de $(1, 0)$ sur $(1, 1)$ et rien sur $(0, 0)$, ou une partie m sur $(1, 1)$ et $\frac{1}{2} - m$ sur $(0, 0)$ ne change rien au coût de transport total.

Notons que le problème de Monge correspondant admet deux solutions.

La question qui se pose maintenant est l'existence de minimiseurs à ces problèmes et leurs caractérisations.

3.3.1.3 Principaux résultats

Nous nous limiterons au cas $X = Y = \mathbb{R}^d$ et $c(x, y) = \|x - y\|^p, 1 < p < \infty$. Les hypothèses présentées ne seront pas toujours les plus fines pour faciliter la compréhension. Le lecteur pourra se référer à [37] pour des hypothèses plus faibles.

Théorème 22 (transport optimal pour un coût strictement convexe)

Soit μ, ν deux mesures de probabilités sur \mathbb{R}^d . On pose $c(x, y) = \|x - y\|^p, p > 1$. Si μ est absolument continue par rapport à la mesure de Lebesgue, alors il existe une unique solution π au problème de Kantorovich. De plus, cette solution est aussi l'unique solution du problème de Monge, en ce sens que $\pi = (\text{Id} \times T)_\# \mu$ où T est la solution du problème de Monge, et les coûts de transports coïncident.

Si $p = 2$, on a $T = \nabla\varphi$ où $\nabla\varphi$ est l'unique ($d\mu$ -presque partout) gradient d'une fonction convexe tel que $\nabla\varphi_\# \mu = \nu$.

3.3 Distance de Wasserstein

On se permettra donc maintenant de parler de solution du problème de Monge-Kantorovich.

Remarque 23

Historiquement, un premier résultat a été démontré dans le cas $p = 2$ sous des hypothèses un peu plus contraignantes, indépendamment dans les années 1985 par Knott et Smith, ainsi que par Yann Brenier ([10]). C'est une dizaine d'années plus tard que les travaux de Gangbo et McCann ont amené un résultat plus général ([26], [18]).

Disons un mot du cas $p = 1$. La perte de stricte convexité de la fonction coût rend la situation très compliquée, les techniques de démonstration reposant pour une part importante sur de l'analyse convexe. Disons simplement que si μ et ν sont absolument continues par rapport à la mesure de Lebesgue, alors il existe des solutions aux problèmes de Monge qui sont aussi solutions du problèmes de Kantorovich, mais qu'il n'y a pas d'unicité. Le lecteur intéressé pourra consulter la page 100 de [37].

3.3.1.4 Espace de Wasserstein

Les outils du transport optimal peuvent servir pour une étude métrique d'espaces de mesures. Nous allons nous contenter ici d'introduire les premières définitions et résultats qui nous serviront par la suite,.

Théorème (distance de Wasserstein)

Soit $\mathcal{P}_p(\mathbb{R}^d)$ l'ensemble des mesures de probabilités sur \mathbb{R}^d admettant des moments finis jusqu'à l'ordre p . Pour μ et ν appartenant à $\mathcal{P}_p(\mathbb{R}^d)$, on pose $C_p(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \int \|x - y\|^p d\pi$. Alors

$$W_p(\mu, \nu) = (C_p(\mu, \nu))^{\frac{1}{p}}$$

définit une distance sur $\mathcal{P}_p(\mathbb{R}^d)$ qu'on appelle distance de Wasserstein.

Remarque 24

Si l'on compare avec les hypothèses du Théorème 22, nous n'avons pas forcément ici l'unicité d'un plan de transport optimal.

Théorème (Wasserstein métrise la convergence faible)

Soit $(\mu)_k$ une suite de $\mathcal{P}_p(\mathbb{R}^d)$. $\lim_{k \rightarrow \infty} W_p(\mu_k, \mu) \rightarrow 0$ si et seulement si :

1. μ_k converge faiblement vers μ ;
2. il y a convergence des moments d'ordre p : $\int \|x\|^p d\mu_k \rightarrow \int \|x\|^p d\mu$.

En particulier l'espace $\mathcal{P}_p(\mathbb{R}^d)$ est un espace métrique complet.

3.3.1.5 Temps et géodésiques

Jusqu'à maintenant, la mesure de départ était instantanément transportée sur la mesure d'arrivée. Ainsi, la notion de « chemin » pris par la matière pour se déplacer n'avait pas de sens. Nous allons maintenant, en introduisant un paramètre temporel, y donner un sens.

On se placera maintenant dans le cas où $p = 2$.

Définition 6

Soit ρ_0, ρ_1 deux mesures de probabilités absolument continues par rapport à la mesure de Lebesgue. Soit T le transport optimal de ρ_0 sur ρ_1 . On appelle **interpolée de McCann** la courbe reliant ρ_0 à ρ_1 définie sur $[0, 1]$ par

$$\rho_t = \{(1-t)\text{Id} + tT\}_\# \rho_0 \quad .$$

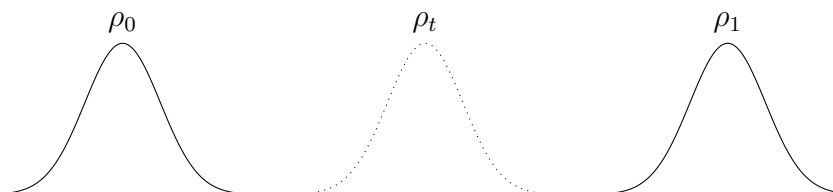
On dit qu'un sous-ensemble X de l'ensemble des mesures de probabilités est **est géodésiquement convexe** si pour tout couple de mesure de X , leur interpolée de McCann est dans X .

Remarque 25

Nous allons comparer l'interpolation introduite ici avec l'interpolation L^1 classique ($\rho_t = (1-t)\rho_0 + t\rho_1$). On considère $\rho_1(x) = \rho_0(x-c)$. On peut alors facilement montrer que $T(x) = x+c$, et donc

$$\begin{aligned} \rho_t &= ((1-t)x + t(x+c))_\# \rho_0 \\ &= (x+tc)_\# \rho_0 \quad . \end{aligned}$$

Le premier dessin représente l'interpolation au sens de McCann, le deuxième l'interpolation classique.



3.3 Distance de Wasserstein



Visuellement, l'interpolation de McCann déplace la densité de départ sur la densité d'arrivée. Cette interpolation est d'ailleurs couramment appelé en anglais, suivant McCann, **displacement interpolation**. C'est une interpolation très naturellement liée au contexte de l'équation de transport.

Remarque 26

On peut montrer que l'interpolation de McCann est la solution d'un problème de transport optimal dépendant du temps et qu'en un sens à préciser, c'est le plus court chemin entre ρ_0 et ρ_1 . C'est ce que nous dit partiellement l'égalité $W_2(\rho_0, \rho_t) = t W_2(\rho_0, \rho_1)$. Nous ne rentrerons pas plus en détail sur ces considérations ici, qui amènent à l'utilisation d'outils théoriques compliquées et d'hypothèses assez fines.

3.3.1.6 Le cas particulier des gaussiennes

Nous allons présenter ici quelques résultats génériques sur le transport optimal et la distance de Wasserstein pour les gaussiennes. On pourra consulter [19] pour une étude de la distance entre deux gaussiennes pour la distance de Wasserstein quadratique. [36] développe cette étude en la complétant notamment sur la question de la courbure de l'espace. [27] a introduit l'interpolation de McCann et s'en sert pour démontrer le caractère géodésiquement convexe de l'espace des gaussiennes.

Le théorème suivant nous fournit un transport optimal entre deux gaussiennes

Théorème 27 ([36])

Soit g_1 et g_2 deux gaussiennes de \mathbb{R}^d , d'espérance μ_1 et μ_2 , de matrices de covariance V_1 et V_2 . On pose

$$W = V_1^{\frac{1}{2}} \left(V_1^{\frac{1}{2}} V_2 V_1^{\frac{1}{2}} \right)^{-\frac{1}{2}} V_1^{\frac{1}{2}} \quad .$$

La fonction T ci dessous est le transport optimal de g_1 sur g_2

$$T(x) = 2W(x - \mu_1) + \mu_2$$

\mathcal{G}_d est géodésiquement convexe et la distance de Wasserstein quadratique entre les deux gaussiennes est donnée par

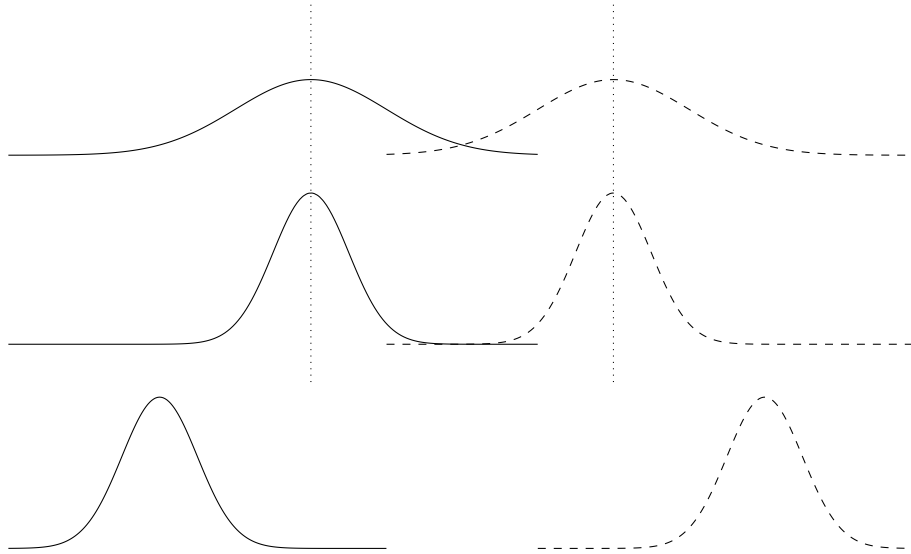
$$W_2(g_1, g_2)^2 = \|\mu_1 - \mu_2\|_2^2 + \text{tr } V_1 + \text{tr } V_2 - 2 \text{tr} \left(V_1^{\frac{1}{2}} V_2 V_1^{\frac{1}{2}} \right)^{\frac{1}{2}} .$$

Remarque 28

Prenons $d = 1$. On a $\sigma_i^2 = V_i$. De là,

$$W_2(g_1, g_2)^2 = |\mu_1 - \mu_2|^2 + |\sigma_1 - \sigma_2|^2$$

et l'on en déduit que (\mathcal{G}_1, W_2) est isométrique au demi plan supérieur ouvert muni de la distance euclidienne. En particulier, si $\sigma_1 = \sigma_2$, la distance ne dépend que de μ_0 et μ_1 . Ainsi, en reprenant la figure ouvrant ce chapitre dans le cas des gaussiennes, la distance est identique dans les deux premières configurations, la troisième distance est plus importante.



3.3.2 L'équation de transport est bien posée pour la distance de Wasserstein

On considère l'équation de transport conservative (2.1) (page 21), avec les hypothèses définies en (2.2) (page 22), sur le champ F . Notons que d'après l'EDO vérifié par la différentielle du flot (Proposition 5, page 19) on a la borne suivante sur la

3.3 Distance de Wasserstein

dérivée du flot à t fixé :

$$\|d\phi(\cdot, t)\|_\infty \leq \exp\left(\int_0^t k(s) ds\right)$$

d'après le lemme de Grönwall.

Théorème 29

L'équation de transport conservative est bien posée sur $(\mathcal{P}_p(\mathbb{R}^d), W_p)$.

Formulé un peu différemment sur le sous-ensemble \mathcal{X}_d^ϕ , le théorème nous donne le corollaire suivant

Corollaire 30

Si $\phi_\# \rho_0$ et $\phi_\# \rho_1$ appartiennent à \mathcal{X}_d^ϕ , on a

$$W_2(\phi_\# \rho_0, \phi_\# \rho_1) \leq \sup_{x \in \mathbb{R}^d} \|d\phi_x\| W_2(\rho_0, \rho_1) \quad .$$

Démonstration

L'existence et l'unicité d'une solution pour toute donnée initiale est déjà connu. $d\phi$ est borné, donc les moments d'ordre p sont finis. Reste à voir la continuité par rapport aux données initiales. Soit ρ_1, ρ_2 deux données initiales appartenant à $\mathcal{P}_p(\mathbb{R}^d)$. On notera ρ_1^t et ρ_2^t les solutions de l'équation de transport au temps t .

Soit T transport optimal de ρ_1 sur ρ_2 . On pose $T_t = \phi(\cdot, t) \circ T \circ \phi(\cdot, -t)$. Alors T_t est un transport de ρ_1^t sur ρ_2^t . En effet, soit $f \in \mathcal{C}_b^0(\mathbb{R}^d)$. On a

$$\begin{aligned} T_{t\#} \rho_1^t(f) &= \int (f \circ T_t) d\rho_1^t \\ &= \int ((f \circ \phi(\cdot, t) \circ T) \circ \phi(\cdot, -t)) d\rho_1^t \\ &= \int (f \circ \phi(\cdot, t) \circ T) d\rho_1 \quad \text{car } \rho_1 = \phi(\cdot, -t)_\# \rho_1^t \\ &= \int (f \circ \phi(\cdot, t)) d\rho_2 \quad \text{car } \rho_2 = T_\# \rho_1 \\ &= \int f d\rho_2^t \quad \text{car } \rho_2^t = \phi(\cdot, t)_\# \rho_2 \quad . \end{aligned}$$

De là, on en déduit que

$$\begin{aligned}
 W_p(\rho_1^t, \rho_2^t)^p &\leq \int \|T_t(x) - x\|^p d\rho_1^t \\
 &\leq \int \|(\phi(\cdot, t) \circ T \circ \phi(\cdot, -t))(x) - x\|^p d\rho_1^t \\
 &\leq \int \|\phi(T(y), t) - \phi(y, t)\|^p d\rho_1 \quad \text{car } \rho_1 = \phi(\cdot, -t)_\# \rho_1^t \\
 &\leq \|d\phi(\cdot, t)\|_\infty^p \int \|T(y) - y\|^p d\rho_1 \quad \text{car } \|d\phi(\cdot, t)\|_\infty \text{ majoré.}
 \end{aligned}$$

Corollaire 31

Soit $\phi_\# \rho_0$ et $\phi_\# \rho_1$ appartenant à \mathcal{X}_1^ϕ . On a

$$\inf |\phi'| W_2(\rho_0, \rho_1) \leq W_2(\phi_\# \rho_0, \phi_\# \rho_1) \leq \sup |\phi'| W_2(\rho_0, \rho_1) \quad .$$

Démonstration

Soit T transport optimal de ρ_1 sur ρ_2 . On pose $T_t = \phi \circ T \circ \phi^{-1}$. On sait d'après la démonstration précédente que T_t est un transport de ρ_1 sur ρ_2 . Il est facile de voir que T_t est le gradient d'une fonction convexe. Donc d'après le Théorème 22 (page 44), T_t est le transport optimal de $\phi_\# \rho_0$ sur $\phi_\# \rho_1$. Donc

$$\begin{aligned}
 W_2(\phi_\# \rho_0, \phi_\# \rho_1) &= \int |T_t(x) - x|^2 d\rho_1^t \\
 &= \int |(\phi \circ T \circ \phi^{-1})(x) - x|^2 d\rho_1^t \\
 &= \int |\phi - \phi(y, t)|^2 d\rho_1 \quad \text{car } \rho_1 = \phi_\#^{-1} \rho_1^t \\
 &= \int \left| \int_y^{T(y)} d\phi_x dx \right|^2 d\rho_1 \\
 &\geq \inf_{x \in \mathbb{R}} |d\phi_x|^2 \int |T(y) - y|^2 d\rho_1 \\
 &\geq \inf_{x \in \mathbb{R}} |d\phi_x|^2 W_2(\rho_0, \rho_1)^2 \quad .
 \end{aligned}$$

Appliquons ce dernier résultat à deux gaussiennes 1d g_0 et g_1 , d'espérance μ_1 et μ_2 , ayant le même écart-type σ . On a alors, d'après les résultats sur la distance de Wasserstein quadratique entre gaussiennes (Théorème 27 page 47 et Remarque 28

3.4 Distance de Fisher

page 48),

$$\inf |\phi'| \sqrt{|\mu_1^2 - \mu_0^2|} \leq W_2(\phi_{\#}g_0, \phi_{\#}g_1) \leq \sup |\phi'| \sqrt{|\mu_1^2 - \mu_0^2|} .$$

La concentration des gaussiennes n'a donc pas d'influence.

En conclusion, la distance de Wasserstein est naturellement adapté au cadre théorique du transport, mais la Remarque 28 et cet exemple illustrent le fait que la distance de Wasserstein ne permet de percevoir les phénomènes de concentration. Nous allons maintenant présenter une distance qui permet de les percevoir, mais d'un usage plus complexe.

3.4 Distance de Fisher

3.4.1 Introduction

La distance de Fisher est un outil issu de la *Géométrie de l'information* ([1], [2]). Celle-ci utilise des outils de géométrie différentielle à des fins statistiques. C'est une branche relativement récente des mathématiques, puisqu'elle s'est développée vers 1985 à partir du Japon. Un nombre de résultats significatifs n'étaient encore accessibles qu'en japonais à l'aube des années 2000.

L'idée centrale est de considérer un espace de densités de probabilités comme une variété riemannienne, permettant ensuite la résolution de problèmes de nature statistique.

Cette approche est fortement basée sur l'estimateur du maximum de vraisemblance, regardant $\rho(\cdot, \theta)$ appartenant à une famille de fonctions $(\rho_\theta)_{\theta \in \Theta}$ comme une fonction en θ . Nous ne reprendrons pas cet estimateur ici, on trouvera en annexe une petite introduction aux statistiques le présentant (section 7.2, page 134). On y trouvera aussi quelques rappels de géométrie différentielle (section 7.3, page 139). Pour une introduction complète aux statistiques, on pourra se reporter à [33], pour la géométrie différentielle et riemannienne à [23], [17].

3.4.2 Matrice de Fisher

3.4.2.1 Cas général

Soit

$$M = \left\{ x \mapsto \rho(x, \theta) \mid \int \rho \, dx = 1, \theta \in \Theta \text{ ouvert de } \mathbb{R}^m \right\}$$

famille de densités de probabilités définies sur \mathbb{R}^d . On notera parfois ρ_θ la densité $\rho(\cdot, \theta)$. On considérera fréquemment

$$L = \{x \mapsto \ln(\rho(x, \theta)) \mid \rho(\cdot, \theta) \in M\}$$

pour des calculs plus simples et on notera $l(\cdot, \theta)$ les éléments de L , ie $l(\cdot, \theta) = \ln(\rho(\cdot, \theta))$. Dans la suite, on fait les hypothèses suivantes :

- $\theta_1 \neq \theta_2 \implies \rho(\cdot, \theta_1) \neq \rho(\cdot, \theta_2)$, c'est à dire que le modèle est identifiable ;
- le support de ρ_θ , $\overline{\{x \in \mathbb{R}^d \mid \rho_\theta(x) > 0\}}$, ne dépend pas de θ .
- pour presque tout x , l'application $\theta \mapsto \rho(x, \theta)$ est différentiable en θ et on peut différentier sous le signe somme.

Définition 7

On appelle **matrice d'information de Fisher** la matrice semi-définie positive définie par

$$I_\theta = \mathbb{E} \left[\frac{\partial}{\partial \theta_i} l(\cdot, \theta) \times \frac{\partial}{\partial \theta_j} l(\cdot, \theta) \right]_{(i,j)}$$

On a le résultat suivant très pratique pour le calcul de cette matrice

Proposition 32

Si pour presque tout x , l'application $\theta \mapsto \rho(x, \theta)$ est deux fois différentiable en θ et on peut différentier deux fois sous le signe somme, alors

$$I_\theta = - \left(\mathbb{E} \left[\partial_{ij}^2 l(\cdot, \theta) \right] \right)_{i,j}$$

Démonstration

Nous allons nous contenter ici de parler du cas 1d pour éclairer un peu la situation et l'importance des hypothèses. On a

$$\int \rho(x, \theta) \, dx = 1 \quad \forall \theta \quad .$$

3.4 Distance de Fisher

Par dérivation sous le signe somme, on a

$$\begin{aligned} 0 &= \int \partial_\theta \rho(x, \theta) dx \\ &= \int \partial_\theta [\ln(\rho(x, \theta))] \rho(x, \theta) dx \quad \text{comme le support ne dépend pas de } \theta. \end{aligned}$$

On dérive une seconde fois sous le signe somme, et on a

$$\begin{aligned} 0 &= \partial_\theta \int [\ln(\rho(x, \theta))] \rho(x, \theta) dx \\ &= \int \partial_\theta^2 [\ln(\rho(x, \theta))] \rho(x, \theta) dx + \int (\partial_\theta [\ln(\rho(x, \theta))])^2 dx \quad . \end{aligned}$$

Et donc

$$I(\theta) = - \int \partial_\theta^2 [\ln(\rho(x, \theta))] \rho(x, \theta) dx \quad .$$

Si la matrice de Fisher existe et si elle est définie-positive, alors on peut munir Θ d'une structure de variété riemannienne \mathcal{C}^1 nous donnant la métrique de Fisher. Nous allons dès maintenant illustrer ce concept sur les gaussiennes 1d et nous verrons sa signification statistique.

3.4.2.2 Expression de la métrique de Fisher pour les gaussiennes 1d pour le paramétrage classique

On travaille sur les gaussiennes 1d

$$\mathcal{G}_1 = \left\{ x \mapsto \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \mid \mu \in \mathbb{R}, \sigma \in \mathbb{R}^{+*} \right\} \quad .$$

Il est naturel de prendre comme paramétrage $\theta = (\mu, \sigma)$ et $\Theta = \mathbb{R} \times \mathbb{R}^{+*}$.

Proposition 33

La matrice de Fisher pour une gaussienne 1d est donnée par

$$I(\mu, \sigma) = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{pmatrix}$$

Démonstration

On a

$$l(x, \theta) = -\ln(2\pi) - \ln(\sigma) - \frac{(x-\mu)^2}{2\sigma^2} \quad .$$

De là, on en déduit que

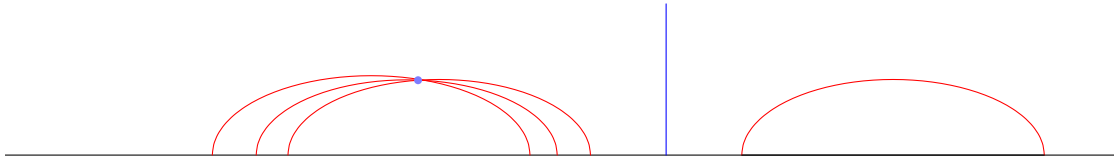
$$\begin{aligned}\partial_{\mu}^2 l(x, \theta) &= -\frac{1}{\sigma^2} \quad , \\ \partial_{\sigma}^2 l(x, \theta) &= \frac{1}{\sigma^2} - 3\frac{(x - \mu)^2}{\sigma^4} \quad , \\ \partial_{\mu\sigma}^2 l(x, \theta) &= -2\frac{x - \mu}{\sigma^3} \quad ,\end{aligned}$$

et ensuite

$$\begin{aligned}\int \partial_{\mu\sigma}^2 l(x, \theta) \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx &= 0 \quad , \\ \int \partial_{\mu}^2 l(x, \theta) \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx &= -\frac{1}{\sigma^2} \quad , \\ \int \partial_{\sigma}^2 l(x, \theta) \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx &= \frac{1}{\sigma^2} - 3 \int \frac{(x - \mu)^2}{\sigma^4} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx \\ &= -\frac{2}{\sigma^2} \quad .\end{aligned}$$

$\mathbb{R} \times \mathbb{R}^{+*}$ muni de cette famille est donc une variété riemannienne et on retrouve donc, à isométrie près, le demi plan de Poincaré. On trouvera en annexe, (section 7.1, page 131), une brève introduction sur ce sujet. Donnons simplement les informations essentielles ici.

On trouvera sur la figure suivante les courbes géodésiques. Le facteur en $\frac{1}{\sigma^2}$ rend coûteux le fait de s'approcher du bord, ce que montre bien la figure



En reprenant la figure du début du chapitre (page 39) en regardant les courbes comme des gaussiennes, cette distance nous dit que les trois situations sont différentes, allant de la plus proche (en haut) à la plus éloignée (en bas). Le théorème suivant formalise ces observations et donne une formule explicite pour la distance.

3.4 Distance de Fisher

Théorème 34 ([4])

Soit $\theta_1 = (\mu_1, \sigma_1)$ et $\theta_2 = (\mu_2, \sigma_2)$ points du demi-plan. On a

$$d(\theta_1, \theta_2) = \sqrt{2} \operatorname{arccosh} \left(1 + \frac{(\mu_1 - \mu_2)^2 + 2(\sigma_1 - \sigma_2)^2}{4\sigma_1\sigma_2} \right) .$$

Les courbes géodésiques sont de la forme

$$\begin{pmatrix} \mu(t) \\ \sigma(t) \end{pmatrix} = \begin{pmatrix} c_1 + c_2 2 \tanh \left(\frac{t}{\sqrt{2}} + c_3 \right) \\ \sqrt{2} c_2 \operatorname{arccosh} \left(\frac{t}{\sqrt{2}} + c_3 \right) \end{pmatrix} \quad \text{quand } \mu_1 \neq \mu_2$$

et

$$\begin{pmatrix} \mu(t) \\ \sigma(t) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \exp \left(\frac{t}{\sqrt{2}} + c \right) \end{pmatrix} \quad \text{quand } \mu_1 = \mu_2$$

Disons un mot du sens statistique de cette distance. Commençons pour cela par illustrer l'information de Fisher. Nous sommes toujours dans le cas gaussien. Le théorème de Cramer-Rao (Théorème 67, 137) nous dit que l'information de Fisher nous donne une borne inférieure sur la précision d'un estimateur. Soit (X_1, \dots, X_n) un échantillon indépendant identiquement distribué de loi $\mathcal{N}(\mu, \sigma)$. On a par propriété (Proposition 66, page 137) de l'information de Fisher

$$I_{(X_1, \dots, X_n)} = nI_{X_1} = n \frac{1}{\sigma^2} \quad :$$

plus on a d'observations, plus on a d'informations. Supposons que l'on se donne un estimateur $\hat{\mu}_n(X_1, \dots, X_n)$ pour l'espérance. Alors l'erreur commise par l'estimateur est borné inférieurement d'après le théorème de Cramer-Rao (Théorème 67, page 137) :

$$\mathbb{E}_\theta \left[(\hat{\mu}_n - \mu(\theta))^2 \right] \geq \frac{\sigma^2}{n} .$$

Ainsi, plus la gaussienne sera resserrée, plus la borne inférieure sera petite.

Supposons maintenant que l'on a deux ensemble de points (x_1, \dots, x_n) et (y_1, \dots, y_n) correspondant à n tirages d'une variable aléatoire X et d'un variable aléatoire Y , de densités respectives les gaussiennes g_0 et g_1 . Si nous n'avons pas d'informations préalables sur l'appartenance des points à chaque ensemble, la distance de Fisher nous renseigne sur le risque de considérer un point x_i comme faisant partie de la famille (y_j) . Ou, de manière similaire, de reconstruire la gaussienne g_1 à partir des informations sur g_0 .

Notons pour finir que d'autres paramétrages sont possibles pour les gaussiennes 1d. Nous en mentionnons deux, par les moments $\theta = (\mu, \mu^2 + \sigma^2)$ et surtout sous la forme d'une famille exponentielle, $\theta = \left(\frac{\mu}{\sigma^2}, -\frac{1}{2\sigma^2}\right) \in \mathbb{R} \times]-\frac{1}{2}, -\infty[$. On pourra se référer à [1] pour une présentation des familles exponentielles et à [13] pour une présentation de la métrique de Fisher et des géodésiques suivant ces différents paramétrages.

3.4.3 α -connexion

On trouvera en annexe (section 7.3.2, page 141) une petit introduction aux connexions. On pourra se reporter à [1] pour plus de détails sur les résultats présentés ici.

Définition 8

Soit $\alpha \in \mathbb{R}$. On définit une connexion $\nabla^{(\alpha)}$ sur M au moyen des coefficients

$$\Gamma_{ijk} = \mathbb{E} \left[\left(\partial_i \partial_j l(x, \theta) + \frac{1-\alpha}{2} \partial_i l(x, \theta) \partial_j l(x, \theta) \right) \partial_k l(x, \theta) \right]$$

où $\Gamma_{ijk} = \sum_l \Gamma_{ij}^l g_{kl}$

Remarque 35

Le choix du paramètre α va dépendre de ce que l'on souhaite observer. Supposons que l'on a deux densité de probabilités ρ_0 et ρ_1 , la situation nous amenant à considérer comme naturel de les relier par

$$\rho(x, t) = (1 - t)\rho_0(x) + t\rho_1(x) \quad .$$

Alors cette courbe est une géodésique pour la -1 -connexion ([1]).

si l'on relie les deux densités au moyen de la famille exponentielle, c'est à dire $\rho(x, t) = \exp((1 - t)l_0(x) + tl_1(x) - c(t))$ où $l_i(x) = \ln(\rho_i(x))$, alors c'est géodésique pour la 1 -connexion ([1]).

Cet outil nous servira pour un calcul numérique de la distance de Fisher entre les gaussiennes multivariées.

Proposition 36 ([1])

La connexion de Levi-Civita sur M pour la métrique de Fisher est donnée par la 0 -connexion.

3.4.4 Les gaussiennes multidimensionnelles

Pour ce qui concerne l'étude des gaussiennes multidimensionnelles pour la distance de Fisher et les démonstrations de cette section, on pourra se reporter à [34]. Nous allons dire quelques mots de la distance de Fisher sur les gaussiennes multidimensionnelles. Nous allons assez peu détailler les résultats, mais mettre en avant les importantes différences avec le cas des gaussiennes 1d et les questions que cela soulève dans une perspective d'applications.

Soit S_d l'ensemble des matrices symétriques de dimension d et S_d^{++} le sous-ensemble des matrices symétriques définies positives. On pose

$$E_{ij} = \begin{cases} (\delta_{ik}\delta_{il})_{k,l} & \text{si } i = j \\ (\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})_{k,l} & \text{si } i \neq j \end{cases}$$

et la famille des $(E_{ij})_{i,j}$ nous définit une base de l'ensemble des matrices symétriques.

L'ensemble \mathcal{G}_d des gaussiennes de dimension d s'identifie avec $\mathbb{R}^d \times S_d^{++}$. On obtient alors une identification avec un ouvert Θ de $\mathbb{R}^{d+\frac{1}{2}d(d+1)}$ en posant

$$\Theta = \left\{ \left(\theta_1, \dots, \theta_{d+\frac{1}{2}d(d+1)} \right) = \left(\mu_1, \dots, \mu_d, (v_{ij})_{1 \leq i \leq j \leq d} \right) \mid \sum_{i,j} v_{ij} E_{ij} \in S_d^{++} \right\} .$$

La proposition suivante nous donne la matrice de Fisher

Proposition 37 ([34])

$$\left(\begin{array}{c|c} V^{-1} & 0 \\ \hline 0 & \left(\frac{1}{2} \text{tr}(V^{-1} E_{ij} V^{-1} E_{kl}) \right)_{(i,j),(k,l)} \end{array} \right) .$$

où l'indice $(i, j), (k, l)$ correspond aux coordonnées (v_{ij}, v_{kl}) .

Démonstration

Nous allons présenter ici une méthode différente du cas 1d pour obtenir ces résultats, basée sur des outils statistiques.

On décompose la matrice de Fisher en quatre blocs (de tailles différentes)

$$\left(\begin{array}{c|c} I_{\mu\mu} & I_{\mu V} \\ \hline I_{\mu V} & I_{VV} \end{array} \right)$$

qui vont être calculés séparément. Nous allons nous contenter de montrer le calcul de $I_{\mu\mu}$.

Soit (X_1, \dots, X_n) un n -échantillon de loi $\mathcal{N}(\mu, V)$. Calculons $\hat{\mu}$ l'estimateur du maximum de vraisemblance (Définition 18, page 136). On pose L la log-vraisemblance. On a

$$L(\mu, V) = - \sum_{i=1}^n \ln(2\pi) + \frac{1}{2} \ln(\det V) + {}^t(X_i - \mu) \frac{V^{-1}}{2} (X_i - \mu) \quad .$$

On souhaite maximiser L par rapport à μ . On a, à V fixé,

$$\nabla_{\mu} L(\mu) = \sum_{i=1}^n \frac{1}{2} 2V^{-1} (X_i - \mu) \quad ,$$

donc

$$\begin{aligned} \nabla_{\mu} L(\theta) = 0 &\iff \sum_{i=1}^n v^{-1} (X_i - \mu) = 0 \\ &\iff V^{-1} \sum_{i=1}^n (X_i - \mu) = 0 \\ &\iff \sum_{i=1}^n (X_i - \mu) = 0 \\ &\iff \mu = \frac{1}{n} \sum_{i=1}^n X_i \quad . \end{aligned}$$

On a donc

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i \quad .$$

C'est un estimateur non-biaisé, on a donc d'après le théorème de Cramer-Rao (Théorème 68, page 138)

$$\text{Var}(\hat{\mu}) \geq \frac{1}{n} I_{\mu\mu}^{-1}$$

où $\text{Var}(\hat{\mu})$ désigne la matrice de variance-covariance et l'inégalité $A \geq B$ est définie par $A - B$ semi définie positive. On en déduit que

$$\text{Var}(\sqrt{n}\hat{\mu}) \geq I_{\mu\mu}^{-1} \quad .$$

L'estimateur du maximum de vraisemblance est asymptotiquement efficace ([33]), c'est à dire qu'il devient efficace (borne de Cramer-Rao atteinte effectivement par

3.4 Distance de Fisher

l'estimateur) de manière asymptotique, donc

$$\lim_{n \rightarrow \infty} \text{Var}(\sqrt{n}\hat{\mu}) = I_{\mu\mu}^{-1} \quad .$$

On a

$$\begin{aligned} \text{Var}(\sqrt{n}\hat{\mu}) &= n \mathbb{E}[(\hat{\mu} - \mathbb{E}[\hat{\mu}])^t (\hat{\mu} - \mathbb{E}[\hat{\mu}])] \\ &= n \mathbb{E}[(\hat{\mu} - \mu)^t (\hat{\mu} - \mu)] \\ &= n \mathbb{E}\left[\left(\frac{1}{n} \sum_i (X_i - \mu)\right)^t \left(\frac{1}{n} \sum_j (X_j - \mu)\right)\right] \\ &= n \frac{1}{n^2} \sum_i \sum_j \mathbb{E}[(X_i - \mu)^t (X_j - \mu)] \\ &= \frac{1}{n} \sum_i \mathbb{E}[(X_i - \mu)^t (X_i - \mu)] \quad \text{car les } X_i \text{ sont indépendants} \\ &= \frac{1}{n} \sum_i \text{Var}[X_i] \\ &= V \quad \text{car } \text{Var}[X_i] = V \quad , \end{aligned}$$

donc $I_{\mu\mu} = V^{-1}$.

Abordons maintenant la question des différences entre le cas des gaussiennes 1d et les gaussiennes multidimensionnel. Alors que dans le cas 1d, la géométrie de l'espace est très bien connue (géodésiques, géodésiques pour les connexions, expression explicite de la distance etc.), dans le cas multidimensionnel très peu de choses peuvent être dites. Cela n'est pas uniquement dû à des calculs plus compliqués, un changement important a lieu dans la géométrie de l'espace. A titre d'exemple, alors qu'en dimension 1, l'espace des gaussiennes est de courbure constante égale à $-\frac{1}{2}$, ici la courbure n'est pas constante ([34]). De manière plus simple, on peut se contenter de regarder l'espace de travail : $\mathbb{R} \times \mathbb{R}^{++}$ dans le cas 1d, $\mathbb{R}^d \times S^{++}$ dans le cas multidimensionnels.

Nous allons cependant voir dans la partie sur les applications (chapitre 5, page 87) une technique pour calculer numériquement la distance entre deux gaussiennes, qui reposera sur le calcul de la géodésique à partir d'une EDO avec des conditions aux limites s'appuyant sur une expression à partir des connexions.

Nous avons traité ici uniquement le cas gaussien. L'application à des gaussiennes

transportées nous a conduit à d'importantes questions théoriques et techniques actuellement sans réponses. Nous n'aborderons donc pas cette question ici. Disons simplement que la méthode la plus naturelle pour le faire est de se placer dans le cadre des familles exponentielles ([2]), définition naturelle pour une gaussienne transportée.

Une seconde question, plus vaste, est celle du cas non-paramétrique. C'est un sujet encore très neuf, on trouve diverses approches dans la littérature pour généraliser la théorie. La plus ancienne à notre connaissance est [1], qui en vient à identifier l'ensemble des densités de probabilités comme un sphère L^2 . Elle s'inscrit dans un ensemble d'approche variant autour de l'idée suivante : considérer le plongement $\rho \mapsto \sqrt{\rho}$ des densités de probabilités dans l'espace L^2 . Une approche basée sur des outils théoriques complexes est proposée dans [31]. On trouvera à la fin de [30] un bref aperçu des différentes approches.

3.5 Distance de Wasserstein pondérée en 1d

Nous avons vu que la distance de Fisher est un outil intéressant car il permet de percevoir aussi bien l'écartement que la concentration des densités, comme l'illustre le cas des gaussiennes 1d (voir section précédente), phénomènes qui ne sont pas perçus simultanément par les distances L^1 et W_p . Néanmoins, à part dans le cas des gaussiennes 1d, et multidimensionnelles en utilisant des outils numériques (Chapitre 5, page 87), c'est un outil malaisé à étudier, notamment dans le contexte du transport.

Nous allons donc partir ici des distances de Wasserstein, adaptées au cadre théorique de l'équation de transport, que nous allons, suivant une approche que l'on peut trouver dans [9] et nous inspirant de la construction du demi-plan de Poincaré, pondérer par une fonctionnelle. Cela nous permettra de percevoir les phénomènes de concentration.

Nous allons commencer par quelques généralités sur de telles distances. Nous l'appliquerons alors à la distance entre deux gaussiennes, une gaussienne transportée et son approximation, et deux gaussiennes transportées.

3.5.1 Généralités sur les distances construites à partir d'une fonctionnelle sur un espace métrique

On considère un espace métrique (X, d) et une fonctionnelle $J : X \rightarrow [0, \infty]$

Définition 9

Soit γ une courbe de X . On appelle, si elle existe, la quantité suivante la dérivée métrique de γ en t

$$|\gamma|'(t) = \lim_{s \rightarrow t} \frac{d(\gamma(s), \gamma(t))}{|s - t|} .$$

On a comme conséquence du théorème de Rademacher la proposition suivante.

Proposition 38 ([3])

si γ est lipschitzienne, alors la dérivée métrique existe presque partout par rapport à la mesure de Lebesgue

Définition 10

Soit $J : X \rightarrow]0, +\infty]$ une fonctionnelle continue, γ une courbe de X définie sur $[a, b]$ admettant une dérivée métrique. On définit la longueur de γ par rapport à J , si elle existe, par

$$l_J(\gamma) = \int_a^b J(\gamma(t)) |\gamma|'(t) dt .$$

Proposition 39 ([9])

On suppose X connexe par arcs lipschitzien. On pose

$$\Gamma_{x,y} = \{ \gamma \mid \gamma \text{ lipschitzienne, } \gamma \text{ d'extrémités } x \text{ et } y \} .$$

Si pour tout x, y il existe une courbe de longueur finie reliant x à y , alors

$$d(x, y) = \inf_{\gamma \in \Gamma} l_J(\gamma)$$

définit une distance sur X .

Remarque 40

L'existence de la distance est possible sous des hypothèses beaucoup plus générales. On pourra par exemple consulter [9] sur ce sujet.

3.5.2 Application aux gaussiennes 1d transportées

Dans la suite, ϕ désignera un \mathcal{C}^2 -difféomorphisme de \mathbb{R} sur \mathbb{R} . On supposera de plus que :

- $\phi' > \varepsilon > 0$;
- ϕ'' bornée.

On notera ψ la réciproque de ϕ .

Soit X_ρ une variable aléatoire réelle de densité ρ . On pose

$$J(\rho) = \frac{1}{\text{Var}X_\rho} \quad . \quad (3.1)$$

c'est cette fonctionnelle J qui servira dans l'applications des outils définis précédemment. On peut déjà intuitivement voir que cette fonctionnelle va être d'autant plus grande que la densité aura une petit variance.

3.5.2.1 L'ensemble des gaussiennes 1d

On considère l'ensemble des gaussiennes 1d \mathcal{G}_1 qu'on identifie avec le demi-plan supérieur. On prend pour d la distance de Wasserstein quadratique. On sait, d'après le Théorème 27 (page 47), que dans ce cas elle s'identifie avec la distance euclidienne. Sur \mathcal{G}_1 , la fonctionnelle J définie en (3.1) nous donne

$$J(\mu, \sigma) = \frac{1}{\sigma^2} \quad .$$

Nous pouvons en déduire la longueur d'une courbe suivant la Définition 10 (page 61). Soit

$$\gamma(t) = \begin{pmatrix} \mu(t) \\ \sigma(t) \end{pmatrix}$$

une courbe lipschitzienne de \mathcal{G}_1 . Comme γ est dérivable presque partout, on a presque partout $|\gamma|'(t) = \|\gamma'(t)\|$. On en déduit que

$$l_J(\gamma) = \int_a^b \frac{1}{\sigma(t)^2} \|\gamma'(t)\| dt \quad .$$

On en déduit que l'espace des gaussiennes 1d muni de la distance de Wasserstein pondéré par la fonctionnelle $J(\mu, \sigma) = \frac{1}{\sigma^2}$ s'identifie au demi-plan de Poincaré (voir section 7.1, page 131). On notera d_g cette distance et on a (Théorème 64, page 133,

3.5 Distance de Wasserstein pondérée en 1d

d'après [35])

$$d_g(g_0, g_1) = \operatorname{arccosh} \left(1 + \frac{(\mu_1 - \mu_0)^2 + (\sigma_1 - \sigma_0)^2}{2\sigma_0\sigma_1} \right) .$$

3.5.2.2 L'ensemble des gaussiennes transportées

On considère \mathcal{G}_1^ϕ l'ensemble des gaussiennes 1d transportées par ϕ , qu'on munit de la distance de Wasserstein quadratique. On note d_ϕ la distance associée à la fonctionnelle J (Proposition 39, page 61).

Théorème 41

Soit $\rho_0 = \phi_{\#}g_0$ et $\rho_1 = \phi_{\#}g_1$ deux gaussiennes transportées. On notera (μ_0, σ_0) et (μ_1, σ_1) leurs paramètres et on supposera les écarts-types assez petits. On a

$$\frac{\inf |\phi'|}{2 \sup |\phi'|^2} d_g(g_0, g_1) \leq d_\phi(\rho_0, \rho_1) \leq 2 \frac{\sup |\phi'|}{\inf |\phi'|^2} d_g(g_0, g_1)$$

Remarque 42

On a donc une formule d'encadrement explicite, puisque l'on a (Théorème 64, page 133) l'expression $d_g(g_0, g_1) = \operatorname{arccosh} \left(1 + \frac{(\mu_1 - \mu_0)^2 + (\sigma_1 - \sigma_0)^2}{2\sigma_0\sigma_1} \right)$

Avant de démontrer ce résultat, commençons par un résultat préliminaire sur le contrôle de la fonctionnelle J .

Lemme 43

soit $\rho = \phi_{\#}g$ une gaussienne transportée, avec g de paramètre (μ, σ) . Si $\sigma \leq K \frac{\inf |\phi'|}{\sup |\phi'|}$ où K est une constante ne dépendant pas des données du problème, alors

$$\frac{1}{2 \sup |\phi'|^2 \sigma^2} \leq \frac{1}{2 \phi'(\mu)^2 \sigma^2} \leq J(\rho) \leq \frac{2}{\phi'(\mu)^2 \sigma^2} \leq \frac{2}{\inf |\phi'|^2 \sigma^2}$$

Démonstration

soit $\rho = \phi_{\#}g$ une gaussienne transportée, avec g de paramètre (μ, σ) . On pose $V = \operatorname{Var} X_\rho$.

Par formule de Taylor, on a $\phi(\mu + \sigma x) = \phi(\mu) + \phi'(\mu)\sigma x + R_2(x)$. De là, on a

que :

$$\begin{aligned}
 V &= \int \phi(\mu + \sigma x)^2 g(x) dx - \left(\int \phi(\mu + \sigma x) g(x) dx \right)^2 \\
 &= \phi(\mu)^2 + (\phi'(\mu)\sigma)^2 + \int R_2(x)^2 g(x) dx + 2\phi(\mu) \int R_2(x)g(x) dx + \\
 &\quad 2\sigma\phi'(\mu) \int xR_2(x)g(x) dx - \phi(\mu)^2 - \left(\int R_2(x)g(x) dx \right)^2 - 2\phi(\mu) \int R_2(x)g(x) dx \\
 &= (\phi'(\mu)\sigma)^2 + 2\sigma\phi'(\mu) \int xR_2g(x) dx + \text{Var}(R_2(x)) \quad .
 \end{aligned}$$

On a :

- $|R_2(x)| \leq \sup |\phi''| \sigma^2$;
- $|\text{Var}(R_2(x))| \leq 2 \sup |\phi''|^2 \sigma^4$.

Et donc

$$|V - (\phi'(\mu)\sigma)^2| \leq C \sup |\phi''| (\phi'(\mu) + \sup |\phi''| \sigma) \sigma^3 \quad .$$

Si $\sigma \leq \frac{\inf \phi'}{\sup |\phi''|}$, on a alors

$$|V - (\phi'(\mu)\sigma)^2| \leq C \sup |\phi''| |\phi'(\mu)| \sigma^3 \quad \text{avec } C = 2C \quad ,$$

et si $\sigma \leq \frac{\inf \phi'}{2C \sup |\phi''|}$, alors on a

$$|V - (\phi'(\mu)\sigma)^2| \leq \frac{1}{2} (\phi'(\mu)\sigma)^2 \quad .$$

Donc

$$\frac{1}{2} (\phi'(\mu)\sigma)^2 \leq V \leq \frac{3}{2} (\phi'(\mu)\sigma)^2 \quad .$$

On en conclut que

$$\frac{2}{(\inf |\phi'| \sigma)^2} \geq J(\rho) \geq \frac{1}{2(\sup |\phi'| \sigma)^2} \quad \text{car } \frac{1}{2} < \frac{2}{3}$$

Passons maintenant à la preuve du théorème.

Démonstration

Soit $\rho_0 = \phi_{\#}g_0$ et $\rho_1 = \phi_{\#}g_1$ deux gaussiennes transportées. On souhaite contrôler leur distance. Commençons par la majoration

Pour $t \in [0, 1]$, on pose g_t la courbe minimisante dans \mathcal{G}_1 entre g_0 et g_1 , et

3.5 Distance de Wasserstein pondérée en 1d

$\rho_t = \phi_{\#} g_t$. On a

$$\begin{aligned}
 d_{\phi}(\rho_0, \rho_1) &\leq \int J(\rho_t) |\rho'| (t) dt \\
 &\leq \int \frac{2}{\inf |\phi'|^2 \sigma_t^2} |\rho'| (t) dt \\
 &\leq \frac{2}{\inf |\phi'|^2} \int \frac{1}{\sigma_t^2} |\rho'| (t) dt \\
 &\leq 2 \frac{\sup |\phi'|}{\inf |\phi'|^2} \int \frac{1}{\sigma_t^2} |g'| (t) dt \\
 &\leq 2 \frac{\sup |\phi'|}{\inf |\phi'|^2} d_g(g_0, g_1) \quad .
 \end{aligned}$$

Passons à la minoration. Soit ρ_t une courbe minimisante. On a

$$\begin{aligned}
 d_{\phi}(\rho_0, \rho_1) &= \int J(\rho_t) |\rho'| (t) dt \\
 &\geq \frac{1}{2 \sup |\phi'|^2} \int \frac{1}{\sigma_t^2} |\rho'| (t) dt \quad (\text{Lemme 43}), \text{ page 63.}
 \end{aligned}$$

Contrôlons la dérivée métrique :

$$\begin{aligned}
 |\rho'| (t) &= \lim_{h \rightarrow 0} \sqrt{\int \left(\frac{\phi(\mu(t+h) + \sigma(t+h)x) - \phi(\mu(t) + \sigma(t)x)}{h} \right)^2 g_0(x) dx} \\
 &= \sqrt{\int [(\mu'(t) + \sigma'(t)x) \phi'((\mu(t) + \sigma(t)x))]^2 g_0(x) dx} \\
 &\geq \inf |\phi'| \sqrt{\int [(\mu'(t) + \sigma'(t)x)]^2 g_0(x) dx} \\
 &\geq \inf |\phi'| |g'| (t) \quad ,
 \end{aligned}$$

donc

$$d_{\phi}(\rho_0, \rho_1) \geq \frac{\inf |\phi'|}{2 \sup |\phi'|^2} \int \frac{1}{\sigma_t^2} |g'| (t) dt \geq \frac{\inf |\phi'|}{2 \sup |\phi'|^2} d_g(g_0, g_1)$$

Remarque 44

Présentons un formalisme différent. On pose \mathcal{P}_2^{ac} le sous-ensemble de $\mathcal{P}_2(\mathbb{R})$ contenant les mesures qui sont de plus absolument continues par rapport à la mesure de Lebesgue. C'est un ouvert de $(\mathcal{P}_2(\mathbb{R}), W_2)$.

En considérant la fonctionnelle J sur \mathcal{P}_2^{ac} , $J: \rho \in \mathcal{P}_2^{ac} \rightarrow \frac{1}{\text{Var} X_{\rho}}$, on peut

définir une distance d sur \mathcal{P}_2^{ac} .

Sur le sous-espace \mathcal{G}_1 de \mathcal{P}_2^{ac} on a $d(g_0, g_1) \leq \operatorname{arccosh}\left(1 + \frac{(\mu_1 - \mu_0)^2 + (\sigma_1 - \sigma_0)^2}{2\sigma_0\sigma_1}\right)$. En effet, le plus court chemin dans \mathcal{P}_2^{ac} est plus court que celui dans \mathcal{G}_1 , qui nous donne la distance induite par le demi-plan de Poincaré (Théorème 64, page 133).

Sur \mathcal{G}_1^ϕ , suivant la démonstration du Théorème 41 on perd la minoration mais la majoration $d(\rho_0, \rho_1) \leq 2 \frac{\sup |\phi'|}{\inf |\phi'|^2} d(g_0, g_1)$ reste valable.

La distance construite nous permet de retrouver les caractéristiques de la distance de Fisher qui nous intéressaient, mais dans un cadre propice au transport, nous permettant d'obtenir naturellement des estimations entre deux gaussiennes transportées. Notons d'ailleurs que la preuve du Théorème 41 est valable pour des densités de \mathcal{X}_1 et que nous l'avons formulé pour les gaussiennes uniquement pour des raisons de continuité avec la présentation de la distance de Fisher permettant la comparaison entre les deux distances.

Concernant l'extension au cas multidimensionnel, la difficulté n'est pas technique, les outils employés en 1d s'utilisant de manière similaire en dimension supérieure. Elle est dans le choix de la fonctionnelle J permettant de saisir les phénomènes de concentration. Des essais ont été menés, dans le cas gaussien, avec la fonctionnelle $J(\mu, V) = \frac{1}{\det V}$ mais ne sont pas satisfaisants.

3.6 conclusion du chapitre

Nous avons vu au chapitre précédent que l'utilisation de la distance L^1 pour l'étude de l'écart entre une solution concentrée et son approximation (Théorème 8, page 26) soulève des interrogations. Nous avons mentionné en introduction que dans l'étude de phénomènes physiques (écart entre des polluants), elle n'est pas toujours satisfaisante, la question plus générale étant le choix d'une distance suivant les phénomènes à étudier.

Nous avons commencé par introduire la distance de Wasserstein. Naturelle dans le contexte théorique des équations de transport, elle est satisfaisante tant que le contexte de concentration joue de manière marginale.

Dans le cas contraire, nous nous sommes intéressés à la distance de Fisher. C'est un outil permettant de percevoir des phénomènes nous intéressant pour l'étude des

3.6 conclusion du chapitre

densités concentrées utilisées au chapitre précédent (densité initiale (2.3), page 22 et solution exacte (2.4), page 22). Il reste néanmoins encore assez peu connu théoriquement et il n'est en l'état pas possible de l'utiliser en dehors des gaussiennes.

Néanmoins, en s'inspirant du demi-plan de Poincaré que nous donne la métrique de Fisher, il est possible de construire à partir de la distance de Wasserstein une distance qui est en même temps adapté au cadre théorique du transport, et permet de percevoir les phénomènes de concentration si ces derniers sont d'importances.

CHAPITRE 3 : *Comparaison de distances*

Chapitre 4

Courbes

4.1 Rappel sur les courbes et notations

Dans la suite, on appellera **arc** du plan de classe \mathcal{C}^k un couple (I, γ) où I est un intervalle de \mathbb{R} et $\gamma \in \mathcal{C}^k(I, \mathbb{R}^2)$. Par abus de notation, on fera fréquemment référence à un arc sans préciser l'intervalle. On appellera **courbe** l'image d'un arc. On emploiera parfois le mot **interface** pour désigner une courbe et fréquemment le mot **paramétrage** comme synonyme d'arc. On utilisera des lettres minuscules (typiquement γ) pour désigner un arc, et des majuscules (typiquement Γ) pour désigner la courbe associée.

On rappelle qu'on a une relation d'équivalence sur les arcs \mathcal{C}^k donnée par

$$(I, \gamma) \sim (J, \delta) \iff \text{il existe } \theta \text{ un } \mathcal{C}^k\text{-difféo de } I \text{ sur } J \text{ tel que } \gamma = \delta \circ \theta \quad .$$

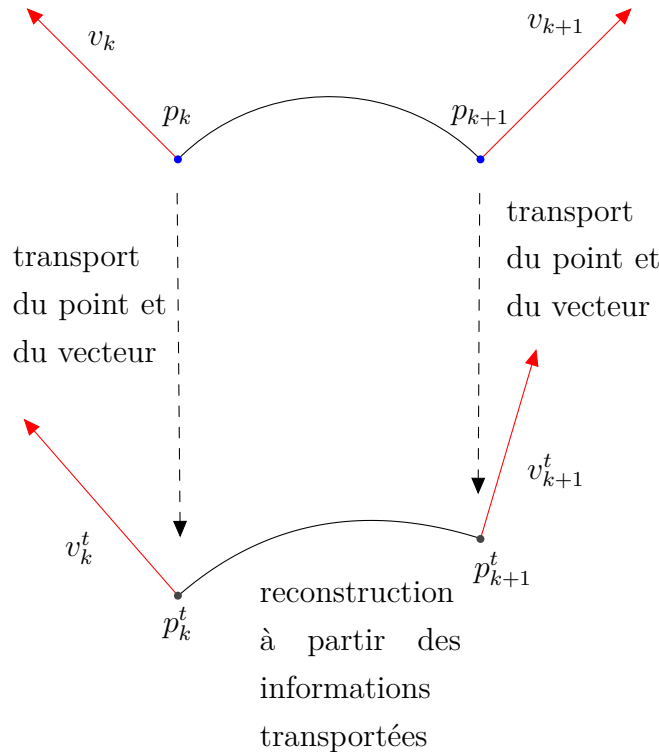
Les éléments d'une classe d'équivalence ont pour image la même courbe, mais la réciproque est fautive : une courbe peut admettre des paramétrages dans différentes classes d'équivalence, considérer par exemple l'axe des abscisses paramétré par $[s, 0]$ ou par $[s^3, 0]$. Suivant [6], on appellera **propriété géométrique** une propriété ne dépendant que de la classe d'équivalence choisie et pas du représentant choisi.

On dit qu'un point est singulier si $\gamma'(s) = 0$. C'est une propriété géométrique. On rappelle qu'il y a quatre sous cas possibles : point d'inflexion (la courbe traverse sa tangente), point ordinaire, point de rebroussement de première espèce, point de rebroussement de deuxième espèce.

4.2 Introduction et remarques préalables

4.2.1 Introduction

Nous considérons une courbe $\Gamma \subset \mathbb{R}^2$ et ϕ un \mathcal{C}^∞ -difféomorphisme. On pose $\Gamma_t = \phi(\Gamma)$ la transportée de Γ par ϕ . On souhaite calculer numériquement Γ_t . Nous commençons pour cela par nous donner une discrétisation $(p_k)_k$ de Γ que nous allons transporter par ϕ . On posera $p_k^t = \phi(p_k)$. Suivant l'esprit de la méthode introduite au chapitre 2, nous allons en utilisant la différentielle de ϕ , transporter un vecteur v_k associée à chaque point permettant une meilleure reconstruction de la courbe. Nous appellerons v_k^t ce vecteur transporté.



Dans un premier temps la courbe Γ sera discrétisée en associant à chaque point p_k le vecteur normal unitaire extérieur n_k . Nous montrerons que nous avons alors une unique reconstruction de Γ_t par une courbe paramétrée par morceaux par des polynômes P_k de degré 2 entre p_k^t et p_{k+1}^t , et nous présenterons une méthode raffinant la discrétisation nécessaire à l'efficacité de la méthode.

Nous nous intéresserons ensuite au cas où Γ est paramétrée par un arc γ . Nous effectuerons alors une discrétisation au moyen de splines cubiques d'Hermite, et nous transporterons les vecteurs tangents.

4.3 Utilisation de la normale et reconstruction de degré 2

On trouvera dans le chapitre 5 une application à un problème de Stokes diphasique.

4.2.2 Points singuliers de Γ_t

On suppose que Γ est paramétrée par γ . $\gamma_t = \phi \circ \gamma$ est un paramétrage de Γ_t . On a alors la proposition suivante

Proposition 45

$\gamma'_t(s)$ est un point singulier si et seulement si $\gamma'(s)$ est un point singulier. De plus, ils sont du même type.

Démonstration

$$\begin{aligned}\gamma'_t(s) = 0 &\iff d\phi_{\gamma(s)}\gamma'(s) = 0 \\ &\iff \gamma'(s) = 0 \quad \text{car } d\phi \text{ est un difféomorphisme} \quad .\end{aligned}$$

Intéressons maintenant à la nature des points singuliers. Soit $\gamma(s)$ un point singulier. Comme $\gamma'(s) = 0$ on a $\gamma_t^{(k)}(s) = d\phi_{\gamma(s)}\gamma^{(k)}(s)$. De là on en déduit immédiatement que les points singuliers ont la même nature.

4.3 Utilisation de la normale et reconstruction de degré 2

Nous considérons ici le cas où à chaque point p_k est associé un vecteur normal unitaire. Afin de pouvoir donner une orientation, il sera considéré extérieur. Après quelques remarques préliminaires, Nous allons montrer que nous pouvons dans ce cas reconstruire de manière unique une courbe paramétrée localement par des polynômes de degré 2.

4.3.1 Remarques préliminaires

Soit

$$P(t) = t^2 \begin{pmatrix} a_x \\ a_y \end{pmatrix} + t \begin{pmatrix} b_x \\ b_y \end{pmatrix} + \begin{pmatrix} c_x \\ c_y \end{pmatrix}$$

un polynôme de degré 2. Il existe t tel que $P'(t) = 0$, si et seulement si

$$\frac{b_x}{a_x} = \frac{b_y}{a_y}$$

et donc si et seulement si a et b sont colinéaires ou $b = 0$. Donc

$$P(t) = (t^2 + \lambda t) \begin{pmatrix} a_x \\ a_y \end{pmatrix} + \begin{pmatrix} c_x \\ c_y \end{pmatrix}$$

et le polynôme paramètre une droite. Ainsi, P a un point singulier si et seulement s'il paramètre une droite et c'est alors un point ordinaire. On a $P''(t) = 2a$ qui est non nul, donc il ne peut y avoir de point d'inflexion.

Abordons maintenant la question du transport du vecteur normal. Soit ϕ un difféomorphisme. On définit les transportés de p_0, p_1 par $p'_0 = \phi(p_0)$, $p'_1 = \phi(p_1)$. On souhaite transporter conjointement les vecteurs normaux n_0 et n_1 . En appliquant la définition du transport d'un vecteur vu au chapitre 2, on a $n'_0 = d\phi_{p_0}n_0$, $n'_1 = d\phi_{p_1}n_1$. Cependant, la définition préserve les vecteurs tangents et rien ne garantit qu'elle préserve les vecteurs normaux. En fait, la condition naturelle pour préserver un vecteur normal est même assez forte comme le montre la proposition suivante

Proposition 46

Soit u et v deux vecteurs orthogonaux du plan basé en point p quelconque. On pose u_t et v_t les transportés de u et v . u_t et v_t sont orthogonaux indépendamment du choix de p si et seulement si $d\phi$ est la composé d'un endomorphisme orthogonal et d'une homotéthie.

Démonstration

$$\begin{aligned} \langle u_t, v_t \rangle = 0 &\iff \langle d\phi_p u, d\phi_p v \rangle = 0 \\ &\iff \langle {}^t d\phi_p d\phi_p u, v \rangle = 0 \\ &\iff {}^t d\phi_p d\phi_p = \lambda \text{Id}, \lambda \neq 0 \end{aligned}$$

On peut cependant très simplement effectuer l'advection en repassant par un vecteur tangent :

- appliquer une rotation d'angle $\frac{\pi}{2}$ au vecteur n_k pour obtenir un vecteur tangent v_k ;
- transporter v_k . On obtient un vecteur v_k^t qui est tangent ;

4.3 Utilisation de la normale et reconstruction de degré 2

- appliquer une rotation d'angle $-\frac{\pi}{2}$ à v_k^t pour réobtenir le vecteur normal unitaire extérieur n_k^t .

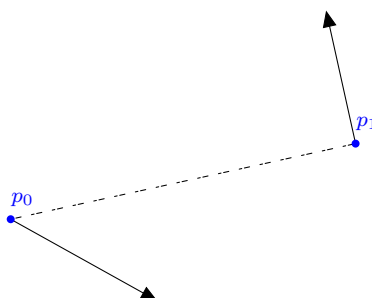
4.3.2 Existence et unicité du polynôme d'interpolation

Théorème 47

Soit p_0 et p_1 deux points du plan, $n_0 \neq 0$ et $n_1 \neq 0$ deux vecteurs non colinéaires tel que $(p_0 - p_1) \cdot n_i \neq 0$. Il existe une unique courbe passant par p_0 et p_1 , normales en p_0 et p_1 à n_0 et n_1 , admettant un paramétrage par un polynôme de degré 2.

Remarque 48

Commençons par une remarque préalable qui permettra de mieux saisir la démonstration, remarque concernant l'hypothèse $(p_0 - p_1) \cdot n_i \neq 0$. Considérons la figure suivante présentant une configuration de ce type



Une courbe paramétrée par un polynôme de degré 2 ne pouvant avoir de point d'inflexion, la seule façon de relier p_0 à p_1 par un polynôme de degré 2 est d'avoir pour image une ligne droite, ne satisfaisant pas les conditions requises sur les vecteurs normaux.

Démonstration

On pose

$$p_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \quad p_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \quad n_0 = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}, \quad n_1 = \begin{pmatrix} u_1 \\ v_1 \end{pmatrix}.$$

Soit t_0 et t_1 deux réels tel que $t_0 < t_1$. On pose $dt = t_1 - t_0$. Soit $Q(t) = (t - t_0)^2 a + (t - t_0)b + c$ un polynôme à valeurs dans le plan. Dire que Q a pour image une courbe tel que défini dans le théorème, cela implique :

- $Q(t_0) = p_0$;
- $Q(t_1) = p_1$;
- $Q'(t_0) \cdot n_0 = 0$;
- $Q'(t_1) \cdot n_1 = 0$.

Notons que nous n'avons pas équivalence. Il est en effet possible d'avoir $Q'(t_i) = 0$, donc $Q'(t_i) \cdot n_i = 0$ sans pour autant que le vecteur soit normal à la courbe (cf Remarque 48).

Nous déduisons des conditions le système suivant

$$\begin{pmatrix} dt^2 & dt & 0 & 0 & 1 & 0 \\ 0 & 0 & dt^2 & dt & 0 & 1 \\ 0 & u_0 & 0 & v_0 & 0 & 0 \\ 2dtu_1 & u_1 & 2d tv_1 & v_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_x \\ b_x \\ a_y \\ b_y \\ c_x \\ c_y \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \\ 0 \\ 0 \\ x_0 \\ y_0 \end{pmatrix} .$$

Notons M la matrice associé au système. On a $\det(M) = -dt^4(u_0v_1 - v_0u_1)$. Le déterminant est nul si et seulement si les vecteurs n_0 et n_1 sont colinéaires, ce qui n'est pas le cas. Il y a donc une unique solution. Après résolution, on obtient que

$$a_x = -\frac{(u_0v_1x_0 + u_1v_0x_0 - u_0v_1x_1 - u_1v_0x_1 + 2v_0v_1y_0 - 2v_0v_1y_1)}{(dt^2(u_0v_1 - u_1v_0))} ,$$

$$a_y = \frac{(2u_0u_1x_0 - 2u_0u_1x_1 + u_0v_1y_0 + u_1v_0y_0 - u_0v_1y_1 - u_1v_0y_1)}{(dt^2(u_0v_1 - u_1v_0))} ,$$

$$b_x = \frac{(2v_0(u_1x_0 - u_1x_1 + v_1y_0 - v_1y_1))}{(dt(u_0v_1 - u_1v_0))} ,$$

$$b_y = -\frac{(2u_0(u_1x_0 - u_1x_1 + v_1y_0 - v_1y_1))}{(dt(u_0v_1 - u_1v_0))} ,$$

et $c = p_0$.

Prouvons maintenant l'unicité. Pour simplifier on suppose que $t_0 = 0$. Soit s_1 un réel strictement positif. On note P le polynôme associé à l'intervalle $[0, s_1]$. Vérifions que P et Q ont le même graphe. On a $Q(t) = t^2a + tb + c$ et $P(s) = s^2a' + sb' + c'$.

4.3 Utilisation de la normale et reconstruction de degré 2

On pose $l = \frac{t_1}{s_1}$. On alors une bijection de $[0, s_1]$ sur $[0, t_1]$ définie par ls . On a

$$\begin{aligned} Q(t) &= t^2a + tb + c \\ &= (ls)^2a + lsb + c \\ &= s^2l^2a + slb + c \quad . \end{aligned}$$

Il faut maintenant vérifier que $l^2a = a', lb = b', c' = c$. Par résolution du système, on a :

$$a_x = -\frac{(u_0v_1x_0 + u_1v_0x_0 - u_0v_1x_1 - u_1v_0x_1 + 2v_0v_1y_0 - 2v_0v_1y_1)}{(t_1^2(u_0v_1 - u_1v_0))} \quad ,$$

$$a_y = \frac{(2u_0u_1x_0 - 2u_0u_1x_1 + u_0v_1y_0 + u_1v_0y_0 - u_0v_1y_1 - u_1v_0y_1)}{(t_1^2(u_0v_1 - u_1v_0))} \quad ,$$

et

$$a'_x = -\frac{(u_0v_1x_0 + u_1v_0x_0 - u_0v_1x_1 - u_1v_0x_1 + 2v_0v_1y_0 - 2v_0v_1y_1)}{(s_1^2(u_0v_1 - u_1v_0))} \quad ,$$

$$a'_y = \frac{(2u_0u_1x_0 - 2u_0u_1x_1 + u_0v_1y_0 + u_1v_0y_0 - u_0v_1y_1 - u_1v_0y_1)}{(s_1^2(u_0v_1 - u_1v_0))} \quad .$$

On a donc bien $l^2a = a'$. De même on a

$$b_x = \frac{(2v_0(u_1x_0 - u_1x_1 + v_1y_0 - v_1y_1))}{(t_1(u_0v_1 - u_1v_0))} \quad ,$$

$$b_y = -\frac{(2u_0(u_1x_0 - u_1x_1 + v_1y_0 - v_1y_1))}{(t_1(u_0v_1 - u_1v_0))} \quad ,$$

et

$$b'_x = \frac{(2v_0(u_1x_0 - u_1x_1 + v_1y_0 - v_1y_1))}{(s_1(u_0v_1 - u_1v_0))} \quad ,$$

$$b'_y = -\frac{(2u_0(u_1x_0 - u_1x_1 + v_1y_0 - v_1y_1))}{(s_1(u_0v_1 - u_1v_0))} \quad .$$

On a donc bien $lb = b'$. Et enfin on a immédiatement $c = c'$.

Reste maintenant à vérifier que n_0 et n_1 sont bien des vecteurs normaux. On a vu que c'était le cas si $Q'(t_0) \neq 0$ et $Q'(t_1) \neq 0$. Traitons le cas $Q'(t_0) \neq 0$.

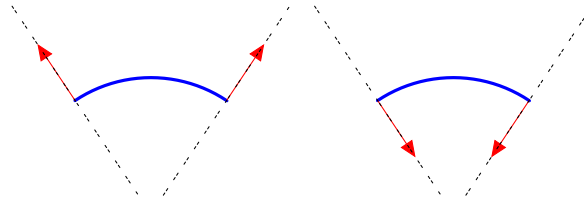
$$\begin{aligned} Q'(t_0) = 0 &\iff b = 0 \\ &\iff u_1x_0 - u_1x_1 + v_1y_0 - v_1y_1 = 0 \quad \text{car } (u_0, v_0) \neq (0, 0) \\ &\iff (p_0 - p_1) \cdot n_1 = 0 \quad . \end{aligned}$$

Par hypothèse, $(p_0 - p_1) \cdot n_1 \neq 0$ donc $Q'(t_0) \neq 0$ et n_0 est bien normal à la courbe

en p_0 . On traite de la même façon l'extrémité p_1 .

Remarque 49

L'interpolation ne nécessite pas un vecteur, mais simplement d'une direction. Comme l'illustre le schéma ci dessous, remplacer la normale par son opposé ne change rien.



Par contre, l'utilisation d'un vecteur permet de se fixer une orientation, permettant de distinguer l'intérieur de l'extérieur de la courbe lorsque celle-ci est fermée.

4.3.3 Transport numérique et discrétisation adaptative

Nous allons présenter une méthode de raffinement de la discrétisation permettant une meilleure approximation de l'interface transportée. Elle reposera sur la détection d'apparition de points d'inflexion sur la courbe exacte. Nous commencerons par aborder ce point au travers d'un exemple.

4.3.3.1 Apparition de points d'inflexion

On considère comme interface initiale le cercle de centre $(2, 2)$ de rayon 2, qu'on souhaite advecter par le difféomorphisme

$$\phi(x, y) = \begin{pmatrix} x \cos(\arctan(x^2 + y^2)) - y \sin(\arctan(x^2 + y^2)) \\ x \sin(\arctan(x^2 + y^2)) + y \cos(\arctan(x^2 + y^2)) \end{pmatrix} .$$

C'est le flot pris en $t = 1$ du champ de vecteur

$$\arctan(x^2 + y^2) \begin{pmatrix} -y \\ x \end{pmatrix} .$$

On discrétise l'interface initiale avec 22 points équirépartis et les normales correspondantes. On advecte ensuite exactement et avec notre méthode.

4.3 Utilisation de la normale et reconstruction de degré 2

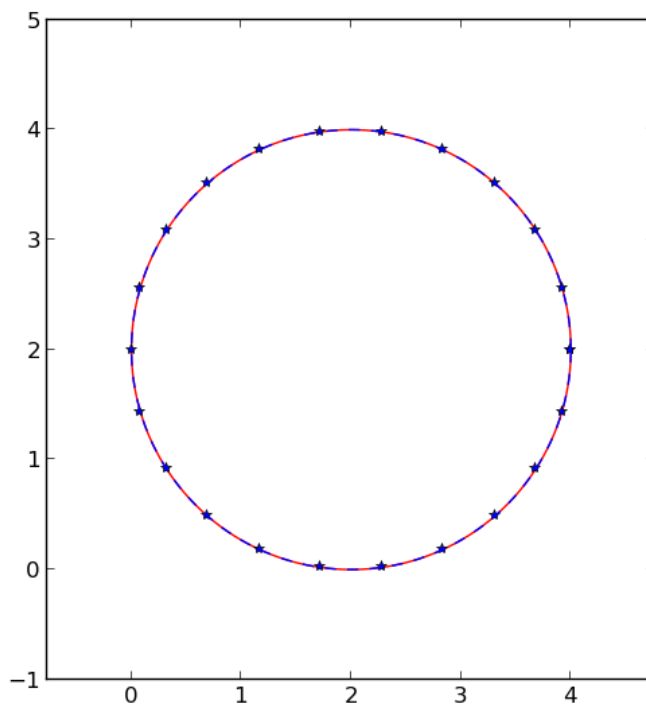


FIGURE 4.1 Interface initiale, exacte (rouge) et approchée (bleu)

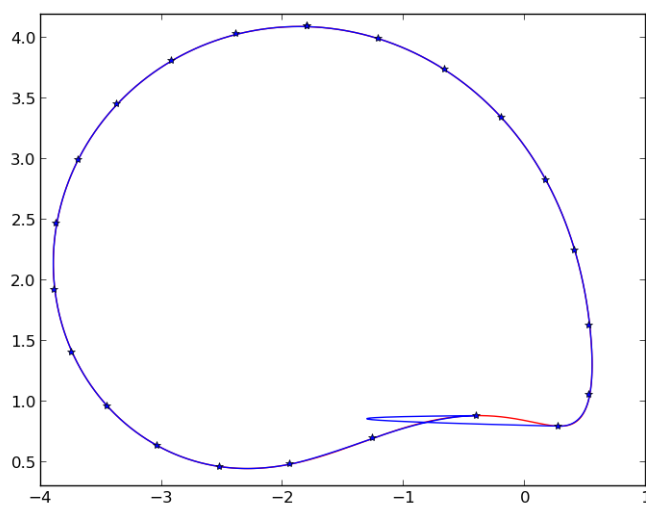


FIGURE 4.2 Interface advectée, exacte (rouge) et approchée (bleu)

Il y a apparition d'un point d'inflexion sur la courbe exacte que ne peut capter un polynôme de degré 2, entraînant l'apparition d'un point de rebroussement sur

la courbe approchée. Rajoutons, suivant ce qui a été abordé précédemment (voir Remarque 49, page 76), que l'interpolation ne se préoccupant pas de l'orientation, on a ici perte de l'orientation ce qui pose un problème pour détecter l'intérieur de l'extérieur.

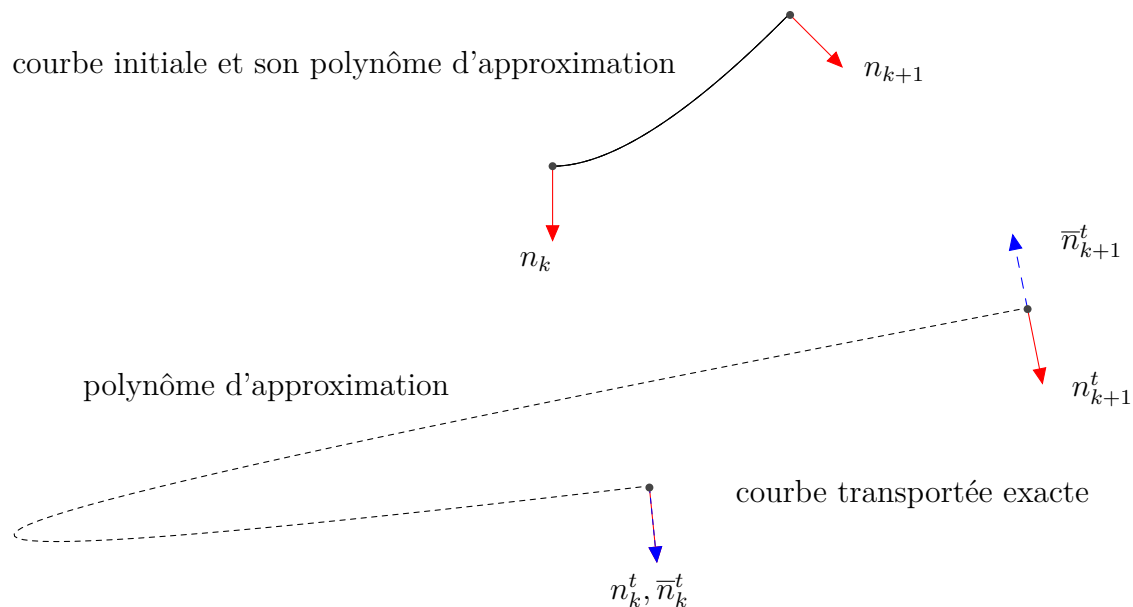
A partir de cet exemple, nous pouvons introduire la technique de raffinement.

4.3.3.2 Une technique de raffinement

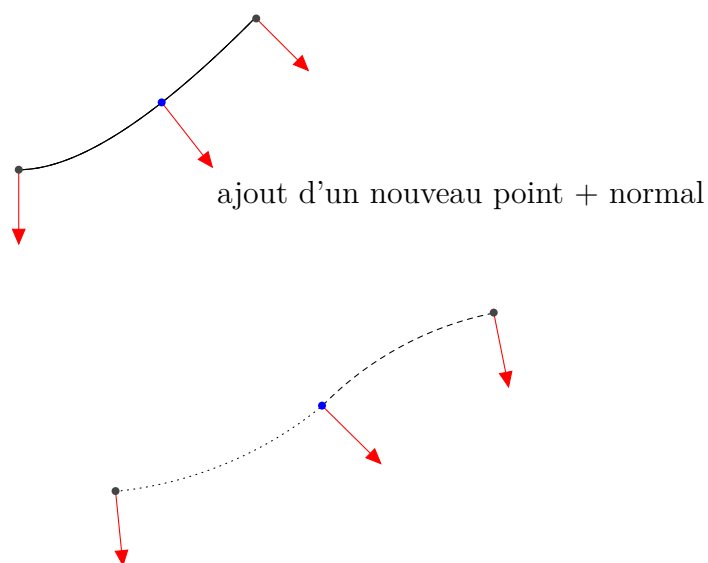
Nous allons ici présenter une méthode simple de raffinement. On trouvera un schéma explicatif sur la page suivante :

1. transporter numériquement les points p_k et les vecteurs n_k . On appellera p_k^t et n_k^t les transportés ;
2. calculer les coefficients des polynomes ;
3. en déduire les vecteurs normaux extérieurs à partir des coefficients polynomiaux (rotation d'angle $\frac{\pi}{2}$ des vecteurs tangents recalculés). On appellera \bar{n}_k^t ces vecteurs ;
4. détecter les changements de sens en calculant le produit scalaire $\langle n_k^t, \bar{n}_k^t \rangle$. S'il est négatif, il y a création d'un point de rebroussement sur la courbe approchée ;
5. pour les sections avec points de rebroussement, rajouter sur la section correspondante de la courbe initiale un point et une normale et les advecter pour raffiner la zone.

4.3 Utilisation de la normale et reconstruction de degré 2



Le produit scalaire $\langle n_{k+1}^t, \bar{n}_{k+1}^t \rangle$ est négatif, on rajoute donc un point.



Pour illustrer la méthode, on reprend l'exemple de la section 4.3.3.1

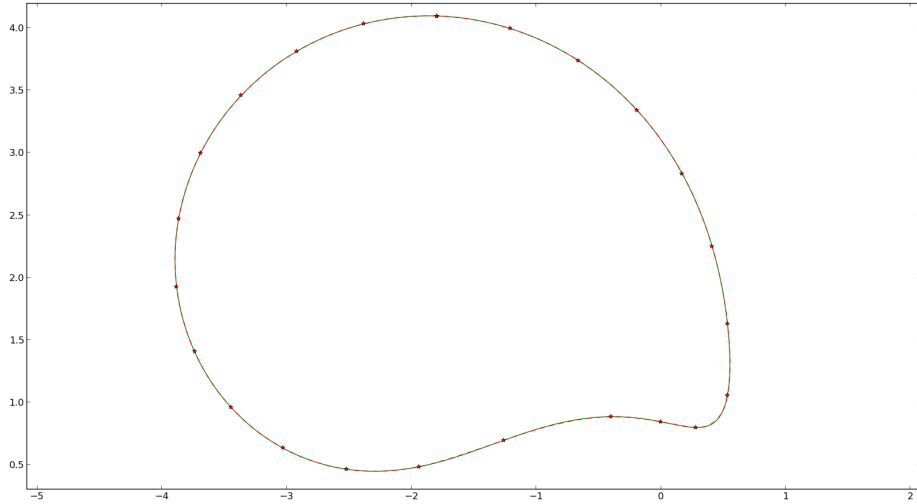


FIGURE 4.3 Interface advectée, exacte (rouge) et approchée (verte)

Dans le cas présent, l'algorithme a bien détecté la zone problématique et capté l'inflexion.

4.4 Utilisation des splines cubiques

4.4.1 Notations

Soit s_i, s_f deux réels tel que $s_i < s_f$, $\delta = s_f - s_i$. On pose :

- $h_{00}(s) = 2s^3 - 3s^2 + 1$;
- $h_{10}(s) = s^3 - 2s^2 + s$;
- $h_{01}(s) = -2s^3 + 3s^2$;
- $h_{11}(s) = s^3 - s^2$.

L'unique spline cubique reliant p_0 et p_1 avec pour tangentes aux extrémités v_0 et v_1 est donnée par

$$P(t) = h_{00}\left(\frac{s-t_i}{\delta}\right)p_0 + \delta h_{01}\left(\frac{s-t_i}{\delta}\right)v_0 + h_{01}\left(\frac{s-t_i}{\delta}\right)p_1 + \delta h_{11}\left(\frac{s-t_i}{\delta}\right)v_1 \quad .$$

On écrira parfois P_{p_0, p_1, v_0, v_1} quand on veut faire référence explicitement aux paramètres de la spline.

4.4 Utilisation des splines cubiques

4.4.2 Principe

On considère une courbe Γ . On la suppose paramétrée par un arc γ sur un intervalle I . On effectue une approximation par spline cubique d'hermite : on choisit une discrétisation $(s_k)_k$ de I qui donne une discrétisation $(p_k)_k$ pour les points de la courbe, avec $p_k = \gamma(s_k)$ et une discrétisation $(v_k)_k$ pour les tangentes, avec $v_k = \gamma'(s_k)$. On adveccte les points et les tangentes en posant $p_k^t = \phi(p_k)$ et $v_k^t = d\phi_{p_k} v_k$. On effectue alors une approximation par spline cubique de l'arc $\gamma^t = \phi \circ \gamma$ défini sur l'intervalle I en posant $\bar{\gamma}^t(s) = P_{p_k^t, p_{k+1}^t, v_k^t, v_{k+1}^t}(s)$ si $s \in [s_k, s_{k+1}]$. On obtient alors une approximation de la courbe $\Gamma^t = \phi(\Gamma)$ en considérant l'image de $\bar{\gamma}^t$.

4.4.3 Dépendance en fonction du paramétrage

Notons que la situation dépend du paramétrage choisi pour la courbe initiale. Supposons que l'on souhaite approcher le quart supérieur droit du cercle unité. On peut le paramétrer sur l'intervalle $[0, 1]$ avec $\gamma(s) = (\cos(\pi/2s), \sin(\pi/2s))$, mais aussi avec $\alpha(s) = (\cos(\pi/4(s^3 + s)), \sin(\pi/4(s^3 + s)))$. On obtient alors deux approximations différentes de la courbe, comme nous le montre la figure 4.4

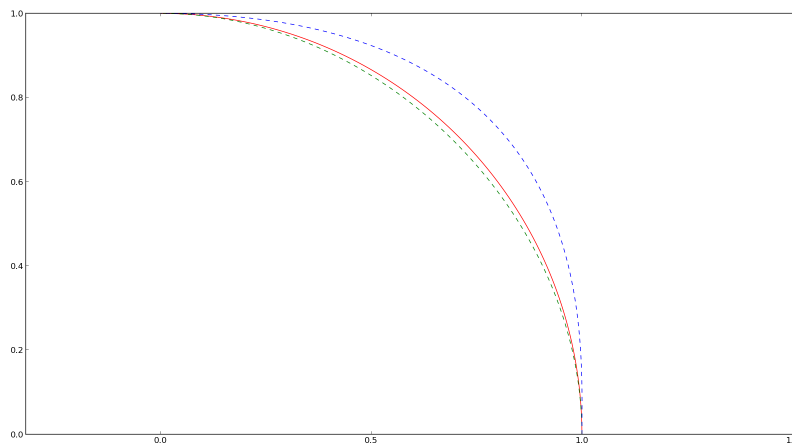


FIGURE 4.4 la même interface (traits pleins) avec deux paramétrages différents sur le même intervalle donne deux approximations différentes (pointillés)

4.4.4 Cas test

On reprend l'exemple de la section 4.3.3.1. On considère comme interface initiale le cercle de centre $(2, 2)$ qu'on paramétrise par $\gamma(s) = (2 \cos(s) + 2, \sin(s) + 2)$ sur $[0, 2\pi]$ et on effectue une discrétisation uniforme avec 12 points. On considère comme précédemment le difféomorphisme

$$\phi(x, y) = \begin{pmatrix} x \cos(x^2 + y^2) - y \sin(x^2 + y^2) \\ x \sin(x^2 + y^2) + y \cos(x^2 + y^2) \end{pmatrix}$$

Le résultat de la simulation est visible sur la figure 4.6. Avec peu de points, on

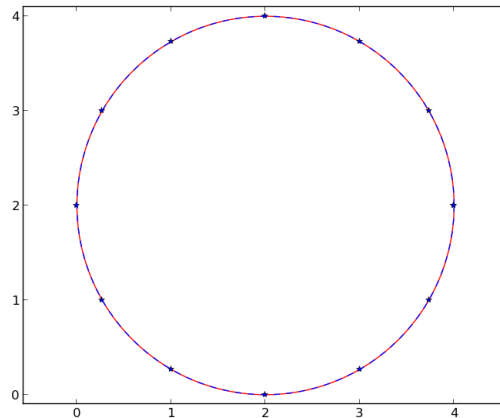


FIGURE 4.5 Interface initiale, exacte (rouge) et approchée (bleu)

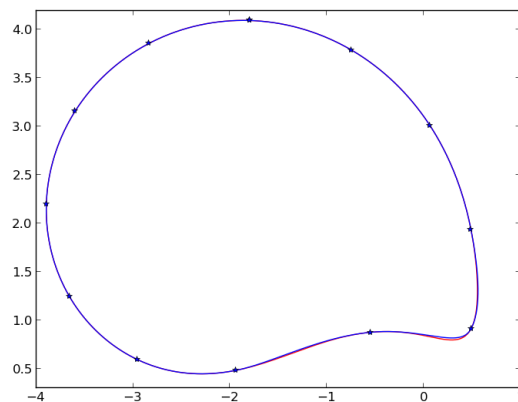


FIGURE 4.6 Interface advectée, exacte (rouge) et approchée (bleu)

obtient une approximation de bonne qualité. La situation est plus robuste qu'avec les polynômes de degré 2.

4.4.5 Erreur géométrique

Nous partons ici de l'erreur entre le paramétrage et son approximation par une spline cubique (erreur entre deux fonctions) pour en faire sortir des quantités géométriques (écarts entre les points de la courbe). Cela permettra de mieux comprendre les problèmes soulevés dans la section 4.4.3 (page 81).

Soit γ un arc défini sur $[s_i, s_f]$, paramétrant une courbe Γ . On note p_0 et p_1 ses extrémités, m_0 et m_1 les tangentes de γ en ces extrémités. On pose $\delta = s_f - s_i$. Soit ϕ un difféomorphisme du plan. On définit la transporté de γ par $\gamma_t = \phi(\gamma)$. On note Γ_t la courbe paramétrée par γ_t .

On peut maintenant définir une approximation de γ_t au moyen d'une spline cubique paramétrée sur $[s_1, s_f]$, d'extrémités $p_0^t = \phi(p_0)$ et $p_1^t = \phi(p_1)$, de tangentes respectives $m_0^t = d\phi_{p_0}m_0$ et $m_1^t = d\phi_{p_1}m_1$. On notera $\bar{\gamma}_t$ cette spline et $\bar{\Gamma}_t$ la courbe qu'elle paramètre.

Théorème 50

L'erreur entre γ_t et son approximation $\bar{\gamma}_t$ admet la majoration suivante

$$\begin{aligned} \|\gamma_t - \bar{\gamma}_t\|_\infty &\leq C(d\phi, d^2\phi, d^3\phi, d^4\phi) \times \\ &\quad \left(\sum_{k=2}^4 \|p_1 - p_0\|^k + \sum_{k=2}^4 \delta^k (\|m_0\| + \|m_1\|)^k + \right. \\ &\quad \left. \left(\|p_1 - p_0\|^2 \delta (\|m_0\| + \|m_1\|) \right) + \left(\|p_1 - p_0\| \delta^2 (\|m_0\| + \|m_1\|)^2 \right) \right) \\ &\quad + (\|p_1 - p_0\| \delta (\|m_0\| + \|m_1\|)) \end{aligned}$$

Remarque 51

Effectuons un changement de coordonnées affine en considérant un paramétrage par $\gamma(as + b)$. Cela nous donne une nouvelle spline cubique d'approximation mais paramétrant toujours $\bar{\Gamma}_t$. La majoration de l'erreur dans le théorème ne change pas, la dilation des vecteurs tangents étant compensée par les termes en δ .

Il est possible de compliquer les choses en considérant $\gamma(\theta(s))$ où θ est un difféomorphisme sur $[s_i, s_f]$. On obtient alors une spline cubique $\tilde{\gamma}_t$ paramétrant une nouvelle courbe $\tilde{\Gamma}_t$, qui en toute généralité est distincte de $\bar{\Gamma}_t$. On peut alors montrer que l'erreur entre γ_t et $\tilde{\gamma}_t$ est gouvernée par l'écart entre les sommets (termes en $\|p_1 - p_0\|$) et la longueur des tangentes (termes en $(\|m_0\| + \|m_1\|)$) pondéré par δ , mais avec une influence du max des dérivées de θ .

Démonstration

$\bar{\gamma}_t$ étant l'approximation par spline cubique de γ^t , on a par propriétés des splines cubiques ([14]) que

$$\|\gamma_t - \bar{\gamma}_t\|_\infty \leq \|d^4\gamma_t\|_\infty \delta^4 \quad .$$

Étudions le terme $\|d^4\gamma_t\|_\infty$. On a $d^4\gamma_t = d^4(\phi \circ \gamma)$. Calculons successivement les différentielles de γ_t .

$$d(\phi \circ \gamma)_s = d\phi_\gamma \gamma' \quad ,$$

$$d^2(\phi \circ \gamma)_s = d^2\phi_\gamma (\gamma')^2 + d\phi_\gamma \gamma'' \quad ,$$

$$\begin{aligned} d^3(\phi \circ \gamma)_s &= d^3\phi_\gamma (\gamma')^3 + 2d^2\phi_\gamma \gamma' \gamma'' + d^2\phi_\gamma \gamma' \gamma'' + d\phi_\gamma \gamma''' \\ &= d^3\phi_\gamma (\gamma')^3 + 3d^2\phi_\gamma \gamma' \gamma'' + d\phi_\gamma \gamma''' \quad , \end{aligned}$$

$$\begin{aligned} d^4(\phi \circ \gamma)_s &= d^4\phi_\gamma (\gamma')^4 + 6d^3\phi_\gamma (\gamma')^2 \gamma'' \\ &\quad + 3d^3\phi_\gamma (\gamma'')^2 + 4d^2\phi_\gamma \gamma' \gamma''' \\ &\quad + d\phi_\gamma \underbrace{\gamma''''}_{=0} \quad . \end{aligned}$$

Dans la suite, on posera $r(s) = \frac{s - s_i}{\delta} \in [0, 1]$. On rappelle qu'on a :

- $h_{00}(r) = 2r^3 - 3r^2 + 1$;
- $h_{10}(r) = r^3 - 2r^2 + r$;
- $h_{01}(r) = -2r^3 + 3r^2$;
- $h_{11}(r) = r^3 - r^2$.

4.4 Utilisation des splines cubiques

Contrôlons maintenant les dérivées de γ . On a

$$\begin{aligned}
\|\gamma'(s)\| &\leq \|(h_{00}(r)p_0 + h_{01}(r)p_1 + \delta(h_{10}(r)m_0 + h_{11}(r)m_1))'(s)\| \\
&\leq \frac{1}{\delta} \|h'_{00}(r)p_0 + h'_{01}(r)p_1 + \delta(h'_{10}(r)m_0 + h'_{11}(r)m_1)\| \\
&\leq \frac{1}{\delta} \|(6r^2 - r)(p_0 - p_1) + \delta((3r^2 - 4r + 1)m_0 + (3r^2 - 2r)m_1)\| \\
&\leq \frac{C}{\delta} \|p_1 - p_0\| + C(\|m_0\| + \|m_1\|) \quad ,
\end{aligned}$$

$$\begin{aligned}
\|\gamma''(s)\| &\leq \|(h_{00}(r)p_0 + h_{01}(r)p_1 + \delta(h_{10}(r)m_0 + h_{11}(r)m_1))''(s)\| \\
&\leq \frac{1}{\delta^2} \|(12r - 6)(p_0 - p_1) + \delta((6r - 4)m_0 + (6r - 2)m_1)\| \\
&\leq \frac{C}{\delta^2} \|p_1 - p_0\| + \frac{1}{\delta} (\|m_0\| + \|m_1\|) \quad ,
\end{aligned}$$

$$\begin{aligned}
\|\gamma'''(s)\| &\leq \|(h_{00}(r)p_0 + h_{01}(r)p_1 + \delta(h_{10}(r)m_0 + h_{11}(r)m_1))'''(s)\| \\
&\leq \frac{1}{\delta^3} \|12(p_0 - p_1) + 6\delta(m_0 + m_1)\| \\
&\leq \frac{C}{\delta^3} \|p_1 - p_0\| + \frac{C}{\delta^2} \|m_0 + m_1\| \quad .
\end{aligned}$$

Revenons à la majoration de $\|\gamma_t - \bar{\gamma}_t\|$. On a

$$\begin{aligned}
\|\gamma_t - \bar{\gamma}_t\| &\leq \delta^4 \|d^4(\phi \circ \gamma)_s\| \\
&\leq \delta^4 \left(\|d^4\phi_\gamma\| \|\gamma'\|^4 + 6 \|d^3\phi_\gamma\| \|\gamma'\|^2 \|\gamma''\| \right. \\
&\quad \left. + 3 \|d^3\phi_\gamma\| \|\gamma''\|^2 + 4 \|d^2\phi_\gamma\| \|\gamma'\| \|\gamma'''\| \right) \\
&\leq C(d^4\phi, d^3\phi, d^2\phi) \delta^4 \left(\|\gamma'\|^4 + \|\gamma'\|^2 \|\gamma''\| + \|\gamma''\|^2 + \|\gamma'\| \|\gamma'''\| \right) \quad .
\end{aligned}$$

En utilisant des propriétés de convexité sur les majorations que nous avons, nous

pouvons majorer $\|\gamma'\|^4$, $\|\gamma'\|^2$, $\|\gamma''\|^2$ et on a alors

$$\begin{aligned}
 \|d^4(\phi \circ \gamma)_s\| &\leq C(d\phi, d^2\phi, d^3\phi, d^4\phi)\delta^4(\\
 &\quad \frac{1}{\delta^4}\|p_1 - p_0\|^4 + (\|m_0\| + \|m_1\|)^4 + \\
 &\quad \left(\frac{1}{\delta^2}\|p_1 - p_0\|^2 + (\|m_0\| + \|m_1\|)^2\right) \left(\frac{1}{\delta^2}\|p_1 - p_0\| + \frac{1}{\delta}(\|m_0\| + \|m_1\|)\right) + \\
 &\quad \frac{1}{\delta^4}\|p_1 - p_0\|^2 + \frac{1}{\delta^2}(\|m_0\| + \|m_1\|)^2 + \\
 &\quad \left(\frac{1}{\delta}\|p_1 - p_0\| + (\|m_0\| + \|m_1\|)\right) \left(\frac{1}{\delta^3}\|p_1 - p_0\| + \frac{1}{\delta^2}(\|m_0\| + \|m_1\|)\right) \\
 &\leq \sum_{k=2}^4 \|p_1 - p_0\|^k + \sum_{k=2}^4 \delta^k (\|m_0\| + \|m_1\|)^k \\
 &\quad \left(\|p_1 - p_0\|^2 \delta (\|m_0\| + \|m_1\|)\right) + \left(\|p_1 - p_0\| \delta^2 (\|m_0\| + \|m_1\|)^2\right) \\
 &\quad + (\|p_1 - p_0\| \delta (\|m_0\| + \|m_1\|)) \quad .
 \end{aligned}$$

Ce qui conclue la démonstration.

Chapitre 5

Applications

5.1 Calcul numérique de la distance de Fisher pour les gaussiennes 2D

Nous avons vu que l'étude théorique de la variété riemannienne des gaussiennes multivariées munie de la métrique de Fisher est un sujet théoriquement compliqué et que peu de résultats sont connus en dehors du cas 1d. En particulier, il n'existe pas de formules explicites pour la distance. Nous allons ici proposer une technique numérique de calcul que nous présenterons dans le cas des gaussiennes 2d. La méthode est théoriquement applicable indépendamment de la dimension, mais la lourdeur des calculs préparatoires la rendra en pratique difficile à utiliser.

On rappelle qu'on identifie les gaussiennes multivariées 2d avec

$$\Theta = \left\{ (\mu_1, \mu_2, v_{11}, v_{12}, v_{22}) \in \mathbb{R}^5 \mid \sum_{i,j} v_{ij} E_{ij} \in S_2^{++} \right\} .$$

Soit θ_0 et θ_1 deux vecteurs de paramètres représentant deux gaussiennes g_0 et g_1 . Supposons qu'on connaisse $\theta(t)$ la géodésique définie sur $[0, 1]$ reliant g_0 et g_1 , qu'on écrira parfois (μ_t, V_t) . Si on note M la matrice de Fisher et d la distance associée, alors on a

$$d(g_0, g_1) = \int_0^1 \sqrt{{}^t\theta'(t) M(\theta(t)) \theta'(t)} dt,$$

quantité qu'on peut calculer au moyen d'une formule de quadrature. Le calcul de la géodésique nous permet donc d'obtenir la distance. Présentons donc une technique

de calcul de la géodésique.

5.1.1 Présentation de la méthode

D'après la Proposition 36 (page 56), la géodésique est aussi une géodésique pour la 0-connexion (section 7.3.2.3, page 143). Nous allons donc calculer la géodésique pour la connexion, au moyen de la Proposition 69 (page 144) qui nous donne l'équation différentielle vérifiée par une connexion. Elle est calculé pour la 0-connexion dans le cas des gaussiennes multivariées dans [34] et nous donne le système suivant :

$$\begin{cases} \mu'' - V'V^{-1}\mu' & = 0 \\ V'' + \mu' {}^t\mu' - V'V^{-1}V & = 0 \end{cases} .$$

La démonstration est purement calculatoire et repose sur le calcul des coefficients Γ_{ij}^k .

La solution du système d'EDO avec conditions aux limites suivant

$$\begin{cases} \mu'' - V'V^{-1}\mu' & = 0 \\ V'' + \mu' {}^t\mu' - V'V^{-1}V & = 0 \\ (\mu(0), V(0)) & = (\mu_0, V_0) \\ (\mu(1), V(1)) & = (\mu_1, V_1) \end{cases} .$$

nous donne donc la géodésique reliant (μ_0, V_0) à (μ_1, V_1) . Présentons formellement la technique de résolution qui sera basé sur une technique de tir.

On déduit du système précédent un problème de Cauchy d'ordre 1 qui nous servira pour la mise en oeuvre de la méthode de tir

$$\begin{cases} Y'(t) & = F(Y(t)) \\ Y(0) & = Y_0 \end{cases}$$

où $Y(t) = (\mu(t), V(t), \mu'(t), V'(t))$. On posera

$$y(t) = (\mu(t), V(t), \mu'(t), V'(t)), \quad y_0 = (\mu_0, V_0) \quad y_1 = (\mu_1, V_1) \quad .$$

5.1 Calcul numérique de la distance de Fisher pour les gaussiennes 2D

On pose $\Phi(y, p, t)$ la solution au temps t ayant pour données initiales (y, p) . On pose alors

$$\begin{aligned} \varphi: \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ p &\mapsto \Phi(y_0, p, 1) \end{aligned}$$

Nous sommes donc à la recherche du zéro (unique car Φ difféomorphisme) de $\varphi(p) - y_1$, zéro que nous allons trouver au moyen d'une méthode de Newton. $d\varphi_p$ se déduit en extrayant une sous-matrice de $d\Phi$ et cette dernière vérifie l'équation sur la différentielle du flot (Proposition 5, page 19)

$$\partial_t d\Phi(\cdot, \cdot, t)_{y,p} = dF_{\Phi(y,p,t)} \circ d\Phi(\cdot, \cdot, t)_{y,p}$$

Nous pouvons donc maintenant présenter l'algorithme de résolution. Nous considérons donc le système

$$\left\{ \begin{array}{l} Y'(t) = F(Y(t)) \\ \partial_t d\Phi(Y, t) = dF_{Y(t)} \circ d\Phi(Y(t), t) \\ Y(0) = (\mu_0, V_0, \mu'_0, V'_0) \\ d\Phi(Y_0, 0) = \text{Id} \end{array} \right.$$

et nous effectuons les étapes suivantes :

1. choix des paramètres de tir μ'_0 et d'un V'_0 ;
2. résolution de l'EDO. Si le point d'arrivée $y(1)$ est à la précision voulue de y_1 , alors on arrête ;
3. Sinon, utilisation de la méthode de Newton et obtention de nouvelles conditions initiales μ'_0 et d'un V'_0 ;
4. retour à l'étape 2.

5.1.2 Détail du champ et matrice de Fisher

On a $F : \mathbb{R}^{10} \rightarrow \mathbb{R}^{10}$ donné par, si l'on appelle les variables $(\mu_0, \mu_1, a, b, c, \mu'_0, \mu'_1, a', b', c')$,

$$\begin{aligned}
 F_1 &= \mu'_0 \quad , \\
 F_2 &= \mu'_1 \quad , \\
 F_3 &= a' \quad , \\
 F_4 &= b' \quad , \\
 F_5 &= c' \quad , \\
 F_6 &= -\frac{1}{ac-b^2} (bb' - a'c - ab'\mu'_1 + a'b\mu'_1) \quad , \\
 F_7 &= \frac{1}{ac-b^2} (ac'\mu'_1 - bb'\mu'_1 - bc'\mu'_0 + b'c\mu'_0) \quad , \\
 F_8 &= \frac{1}{ac-b^2} (c(a')^2 - 2a'bb' + b^2(\mu'_0)^2 + a(b')^2 - ac(\mu'_0)^2) \quad , \\
 F_9 &= -\frac{1}{ac-b^2} (b(b')^2 - ab'c' + a'bc - b^2\mu'_0\mu'_1 + ac\mu'_0\mu'_1) \quad , \\
 F_{10} &= \frac{1}{ac-b^2} (b^2(\mu'_1)^2 - 2bb'c' + c(b')^2 + a(c')^2 - ac(\mu'_1)^2) \quad ,
 \end{aligned}$$

On en déduit que dF est donné par

$$\frac{1}{\delta^2} \begin{pmatrix}
 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
 0 & 0 & f_{63} & f_{64} & f_{65} & 0 & \delta s_5 & \delta(c - b'\mu'_1) & \delta(a\mu'_1 - 1) & 0 \\
 0 & 0 & s_3 s_8 & f_{74} & s_4 s_7 & -\frac{1}{\delta} & \delta s_4 & 0 & -\delta s_8 & \delta s_7 \\
 0 & 0 & -s_6^2 & -2s_5 s_6 & -s_5^2 & -2\mu'_0 & 0 & 2\delta(a'c - bb') & 2\delta(ab' - a'b) & 0 \\
 0 & 0 & s_6 s_3 & f_{94} & -s_5 s_4 & \mu'_1 & -\mu'_0 & -\frac{1}{\delta} & \frac{1}{\delta}(ac' + a'c - 2bb') & \delta s_5 \\
 0 & 0 & -s_3^2 & 2s_4 s_3 & -s_4^2 & 0 - 2\mu'_1 & 0 & -2\delta(bc' - b'c) & 2\delta(ac' - bb') & 0
 \end{pmatrix}$$

5.1 Calcul numérique de la distance de Fisher pour les gaussiennes 2D

avec

$$\begin{aligned}
 \delta &= (ac - b^2) \\
 s_3 &= bc' - b'c \quad , \\
 s_4 &= ac' - bb' \quad , \\
 s_5 &= ab' - a'b \quad , \\
 s_6 &= a'c - bb' \quad , \\
 s_7 &= a\mu'_1 - b\mu'_0 \quad , \\
 s_8 &= b\mu'_1 - c\mu'_0 \quad , \\
 f_{63} &= bc + b'c - a'c^2 - b^2b'\mu'_1 + a'bc\mu'_1 \quad , \\
 f_{64} &= ac + 2bb' + b^2 - 2a'bc + a'b^2\mu'_1 + aa'c\mu'_1 - 2abb'\mu'_1 \quad , \\
 f_{65} &= ab + ab' - a'b^2 - a^2b'\mu'_1 + aa'b\mu'_1 \quad , \\
 f_{74} &= b^2b'\mu'_1 + b^2c'\mu'_0 - 2abc'\mu'_1 + ab'c\mu'_1 + acc'\mu'_0 - 2bb'c\mu'_0 \quad , \\
 f_{94} &= -(b^2(b')^2 + a(b')^2c + a'b^2c' + aa'cc' - 2abb'c' - 2a'bb'c) \quad ,
 \end{aligned}$$

La matrice de Fisher est donnée par

$$\frac{1}{\delta} \begin{pmatrix} v_{22} & -v_{12} & 0 & 0 & 0 \\ -v_{12} & v_{11} & 0 & 0 & 0 \\ 0 & 0 & \frac{v_{22}^2}{2\delta} & -\frac{v_{12}v_{22}}{\delta} & \frac{v_{12}^2}{2\delta} \\ 0 & 0 & -\frac{v_{12}v_{22}}{\delta} & \frac{v_{11}v_{22} + v_{12}^2}{\delta} & -\frac{v_{11}v_{12}}{\delta} \\ 0 & 0 & \frac{v_{12}^2}{2\delta} & -\frac{v_{11}v_{12}}{\delta} & \frac{v_{11}^2}{2\delta} \end{pmatrix}$$

5.1.3 Résultats numériques

5.1.3.1 Validation du code

On va valider le code en se ramenant à une situation 1d pour laquelle nous avons une formule explicite pour la distance. On prend

$$\mu_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

et

$$\mu_1 = \begin{pmatrix} 0 \\ 1.5 \end{pmatrix},$$

$$V_0 = V_1 = 0.5 \text{Id}.$$

On est dans ce cas ramené à la distance entre

$$\frac{1}{\sqrt{2\pi} \times \sqrt{0.5}} \exp\left(-\frac{x^2}{2 \times \sqrt{0.5}^2}\right)$$

et

$$\frac{1}{\sqrt{2\pi} \times \sqrt{0.5}} \exp\left(-\frac{(x - 1.5)^2}{2 \times \sqrt{0.5}^2}\right)$$

Que l'on peut calculer exactement d'après les résultats sur les gaussiennes 1d (Théorème 34, page 55).

- distance exacte 1.96051628 , distance trouvé par l'algorithme 1.96051629628, tolérance 10^{-4} demandé au code ;
- 8 itérations de Newton.
- temps de calcul : 0.9 secondes.

La figure 5.1 représente quelques gaussiennes de la géodésique 1d correspondante à partir des paramètres calculés par notre code. Les résultats sont donc cohérents avec ce que nous savons du cas 1d.

5.1.3.2 Test sur un cas purement 2d

On prend $\mu_0 = (0, 0)$ et $\mu_1 = (16, 0)$ et

$$V_0 = \begin{pmatrix} 0.7 & 0 \\ 0 & 3 \end{pmatrix}, \quad V_1 = {}^tR \circ \begin{pmatrix} 0.7 & 0 \\ 0 & 3 \end{pmatrix} \circ R, \quad \text{où } R = \begin{pmatrix} \cos\left(\frac{\pi}{3}\right) & \sin\left(\frac{\pi}{3}\right) \\ -\sin\left(\frac{\pi}{3}\right) & \cos\left(\frac{\pi}{3}\right) \end{pmatrix}.$$

5.1 Calcul numérique de la distance de Fisher pour les gaussiennes 2D

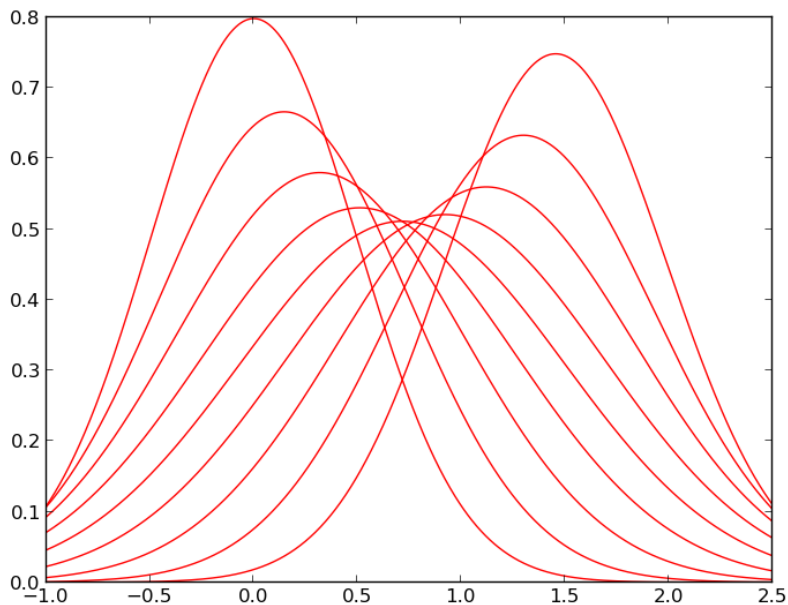


FIGURE 5.1 Quelques gaussiennes de la géodésique 1d

L'initialisation de la première direction de tir est ici délicate et peut fréquemment conduire à passer par des gaussiennes très resserrées puis sortir du domaine. En l'absence de résultats théoriques généraux sur ce choix, nous avons effectué le constat suivant dont nous avons déduit une méthode assez générale mais augmentant les coûts de calcul : si on choisit comme première direction de tir $(\mu_1 - \mu_0, V_1 - V_0)$ (l'espace tangent de S^{++} étant l'ensemble des matrices symétriques) et qu'on résout sur l'intervalle $[0, 0.1]$, nous ne sortons pas du domaine et convergions pour des densités proches.

Voilà le principe de la méthode que nous en déduisons : nous partons de (μ_0, V_0) et nous commençons par essayer d'atteindre une densité proche en utilisant notre méthode. Nous conservons la direction de tir que nous obtenons après convergence. Nous partons de nouveau de (μ_0, V_0) avec cette nouvelle direction de tir pour essayer d'atteindre une densité plus éloignée. Nous itérons le procédé jusqu'à arriver à (μ_1, v_1) .

On trouvera sur les figures 5.2 et 5.3 huit gaussiennes tirées de la géodésique . On trouvera sur 5.4 l'évolution des valeurs propres et du déterminant

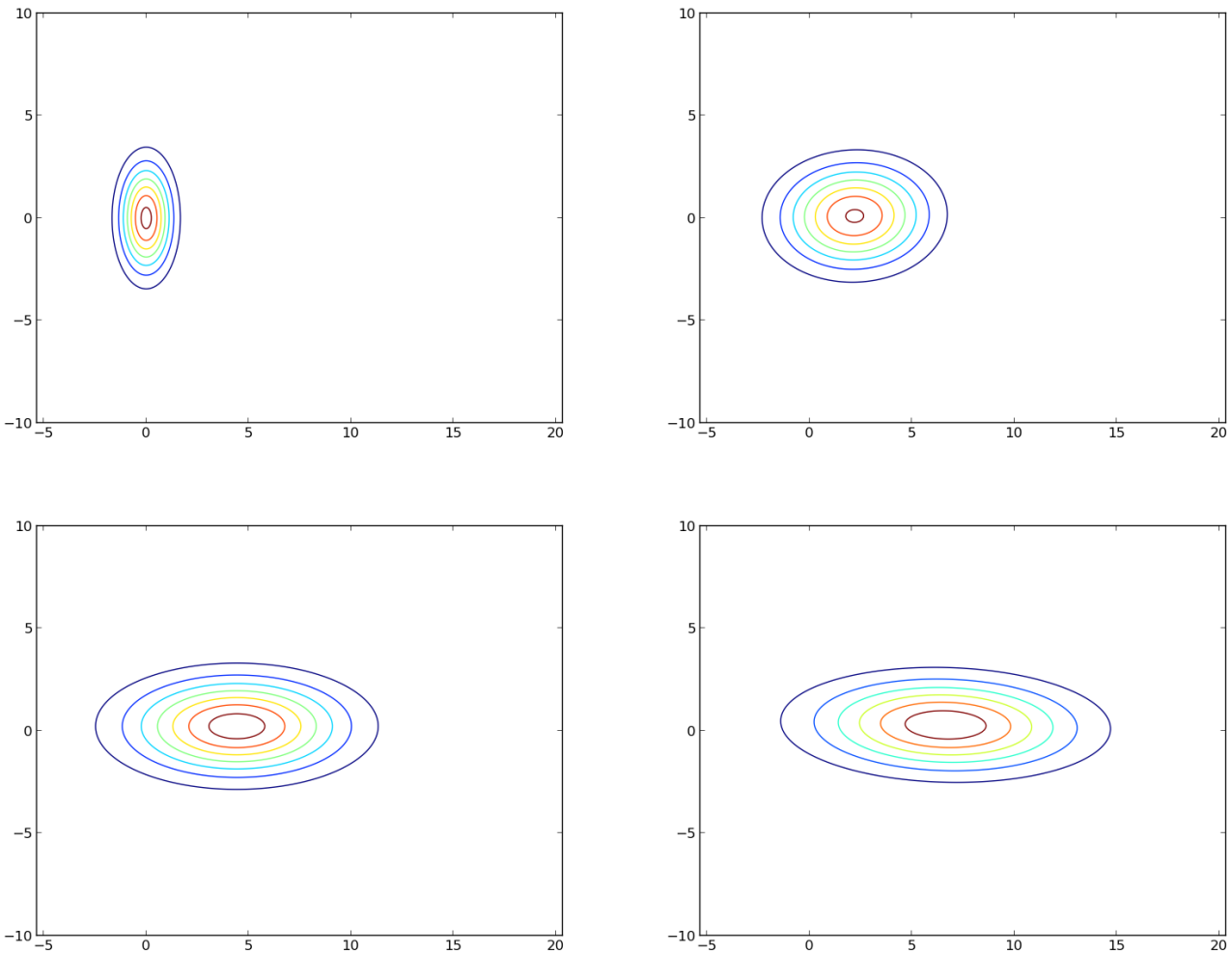


FIGURE 5.2 4 éléments de la géodésique , 1

5.2 Élaboration d'un programme de mise en oeuvre de la méthode pour l'équation de transport conservative

5.2.1 Considérations générales sur le programme

5.2.1.1 Introduction

On présente ici un code développé pour l'utilisation de la méthode dans le cadre du transport 2d, plus précisément pour traiter l'équation de transport conservative

$$\begin{cases} \partial_t \rho + \operatorname{div}(\rho F) = 0 \\ \rho(x, 0) = \rho_0(x) \end{cases}$$

5.2 Élaboration d'un programme de mise en oeuvre de la méthode pour l'équation de transport conservative

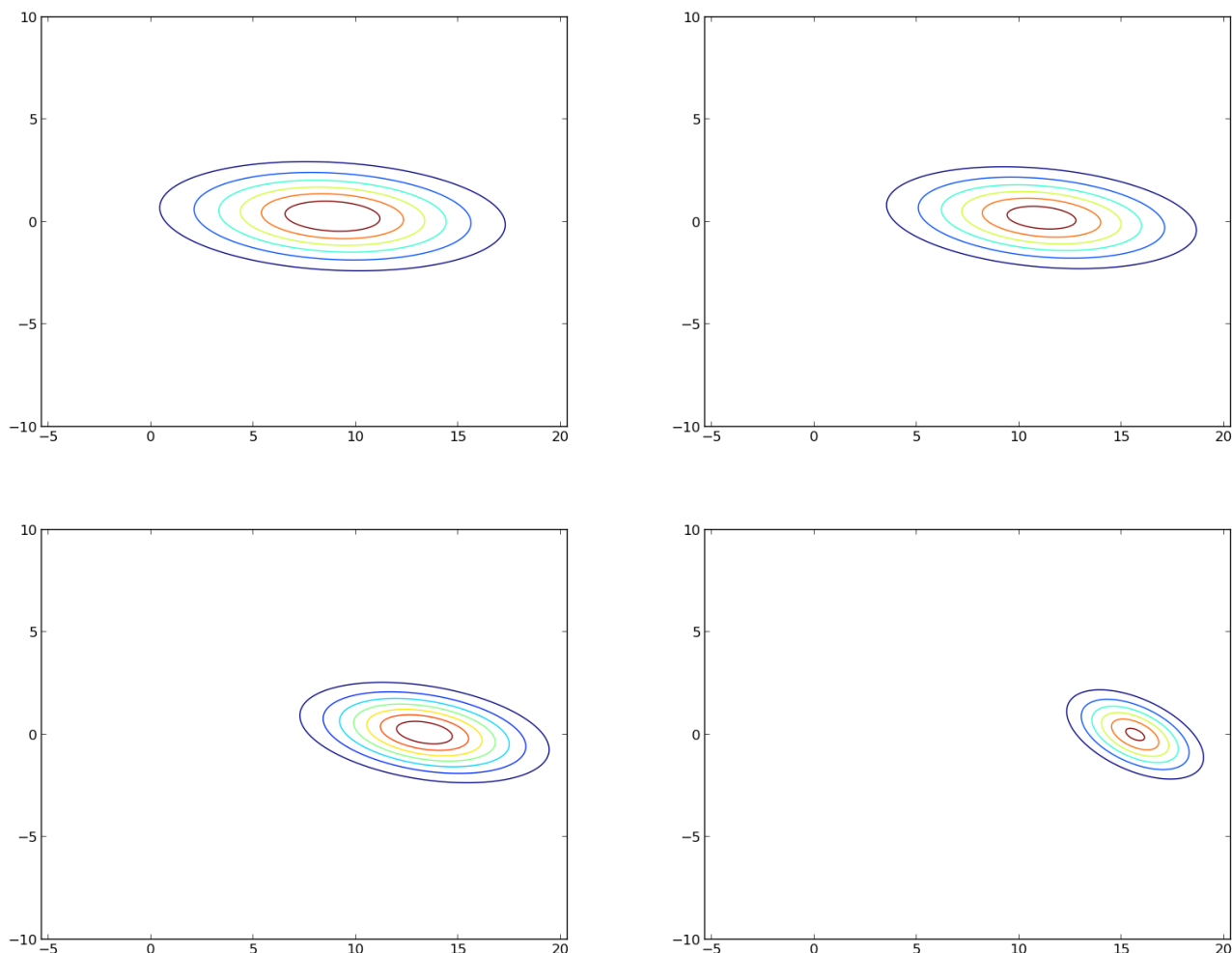


FIGURE 5.3 4 éléments de la géodésique , 2

Dans des situations très concrètes et variées, où le champ F peut-être donné exactement, sur un maillage (par exemple s'il provient d'un solveur de mécanique des fluides) ou sur d'autres formes de supports et à partir de fichiers particuliers (par exemple des données ponctuelles comme ça sera le cas pour perspectives d'applications à l'océanographie).

On suppose que la discrétisation de ρ_0 en particules a été effectuée, c'est à dire que ρ_0 est de la forme

$$\rho_0(x) = \sum_{i=0}^n \alpha_i \frac{1}{\sqrt{|\det V_i|}} \varphi_i \left({}^t(x - \mu_i) V_i^{-1} (x - \mu_i) \right)$$

Où V_i est une matrice symétrique définie positive, μ_i un point de \mathbb{R}^2 , α_i un poids appartenant à \mathbb{R}^{+*} , φ_i est une densité de probabilité, typiquement une fonction

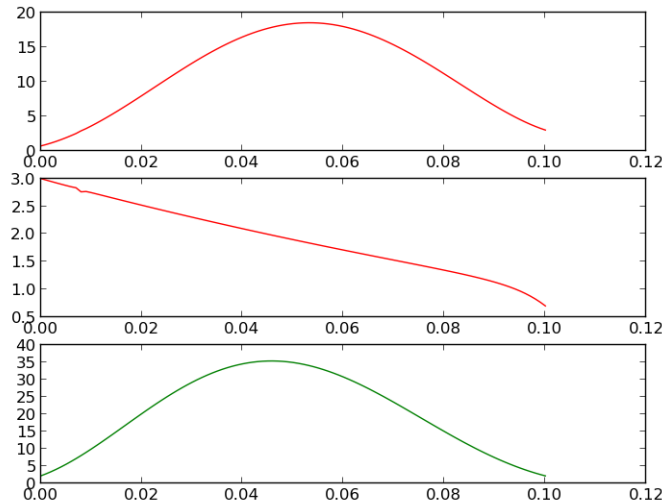


FIGURE 5.4 évolution des valeurs propres (deux premiers graphiques) et du déterminant (dernier)

« bosse » : gaussienne, spline, gaussienne ou spline tronquée, indicatrice ...

On a

$$\rho(x, t) = \sum_{i=0}^n \alpha_i \frac{1}{\sqrt{|V_i^t|}} \varphi_i \left({}^t(x - \mu_i^t) (V_i^t)^{-1} (x - \mu_i^t) \right)$$

Où $\mu_i^t = \phi(\mu, t)$ et $V_i^t = {}^t d\phi(\cdot, t)_\mu \circ V_i \circ d\phi(\cdot, t)_\mu$. Très schématiquement, l'algorithme de résolution consiste donc à :

1. charger le champ F et calculer sa différentielle ;
2. transporter les points μ_i et les tenseurs V_i ;
3. reconstruire la solution.

5.2.1.2 Caractéristique du projet et choix du langage Python

Présentons quelques caractéristiques du projet qui nous amènerons au choix d'un développement en Python.

Le langage doit permettre de traiter de gérer des maillages (dont nous verrons plus loin que les outils sont écrit en langage C), des fichiers tirés de l'océanographie (format de données netCDF).

De manière plus générale, le programme n'a pas vocation à traiter un problème très précis, mais doit rester évolutif et ouvert. Ouvert « horizontalement » en pou-

5.2 *Élaboration d'un programme de mise en oeuvre de la méthode pour l'équation de transport conservative*

vant traiter de nouveaux formats de fichiers correspondants à des champs de vecteurs venant de divers domaines d'applications (météorologie, fichiers issus de l'industrie, etc.), ouvert « verticalement » en pouvant être interfacé avec des outils externes (solveurs de mécanique des fluides, outils de gestion de gestion maillages rectangulaires, etc.)

Cet aspect exploratoire et ouvert du code doit permettre la réalisation naturelle de scripts et de cas tests, tout en permettant, lors de la phase de consolidation du programme, l'utilisation de la machine à bas niveau si nécessaire pour un programme efficace.

Le langage Python est donc naturellement adapté. Premièrement, il peut fonctionner comme un langage de script, permettant rapidement la mise en place de test. En même temps, son orientation objet permet le développement de programme important.

Deuxièmement, il possède un très grand nombre d'extensions. Notamment NumPy qui permet la gestion immédiate des structure inhérentes à l'analyse numérique et Matplotlib l'utilisation de graphique, nous dotant immédiatement d'un outil complet pour le calcul scientifique. Mais aussi des extensions pour la gestion de différents types de fichiers, qui vont nous permettre d'utiliser des fichiers tirés de l'océanographie et de manière générale nous garantissent, en cas de développement ultérieurs amenant à être confronté à un format de fichiers atypiques, d'être bien armés pour traiter la situation.

Troisièmement, il est facilement interfaçable avec d'autres langages. Ainsi, nous pourrions travailler avec des outils de gestion de maillage codés en C, et d'autres si cela devait s'avérer nécessaire. De plus, posé les bases du code, si certains algorithmes devaient s'avérer chronophage, il serait possible de les descendre en un langage objet de plus bas-niveau (type C++) et d'interfacier cette brique avec le reste du code.

5.2.1.3 **Première présentation du code**

Quelques rappels sur Python. Le typage est dynamique, la gestion de la mémoire est automatique par ramasse-miettes et il y a un système de gestions d'exceptions. Pour ce qui concerne le calcul scientifique, les matrices et les vecteurs seront gérés par le type **array** de l'extension NumPy

Le code est orienté objet. Nous allons présenter dans leurs grandes lignes les

différentes classes et leurs interactions, on trouvera un schéma explicatif sur la figure 5.5, page 111.

Les objets de la classe **Particule** vont représenter nos couples particule+tenseur. Un objet particule est constitué de :

- **indice** : 0 pour la première particule créée, 1 pour la deuxième, etc. ;
- **position** : un array représentant le vecteur de la position courant de la particule ;
- **Dposition** : un array représentant la matrice de la différentielle du flot au point donné par position ;
- **tenseur_initial** : le tenseur initial ;
- **typeparticule** : un entier codant pour le « blob » utilisé (gaussienne, spline, etc.) ;
- **poids** : le poids de la particule.

La classe contient principalement la méthode **tenseur**, ne prenant aucun paramètre et renvoyant le tenseur actuel de la particule appelant la méthode, calculé à partir de **Dposition** et de **tenseur_initial**.

La classe **Champ** est centrale. Un objet de la classe va représenter le problème, c'est à dire le domaine, le « fluide » qui s'écoule à l'intérieur, les particules transportées, etc. C'est principalement une classe-mère constituée de prototypes, qui seront spécifiées dans les classes filles. Parmi les attributs, on a :

- **temps_initial** : le temps initial ;
- **temps_actuel** : le temps actuel ;
- **ensemble_particules** : un ensemble d'objets particules représentant ρ_0 .

Les méthodes essentielles de la classe sont :

- **calcul_vitesse** : prototype. Renvoie la vitesse ainsi que sa différentielle en un point ;
- **advection** : avance l'ensemble des particule d'un pas de temps dt en utilisant un solveur EDO ;
- **evolution** : prototype. mets à jour les valeurs du champ de vitesse au temps actuel, c'est à dire passe de $F(\cdot, t^n)$ à $F(\cdot, t^{n+1})$.

La classe **Champ** possède trois classes filles. La classe **ChampMaillage** qui correspond au cas où le champ est donné par un maillage triangulaire quelconque, la classe **ChampExacte** qui correspond au cas où le champ est donné de manière exacte et la classe **ChampSCCOOS** qui correspond au cas où le champ est donné par un fichier océanographique.

5.2 Élaboration d'un programme de mise en oeuvre de la méthode pour l'équation de transport conservative

La classe **MethodesAdvection** contient les solveurs EDO. Il aurait été possible de les intégrer dans la classe **Champ**. Cependant, leurs nombres pouvant augmenter, avec par exemple la mise en place de pas adaptatif, il était préférable d'avoir une classe dédiée au sujet, l'intégration se faisant lors de la construction de l'objet **Champ** au moyen des décorateurs Python. Actuellement, la classe contient deux solveurs, **euler** et **RK4**. Chacune des deux fonctionne de la même manière : elle prend un objet **champ**, une position et un array correspondant à la différentielle du flot et avance d'un pas dt la position et la différentielle du flot.

Pour finir, rajoutons l'existence d'une classe **MesOutils** contenant divers outils, et surtout d'une classe **VolumesFinis**. Elle va prendre un objet **champ** et permettre la résolution du problème associé par méthode des volumes finis sur grille cartésienne avec *splitting* de Strang. Il paraîtrait peut-être plus naturelle à première vue d'intégrer dans la classe **Champ** ce solveur, afin d'avoir deux façons différentes de résoudre le problème, mais l'utilisation d'une grille cartésienne amène un cadre particulier amenant plus naturellement un nouvel objet. La classe **VolumesFinis** a trois classes filles, la classes **Upwind2d**, **Vanleer2d** (schéma de Lax-Wendroff avec limiteurs de flux, [25]) et la classe **DL2d** (schéma de Desprès-Lagoutière, décentré aval sous contrainte amont : ordre un limitant la diffusion, [24]).

5.2.2 La classe **ChampSCCOOS**

5.2.2.1 Les fichiers océanographiques et leur traitement

Le nom de la classe fait au **SCOOS**, acronyme pour *Southern California Coastal Ocean Observing System* puisque c'est de cet organisme que sont venus les premiers fichiers océanographiques que nous avons essayé de traiter.

Ces fichiers sont au format **netCDF**, format de données permettant le partage de données scientifiques stockées sous forme de tableaux, utilisé et développé notamment dans le domaine météorologique et océanographique. Ce format de fichiers est complété d'un ensemble de bibliothèques logicielles développé en C. Leur gestion est possible en Python au moyen de l'extension **netCDF4**.

Précisons immédiatement une erreur de conception dans notre programme. Il s'avère à la pratique que ces fichiers ne sont pas tout à fait standards et que les noms de variables peuvent différer légèrement d'un organisme à l'autre. Il semblerait finalement plus pertinent de créer une classe **ChampPonctuelle**, définissant un

champ dont la vitesse est donnée sur une grille de points (nous détaillerons ce point plus loin), ayant une classe fille ChampMeteo, elle même ayant des classes filles ChampSCCOOS, etc. qu'il serait possible de créer à la volée au moyen des méta-classes Python.

A quelques différences près suivant les organismes, le domaine d'applications, les instruments de mesure concernés, voici les éléments essentiels que contient un fichier nc :

- un jeu de coordonnées sous forme longitude/latitude ;
- un vecteur vitesse associé en chacun de ses points ;
- la date et l'heure des mesures ;
- une incertitude sur la mesure en chacun des points.

La vitesse est donné en mètres/seconde, la longitude et la latitude sont donnés en degrés, il convient donc d'opérer une conversion de la vitesse en degré/seconde.

5.2.2.2 champ à données ponctuelles

Voici la situation sur les vitesses traduit en termes mathématiques :

- une discrétisation $(x_i, y_j)_{(i,j)}$ de pas h_x en abscisse, h_y en ordonnées ;
- un ensemble de vecteurs vitesse $(F_{(i,j)})$.

Nous allons à partir de ces éléments construire un champ défini sur tous le domaine, en s'inspirant d'outils utilisés en PUFEM (*Partition of Unity Finite Elements Methods*).

On se donne deux rayons d'interactions, r_x et r_y , de manière à ce que les rectangles $R_{i,j} = [x_i - r_x, x_i + r_x] \times [y_j - r_y, y_j + r_y]$ forment un recouvrement de \mathbb{R}^2 . On pose

$$\varphi_{i,j}(x, y) = \varphi\left(\frac{x - x_i}{r_x}\right) \varphi\left(\frac{y - y_i}{r_y}\right)$$

et

$$\psi_{i,j}(x, y) = \frac{\varphi_{i,j}(x, y)}{\sum_{k,l} \varphi_{k,l}(x, y)}$$

Les $\psi_{i,j}$ sont bien définis car les $R_{i,j}$ forment un recouvrement de \mathbb{R}^2 .

Si $F(x, y)$ est une fonction de \mathbb{R}^2 , on définit alors son approximation par

$$\bar{F}(x, y) = \sum_{i,j} F(x_i, y_j) \psi_{i,j}(x, y)$$

5.3 Premier cas test

On a alors le théorème suivant nous garantissant un contrôle intéressant sur l'approximation de F ainsi que de son gradient, nécessaire à notre méthode

Théorème 52

Soit F fonction \mathcal{C}^2 bornée à dérivées partielles bornées. Si $r_x = 2^{j_x} h_x$ et $r_y = 2^{j_y} h_y$, alors on a

$$\|F - \bar{F}\|_\infty \leq O(h_x^2) + O(h_y^2)$$

Si de plus F est \mathcal{C}^3 , alors on a aussi

$$\|dF - d\bar{F}\|_\infty \leq O(h_x^2) + O(h_y^2)$$

On trouvera en annexe, en 6.1 page 119, une présentation détaillée de cette méthode de reconstruction ainsi que les démonstrations liées, qui repose sur les phénomènes de surconvergence de la méthode des rectangles lors de l'intégration d'une fonction régulière à support compacte

5.2.3 La classe ChampMaillage

Cette classe utilise un code en C développé par Pascal Frey et Charles Dapogny à partir d'un code de Cuc Bui, pour la mise en oeuvre d'optimisation de forme par level set avec adaptation de maillage, code utilisant lui-même la librairie libmesh. Il est interfacé avec Python au moyen de Swig.

Ce code inclue notamment la méthodes des caractéristiques sur maillage non-structuré avec champ de vitesse P1. Il a été modifié pour permettre l'advection de tenseur en utilisant la différentielle du champ de vitesse, qui est une fonction P0.

5.3 Premier cas test

On considère sur \mathbb{R}^2 l'équation

$$\begin{cases} \partial_t \rho + \operatorname{div}(\rho F) & = 0 \\ \rho(x, 0) & = \rho_0(x) \end{cases}$$

avec $F(x, y) = (x, xy)$. Un calcul élémentaire nous montre qu'on a

$$\phi(x, y, t) = \begin{pmatrix} e^t x \\ ye^{x(e^t-1)} \end{pmatrix}$$

et donc

$$d\phi(\cdot, t)_{(x,y)} = \begin{pmatrix} e^t & 0 \\ y(e^t - 1) & e^{x(e^t-1)} \end{pmatrix}$$

On prend $\rho_0(x)$ spline cubique tensorisée de paramètre μ, V , c'est à dire, $\rho_0(x, y) = \frac{1}{\sqrt{\det V}} \varphi(u)\varphi(v)$ où

$$\begin{pmatrix} u \\ v \end{pmatrix} = \sqrt{V^{-1}} \left(\begin{pmatrix} x \\ y \end{pmatrix} - \mu \right)$$

On posera $\Sigma = \sqrt{V}$. Dans la suite on considérera une résolution à $t = 0.5$, avec $\mu = (1, 1)$.

5.3.1 Première situation : V petit

On considère $V = 10^{-2} \text{Id}$. Cette configuration est a priori défavorable pour les volumes finis puisqu'elle va nécessiter un maillage fin en espace pour capter la concentration et en conséquence imposer par CFL un pas de temps petit. Et favorable à notre méthode par concentration de la densité.

Résolution par notre méthode au moyen d'un schéma de RK4 avec un pas de temps de 10^{-1} :

- temps de calcul : 10^{-3} secondes ;
- erreur L^1 : 3×10^{-2} ;
- erreur L^∞ : 3×10^{-1} .

Résolution par un schéma de Vanleer sur $[-0.5, 3.5] \times [0, 4]$ avec $dx = 0.01$ et $dy = 0.01$:

- temps de calcul : environ 5 minutes ;
- erreur L^1 : 4×10^{-2} ;
- erreur L^∞ : 1.33.

On se gardera cependant de dire que la méthode basée sur les caractéristiques est plus efficace. En effet, les volumes finis permettent d'atteindre, théoriquement,

5.3 Premier cas test

une précision arbitraire. Nous avons ici très rapidement une approximation de la solution, mais nous n'avons pas de convergence puisque V est fixé. C'est donc une méthode adaptée pour obtenir rapidement une estimation de la solution.

Cependant, nous allons voir maintenant que quand V est plus large, on peut quand même au moyen d'une sous-discrétisation obtenir des résultats convenables.

5.3.2 Deuxième situation : V moins petit

On prend $V = 0.25\text{Id}$. La situation est ici plus favorable au volumes fins, puisqu'une discrétisation du même ordre que précédemment permettra de capter plus finement la densité.

La matrice de variance étant assez peu concentrée, nous allons présenter une technique de sous-discrétisation pour notre méthode. Si φ est la spline cubique 1d standard, on a

$$\varphi(x) = \frac{1}{8}\varphi(2x-2) + \frac{1}{2}\varphi(2x-1) + \frac{3}{4}\varphi(2x) + \frac{1}{2}\varphi(2x+1) + \frac{1}{8}\varphi(2x+2)$$

Ce qui nous permet de la sous-discrétiser en sommes de 5 splines. On en déduit donc une sous-discrétisation de ρ_0 au moyen de 25 particules.

Résolution par notre méthode :

- temps de calcul : 2×10^{-2} secondes ;
- erreur L^1 : 10^{-2} ;
- erreur L^∞ : 5×10^{-3} .

A titre de comparaison, l'erreur avec une seule particule donne une erreur L^1 de 0.1 et une erreur L^∞ de 7×10^{-2}

Résolution par un schéma de Vanleer sur $[-0.5, 3.5] \times [0, 4]$ avec $dx = 0.01$ et $dy = 0.01$:

- temps de calcul : environ 5 minutes ;
- erreur L^1 : 4×10^{-4} ;
- erreur L^∞ : 10^{-3} .

La sous-discrétisation nous permet donc de capter la solution, malgré le caractère plus étalé de la densité de départ.

5.4 Application à des fichiers océanographiques

5.4.1 Carte océanographique

OSCAR (*Ocean Surface Current Analysis third degree Resolution*) est un produit développé par des membres de Earth and Space Research. Il permet notamment d'obtenir des courants de surface à l'échelle du globe. Plus précisément, il fournit des fichiers nc contenant :

- une discrétisation de la terre en longitude/latitude (jusqu'à +80/-80 degrés en latitude) avec un pas de 0.33 degrés ;
- une discrétisation de l'année en 72 champs de vitesse, soit un champ pour une période de cinq jours.

Ces champs de vitesse sont calculés à partir de d'équations de mécanique des fluides ([7]) à partir de données (hauteur de la surface, contrainte de cisaillement exercé par le vent, température de la surface) obtenues par satellites et instruments *in-situ*.

D'après [8], ce champ de vitesse est plus adapté pour les échelles de temps large (plus de 20 jours) ou pour les petites variations proches de l'équateur. Nous allons tenter de nous inspirer d'une situation réelle pour étudier le sujet.

5.4.2 Bateau à la dérive

Tauaea Raioaoa, dit « Tavae », parti de Tahiti pour pêcher au sud de Moorea le 15 mars 2002, subit une panne moteur et dérive pendant 118 jours avant de s'échouer sur une plage de l'île Aitutaki, dans les îles Cook, à près de 1200 km de son point de départ.

Nous allons à partir des données que nous avons voir les différences entre son trajet et le suivi des courants par une particule + tenseur représentant l'incertitude sur sa position de départ. La figure 5.9 nous montre les points de départ et d'arrivée de Tavae ainsi que les résultats de notre simulation.

Le résultat ne correspond pas au trajet réel. C'est un sujet complexe mais on va voir qu'on peut, après avoir éliminé un certain nombre de sujets annexes, en tirer une réflexion apportant un éclairage sur le code.

Premièrement, nous ne prenons pas en compte l'influence du vent. Or si l'on considère que dans les modèles de dérive de nappes de pétrole, celles-ci sont poussées

5.5 Champ de constriction

par le vent à 3% de sa vitesse, pour un bateau rigide offrant une prise au vent non négligeable l'influence ne peut-être négligée. Deuxièmement, Tavae a affronté une très grosse tempête pendant une semaine, phénomène que ne peut capter proprement OSCAR et ayant une influence importante sur son trajet. De manière plus générale, nous pouvons nous intéresser à ces questions ainsi qu'à la fiabilité des mesures, mais avec la limite forte de ne pas en être spécialiste donc d'être très vite bloqué.

Par contre, son trajet est perçu comme relativement court. Or le trajet que nous obtenons est plus court, ce qui signale un problème potentiel plus important.

Pour éclaircir cela, nous allons suivre le gulfstream. La figure 5.10 représente le trajet d'une particule immergée dans le gulf-stream sur une durée de 239 jours. La particule semble sortir du Gulf Stream, encore qu'elle puisse être à l'extrémité de sa partie la plus occidentale, amplifié par des erreurs de mesures. Le trajet implique au mieux une vitesse moyenne de l'ordre 0.14 mètres par seconde, ou 12.5 kilomètres par jour, valeurs qui ne semblent pas aberrantes.

En l'état, il est difficile de se prononcer sur la validité de phénomènes de transport au moyen de ces fichiers.

5.5 Champ de constriction

On présente ici un cas test dont le but est de mieux comprendre le comportement du raffinement. On considère un écoulement de Poiseuille sur $\mathbb{R} \times [-1, 1]$ donné par

$$F_0(x, y) = 3 \begin{pmatrix} (1 - y)^2 \\ 0 \end{pmatrix}$$

On le déforme par un difféomorphisme pour resserrer l'écoulement sur une section. Pour cela, on pose

$$f(x) = 1 - \frac{1}{3}\varphi(3x) \quad \text{où } \varphi \text{ désigne la spline cubique}$$

on considère le difféomorphisme

$$\psi(x, y) = \begin{pmatrix} x \\ f(x)y \end{pmatrix}$$

Et on pose alors

$$F(x, y) = d\psi_{\psi^{-1}(x,y)} \circ F_0(\psi^{-1}(x, y))$$

On considère la densité initiale paramétrée par $\mu = (-2, 0.5)$ et

$$\Sigma = {}^t \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \circ \begin{pmatrix} 0.1 & 0 \\ 0 & 0.2 \end{pmatrix} \circ \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

Dont on peut trouver la représentation sur la figure 5.11

Résolution par un schéma de Vanleer :

- discrétisation de $[-3, 3] \times [-1, 1]$ avec 120 points en abscisse et 61 en ordonnées ;
- temps de calcul : 28.3 secondes.

Résolution par notre méthode :

- sous-discrétisation de la spline ;
- temps de calcul : 4 secondes, graphiques compris.

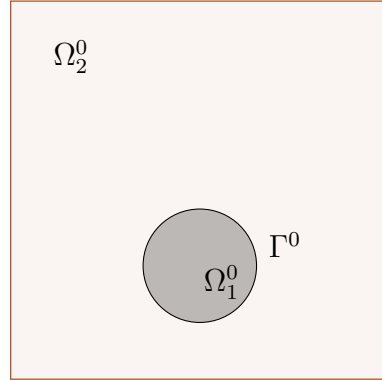
La déformation est captée, quoique de manière assez approximative, pour un temps de calcul modeste.

5.6 Transport pour les courbes appliquées à un problème de Stokes diphasique

5.6.1 Présentation du problème

Soit $\Omega = [0, 1]^2$ et $\Omega_1^0 \subset \Omega$ le cercle de centre $c = (0.5; 0.3)$ et de rayon $r = 0.15$. On pose $\Omega_2^0 = \Omega \setminus \Omega_1^0$. On notera Γ^0 l'interface entre Ω_1^0 et Ω_2^0 .

5.6 problème de Stokes diphasique



Dans le problème considéré, Ω_i^0 est rempli d'un fluide visqueux incompressible caractérisé par sa masse volumique ρ_i et sa viscosité dynamique μ_i . On considère le problème de Stokes diphasique sans tension surfacique avec conditions de Dirichlet donné par

$$\left\{ \begin{array}{l} -\mu_i \Delta u_i + \nabla p_i = \rho_i g \text{ sur } \Omega_i^t \\ \operatorname{div} u_i = 0 \\ u_1 - u_2 = 0 \text{ sur } \Gamma^t \\ (\sigma_1 - \sigma_2) \cdot n = 0 \text{ (pas de tension surfacique)} \\ \Omega_1^t = \phi(\Omega_1, t) \text{ où } \phi \text{ désigne le flot associé à } u \\ u(x, y) = 0 \text{ sur } \partial\Omega \end{array} \right.$$

où g désigne le vecteur gravité.

5.6.2 Résolution numérique du problème

On va effectuer une résolution numérique de ce problème par éléments finis avec un suivi d'interface faisant appel à la méthode présentée dans le chapitre sur les courbes

5.6.2.1 Généralités

Par symétrie, on considère le problème sur $\left[0, \frac{1}{2}\right] \times [0, 1]$ avec la condition $u_1\left(\frac{1}{2}, y\right) = 0$. Par abus de notation, Ω désignera le demi-domaine, Ω_1 le demi-cercle, toutes les autres notations ensemblistes feront références à l'objet correspondant naturellement sur le demi-domaine.

On maille le domaine Ω . On paramétrise Ω_1^0 par

$$\gamma : s \in [1/4, 1/4 + 1/2] \longrightarrow r(\cos(s2\pi), \sin(s2\pi)) + c$$

ce qui va nous permettre d'effectuer ensuite une discrétisation par splines cubiques : on fixe N , on pose $\delta = \frac{1}{2(N-1)}$ et on définit $s_k = k\delta + 1/4$, $p_k = \gamma(s_k)$, $v_k = \gamma'(s_k)$ avec $k \in \{0, \dots, N-1\}$. On pose alors

$$\begin{aligned} \bar{\gamma}(s) = \sum_{k=0}^{N-2} & \left(h_{00} \left(\frac{s-s_k}{\delta} \right) p_k + \delta h_{01} \left(\frac{s-s_k}{\delta} \right) v_k \right. \\ & \left. + h_{10} \left(\frac{s-s_k}{\delta} \right) p_{k+1} + \delta h_{11} \left(\frac{s-s_k}{\delta} \right) v_{k+1} \right) \mathbb{1}_{[s_k, s_{k+1}]}(s). \end{aligned}$$

Voici maintenant comment l'algorithme de résolution passe de t^n à t^{n+1} :

1. on détermine les points du maillage qui sont à l'intérieur et à l'extérieur de l'interface Γ^n pour mettre à jour ρ^n et μ^n ;
2. on résoud le problème de Stokes par éléments finis (P1 bulle / P1) pour obtenir le champ de vitesse u^n ;
3. on calcule la différentielle du champ ddu^n (P0) (voir ci-dessous) ;
4. on utilise un solveur EDO (ici méthode du point milieu) pour advecter les points p_k et les vecteurs v_k et obtenir la nouvelle interface Γ^{n+1} ;
5. adaptation des points de la nouvelle interface.

Notons que le champ ddu^n n'est a priori pas différentiable. Néanmoins, s'agissant d'un déplacement normal à l'interface, il sera régulier tangentiellement, direction suivant laquelle nous advectons nos tangentes.

5.6.2.2 Intérieur/extérieur

Le calcul de l'intérieur de l'extérieur de l'interface doit être effectué rapidement. Ainsi, nous avons écarté une détermination basée sur un calcul d'indice, qui garantirait une méthode "clé en main", mais nécessiterait une reconstruction complète de l'interface, chronophage.

Nous présentons ici une méthode ne nécessitant que la reconstruction d'une section de spline.

Dans un premier temps, nous cherchons le segment $[p_k^t, p_{k+1}^t]$ le plus proche. La

5.6 problème de Stokes diphasique

proposition suivante nous permet de calculer la distance entre un point et la droite $(p_k^t p_{k+1}^t)$.

Proposition 53

Soit $(u, v) \in \mathbb{R}^2$. Soit (x_0, y_0) et (x_1, y_1) deux points distincts du plan et \mathcal{D} la droite qu'ils définissent. Le projeté orthogonal de (u, v) sur \mathcal{D} est donné par :

$$(x_1, v) \text{ si } x_0 = x_1 \quad ,$$

$$(u, y_1) \text{ si } y_0 = y_1 \quad ,$$

et sinon par

$$\left(\frac{ux_0^2 - 2ux_0x_1 - x_0y_0y_1 + vx_0y_0 + x_0y_1^2 - vx_0y_1 + ux_1^2 + x_1y_0^2 - x_1y_0y_1 - vx_1y_0 + vx_1y_1}{(x_0 - x_1)^2 + (y_0 - y_1)^2}, \right. \\ \left. \frac{x_0^2y_1 - x_0x_1y_0 - x_0x_1y_1 + ux_0y_0 - ux_0y_1 + x_1^2y_0 - ux_1y_0 + ux_1y_1 + vy_0^2 - 2vy_0y_1 + vy_1^2}{(x_0 - x_1)^2 + (y_0 - y_1)^2} \right) .$$

Nous effectuons ensuite un passage en coordonnées barycentriques à partir de la proposition suivante :

Proposition 54

Soit $p_0 = (x_0, y_0)$ et $p_1 = (x_1, y_1)$ deux points distincts du plan et \mathcal{D} la droite qu'ils définissent. Soit $p = (x, y) \in \mathcal{D}$. Si $x_0 \neq x_1$, p a pour coordonnées barycentriques par rapport à (p_0, p_1)

$$\left(\frac{x - x_1}{x_0 - x_1}, 1 - \frac{x - x_1}{x_0 - x_1} \right)$$

Si $x_0 = x_1$, alors, les coordonnées sont donnés par

$$\left(\frac{y - y_1}{y_0 - y_1}, 1 - \frac{y - y_1}{y_0 - y_1} \right)$$

La lecture des coordonnées barycentriques nous permet immédiatement de situer la position du projeté par rapport au segment de calculer la distance du point au segment.

Une fois le segment le plus proche trouvé, nous construisons la spline correspondante. Nous déterminons son élément le plus proche du point considéré et concluons par produit scalaire pour savoir si le point est à l'intérieur ou à l'extérieur.

5.6.2.3 Adaptation

Présentons la procédure d'adaptation. Elle vise à limiter la concentration des points vers le bas de l'interface. Heuristiquement il s'agit, en partant du haut de l'interface, de faire remonter les points qui glisseraient trop vers le bas.

On suppose la discrétisation initiale uniforme et on se fixe $\alpha = \|p_{k+1}^0 - p_k^0\|$, la distance entre les points initiaux dont on se sert pour définir une tolérance haute α_+ et une tolérance basse α_- sur la longueur des morceaux de courbes. On approchera la longueur entre les morceaux de courbes par la distance entre ses extrémités :

1. on effectue un balayage des morceaux de courbes en partant du haut ;
2. si la distance entre les extrémités est comprise entre les deux bornes on ne fait rien ;
3. si la distance est supérieure à α_+ , on construit la spline locale. On balaye ses points en partant du bas jusqu'à trouver celui respectant la borne haute. Celui-ci et sa tangente remplace notre extrémité basse ainsi que l'extrémité haut du morceau de courbe suivant ;
4. si la distance est inférieure à α_- , on construit la spline locale supérieure. On remplace alors la première extrémité et sa tangente par le point+tangente de la spline précédente garantissant le respect de la distance ;
5. une fois l'adaptation finie, on effectue une mise à jour de ρ et de μ .

5.6.2.4 Résultats numériques

On prend $\mu_1 = 0.1$, $\mu_2 = 1$, $\rho_1 = 0$, $\rho_2 = 10$. $dt = 0.5$, un maillage de 25 sur 50. Une courbe à 8 points, pour l'adaptation, $\alpha_+ = 2.5\alpha$ et $\alpha_- = \alpha$.

On trouvera sur les figures 5.14 à 5.17 l'interface de la méthode avec raffinement, de 5.18 à 5.21 la méthode sans raffinement. La méthode avec raffinement préserve une certaine homogénéité dans la répartition des points. Sans raffinement les points se concentrent fortement vers le bas. Cependant, dans le cas sans raffinement, compte tenu à $t = 9$ de l'écart entre le point au sommet et le point suivant, la précision n'est pas si mauvaise. A la vue de cet exemple, probablement qu'une méthode d'adaptation de meilleure qualité devrait gérer un nombre de points de manière dynamique. Elle permettrait d'enlever un certain nombre de points en bas de l'interface, tout en en maintenant quelques uns sur la partie supérieure. Des contraintes liés à la maîtrise de certains outils utilisés rendent cela impossible pour l'instant.

5.6 problème de Stokes diphasique

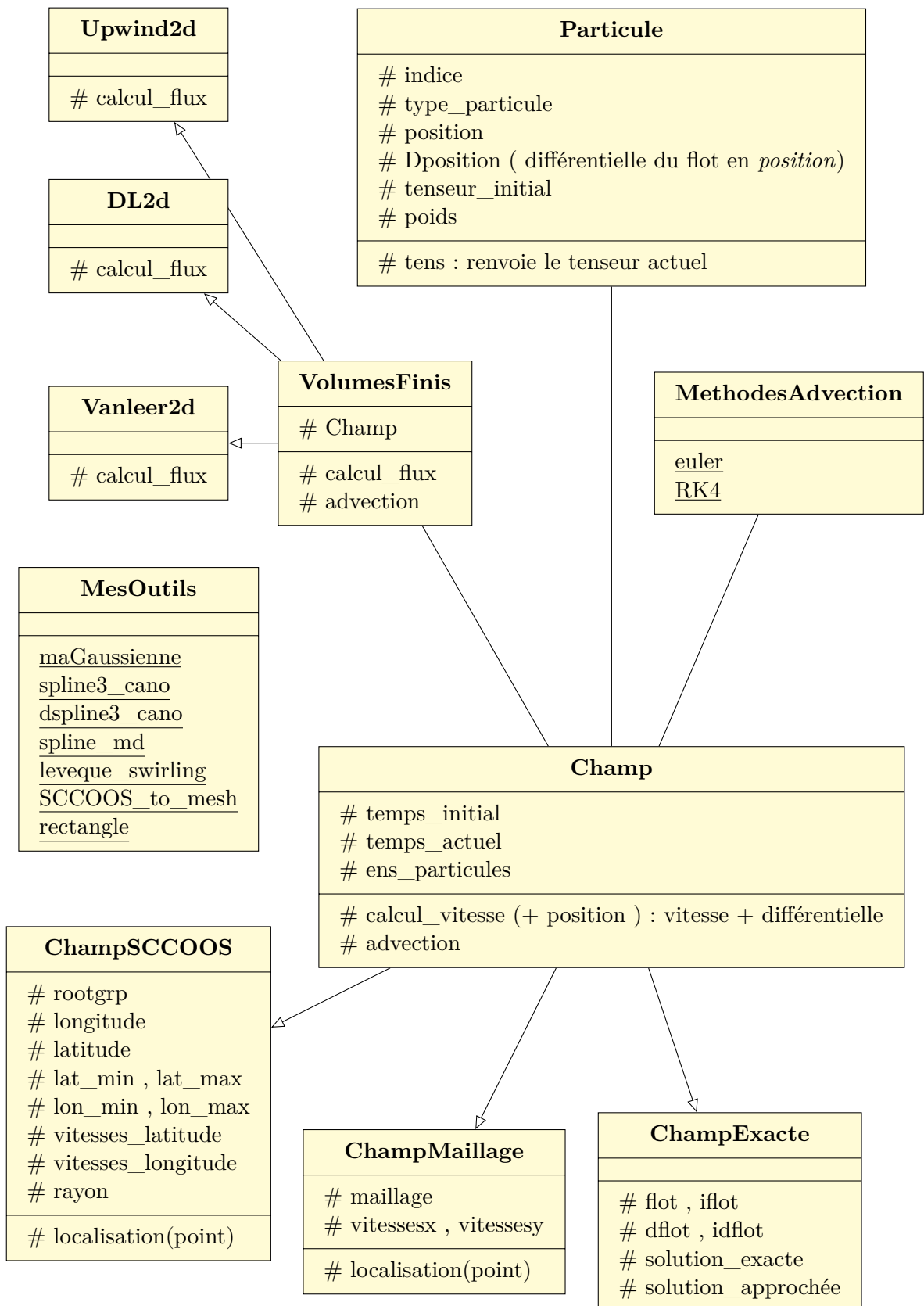


FIGURE 5.5 schéma simplifié des structures du programme

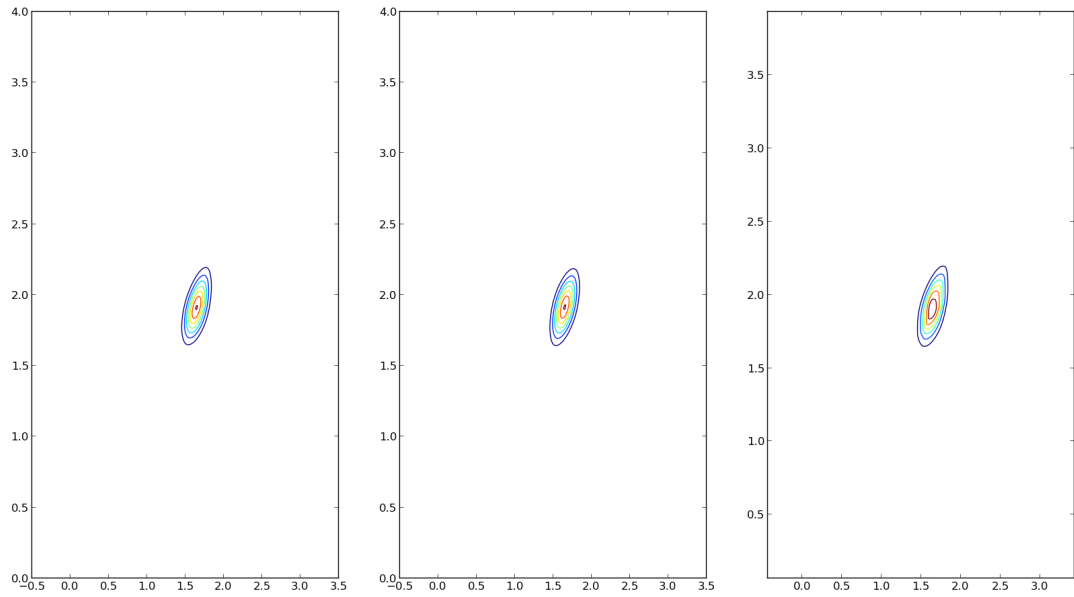


FIGURE 5.6 $V = 10^{-2}$ gauche : Solution exacte. Milieu : notre méthode. Droite : schéma de Vanleer

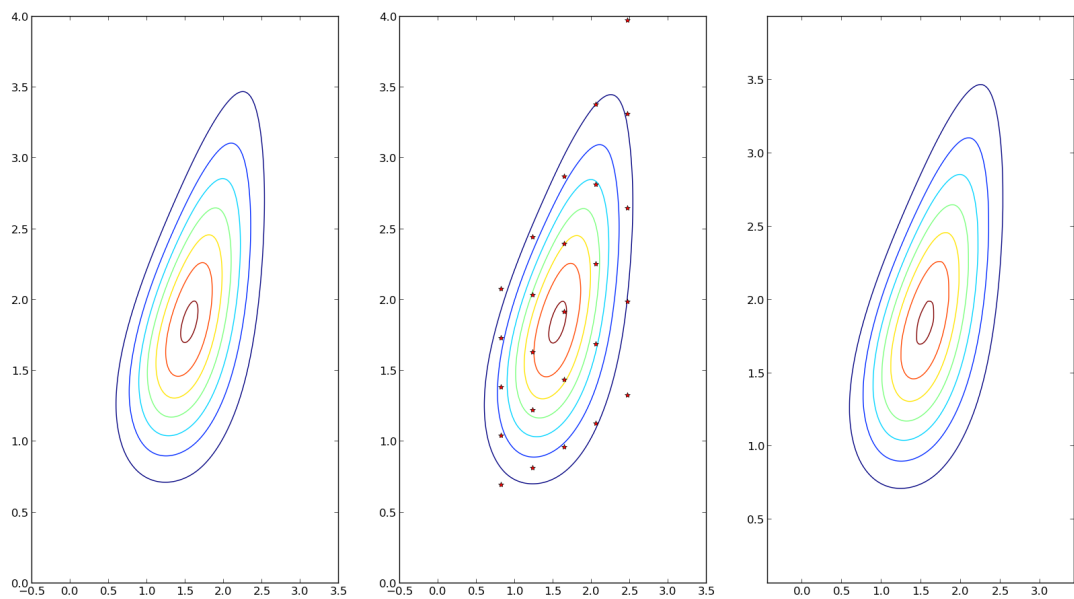


FIGURE 5.7 $V = 0.25 \text{Id}$ gauche : Solution exacte. Milieu : notre méthode avec position des particules. Droite : schéma de Vanleer

5.6 problème de Stokes diphasique

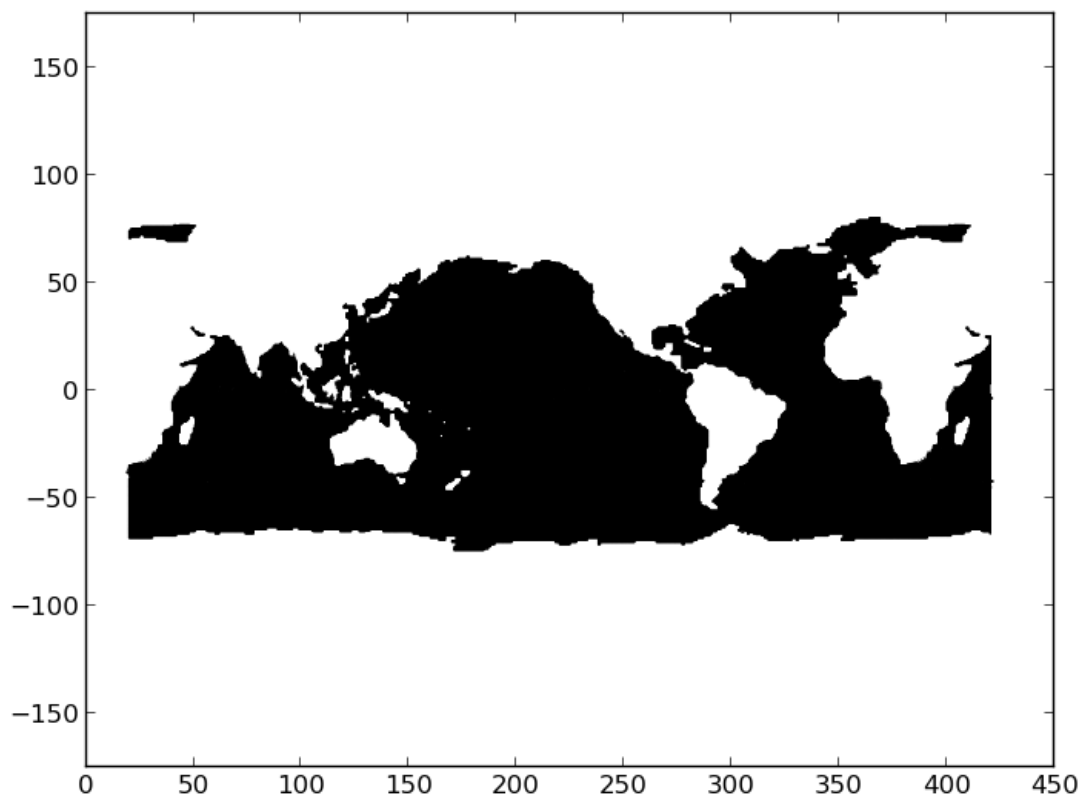


FIGURE 5.8 le domaine de travail disponible

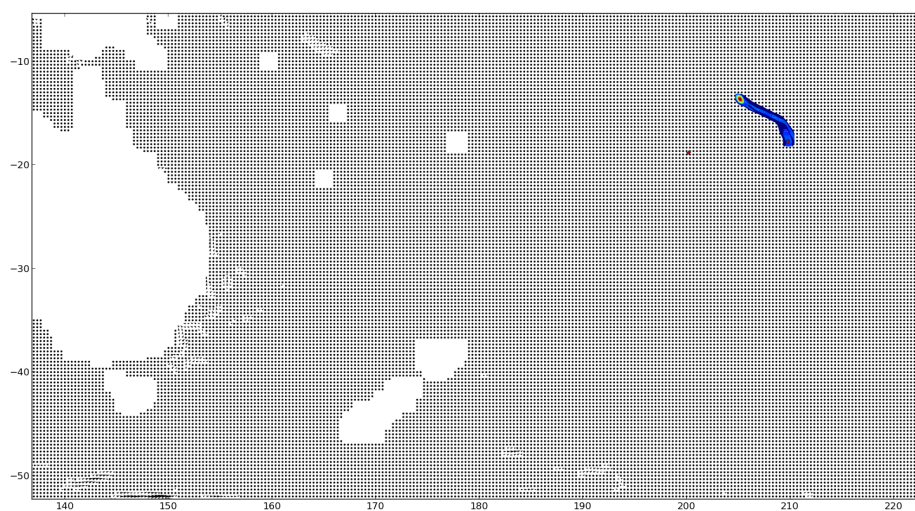


FIGURE 5.9 point rouge : point d'arrivée de Tavae le 10 juillet 2002. Au gauche, Australie et Nouvelle Zélande

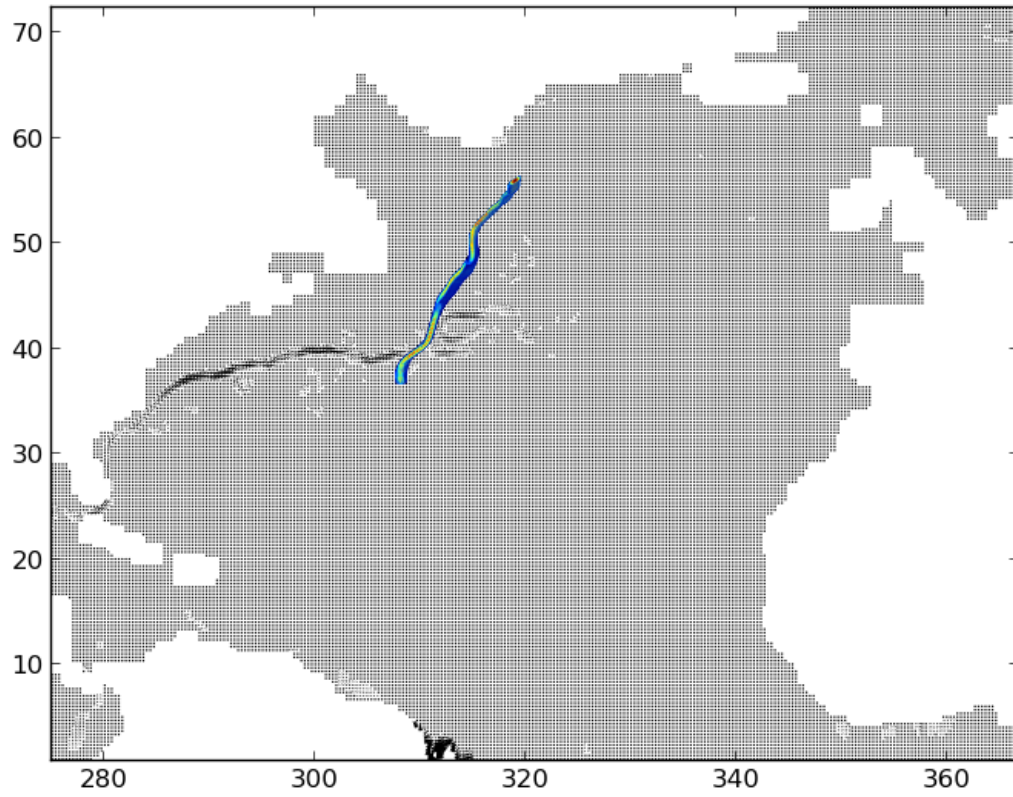


FIGURE 5.10 trajectoire d'une particule plongée dans le gulfstream

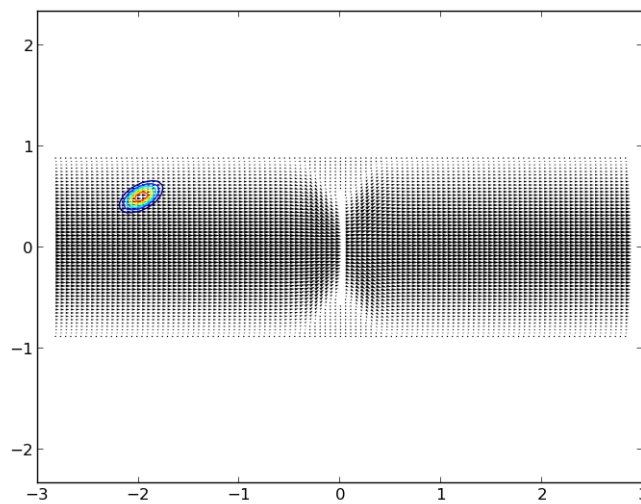


FIGURE 5.11 champ de vecteurs et densité initiale

5.6 problème de Stokes diphasique

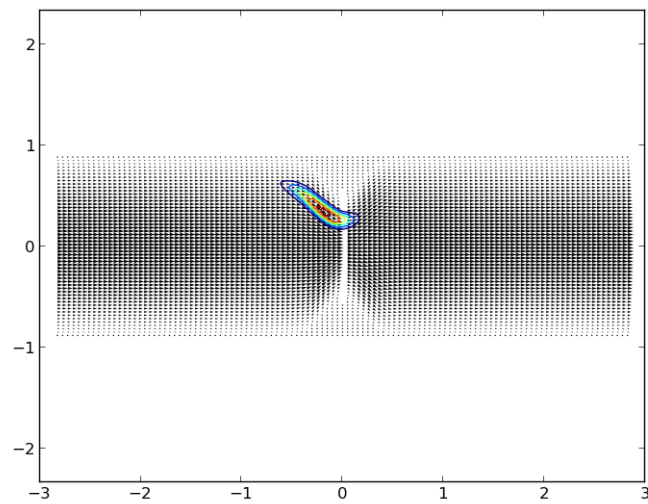


FIGURE 5.12 résolution par un schéma de Vanleer

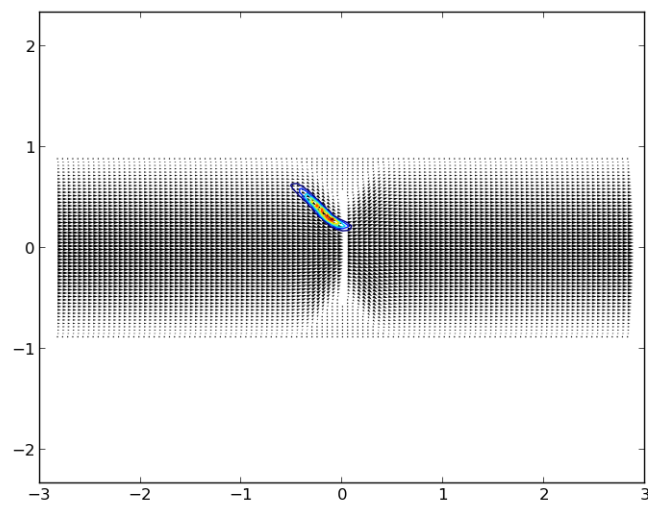


FIGURE 5.13 résolution par notre méthode avec raffinement

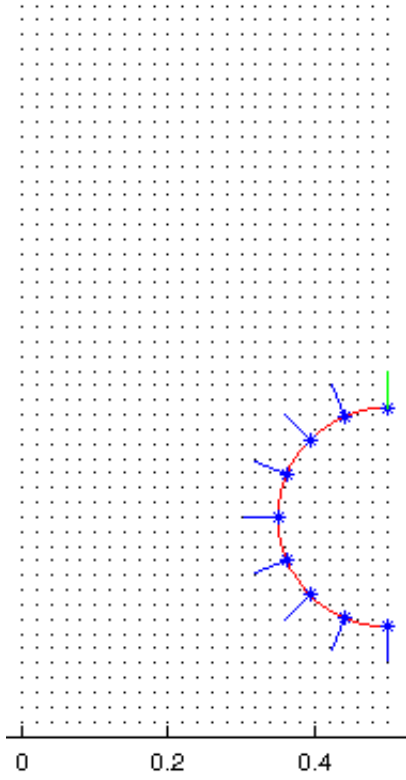


FIGURE 5.14 $t = 0$, avec raffinement

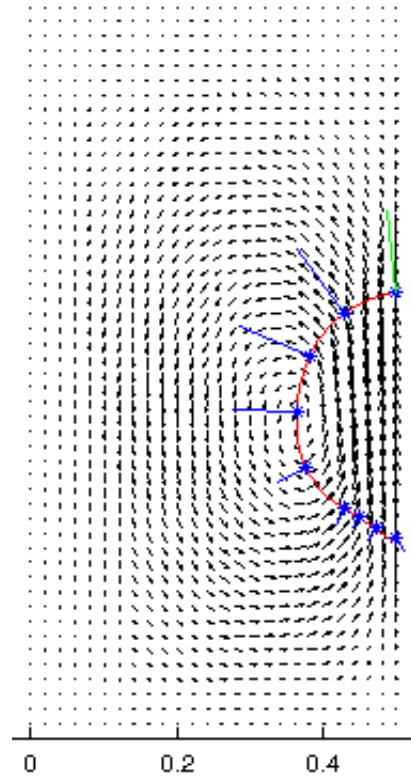


FIGURE 5.16 $t = 3$, avec raffinement

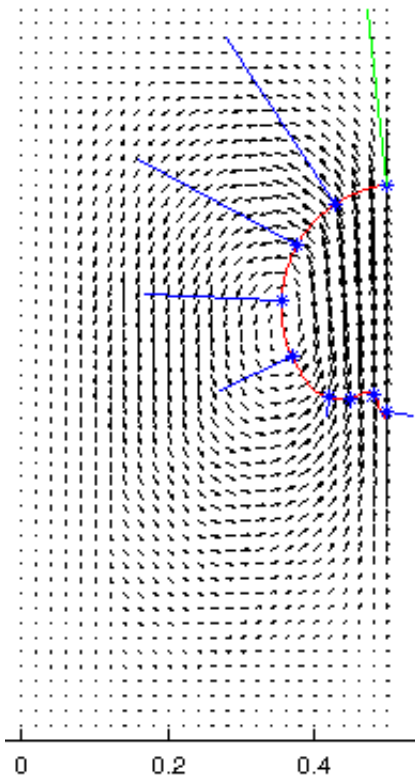


FIGURE 5.15 $t = 6$, avec raffinement

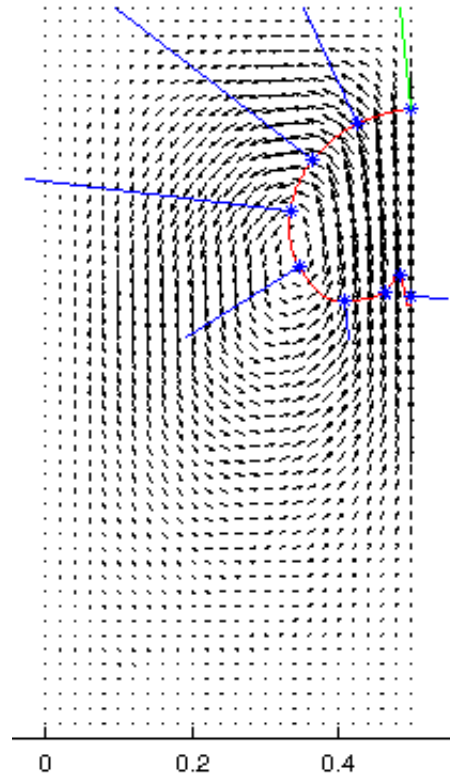


FIGURE 5.17 $t = 9$, avec raffinement

5.6 problème de Stokes diphasique

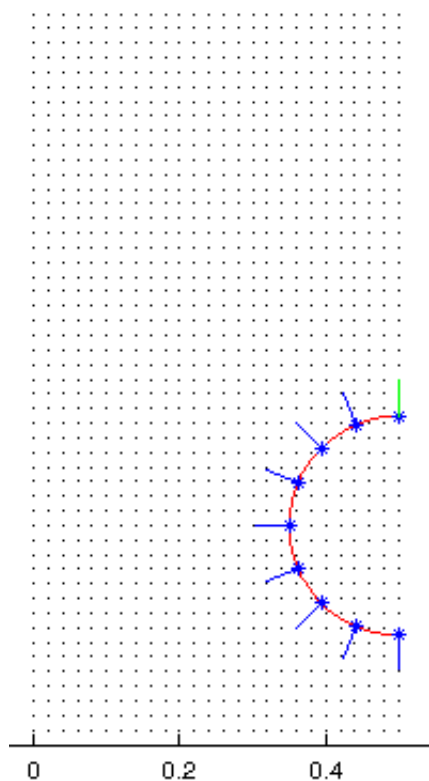


FIGURE 5.18 $t = 0$, sans raffinement

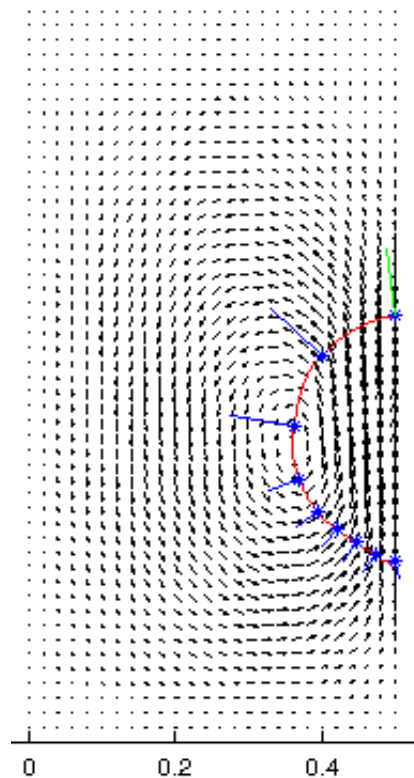


FIGURE 5.20 $t = 3$, sans raffinement

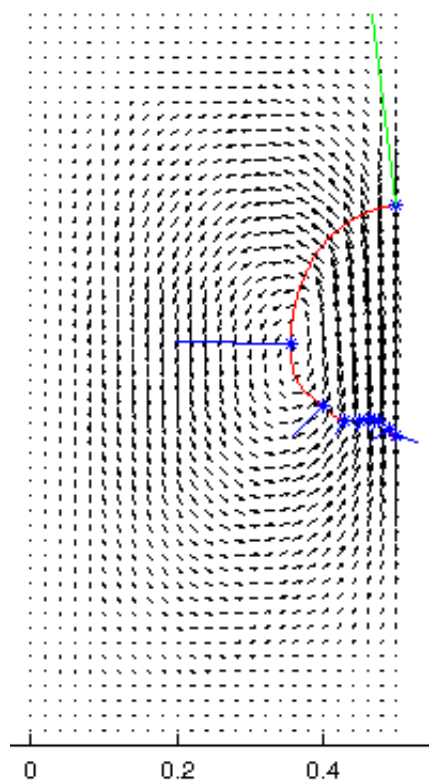


FIGURE 5.19 $t = 6$, sans raffinement

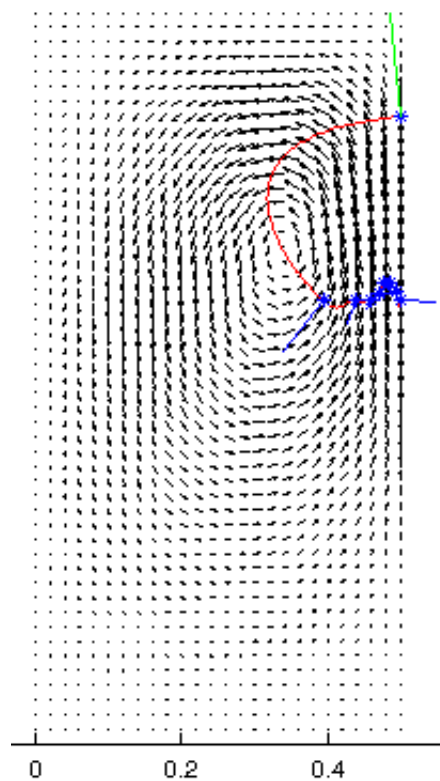


FIGURE 5.21 $t = 9$, sans raffinement

Chapitre 6

Annexes 1

6.1 Fonctions définies par une discrétisation ponctuelle

Nous allons ici, en nous inspirant des outils développés pour les meshless methods et en particulier la PUFEM (Partition of Unity Finite Elements Methods), proposer une méthode de reconstruction de fonction qui permet de tirer profit du phénomènes de surconvergence de la méthode des rectangles pour des fonctions à régulières support compact pour obtenir un bon contrôle sur l'erreur d'approximation ainsi que sur celle de la dérivée.

On pourra consulter [28] pour une introduction aux meshless methods, en particulier le début, et [29] pour la PUFEM. On pourra consulter [22] pour quelques éléments liés aux phénomènes de surconvergence de la méthode des rectangles. Nous n'aurons cependant pas besoin des résultats présentés dans cette dernière référence, puisque nous démontrerons les outils dont nous aurons besoin.

Nous présenterons en détail le cas 1d. le cas 2d se faisant dans les applications à partir de la tensorisation de fonctions 1d, nous nous contenterons de présenter le résultat principal.

6.1.1 Notations, hypothèses et énoncé du théorème

Dans la suite on considère :

- une discrétisation $(x_i)_i$ de \mathbb{R} de pas constant h , avec $x_i = ih$, $i \in \mathbb{Z}$;
- φ spline cubique normalisée à support dans $[-1, 1]$ dont on rappelle qu'elle est définie par

$$\varphi(y) = \frac{1}{3} \times \begin{cases} 8(1 - |x|)^3 & \text{si } \frac{1}{2} \leq x \leq 1 \\ 24|x|^3 - 24x^2 + 4 & \text{si } |x| \leq \frac{1}{2} \end{cases} ;$$

- un rayon d'interaction pour les particules r fixé, supérieur à h de manière à ce que les intervalles $[x_i - r, x_i + r]$ forment un recouvrement de \mathbb{R} .

On pose alors $\varphi_i(x) = \varphi\left(\frac{x-x_i}{r}\right)$ et on définit une approximation de l'unité au moyen des fonctions

$$\psi_i(x) = \frac{\varphi_i(x)}{\sum \varphi_j(x)} ,$$

formule qui a un sens car les intervalles $[x_i - r, x_i + r]$ forment un recouvrement de \mathbb{R} .

Pour f fonction sur \mathbb{R} , on définit une approximation de f par

$$\bar{f}(x) = \sum f(x_i) \psi_i(x) .$$

On souhaite connaître la précision de l'approximation, mais aussi, dans une perspective d'advection par la différentielle, avoir un contrôle de l'erreur entre \bar{f}' et f' . Le théorème suivant que nous allons démontrer va nous les fournir :

Théorème 55

Soit f fonction \mathcal{C}^2 bornée à dérivées bornées. Si $r = 2^j h$, alors on a

$$\|f - \bar{f}\|_{\infty} \leq O(h^2) .$$

Si de plus f est \mathcal{C}^3 , alors on a aussi

$$\|f' - \bar{f}'\|_{\infty} \leq O(h^2)$$

Nous allons maintenant présenter plusieurs lemmes qui vont nous conduire à la démonstration de ce théorème. Le cadre sous-jacent est celui des méthodes de quadrature.

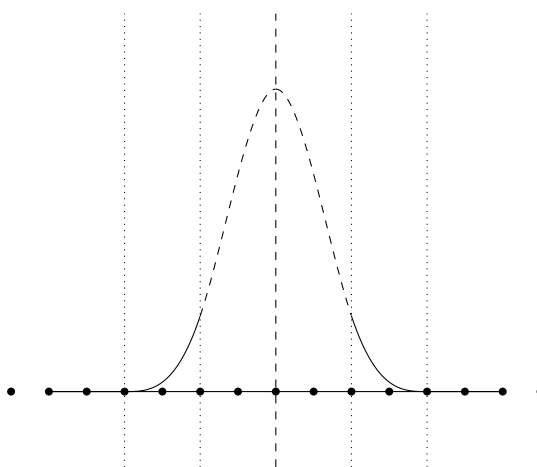
6.1.2 Lemmes préliminaires

Lemme 56

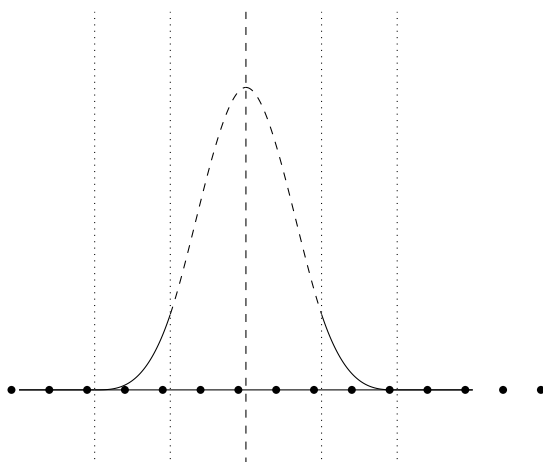
On suppose que $r = 2^j h, j \geq 1$. Soit P un polynôme de degré inférieur ou égal à trois. On a alors :

- la méthode des rectangles à gauche sur les points (x_i) intègre exactement $\int P(y)\varphi\left(\frac{x-y}{r}\right) dy$;
- la méthode des rectangles à gauche sur les points (x_i) intègre exactement $\int P(y)\frac{1}{r}\varphi'\left(\frac{x-y}{r}\right) dy$ et $\int P(y)\varphi'\left(\frac{x-y}{r}\right) dy$.

La démonstration étant calculatoire, elle peut nous faire perdre de vue ce qui se passe. Nous allons donc donner quelques explications avant de l'effectuer. On se place ici dans le cas où l'on regarde $\int \varphi\left(\frac{x-y}{r}\right) dy$. La figure ci-dessous représente la fonction $y \mapsto \varphi\left(\frac{x-y}{r}\right)$ dans le cas où x est un des x_i . Les droites en pointillés matérialisent les points de raccordements de la spline, et la droite en petit trait l'axe de symétrie. Les points noirs représentent les points x_i . Dans ce cas, on aura par formule des rectangles compensations des erreurs deux à deux entre les domaines délimités par les droites, et la quadrature est exacte. Il est important que $r = 2^j h$, ceci pour garantir la présence des points de raccordements dans les x_i .



Maintenant, la figure ci-dessous représente le cas où x n'est pas un point de raccordement de la spline. Il va s'agir de montrer qu'il y a compensation malgré le décalage.



Démonstration

Nous allons nous contenter ici de traiter le cas du polynôme $P(y) = 1$, les deux autres polynômes de la base fonctionnant sur le même principe. On peut supposer que $x \in [0, h[$. On pose

$$I = \int \varphi\left(\frac{x-y}{r}\right) dy \quad .$$

On a $I = r$ et l'approximation de I par les rectangles à gauche est donné par

$$\tilde{I} = \sum_{i \in \mathbb{Z}} \varphi\left(\frac{x-x_i}{r}\right) h \quad .$$

Déterminons les indices pour lesquels les termes de la somme sont non-nuls. Pour x fixé, on a $\varphi\left(\frac{x-y}{r}\right) \geq 0$ si et seulement si $y \in [x-r, x+r]$. Donc

$$\begin{aligned} \varphi\left(\frac{x-x_i}{r}\right) \geq 0 &\iff x - 2^j h \leq x_i \leq x + 2^j h \\ &\iff x - 2^j h \leq ih \leq x + 2^j h \\ &\iff \frac{x}{h} - 2^j \leq i \leq \frac{x}{h} + 2^j \quad . \end{aligned}$$

Comme $0 \leq \frac{x}{h} \leq 1$, on en déduit que

$$\tilde{I} = \sum_{i=-2^j+1}^{2^j} \varphi\left(\frac{x-x_i}{r}\right) h \quad .$$

6.1 Fonctions définies par une discrétisation ponctuelle

Comme $r = 2^j$, on a

$$\begin{aligned}
 -1 \leq \frac{x-x_i}{r} \leq -\frac{1}{2} &\iff \left(2^{j-1} + \frac{x}{h}\right) \leq i \leq \left(2^j + \frac{x}{h}\right) \quad , \\
 -\frac{1}{2} \leq \frac{x-x_i}{r} \leq 0 &\iff \frac{x}{h} \leq i \leq \left(\frac{x}{h} + 2^{j-1}\right) \quad , \\
 0 \leq \frac{x-x_i}{r} \leq \frac{1}{2} &\iff \left(\frac{x}{h} - 2^{j-1}\right) \leq i \leq \frac{x}{h} \quad , \\
 \frac{1}{2} \leq \frac{x-x_i}{r} \leq 1 &\iff \left(\frac{x}{h} - 2^j\right) \leq i \leq \left(\frac{x}{h} - 2^{j-1}\right) \quad .
 \end{aligned}$$

Et de là

$$\begin{aligned}
 \tilde{I} &= \frac{1}{3}h \left(\sum_{i=-2^j+1}^{-2^{j-1}} 8 \left(1 - \left(\frac{x-x_i}{2^j h}\right)\right)^3 \right. \\
 &\quad + \sum_{i=-2^{j-1}+1}^0 24 \left(\frac{x-x_i}{2^j h}\right)^3 - 24 \left(\frac{x-x_i}{2^j h}\right)^2 + 4 \\
 &\quad + \sum_{i=1}^{2^{j-1}} 24 \left(\frac{x_i-x}{2^j h}\right)^3 - 24 \left(\frac{x-x_i}{2^j h}\right)^2 + 4 \\
 &\quad \left. + \sum_{i=2^{j-1}+1}^{2^j} 8 \left(1 - \left(\frac{x_i-x}{2^j h}\right)\right)^3 \right) \quad .
 \end{aligned}$$

On effectue un changement d'indice

$$\begin{aligned}
 \tilde{I} &= \frac{1}{3}h \left(\sum_{i=1}^{2^{j-1}} 8 \left(1 - \left(\frac{x-x_{i-2^j}}{2^j h}\right)\right)^3 \right. \\
 &\quad + \sum_{i=1}^{2^{j-1}} 24 \left(\frac{x-x_{i-2^{j-1}}}{2^j h}\right)^3 - 24 \left(\frac{x-x_{i-2^{j-1}}}{2^j h}\right)^2 + 4 \\
 &\quad + \sum_{i=1}^{2^{j-1}} 24 \left(\frac{x_i-x}{2^j h}\right)^3 - 24 \left(\frac{x-x_i}{2^j h}\right)^2 + 4 \\
 &\quad \left. + \sum_{i=1}^{2^{j-1}} 8 \left(1 - \left(\frac{x_{i+2^{j-1}}-x}{2^j h}\right)\right)^3 \right) \quad .
 \end{aligned}$$

On simplifie

$$\begin{aligned}
 \tilde{I} &= \frac{1}{3}h \left(\sum_{i=1}^{2^{j-1}} 8 \left(\frac{2^j h}{2^j h} - \left(\frac{x - ih + 2^j h}{2^j h} \right) \right)^3 \right. \\
 &\quad + 24 \left(\frac{x - ih + 2^{j-1} h}{2^j h} \right)^3 - 24 \left(\frac{x - ih + 2^{j-1} h}{2^j h} \right)^2 + 4 \\
 &\quad + 24 \left(\frac{x_i - x}{2^j h} \right)^3 - 24 \left(\frac{x - x_i}{2^j h} \right)^2 + 4 \\
 &\quad \left. + 8 \left(\frac{2^j h}{2^j h} - \left(\frac{ih + 2^{j-1} h - x}{2^j h} \right) \right)^3 \right) \\
 &= \frac{1}{3}h \left(\sum_{i=1}^{2^{j-1}} 8 \left(\frac{x_i - x}{r} \right)^3 \right. \\
 &\quad + 24 \left(\frac{x - x_i + \frac{r}{2}}{r} \right)^3 - 24 \left(\frac{x - x_i + \frac{r}{2}}{r} \right)^2 + 4 \\
 &\quad + 24 \left(\frac{x_i - x}{r} \right)^3 - 24 \left(\frac{x - x_i}{r} \right)^2 + 4 \\
 &\quad \left. + 8 \left(\frac{\frac{r}{2} - x_i + x}{r} \right)^3 \right) .
 \end{aligned}$$

On développe

$$\begin{aligned}
 \tilde{I} &= \frac{1}{3}h \left(\sum_{i=1}^{2^{j-1}} \frac{8}{r^3} (x_i - x)^3 \right. \\
 &\quad + \frac{24}{r^3} \left((x - x_i)^3 + 3(x - x_i)^2 \frac{r}{2} + 3(x - x_i) \frac{r^2}{4} \right) \\
 &\quad - \frac{24}{r^2} \left((x - x_i)^2 + 2(x - x_i)r + \frac{r^2}{4} \right) \\
 &\quad + \frac{24}{r^3} (x - x_i)^3 - \frac{24}{r^2} (x - x_i)^2 \\
 &\quad \left. + \frac{8}{r^3} \left((x - x_i)^3 + 3(x - x_i)^2 \frac{r}{2} + 3(x - x_i) \frac{r^2}{4} \right) + 8 \right) .
 \end{aligned}$$

6.1 Fonctions définies par une discrétisation ponctuelle

On regroupe par puissance de $(x_i - x)$

$$\begin{aligned}
 \tilde{I} &= \frac{1}{3}h \left(\sum_{i=1}^{2^{j-1}} (x_i - x)^3 \left(\frac{8}{r^3} - \frac{24}{r^3} + \frac{24}{r^3} - \frac{8}{r^3} \right) \right. \\
 &\quad + (x_i - x)^2 \left(\frac{24}{r^3} \times \frac{3r}{2} - \frac{24}{r^2} - \frac{24}{r^2} + \frac{8}{r^3} \times \frac{3r}{2} \right) \\
 &\quad + (x_i - x) \left(-\frac{24}{r^3} \times \frac{3r^2}{4} + \frac{24}{r^2}r - \frac{8}{r^3} \times \frac{3r^2}{4} \right) \\
 &\quad \left. - \frac{18}{r} + \frac{24}{r} - \frac{6}{r} + 8 - 2 \right) \\
 &= \frac{1}{3}h \sum_{i=1}^{2^{j-1}} (0 + 0 + 0 + 0 + 6)
 \end{aligned}$$

d'où

$$\tilde{I} = \frac{1}{3} \times h \times 6 \times 2^{j-1}h = 2^j h = r = I$$

Lemme 57

Soit f de classe \mathcal{C}^2 bornée à dérivées bornées. On a

$$\frac{1}{r} \sum_i f(x_i) \varphi_i(x) h = f * \frac{1}{r} \varphi \left(\frac{\cdot}{r} \right) + O(r^2) \quad .$$

Démonstration

En regardant $\frac{1}{r} \sum_i f(x_i) \varphi_i(x) h$ comme une formule de quadrature par la méthode des rectangles à gauche, on a

$$\frac{1}{r} \sum_i f(x_i) \varphi_i(x) h = \left(f * \frac{1}{r} \varphi \left(\frac{\cdot}{r} \right) \right) + R \quad .$$

Evaluons le reste. On a

$$R = \frac{1}{r} \sum_i f(x_i) \varphi_i(x) h - \left(f * \frac{1}{r} \varphi \left(\frac{\cdot}{r} \right) \right) \quad .$$

Par utilisation d'une formule de Taylor, on obtient

$$\begin{aligned}
 f * \frac{1}{r} \varphi \left(\frac{\cdot}{r} \right) (x) &= f(x) \int \frac{1}{r} \varphi \left(\frac{x-y}{r} \right) dy + \underbrace{f'(x) \int (x-y) \frac{1}{r} \varphi \left(\frac{x-y}{r} \right) dy}_{=0} \\
 &\quad + \int O((x-y)^2) \frac{1}{r} \varphi \left(\frac{x-y}{r} \right) dy
 \end{aligned}$$

avec

$$O((x-y)^2) \leq \sup |f''| (x-y)^2 \leq \sup |f''| r^2 \quad .$$

Donc on a

$$\begin{aligned} R &= \left| \sum f(x_i) \frac{1}{r} \varphi\left(\frac{x-x_i}{r}\right) h - f(x) \int \frac{1}{r} \varphi\left(\frac{x-y}{r}\right) dy - \int O((x-y)^2) \frac{1}{r} \varphi\left(\frac{x-y}{r}\right) dy \right| \\ &= \left| \sum f(x_i) \frac{1}{r} \varphi\left(\frac{x-x_i}{r}\right) h - f(x) \sum \frac{1}{r} \varphi\left(\frac{x-x_i}{r}\right) h - \int O((x-y)^2) \frac{1}{r} \varphi\left(\frac{x-y}{r}\right) dy \right| \\ &\quad \text{car d'après Lemme 56 on a } \int \frac{1}{r} \varphi\left(\frac{x-y}{r}\right) dy = \sum \frac{1}{r} \varphi\left(\frac{x-x_i}{r}\right) h \\ &\leq \left| \sum (f(x_i) - f(x)) \frac{1}{r} \varphi\left(\frac{x-x_i}{r}\right) h \right| + O(r^2) \\ &\leq \left| \sum (f'(x)(x_i - x) + O((x_i - x)^2)) \frac{1}{r} \varphi\left(\frac{x-x_i}{r}\right) h \right| + O(r^2) \quad . \end{aligned}$$

D'après le Lemme 56, on a

$$f'(x) \sum (x_i - x) \frac{1}{r} \varphi\left(\frac{x-x_i}{r}\right) h = f'(x) \int (y-x) \varphi\left(\frac{x-y}{r}\right) \quad ,$$

donc

$$\begin{aligned} R(h) &\leq \left| \sum O((x_i - x)^2) \frac{1}{r} \varphi\left(\frac{x-x_i}{r}\right) h \right| + O(r^2) \quad \text{car } \int (x-y) \varphi\left(\frac{x-y}{r}\right) = 0 \\ &\leq O(r^2) \quad \text{car } \sum \frac{1}{r} \varphi\left(\frac{x-x_i}{r}\right) h = 1 \text{ d'après Lemme 56} \end{aligned}$$

Lemme 58

Soit f de classe \mathcal{C}^3 a dérivées bornées

$$\sum f(x_i) \frac{1}{r^2} \varphi'\left(\frac{x-x_i}{r}\right) h = f' * \frac{1}{r} \varphi\left(\frac{\cdot}{r}\right) + O(r^2)$$

Démonstration

Le principe est le même que pour la démonstration précédente. Premièrement, on a d'après les propriétés de la convolution

$$f' * \frac{1}{r} \varphi\left(\frac{\cdot}{r}\right) = \int f(y) \frac{1}{r^2} \varphi'\left(\frac{x-y}{r}\right) dy \quad .$$

Le terme de droite est approché par quadrature par

$$\sum f(x_i) \frac{1}{r^2} \varphi'\left(\frac{x-x_i}{r}\right) h$$

6.1 Fonctions définies par une discrétisation ponctuelle

et on veut évaluer l'erreur

$$R = \int f(y) \frac{1}{r^2} \varphi' \left(\frac{x-y}{r} \right) dy - \sum f(x_i) \frac{1}{r^2} \varphi' \left(\frac{x-x_i}{r} \right) h \quad .$$

Pour cela, on effectue dans le membre de gauche le développement de Taylor de $f(y-x+x)$ et dans le membre de droite, celui de $f(x_i-x+x)$. On obtient alors

$$\begin{aligned} R &= \int \left(f(x) + f'(x)(y-x) + \frac{f''(x)}{2}(y-x)^2 + O((y-x)^3) \right) \frac{1}{r^2} \varphi' \left(\frac{x-y}{r} \right) dy \\ &- \sum \left(f(x) + f'(x)(x_i-x) + \frac{f''(x)}{2}(x_i-x)^2 + O((x_i-x)^3) \right) \frac{1}{r^2} \varphi' \left(\frac{x-x_i}{r} \right) h \quad . \end{aligned}$$

On simplifie en utilisant le Lemme 56 et on a

$$|R| \leq Cr^2$$

6.1.3 Démonstration du théorème

A partir de maintenant on suppose la spline cubique normalisée, ce qui facilite les calculs. On commence par l'erreur d'approximation de f :

$$\begin{aligned} |f(x) - \bar{f}(x)| &\leq \left| \sum_i f(x_i) \psi_i(x) - f(x) \right| \\ &\leq \left| \sum_i f(x_i) \frac{\varphi_i(x)}{\sum_j \varphi_j(x)} - f(x) \right| \\ &\leq \left| \sum_i f(x_i) \frac{\varphi_i(x)h}{\sum_j \varphi_j(x)h} - f(x) \right| \\ &\leq \left| \sum_i f(x_i) \frac{1}{r} \varphi_i(x)h - f(x) \right| \quad \text{d'après Lemme 56} \end{aligned}$$

$$\begin{aligned} &\leq \left| f * \frac{1}{r} \varphi \left(\frac{\cdot}{r} \right) - f(x) \right| + O(r^2) \quad \text{d'après Lemme 57} \\ &\leq \| (f * \frac{1}{r} \varphi \left(\frac{\cdot}{r} \right)) - f \|_{\infty} + O(r^2) \\ &\leq O(r^2) \quad \text{d'après les résultats classiques sur la convolution} \\ &\leq O(h^2) \quad \text{car } r = 2^j h \quad . \end{aligned}$$

Passons maintenant à l'erreur pour l'approximation de f' :

$$\begin{aligned}
 |\bar{f}'(x) - f'(x)| &\leq \left| \sum f(x_i) \psi'(x) - f'(x) \right| \\
 &\leq \left| \sum f(x_i) \frac{\varphi'_i(x) \sum_j \varphi_j(x) - \varphi_i(x) \sum_j \varphi'_j(x)}{(\sum_j \varphi_j(x))^2} - f'(x) \right| \\
 &\leq \left| \sum f(x_i) \frac{\varphi'_i(x) \sum_j \varphi_j(x) h - \varphi_i(x) \sum_j \varphi'_j(x) h}{\sum_j \varphi_j(x) \sum_j \varphi_j(x) h} - f'(x) \right| \\
 &\leq \left| \sum f(x_i) \frac{\varphi'_i(x) r - \varphi_i(x) \cdot 0}{\sum_j \varphi_j(x) r} - f'(x) \right| \quad \text{d'après Lemme 56} \\
 &\leq \left| \sum f(x_i) \frac{\varphi'_i(x)}{\sum_j \varphi_j(x)} - f'(x) \right| \\
 &\leq \left| \sum f(x_i) \frac{\varphi'_i(x) h}{\sum_j \varphi_j(x) h} - f'(x) \right| \\
 &\leq \left| \sum f(x_i) \frac{1}{r} \varphi'_i(x) h - f'(x) \right| \quad \text{d'après Lemme 56} \\
 &\leq \left| \sum f(x_i) \frac{1}{r^2} \varphi' \left(\frac{x-x_i}{r} \right) h - f'(x) \right| \quad \text{car } \varphi'_i(x) = \frac{1}{r} \varphi' \left(\frac{x-x_i}{r} \right) \\
 &\leq \left| f' * \frac{1}{r} \varphi \left(\frac{\cdot}{r} \right) - f'(x) + O(r^2) \right| \quad \text{d'après Lemme 58} \\
 &\leq \left\| f' * \frac{1}{r} \varphi \left(\frac{\cdot}{r} \right) - f' \right\|_{\infty} + O(r^2) \\
 &\leq O(r^2) \quad \text{d'après les résultats classiques sur la convolution .}
 \end{aligned}$$

6.2 Le cas 2d

Présentons maintenant la construction du cas 2d. On pose h_x le pas de discrétisation suivant les abscisses, et h_y le pas de discrétisation suivant les ordonnées. On suppose qu'on a $r_x = 2^{j_x} h_x$ et $r_y = 2^{j_y} h_y$, de manière à ce que les rectangles $R_{i,j} = [x_i - r_x, x_i + r_x] \times [y_j - r_y, y_j + r_y]$ forment un recouvrement de \mathbb{R}^2 . On pose

$$\varphi_{i,j}(x, y) = \varphi \left(\frac{x - x_i}{r_x} \right) \varphi \left(\frac{y - y_i}{r_y} \right)$$

et

$$\psi_{i,j}(x, y) = \frac{\varphi_{i,j}(x, y)}{\sum_{k,l} \varphi_{k,l}(x, y)} .$$

Les $\psi_{i,j}$ sont bien définis car les $R_{i,j}$ forment un recouvrement de \mathbb{R}^2 .

6.2 Le cas 2d

Si $F(x, y)$ est une fonction de \mathbb{R}^2 , on définit alors son approximation par

$$\bar{F}(x, y) = \sum_{i,j} F(x_i, y_j) \psi_{i,j}(x, y) \quad .$$

On a le théorème suivant sur la qualité de l'approximation, reposant sur les mêmes principes que le cas 1d :

Théorème 59

Soit F fonction \mathcal{C}^2 bornée à dérivées partielles bornées. Si $r_x = 2^{j_x} h_x$ et $r_y = 2^{j_y} h_y$, alors on a

$$\|F - \bar{F}\|_{\infty} \leq O(h_x^2) + O(h_y^2) \quad .$$

Si de plus F est \mathcal{C}^3 , alors on a aussi

$$\|dF - d\bar{F}\|_{\infty} \leq O(h_x^2) + O(h_y^2) \quad .$$

Chapitre 7

Annexes 2

7.1 Géométrie hyperbolique : le modèle du demi-plan de Poincaré

Nous introduisons ici le demi-plan de Poincaré. Nous nous contenterons ici de présenter la structure d'une des constructions possibles ainsi que les principaux résultats. Nous donnerons à la fin une liste de références commentées.

7.1.1 Construction de la métrique

Soit $E = \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$ le demi-plan supérieur, qu'on munit de la métrique héritée de la norme euclidienne sur \mathbb{R}^2 . Soit $J : E \mapsto \mathbb{R}$ définit par $J(x, y) = \frac{1}{y^2}$.

Définition 11

Soit $\gamma(t)$ courbe de E définit sur $[a, b]$. On définit la longueur de γ par rapport à J par :

$$l_J(\gamma) = \int_a^b \|\gamma'(t)\|_2 J(\gamma(t)) dt \quad .$$

Proposition 60

Soit $p, q \in E$. on définit

$$d(p, q) = \inf_{\gamma} l_J(\gamma) \quad \text{sur les courbes reliant } p \text{ à } q.$$

d est une distance sur E . (E, d) est appelé demi-plan de Poincaré.

7.1.2 Géométrie de l'espace

Définition 12

Soit γ une courbe de E définie sur $I \subset \mathbb{R}$. On dit que γ est une géodésique si pour tout $s_1, s_2 \in I$, on a

$$d(\gamma(s_1), \gamma(s_2)) = l_J(\gamma|_{[s_1, s_2]})$$

Remarque 61

Cela revient à dire que la courbe réalise le plus court chemin entre deux de ses points. Cette notion est géométrique, elle ne dépend pas d'une reparamétrisation de la courbe. On identifiera donc fréquemment la géodésique à son image. On parlera de géodésique maximale lorsque qu'il n'est pas possible de la prolonger.

Théorème 62

Les géodésiques maximales de (E, d) sont les demi-droites (au sens euclidien) perpendiculaires à l'axe des abscisses et les demi-cercles (au sens euclidien) orthogonaux à l'axe des abscisses. (voir figure ci-dessous)

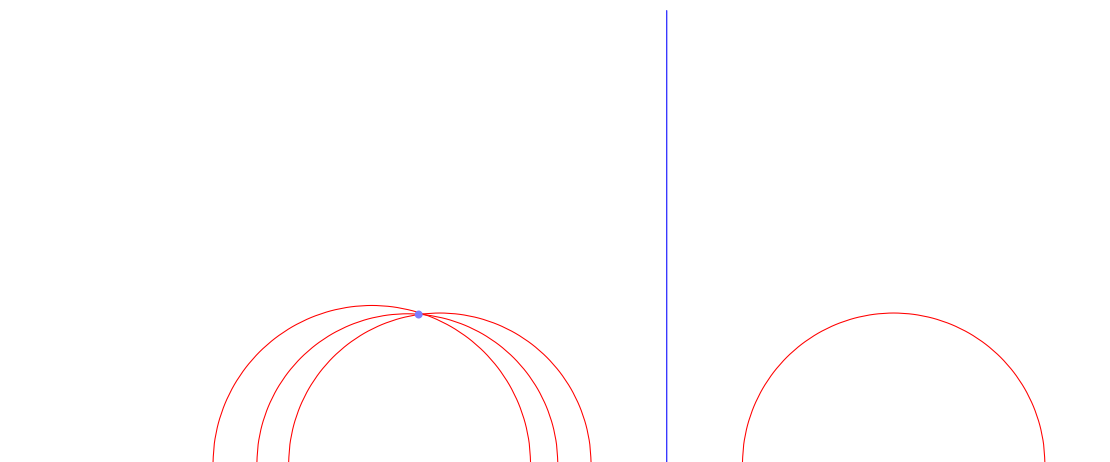
Remarque 63

Donnons les éléments clés de la démonstration de ce théorème, qui est importante : c'est en son sein que se cache les éléments fondamentaux de géométrie, les actions de groupe et les invariants. Les éléments clés sont :

- la détermination des automorphismes (bijections biholomorphes) du demi-plan et leurs propriétés ;
- montrer que la longueur d'une courbe est invariante sous l'action d'un automorphisme et en déduire l'invariance de la métrique sous l'action du groupe des automorphismes ;
- montrer la conservation de l'ensemble des cercles-droites par l'action du groupe des automorphismes.

On trouvera ci dessous quelques exemples de géodésiques. On peut notamment voir qu'il existe une infinité de parallèles à une géodésique donnée passant par un point.

7.1 Géométrie hyperbolique : le modèle du demi-plan de Poincaré



On a une expression explicite de la métrique donnée par le théorème suivant

Théorème 64

La distance de Poincaré entre les points du demi-plan supérieur (x_1, y_1) et (x_2, y_2) est donné par

$$\operatorname{arccosh} \left(1 + \frac{(x_2 - x_1)^2 + (y_2 - y_1)^2}{2y_1 y_2} \right) .$$

7.1.3 Références

On introduit généralement la géométrie hyperbolique au moyen du demi-plan de Poincaré ou du disque de Poincaré. Ces deux espaces sont isométriques et leur construction est similaire dans l'esprit. Voici quelques références :

- le premier chapitre de [21] est une présentation du demi-plan de Poincaré ;
- pour l'approche utilisant les courbes, on trouvera dans la section 6.9 de [16] une présentation assez condensée et complète, mais dans le cadre du disque ;
- il existe une deuxième approche, basée sur la géométrie projective, moins intuitive mais qui permet de mieux saisir les éléments sous-jacent, passant par le birapport. On pourra trouver cette approche dans le modèle du demi-plan dans l'exercice V.5 de [5].
- le calcul explicite de la distance se trouve dans [35].

7.2 Éléments aux statistiques

On pourra se référer à [33] pour une introduction détaillée aux statistiques.

7.2.1 Introduction

Supposons qu'une entreprise utilise de nombreux exemplaires d'un composant électronique donné. Elle souhaite pouvoir estimer le nombre de pannes afin de gérer au mieux ses stocks de composants de rechanges.

Comme toujours en mathématiques appliquées, la première étape est une étape de modélisation. Face à un phénomène aléatoire, c'est à dire qui reproduit dans des conditions standards amène des résultats différents, l'outil naturel est le champ des probabilités/statistiques.

Si on effectue l'approximation courante que la durée de vie d'un composant est un phénomène sans mémoire et donc modélisé par une loi exponentielle, l'entreprise va donc considérer que la durée de vie d'un composant i sera représenté par une variable aléatoire X_i suivant une loi exponentielle de paramètre λ inconnu. L'espace de travail naturel est donc l'ensemble des lois exponentielles, parmi lesquelles on va chercher, à partir des données que nous avons, à trouver celle modélisant la durée de vie d'un composant.

7.2.2 Modèle statistique, estimation

Nous allons maintenant préciser l'intuition, le travail sur un ensemble de lois de probabilités.

Définition 13

Une modèle statistique est la donnée de Ω un espace mesurable, \mathcal{A} une tribu sur Ω , $(\mathbb{P}_\theta)_{\theta \in \Theta}$ une famille de lois de probabilités. On notera $(\Omega, \mathcal{A}, \mathbb{P}_\theta, \theta \in \Theta)$ ce modèle. Il sera dit paramétrique s'il existe $d \in \mathbb{N}^*$ tel que $\Theta \subset \mathbb{R}^d$. Sinon, on parlera de modèle non-paramétrique.

7.2 Éléments aux statistiques

Dans le cas de notre exemple, on peut identifier les lois avec leurs densités et on a le modèle statistique suivant :

$$(\mathbb{R}, \mathcal{B}(\mathbb{R}), x \mapsto \lambda \exp(-\lambda x) \mathbf{1}_{\mathbb{R}^+})$$

La définition suivante permet de repérer les modèles permettant de remonter au paramètre théorique.

Définition 14

Soit $(\Omega, \mathcal{A}, \mathbb{P}_\theta, \theta \in \Theta)$ un modèle statistique. Si l'application $\theta \mapsto \mathbb{P}_\theta$ est injective, on dira que le modèle est identifiable.

Définition 15

Soit $(\Omega, \mathcal{A}, \mathbb{P}_\theta, \theta \in \Theta)$ un modèle statistique. $X = (X_1, \dots, X_n)$ est un n -échantillon si les X_i sont une famille de variable aléatoire iid à valeurs dans Ω dont la loi commune appartient à la famille $(\mathbb{P}_\theta)_{\theta \in \Theta}$.

Dans le cas de l'exemple de départ, cela correspond à la durée de vie de n composants.

On s'intéresse maintenant à la question de l'estimation, c'est à dire retrouver le paramètre théorique ou une valeur en dépendant (espérance, écart-type, etc.). Dans la suite, g sera une fonction de Θ dans \mathbb{R}^p euclidien, pour un certain p dans \mathbb{N}^* .

Définition 16

Soit $(\Omega, \mathcal{A}, \mathbb{P}_\theta, \theta \in \Theta)$ un modèle statistique, X un n -échantillon de loi \mathbb{P}_θ . Un estimateur \hat{g}_n de $g(\theta)$ est une application mesurable en X indépendante de θ . \hat{g}_n est donc de la forme $\hat{g}_n = h(X)$ où h est une application mesurable de Ω^n dans \mathbb{R}^p .

7.2.3 Deux méthodes de constructions d'estimateurs : la méthode des moments et l'EMV

Définition 17

Soit $(\Omega, \mathcal{A}, \mathbb{P}_\theta, \theta \in \Theta)$ un modèle statistique, $X = (X_1, \dots, X_n)$ un échantillon de loi \mathbb{P}_θ , ϕ une fonction telle que $\phi(X_i) \in L^1(\mathbb{P}_\theta)$ pour tout $\theta \in \Theta$. Si

$$g(\theta) = \mathbb{E}_\theta[\phi(X_i)] \quad ,$$

la méthode des moments consiste à définir l'estimateur

$$\hat{g}_n = \frac{1}{n} \sum_{i=1}^n \phi(X_i) \quad .$$

Nous allons maintenant présenter un estimateur important. L'idée sous-jacente est de regarder $\rho(x, \theta)$ comme une fonction en θ , idée qui nous conduira à la définition de l'information de Fisher.

Définition 18

Soit $(\Omega, \mathcal{A}, \mathbb{P}_\theta, \theta \in \Theta)$ un modèle statistique tel que $(\mathbb{P}_\theta)_{\theta \in \Theta}$ soit dominé par une mesure μ . On notera $\rho(x, \theta)$ la densité de \mathbb{P}_θ par rapport à μ .

Soit $X = (X_1, \dots, X_n)$ est un n -échantillon de loi commune \mathbb{P}_θ , dont on note $\rho(x, \theta)$ la densité associée. La vraisemblance V_X est défini par

$$V_X(\theta) = \prod_{i=1}^n \rho(X_i, \theta) \quad .$$

On définit alors l'estimateur du maximum de vraisemblance par

$$\hat{\theta}_n = \max_{\theta \in \Theta} V_X(\theta)$$

Remarque 65

En pratique on maximisera plutôt dans le log de la vraisemblance, appelé log-vraisemblance, pour travailler avec une somme plutôt qu'avec un produit.

7.2.4 Information de Fisher 1d

On suppose maintenant que :

- Θ est un ouvert de \mathbb{R} ;
- les lois du modèle sont dominés ;
- le support de $\rho(\cdot, \theta)$ ne dépend pas de θ ;
- pour presque tout x , l'application $\theta \mapsto \rho(x, \theta)$ est différentiable en θ et on peut dériver sous le signe intégral

Définition 19

Soit $(\Omega, \mathcal{A}, \mathbb{P}_\theta, \theta \in \Theta)$ un modèle statistique tel que $(\mathbb{P}_\theta)_{\theta \in \Theta}$ soit dominée par une mesure μ . On notera $\rho(x, \theta)$ la densité de \mathbb{P}_θ par rapport à μ . On définit

7.2 Éléments aux statistiques

l'information de Fisher de $\rho(x, \theta)$ par

$$I(\theta) = \mathbb{E}_\theta \left[(\partial_\theta \log \rho(x, \theta))^2 \right] \quad .$$

Si X est une variable aléatoire de loi \mathbb{P}_θ , on définit l'information de Fisher de la loi par

$$I_X = I(\theta)$$

Donnons une propriété importante de l'information de Fisher, qui nous dit que plus on a d'observations, plus on a d'informations.

Proposition 66 ([33])

Si (X_1, \dots, X_n) sont des variables aléatoires indépendantes identiquement distribuées de paramètre θ , alors

$$I_{X_1, \dots, X_n} = nI_{X_1}$$

La méthode du maximum de vraisemblance consiste à choisir la méthode maximisant le log de la vraisemblance, donc tel que

$$\mathbb{E} [\partial_\theta \ln(\rho(X\theta))] = 0 \quad .$$

L'information de Fisher est définie comme la variance de la log-vraisemblance. Intuitivement, plus il sera petit, plus il sera probable que l'estimateur du maximum de vraisemblance est proche du paramètre théorique.

Le théorème suivant prolonge ce constant en nous disant de plus que l'information de Fisher nous donne une borne sur la précision que peut atteindre un estimateur suivant la taille de l'échantillon. Gardons nous bien de conclure que l'estimateur du maximum de vraisemblance est le « meilleur », car le sujet est bien plus compliqué que cela, avec ne serait-ce que la question du biais que nous n'avons par abordé ici.

Théorème 67 (borne de Cramer-Rao 1d, ([33]))

Soit $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_\theta, \theta \in \Theta)$ un modèle statistique, (X_1, \dots, X_n) un n -échantillon, $\hat{\theta}_n$ un estimateur de θ . alors on a

$$\mathbb{E}_\theta \left[(\hat{\theta}_n - g(\theta))^2 \right] \geq \frac{1}{nI_{X_1}} \quad .$$

En particulier, si $\hat{\theta}_n$ est sans biais ($\mathbb{E} [\hat{\theta}_n] = \theta$), alors

$$\text{Var} [\hat{\theta}_n] \geq \frac{1}{nI_{X_1}} \quad .$$

Ce théorème admet une formulation en dimension supérieure, utilisant la matrice de Fisher (Définition 7, page 52), qu'on donne dans le cas simplifié des estimateurs sans biais.

Théorème 68 (borne de Cramer-Rao)

Soit $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{P}_\theta, \theta \in \Theta)$ un modèle statistique, (X_1, \dots, X_n) un n -échantillon de paramètre θ , $\hat{\theta}_n$ un estimateur sans biais de θ . alors on a

$$\text{Var} [\hat{\theta}_n] \geq \frac{1}{n} I_{X_1}^{-1} \quad .$$

où $\text{Var} [\hat{\theta}_n]$ désigne la matrice de variance-covariance de $\hat{\theta}$ et l'inégalité $A \geq B$ est définie par $A - B$ définie positive.

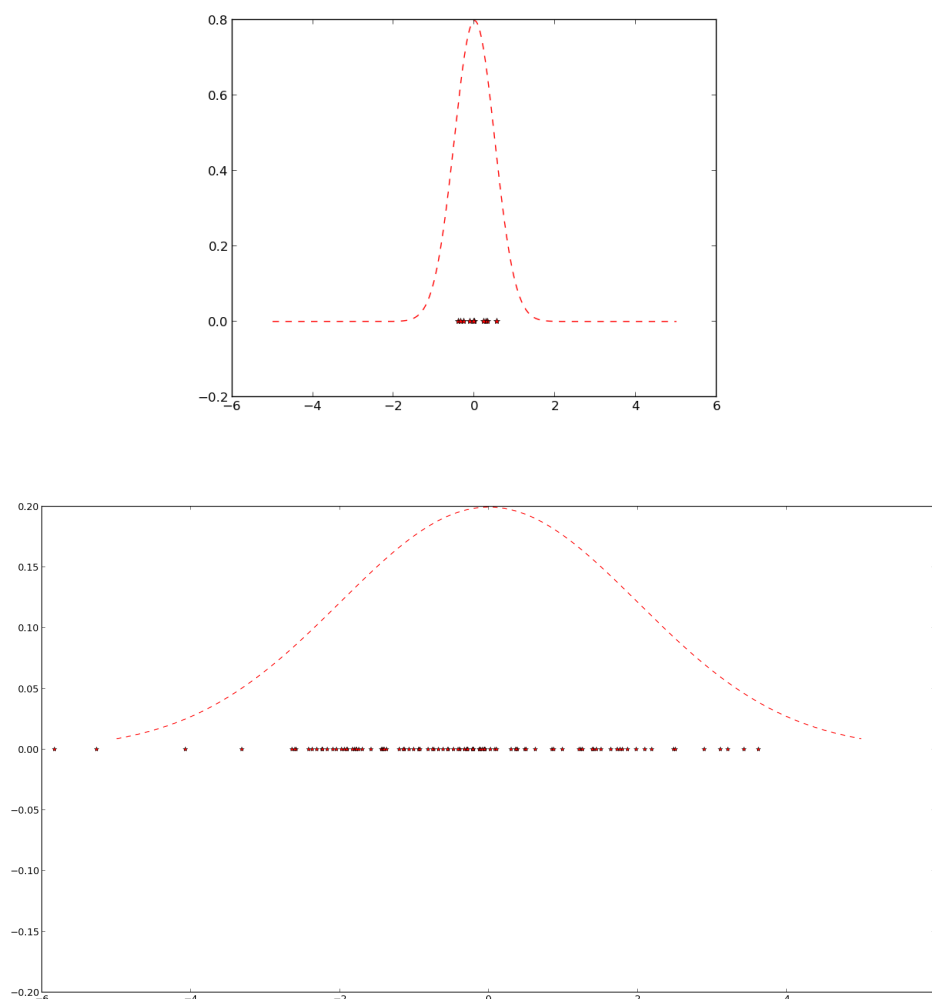
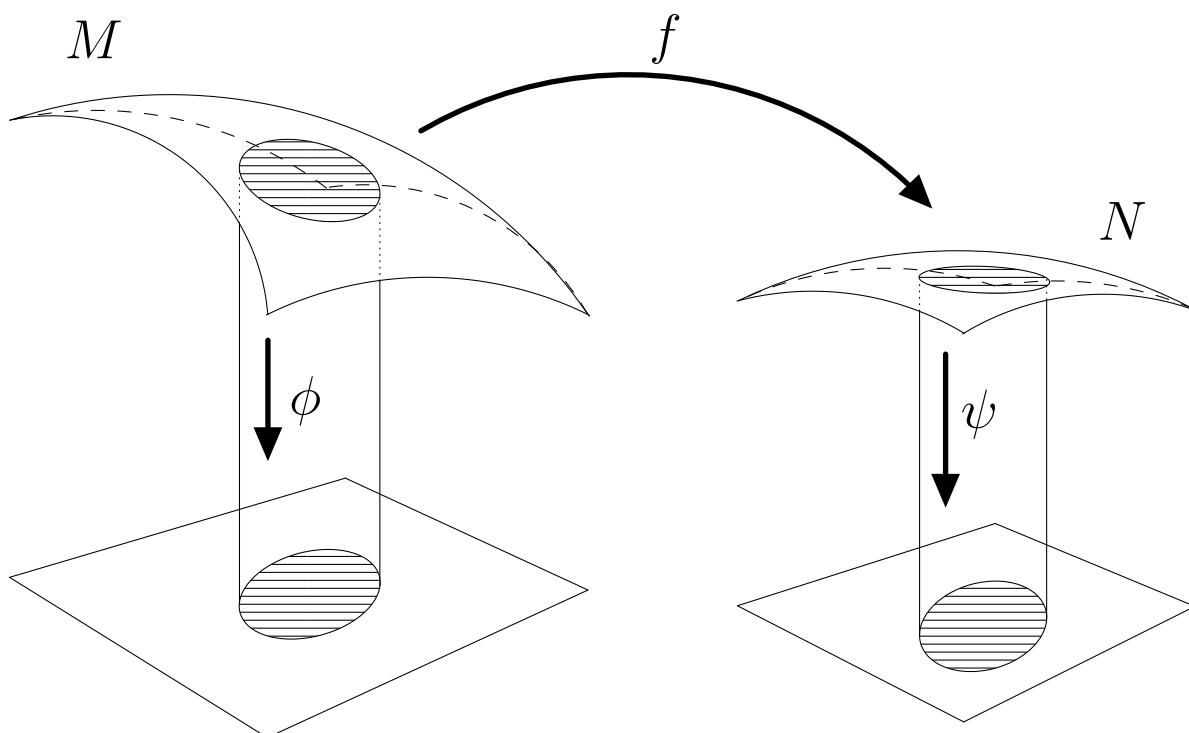


FIGURE 7.1 deux exemples d'échantillons

7.1 donnent une interprétation visuelle de ces résultats, dans le cas des gaussiennes 1d, où μ est connu et fixé et l'on cherche à identifier σ . Dans ce cas, l'information de fisher est donné par $\frac{1}{\sigma^2}$. On voit sur la figure que quand σ est petit, la gaussienne resserrée est perçue avec un petit échantillon, là où une gaussienne plus étalée nécessitent plus de points.

7.3 Géométrie différentielle et riemannienne

7.3.1 Généralités



Nous introduirons les notations et rappellerons les premières définitions au moyen du dessin ci-dessus. Les variétés M et N seront supposées de dimensions respectives m et n :

- ϕ et ψ sont des cartes. Si ϕ_1 et ϕ_2 sont deux cartes de M admettant un ouvert commun dans leur ensemble de définition, on dit qu'elles sont \mathcal{C}^∞ compatibles si $\phi_2 \circ \phi_1^{-1}$ est \mathcal{C}^∞ . On supposera dans la suite que les variétés sont \mathcal{C}^∞ , c'est à dire qu'elles admettent un recouvrement par des cartes \mathcal{C}^∞

compatibles ;

- f est \mathcal{C}^∞ si $\psi \circ f \circ \phi^{-1}$ est \mathcal{C}^∞ .

Dans le but de pouvoir définir explicitement la différentielle de f , il faut définir un espace tangent. Nous rappelons ici les deux approches possibles.

Soit $p \in M$ fixé, Γ_p l'ensemble des courbes \mathcal{C}^1 de M valant p en 0. On définit une relation d'équivalence sur Γ_p en posant

$$c_0 \sim c_1 \iff (\phi \circ c_0)'(0) = (\phi \circ c_1)'(0) \quad .$$

On définit alors l'espace tangent $T_p M$ (de manière ensembliste pour l'instant) comme l'ensemble des classes d'équivalences. La définition ne dépend pas du paramétrage.

On a une bijection naturelle entre $T_p M$ et \mathbb{R}^m : si $\xi \in T_p M$, on pose $\theta_p(\xi) = (\phi \circ c)'(0)$ où c est un représentant de la classe d'équivalence. On se sert ensuite de cette application pour définir une structure d'espace vectoriel sur $T_p M$. On munit ce dernier d'une base canonique en prenant $\theta_p^{-1}(e_i)$, où e_i désigne le i -ième vecteur de la base canonique de \mathbb{R}^m .

Présentons maintenant la deuxième approche. Si on se fixe un point p d'un ouvert de \mathbb{R}^m , à chaque vecteur v de \mathbb{R}^m , on peut associer une application linéaire de $\mathcal{C}^\infty(\mathbb{R}^m, \mathbb{R})$ sur \mathbb{R} en considérant la dérivée directionnelle $\partial_v f(p) = \sum_{i=1}^m v_i \partial_i f(p)$. On peut généraliser cette approche aux variétés. On définit l'espace tangent $T_p M$ comme l'ensemble des applications linéaires de $\mathcal{C}^\infty(M, \mathbb{R})$ à valeurs dans \mathbb{R} vérifiant la règle de Leibniz : $D(fg) = f(p)D(g) + g(p)D(f)$.

Si on a une courbe, on définit alors facilement une dérivée suivant cette courbe. Soit \bar{c} une classe d'équivalence de Γ_p , $f \in \mathcal{C}^\infty(M, \mathbb{R})$. On a

$$f \circ c = \underbrace{(f \circ \phi^{-1})}_{\mathbb{R}^d \rightarrow \mathbb{R}} \circ \underbrace{(\phi \circ c)}_{\mathbb{R} \rightarrow \mathbb{R}^d} \quad .$$

On pose $\tilde{f} = f \circ \phi^{-1}$ et $\gamma = \phi \circ c$. On a

$$(f \circ c)' = \nabla \tilde{f}_{\gamma(t)} \cdot \partial_1 \tilde{f}(\gamma(t))$$

qui ne dépend pas du choix du représentant de la classe d'équivalence. On pose alors

$$D_c(f) = \nabla \tilde{f}_{\gamma(t)} \cdot \partial_1 \tilde{f}(\gamma(t)) \quad .$$

7.3 Géométrie différentielle et riemannienne

On dit que F est un **champ de vecteur** sur M si pour $p \in M$, $F(p) \in T_pM$. En particulier, $p \mapsto (\partial_i)_p$ est un champ de vecteurs.

On dit que M est une **variété riemannienne** si pour tout p , on a un produit scalaire g_p défini sur T_pM tel que pour tous champs de vecteur X, Y , l'application $p \mapsto g_p(X(p), Y(p))$ est \mathcal{C}^∞ . On notera parfois $g(p)$ le produit scalaire correspondant.

Dans la suite, on supposera que M est une variété riemannienne.

7.3.2 Connexions affines

7.3.2.1 Approche intuitive

Comment, si p et q sont deux points appartenant au même ouvert de carte d'une variété M , comparer $v_p = F(p)$ et $v_q = F(q)$ puisqu'ils appartiennent à deux espaces qui n'ont rien à voir, T_pM et T_qM ? De la même manière, si $\gamma(t)$ est une courbe, comment comparer $\gamma'(t)$ et $\gamma'(t + dt)$ qui appartiennent respectivement à $T_{\gamma(t)}M$ et $T_{\gamma(t+dt)}M$? La réponse à cette question est apportée par les connexions qui vont nous permettre de déplacer localement les espaces tangents les uns sur les autres de manière continue. Nous allons commencer par une introduction informelle et intuitive.

Soit p et q deux points « infiniment proches » de M appartenant au même ouvert de la carte ϕ . On pose $d\phi = \phi(p) - \phi(q)$. On écrira parfois formellement (la somme n'ayant pas de sens et les éléments appartenant à deux ensembles différents) $q = p + d\phi$.

On considère une application linéaire $\Pi_{p,q}: T_{p+d\phi}M \rightarrow T_pM$ qui tend vers l'identité quand $d\phi$ tend vers zéro. Donc $\Pi_{p,q}(\partial_j(p + d\phi))$ est envoyé sur un vecteur proche de $\partial_j(p)$. De là, en considérant la différence, on doit donc avoir une expression de la forme

$$\Pi_{p,q}(\partial_j(p + d\phi)) - \partial_j(p) = \sum_{k=1}^m \sum_{i=1}^m d\phi_i \Gamma_{ij}^k(p) \partial_k(p)$$

en négligeant les termes de la forme $d\phi_i d\phi_j$ comme les points sont infiniment proches.

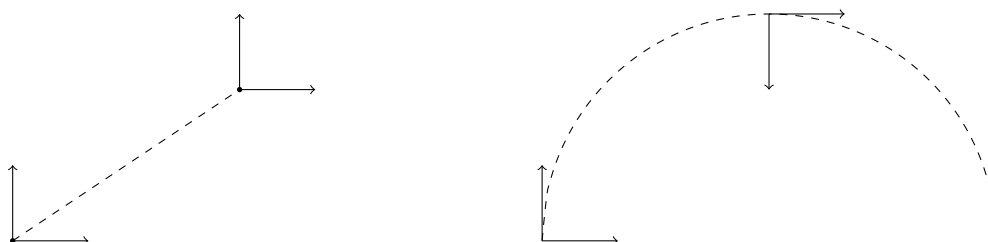
Ainsi, une correspondance continue entre les espaces tangents sur un voisinage de M nous est donné par m^3 fonctions Γ_{ij}^k , appelées **coefficients de la connexion affine**.

Si $d\phi = \varepsilon e_i$, alors on peut écrire

$$\frac{\Pi_{p,q}(\partial_j(p + \varepsilon \partial_i(p))) - \partial_j(p)}{\varepsilon} = \sum_{k=1}^m \Gamma_{ij}^k(p) \partial_k(p) \quad .$$

Le terme de droite peut donc être interprété comme l'évolution du champ de vecteur ∂_j quand p se déplace suivant le champ de vecteur ∂_i , champ de vecteur qu'on notera $\nabla_{\partial_i} \partial_j$

Remarquons qu'il existe plusieurs façon de choisir une connexion, choix décidé par ce que l'on souhaite observer. La figure ci dessous montre deux exemples de déplacement des plans tangents dans le cas du plan.



7.3.2.2 Formalisation

Soit $T(M)$ l'ensemble des champs de vecteurs $\mathcal{C}^\infty(M)$. On dit que

$$\begin{aligned} \nabla: T(M) \times T(M) &\rightarrow T(M) \\ (A, B) &\mapsto \nabla_A B \end{aligned}$$

est une **connexion affine** sur M si c'est une application bilinéaire vérifiant de plus pour toute fonction de $\mathcal{C}^\infty(M, \mathbb{R})$:

- $\nabla_{fA} B = f \nabla_A B$;
- $\nabla_A (fB) = f \nabla_A B + \nabla_A f B$ où $\nabla_A f$ désigne la dérivée de f suivant A .

Intuitivement $\nabla_A B$ correspond à la dérivation du champ de vecteur B suivant A .

7.3 Géométrie différentielle et riemannienne

Si on a une carte et donc des coordonnées, on pose $\nabla_{\partial_j} \partial_k = \sum_i \Gamma_{jk}^i \partial_i$ et on peut alors exprimer la connexion à partir de ses coefficients

$$\begin{aligned}
 \nabla_A B &= \nabla_A (\sum_{k=0}^m B_k \partial_k) \\
 &= \sum_{k=1}^m ((\nabla_A B_k) \partial_k + B_k \nabla_A \partial_k) \\
 &= \sum_{k=1}^m \sum_{j=1}^m ((\nabla_{A_j \partial_j} B_k) \partial_k + B_k \nabla_{A_j \partial_j} \partial_k) \\
 &= \sum_{k=1}^m \sum_{j=1}^m \left(A_j \underbrace{\nabla_{\partial_j} B_k}_{=\partial_j B_k} \right) \partial_k + B_k A_j \underbrace{\nabla_{\partial_j} \partial_k}_{=\sum_i \Gamma_{jk}^i \partial_i} \\
 &= \sum_k \sum_j ((A_j \partial_j B_k) \partial_k + \sum_i B_k A_j \Gamma_{jk}^i \partial_i) \\
 &= \sum_j \sum_k (A_j \partial_j B_k) \partial_k + \sum_j \sum_i \sum_k B_k A_j \Gamma_{jk}^i \partial_i \quad .
 \end{aligned}$$

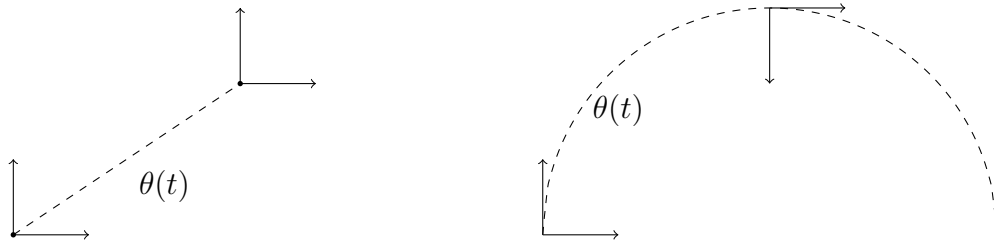
En remplaçant l'indexation en k dans le terme de gauche par une indexation en i , on obtient

$$\nabla_A B = \sum_j \sum_i \left[A_j \partial_j B_i + \left(\sum_k B_k A_j \Gamma_{jk}^i \right) \right] \partial_i \quad .$$

7.3.2.3 Courbes et connexion

Si $\theta(t)$ est une courbe de M , alors $\theta'(t)$ est un champ de vecteur porté par $\theta(t)$. On peut alors considérer $\nabla_{\theta'(t)} \theta'(t)$.

Si $\nabla_{\theta'(t)} \theta'(t) = c(t) \theta'(t)$, qu'on peut ramener à $\nabla_{\theta'(t)} \theta'(t) = 0$ avec un choix de paramétrage convenable, alors on dit que la courbe θ est une **géodésique pour la connexion**. Reprenons nos exemples du plan



Dans les deux situations, θ est une géodésique pour la connexion. Intuitivement, cela signifie que intrinsèquement, le vecteur tangent ne change pas lors du transport.

Remarquons que cette définition peut-être différente de la définition de la géo-

désique entre deux points : dans la deuxième situation, la connexion choisie ne correspond pas à la géodésique pour la métrique euclidienne. Par contre, c'est le cas dans la première situation.

Si (M, g) est une variété riemannienne, on va donc s'intéresser aux connexions pour lesquels les notions de géodésiques coïncident et on appellera une telle connexion **connexion riemannienne** ou **connexion de Levi-Civita**. Nous ne donnerons pas ici une définition précise de cette connexion et renvoyons le lecteur vers [17]. Nous nous contenterons de dire que pour chaque g , elle est uniquement déterminée.

Le résultat suivant, qu'on trouve dans [1], sera utile pour le calcul numérique des géodésiques :

Proposition 69

Soit ∇ une connexion affine muni d'un système de coordonnées. Si θ est une géodésique, alors

$$\theta''_k(t) + \sum_i \sum_j \theta'_i(t) \theta'_j(t) \theta^k_{ij}(\theta(t)) = 0 \quad \forall k \in \{1, \dots, d\} \quad .$$

7.3.3 Rappel sur les méthodes particulières

Nous allons faire quelques rappels sur la résolution de l'équation de transport par les méthodes particulières, en suivant [12].

Soit $h \in \mathbb{R}^{+*}$. Si $k \in \mathbb{Z}^d$, on pose $x_k = kh$. On considère une fonction $\varphi \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$ telle que :

- $\sum_{k \in \mathbb{Z}^d} \varphi(\cdot - k) = 1$;
- il existe $m > 0$ tel que $\sum_{k \in \mathbb{Z}^d} k_1^{p_1} \dots k_d^{p_d} \varphi(x - k) = x_1^{p_1} \dots x_d^{p_d} \quad \forall |p| \leq m - 1$.

On a

$$\rho_0(x) = \sum_{k \in \mathbb{Z}^d} \rho_0(x) \varphi\left(\frac{x - x_k}{h}\right) \quad .$$

De là, en posant

$$\alpha_k = \int_{x_k + [-\frac{h}{2}, \frac{h}{2}]^d} \rho_0(x) \varphi\left(\frac{x - x_k}{h}\right) dx$$

on déduit une approximation de ρ_0 par une somme de masse de Dirac

$$\bar{\rho}_0(x) = \sum_{k \in \mathbb{Z}^d} \alpha_k \delta_{x_k} \quad .$$

7.3 Géométrie différentielle et riemannienne

La solution de l'équation de transport conservative avec pour donnée initiale une masse de dirac $\alpha\delta_{x_0}$ étant donné par $\alpha\delta_{\phi(x_0,-t)}$, on en déduit une première approximation de $\rho(x, t)$ par

$$\sum_{k \in \mathbb{Z}^d} \alpha_k \delta_{\phi(x_k, -t)} \quad .$$

Pour se ramener à une fonction, on se donne une fonction $\psi \in W^{m,1}$ telle que :

- on a

$$\int_{\mathbb{R}^d} \psi(x) \, dx = 1 \quad ;$$

- il existe $r > 0$ tel que

$$\int_{\mathbb{R}^d} x_1^{k_1} \cdots x_d^{k_d} \varphi(x_1, \dots, x_d) \, dx = 0, \quad |k| \leq r - 1 \quad .$$

Alors on se donne $\varepsilon > 0$ et on pose

$$\rho(x, t) = \left(\sum_{k \in \mathbb{Z}^d} \alpha_k \delta_{\phi(x_k, -t)} \right) * \frac{1}{\varepsilon^d} \psi \left(\frac{\cdot}{\varepsilon} \right) (x) = \sum_{k \in \mathbb{Z}^d} \frac{1}{\varepsilon^d} \varphi \left(\frac{x - \phi(x_k, -t)}{\varepsilon} \right) \quad .$$

La question est maintenant le choix de ε en fonction du h que l'on a. Trop petit, la solution s'annulera entre les x_k et l'approximation ne sera pas valable. Trop grand, on n'atteindra pas une précision nécessaire. Le résultat suivant obtenu dans [12], raffinement d'un résultat de [32], nous donne l'estimateur suivant permettant de choisir ε .

Théorème

On suppose que $\rho_0 \in W^{s,p}$, avec $s \geq 0$ et $0 \leq p \leq +\infty$, ainsi que $m \geq \frac{d}{p'}$. Alors il existe une constante $C = C(t, m, r, \varphi, \psi)$ tel que pour tout $0 < h \leq \varepsilon$, on a

$$\|\rho(\cdot, t) - \bar{\rho}(\cdot, t)\|_{L^p} \leq C \left(\varepsilon^r \|\rho_0\|_{W^{r,p}} + \left(\frac{h}{\varepsilon} \right)^m \|\rho_0\|_{L^p} \right)$$

Remarque 70

Si $s \leq r$, alors $\varepsilon = h^{\frac{m}{m+s}}$ nous donne une convergence en $h^{\frac{sm}{m+s}}$. Si l'on peut prendre un m grand, alors on s'approche d'une convergence optimale en h^s , permettant de dépasser la borne $h^{\frac{s}{2}}$ que l'on rencontrait dans [32].

Bibliographie

- [1] Shun-ichi AMARI : *Differential-geometrical methods in statistics*, volume 28 de *Lecture Notes in Statistics*. Springer-Verlag, New York, 1985.
- [2] Shun-ichi AMARI et Hiroshi NAGAOKA : *Methods of information geometry*, volume 191 de *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI; Oxford University Press, Oxford, 2000. Translated from the 1993 Japanese original by Daishi Harada.
- [3] Luigi AMBROSIO et Paolo TILLI : *Selected Topics on « Analysis on Metric Spaces »*.
- [4] Colin ATKINSON et Ann F. S. MITCHELL : Rao's distance measure. *Sankhya : The Indian Journal of Statistics, Series A (1961-2002)*, 43(3):345–365, 1981.
- [5] Michelle AUDIN : *Géométrie*. EDP sciences, 2006.
- [6] Bernard BECK, Isabelle SELON et Christine FEUILLET : *Maths MT Tout en un : cours avec exercices corrigés*. 2006.
- [7] Fabrice BONJEAN et Gary LAGERLOEF : Diagnostic model and analysis of the surface currents in the tropical pacific ocean. *J. Phys. Oceanogr.*, 32, 2002.
- [8] Fabrice BONJEAN, Gary LAGERLOEF, E. JOHNSON et J. GUNN : Validation and error analysis of oscar sea surface currents. 2007.
- [9] A. BRANCOLINI, G. BUTTAZZO et F. SANTAMBROGIO : Path functionals over Wasserstein spaces. *J. Eur. Math. Soc. (JEMS)*, 8(3):415–434, 2006.
- [10] Yann BRENIER : Polar factorization and monotone rearrangement of vector-valued functions. *Communications on Pure and Applied Mathematics*, 44(4): 375–417, 1991.
- [11] Martin CAMPOS-PINTO : Towards smooth particle methods without smoothing. soumis en avril 2014.
- [12] Albert COHEN et Benoit PERTHAME : Optimal approximations of transport equations by particle and pseudoparticle methods. *SIAM Journal on Mathematical Analysis*, 32(3):616–636, 2000.

- [13] S. COSTA, S. SANTOS et Strapasson J. : Fisher information distance : a geometrical reading. dernière date de mise à jour : 13/01/2014.
- [14] Carl de BOOR : *A practical guide to splines*, volume 27 de *Applied Mathematical Sciences*. Springer-Verlag, New York, revised édition, 2001.
- [15] Jean-Pierre DEMAILLY : *Analyse numérique et équations différentielles*. Grenoble sciences. Presses universitaires de Grenoble, 1991 (07-Aubenas, Grenoble).
- [16] Adrien DOUADY et Régine DOUADY : *Algèbre et théories galoisiennes*.
- [17] Sylvestre GALLOT, Dominique HULIN et Jacques LAFONTAINE : *Riemannian geometry*. Universitext. Springer-Verlag, Berlin, third édition, 2004.
- [18] Wilfrid GANGBO et Robert J. MCCANN : The geometry of optimal transportation. *Acta Mathematica*, 177(2):113–161, 1996.
- [19] Clark R. GIVENS et Rae Michael SHORTT : A class of wasserstein metrics for probability distributions. *The Michigan Mathematical Journal*, 31(2):231–240, 1984.
- [20] J. M. HAMMERSLEY et D. C. HANDSCOMB : *Monte Carlo methods*. Methuen & Co., Ltd., London ; Barnes & Noble, Inc., New York, 1965.
- [21] Svetlana KATOK : *Fuchsian groups*.
- [22] Rainer KRESS : Numerical integration. In *Numerical Analysis*, volume 181 de *Graduate Texts in Mathematics*, pages 189–224. Springer New York, 1998.
- [23] Jacques LAFONTAINE : *introduction aux variétés différentielles*. Grenoble Sciences, 2010.
- [24] Frédéric LAGOUTIÈRE : *modélisation mathématique et résolution numérique de problèmes de fluides compressibles à plusieurs constituants*. Thèse de doctorat, Université Pierre et Marie Curie, 2000.
- [25] Randall J. LEVEQUE : *Finite Volume Methods for Hyperbolic Problems*. Cambridge University Press, 2002.
- [26] Robert J. MCCANN : Existence and uniqueness of monotone measure-preserving maps. *Duke Mathematical Journal*, 80(2):309–323, 11 1995.
- [27] Robert J. MCCANN : A convexity principle for interacting gases. *Advances in Mathematics*, 128(1):153 – 179, 1997.
- [28] JensMarkus MELENK : On approximation in meshless methods. In James F. BLOWEY et Alan W. CRAIG, éditeurs : *Frontiers of Numerical Analysis*, Universitext, pages 65–141. Springer Berlin Heidelberg, 2005.

BIBLIOGRAPHIE

- [29] J.M. MELENK et I. BABUŠKA : The partition of unity finite element method : Basic theory and applications. *Computer Methods in Applied Mechanics and Engineering*, 139(1–4):289 – 314, 1996.
- [30] Giovanni PISTONE : Nonparametric information geometry. Submitted for publication in the Proceedings of GSI2013 Aug 28-30 2013 Paris.
- [31] Giovanni PISTONE : Examples of the application of nonparametric information geometry to statistical physics. *Entropy*, 15(10), 2013.
- [32] P.-A. RAVIART : An analysis of particle methods. In *Numerical methods in fluid dynamics (Como, 1983)*, volume 1127 de *Lecture Notes in Math.*, pages 243–324. Springer, Berlin, 1985.
- [33] Vincent RIVOIRARD et Gilles STOLTZ : *Statistique en action : cours & problèmes corrigés : master & agrégation de mathématiques*. Mathématiques concrètes. Vuibert, Paris, 2009. Compléments et annexes en ligne dont simulations sous Matlab <http://www.dma.ens.fr/statenaction>.
- [34] Lene Theil SKOVGAARD : A Riemannian geometry of the multivariate normal model. *Scand. J. Statist.*, 11(4):211–223, 1984.
- [35] J. J. STOKER : *Differential geometry*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1989. Reprint of the 1969 original, A Wiley-Interscience Publication.
- [36] Asuka TAKATSU : Wasserstein geometry of gaussian measures. *Osaka Journal of Mathematics*, 48(4):1005–1026, 12 2011.
- [37] Cédric VILLANI : *topics in optimal transportation*. American Mathematical Society, 2003.