



HAL
open science

Analyse mathématique et convergence d'un algorithme pour le transport optimal dynamique : cas des plans de transports non réguliers, ou soumis à des contraintes

Romain Hug

► **To cite this version:**

Romain Hug. Analyse mathématique et convergence d'un algorithme pour le transport optimal dynamique : cas des plans de transports non réguliers, ou soumis à des contraintes. Optimisation et contrôle [math.OC]. Université Grenoble Alpes, 2016. Français. NNT : 2016GREAM077 . tel-01680973v1

HAL Id: tel-01680973

<https://theses.hal.science/tel-01680973v1>

Submitted on 11 Jan 2018 (v1), last revised 11 Jan 2018 (v2)

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.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Mathématiques Appliquées**

Arrêté ministériel : 7 août 2006

Présentée par

Romain HUG

Thèse dirigée par **Emmanuel MAITRE**
et codirigée par **Nicolas PAPADAKIS**

préparée au sein du **Laboratoire Jean Kuntzmann**
et de **L'École Doctorale Mathématiques, Sciences et Technologies de l'Information, Informatique**

Analyse mathématique et convergence d'un algorithme pour le transport optimal dynamique.

Cas des plans de transports non réguliers, ou soumis à des contraintes

Thèse soutenue publiquement le **09 décembre 2016**,
devant le jury composé de :

M. Didier BRESCH

Directeur de recherche CNRS, Université Savoie-Mont-Blanc, Président

M. Stéphane DESCOMBES

Professeur, Université de Nice, Rapporteur

M. Filippo SANTAMBROGIO

Professeur, Université Paris-Sud XI, Rapporteur

M. Vincent DUVAL

Chargé de recherche INRIA, INRIA Paris, Examineur

M. Emmanuel MAITRE

Professeur, Grenoble-INP, Directeur de thèse

M. Nicolas PAPADAKIS

Chargé de recherche CNRS, Université Bordeaux I, Co-Directeur de thèse



Résumé

ANALYSE MATHÉMATIQUE ET CONVERGENCE D'UN ALGORITHME POUR LE TRANSPORT
OPTIMAL DYNAMIQUE

Cas des plans de transports non réguliers, ou soumis à des contraintes

Au début des années 2000, J. D. Benamou et Y. Brenier ont proposé une formulation dynamique du transport optimal basée sur la recherche en espace-temps d'une densité et d'une quantité de mouvement minimisant une énergie de déplacement entre deux densités. Ils ont alors proposé, pour la résolution numérique de ce problème, d'écrire ce dernier sous la forme d'une recherche de point selle d'un certain lagrangien via un algorithme de lagrangien augmenté. Nous étudierons, à l'aide de la théorie des opérateurs non-expansifs, la convergence de cet algorithme vers un point selle du lagrangien introduit, et ceci dans les conditions les plus générales possibles, en particulier dans les cas où les densités de départ et d'arrivée s'annulent sur certaines zones du domaine de transport. La principale difficulté de notre étude consistera en la preuve de l'existence d'un point selle, et surtout de l'unicité de la composante densité-quantité de mouvement dans de telles conditions. En effet, de tels conditions impliquent de devoir traiter avec des plans de transport optimaux non réguliers : c'est pourquoi une importante partie de nos travaux aura pour objet une étude approfondie de la régularité d'un champ de vitesse associé à de tels plans de transport. Nous tenterons également de caractériser les propriétés d'un champ de vitesse associé à un plan de transport optimal dans l'espace quadratique. Pour finir, nous explorerons différentes approches relatives à l'introduction de contraintes physiques dans la formulation dynamique du transport optimal, basées sur une pénalisation du domaine de transport ou du champ de vitesse.

Mots clés : transport optimal, convergence et consistance d'un algorithme, problèmes multi-physiques

Abstract

MATHEMATICAL ANALYSIS AND CONVERGENCE OF AN ALGORITHM FOR OPTIMAL
TRANSPORT PROBLEM

Case of non regular transportation maps, or subjected to constraints

In the beginning of the 2000 years, J. D. Benamou and Y. Brenier have proposed a dynamical formulation of the optimal transport problem, corresponding to the time-space search of a density and a momentum minimizing a transport energy between two densities. They proposed, in order to solve this problem in practice, to deal with it by looking for a saddle point of some Lagrangian by an augmented Lagrangian algorithm. Using the theory of non-expansive operators, we will study the convergence of this algorithm to a saddle point of the Lagrangian, in the most general conditions, particularly in cases where initial and final densities cancel on some areas of the transportation domain. The principal difficulty of our study will consist of the proof, in these conditions, of the existence of a saddle point, and especially in the uniqueness of the density-momentum component. Indeed, these conditions imply to have to deal with non-regular optimal transportation maps : that is why an important part of our works will have for object a detailed study of the properties of the velocity field associated to an optimal transportation map in quadratic space. To finish, we will explore different approaches for introducing physical priors in the dynamical formulation of optimal transport, based on penalization of the transportation domain or of the velocity field.

Keywords : optimal transport, convergence and consistency of an algorithm, multi-physics problems

Remerciements

Je tiens tout d'abord à remercier mon jury de thèse, pour avoir accepté de se pencher sur mon travail, de faire le déplacement jusqu'à Grenoble, et pour les nombreuses questions pertinentes posées pendant ma soutenance. Merci à mes rapporteurs Filippo Santambrogio et Stéphane Descombes pour la lecture de mon manuscrit. Ce fut pour moi un très grand honneur. Merci pour leurs critiques et remarques très constructives. Merci également à Vincent Duval pour s'être intéressé à mon travail et avoir accepté le rôle d'examinateur. J'ai une pensée particulière pour Didier Bresch, président de mon jury, mais qui fut également, et surtout, mon maître de stage de M2 au laboratoire LAMA à Chambéry, et à qui je dois la rencontre avec mes directeurs de thèse à Grenoble, et qui fut durant cette période plein de bons conseils et de bienveillance à mon égard.

Cette thèse n'aurait pas pu aboutir et aller bien loin sans mes deux formidables directeurs, Emmanuel Maître et Nicolas Papadakis ! Ils auront été absolument parfaits sur tous les aspects, professionnellement bien sûr, mais aussi humainement (ce qui compte le plus à mes yeux). Bienveillants à toute épreuve, et surtout dans les moments difficiles, très à l'écoute et de bons conseils. J'ai pu sentir chez eux le désir sincère de rechercher le meilleur pour moi et ma carrière. Je les remercie également de m'avoir fait confiance, et de m'avoir suivi et accompagné dans mes diverses pérégrinations au fil de ces travaux de thèse, lesquelles ne correspondaient pas toujours (voir même pas souvent) au travail qu'ils attendaient. J'ai ce gros défaut de n'en faire qu'à ma tête...

Bien plus que la thèse, la recherche, les maths et les montagnes, je me souviendrai avant toute chose de ces années à Grenoble, d'une superbe troupe de LJKiens. Certains m'ont vu arriver et m'ont accueilli ; j'en ai vu partir avec émotion ; j'en ai moi-même vu arriver (et j'espère qu'avec les autres, nous les avons bien accueillis) ! Un grand merci à Vincent, Ester, Bertrand, Madison, Federico, Federico (devinez qui est qui), Thomas, Euriell, Eric, Amin, Lukas, Jean-Matthieu, Kole, Nhu, Patricia, Kevin, Burak, Anastasios, Fairouz, Charles, Rémi, Margaux, Jean-Baptiste, Adrien, Alexandre, Cécile, Lionel, Arnaud, Olivier, Charlotte, Laurence, Juana, Cathy,...

Je tiens à remercier tout particulièrement ceux qui constituent pour moi le noyau dur de tous ces compagnons de route, ceux qui seront devenus de véritables amis, dont l'amitié, la bienveillance et l'empathie auront été, je le pense vraiment, l'un de mes principaux soutiens durant toutes ces années, parfois difficiles. Je n'oublierai jamais ces sorties, ces soirées jeux, ces randonnées, ces sessions de ski, ces week-end de vacances,... ! Je remercie donc très chaleureusement Matthias, Chloé, Pierre-Olivier, Nelson, Meriem, Pierre-Jean et Roland. Une mention très particulière pour Morgane dont l'amitié, dès le début et tout au long de cette thèse, fut d'une très grande valeur pour moi. Elle fut toujours présente, dans les bons moments comme dans les moins bons. Un grand merci également à Pierre (Mecha Meister), à Hervé la souris, et à Pépito le chat !

Vous êtes tous des gens exceptionnels et merveilleux !

Grenoble sera synonyme aussi de moments très enrichissants passés avec ces joyeux drilles d'histoire de l'art ! Un grand merci à Pierre et Laure-Anne pour toutes ces sorties, randonnées, périples culturelles, ces cours particuliers d'architecture ancienne... Grâce à eux, j'ai appris énormément dans un domaine où je ne connaissais rien.

J'ai eu également une vie avant la thèse, ailleurs qu'à Grenoble, et de nombreux amis qui auront tous beaucoup compté pour moi. Je remercie mon super groupe du lycée : Aurore, Joanna, Joannie, Denis, Clotilde & Greg, Mélanie.

Mes non moins supers compères de la fac d'Orsay : Matthias, Cédric, Benoît et Jean-Baptiste (merci pour toutes ces prises de tête politiques).

Mes très théâtraux compagnons de jeux de rôles : Hugo, Charlie, Basile, Vincent, Alexandre, Loïc et Dimitri.

Ainsi que ces quelques personnes qui ont le grand mérite d'être de ce monde : Louise, Jean et Florent.

Qu'un camion livreur de vieux fromage me renverse à l'instant même si j'oubliais, ne serait-ce qu'une seconde, d'évoquer ces deux vieux rena... je veux dire ces individus que sont Antonin et Céline, sans l'existence desquels, j'en suis sûr, les baleines à bosses auraient probablement disparues depuis quinze ans déjà, Apollo 11 n'aurait jamais décollé, la tartiflette aurait été un plat à base de vache qui rit, et Robespierre un valet de chambre de Louis XVI (si quelqu'un n'est pas d'accord, qu'il vienne me prouver le contraire!). Pour toutes ces années, ces voyages, ces points de joie ou de santé mentale restaurés (etc), je veux vous exprimer, mes chers amis, un grand "whrr" (le mot "merci" n'étant pas assez fort)! Et je n'oublie pas Ril (le chat) bien sûr!

Est-il vraiment utile de préciser que je remercie chaleureusement toute ma famille, mes très chers parents (Catherine et Vincent), ma soeur Lola, mes tantes/oncles/cousines/cousins/grandes-tantes/grands-oncles (il y en a trop pour tous les citer), et mes chats Delhi et Hôï An.

Un petit mot plein de tendresse pour mes grands-parents et arrière-grands-parents, présents ou dans mes pensées, Françoise et Pierre, Gabriel et Paule, Renée, Roger.





Table des matières

I	Introduction et présentation des résultats	1
I.1	Introduction au transport optimal et ses applications	1
I.1.1	Introduction-fiction	1
I.1.2	Le problème des déblais et des remblais de Gaspard Monge	3
I.1.3	Le problème économique d'allocation optimale de ressources de Leonid Kantorovitch	9
I.1.4	Le Théorème de Brenier : un transport optimal euclidien liant les problème de Monge et de Kantorovitch	14
I.1.5	Une formulation dynamique du problème de transport optimal	22
I.1.6	Géodésique convexité	31
I.2	Algorithmes et transport optimal dynamique	33
I.2.1	Méthodes de "splitting proximal"	35
I.2.2	Application au problème de transport optimal dynamique	39
I.3	Transport Optimal, Applications et Variantes	42
I.3.1	Applications au traitement d'images	42
I.3.2	Flots de gradient dans les espaces de mesures de probabilité	43
I.3.3	Le transport branché	46
I.4	Présentation synthétique des résultats obtenus	47
I.4.1	État de l'art (à notre connaissance) sur la convergence de l'algorithme de Benamou-Brenier	47
I.4.2	Notre contribution à l'étude de consistance de l'algorithme de Benamou-Brenier	48
I.4.3	Étude du champ vitesse pour un transport optimal dans L^2	49
I.4.4	Applications du transport optimal aux problèmes multi-physiques	49
II	Étude de l'algorithme de Benamou-Brenier	51
II.1	Formulation du problème et présentation de l'algorithme	51
II.1.1	Le problème de Monge-Kantorovich dans \mathbb{R}^d pour un coût quadratique	51
II.1.2	Formulation convexe et lagrangien augmenté	52
II.1.3	Algorithme de Benamou-Brenier	54
II.2	Caractérisation des points selles	56
II.3	Quelques résultats de convergence pour les opérateurs non-expansifs	63
II.3.1	Opérateur non-expansif \mathbf{M}	63
II.3.2	Convergence faible de l'algorithme itératif $x_{n+1} = \mathbf{M}x_n$	64
II.3.3	Convergence forte d'une relaxation de l'algorithme $x_{n+1} = \mathbf{M}x_n$	65
II.4	Algorithme B.-B. en terme d'opérateur non-expansif	67
II.5	Conclusion du chapitre	71

III	Préambule et régularité du champ de vitesse	73
III.1	Introduction et notations	74
III.2	Propriétés du potentiel du plan de transport	80
III.3	Propriétés du champ de vitesse	85
III.3.1	Quelques propriétés utiles des sous-gradients de fonctions Lipschitziennes	86
III.3.2	Rappels sur les conjuguées convexes	87
III.3.3	Opérateur Proximal et enveloppe de Moreau :	87
III.3.4	Définition et premières propriétés du champ de vitesse	89
III.3.5	Contrôle du champ de vitesse	92
III.3.6	Régularité lipschitzienne en espace-temps du champ de vitesse	96
III.3.7	Préambule sur la régularité du champ de vitesse	97
III.3.8	Equation de Burgers et résultat de régularité : cas (Γ_2)	99
III.3.9	Equation de Burgers et résultat de régularité : cas (Γ_1)	108
III.3.10	Résultats complémentaires pour les résultats d'unicité à venir	117
III.3.11	Un résultat indépendant mais notable	119
III.4	Résumé et conclusion du chapitre	123
IV	Existence et unicité des points selles	125
IV.1	Construction d'un point selle du Lagrangien	126
IV.1.1	Propriétés de régularité temporelle de la densité	129
IV.1.2	Propriété de conservation de la masse	135
IV.1.3	Existence d'un point selle	136
IV.2	Résultats d'unicité	140
IV.2.1	Non-unicité des points selles de \mathbf{L}	141
IV.2.2	Unicité du champ de vitesse sur le support de la densité	142
IV.2.3	Unicité de la densité dans L^2	145
IV.2.4	Préambule à la démonstration de la Proposition 4.2-5	147
IV.2.5	Résultats préalables et preuve de la Proposition 4.2-5	149
IV.2.6	Retour à la preuve du Théorème 4.2-1	156
IV.3	Caractérisation d'un champ de vitesse de transport optimal	158
IV.4	Conclusion du chapitre	170
V	Transport optimal multi-physique	173
V.1	Formulation dynamique et méthodes de splitting proximal	174
V.1.1	Petit rappel récapitulatif sur le transport optimal dynamique	174
V.1.2	Optimisation avec la méthode de splitting proximal	175
V.1.3	Application au problème (5.1-7)	176
V.1.4	Limitations et motivations	178
V.2	Transport Optimal Anisotrope	179
V.2.1	Cas d'une anisotropie constante	180
V.2.2	Existence d'une solution dans un cas anisotrope en espace seulement	182
V.2.3	Discrétisation, convexification et optimisation	185
V.2.4	Illustration	186
V.3	Transport optimal avec contraintes physiques	188
V.3.1	Couplage non convexe	189
V.3.2	Existence d'un minimiseur	189
V.3.3	Optimisation	192
V.4	Expériences et tests	194
V.4.1	Tests synthétiques	194

V.4.2 Interpolation d'images en océanographie	195
V.5 Conclusion du chapitre	196
VI Conclusion et perspectives	199
A Liste de Théorèmes généraux	201
A.1 Convolutions, distributions, espaces L^p et $W^{1,p}$	201
A.2 Éléments de théorie de la mesure	203
B Convergence des opérateurs non-expansifs	209
B.1 Convergence faible	209
B.2 Convergence forte	210
C Preuves de la section III.2	215
D Calcul d'un opérateur proximal	219
E Preuves de la section IV.1	221
F Autour de la poussée de mesure	231
G Démonstration du Lemme 4.2-4	241

Chapitre I

Introduction et présentation des résultats

I.1 Introduction au transport optimal et ses applications

I.1.1 Introduction-fiction

480 av J.-C., Athènes vient d'être prise et mise à sac par la grande armée perse menée par son roi Xerxès. La population a été évacuée, et une grande partie de celle-ci s'est réfugiée sur l'île de Salamine, bientôt rejointe par la flotte des cités grecques coalisées qui s'y installe. Cette flotte, forte de ses quatre cents trières, est maintenant sur le point d'affronter la flotte de Xerxès bien plus nombreuse, et bien plus forte en tonnage. L'histoire officielle décrit la stratégie grecque comme ayant consisté à enfermer la flotte perse dans une rade, mais nous vous proposons une autre variante. Dans cette version, la bataille se déroule un peu plus loin, en haute mer : impossible pour les grecs de s'appuyer sur la topologie de la côte.

Le Stratège athénien Thémistocle et le Général spartiate Eurybiade concoctent une stratégie pour tenter de l'emporter malgré tout. La situation ne semble à première vue pas du tout à leur avantage. Mais ils disposent néanmoins d'un atout majeur sur leur ennemi : le soudoiment d'un Général perse leur a permis de connaître en détail les plans de leurs adversaires. La flotte perse a l'intention, profitant de la supériorité de tonnage de ses navires sur les trières grecques, de percer la ligne de défense grecque par une attaque centrée, espérant ainsi diviser la flotte grecque en deux parties, détacher un tiers de ses bateaux pour un débarquement à Salamine pendant que le reste se redéploiera derrière les lignes grecques, afin de couper sa retraite et d'ainsi protéger le détachement de débarquement.

La stratégie des grecs mise au point par Thémistocle et Eurybiade consistera donc à faire mine, dans un premier temps, de subir l'assaut perse, puis au moment opportun, à redéployer leur flotte en une manoeuvre d'encercllement rapide et foudroyante pour prendre en tenaille les assaillants.

Nos deux Généraux se faisant une idée assez précise de la situation de leur flotte à la suite de la manoeuvre perse (quitte pour cela à discrètement imposer cette formation au moment du déploiement de la flotte adverse, sans que cela n'y paraisse), vont devoir assigner une trajectoire à chacune de leur trière (qui sera connue à l'avance par les capitaines de ces dernières) pour que la formation d'encercllement s'effectue le plus rapidement possible. Afin de maximiser la vitesse d'accomplissement de la manoeuvre tout en ménageant au maximum la force de leurs rameurs (ceux-ci en auront bien besoin pour la suite de la bataille), ils décident de faire en sorte de minimiser le déplacement de chacun de leur bateau.

Ils peuvent supposer, pour simplifier les données du problème, que leurs trières sont toutes à peu près identiques, ayant le même tonnage et la même force motrice (les trières étant principalement manoeuvrées à la rame pendant les batailles navales), et pouvant donc se déplacer à une même vitesse maximum (négligeant ainsi certaines contraintes, telles que le vent et la houle). Pour la plupart de leurs unités, il est inutile voir même dangereux que celles-ci se retrouvent à leur point de destination avant la fin de la manoeuvre : les bateaux ayant les trajectoires les plus courtes pourront ainsi se permettre d'aller plus lentement et ainsi préserver les forces de leurs hommes.

Malheureusement, en cherchant à répartir leurs bateaux sur leur carte, ils s'aperçoivent rapidement que très souvent, la réduction de longueur de la trajectoire d'une trière s'accompagne en général d'une augmentation de la longueur de trajectoire d'autres trières. Le problème semblant décidément insoluble, Thémistocle décide de le soumettre à son plus brillant géomètre : un esclave gaulois nommé Gaspardakis Mongonos . Celui-ci reformule le problème ainsi : le tout ne serait pas de chercher à optimiser la trajectoire de chaque navire, celle-ci se faisant presque toujours au détriment d'une autre, mais plutôt de chercher à **minimiser la somme de toutes les trajectoires**. De ce problème, se détache, selon lui, la nécessité de respecter deux contraintes simples :

- Les bateaux doivent chercher autant qu'ils le peuvent à se déplacer en **ligne droite**.
- Les tracés des routes suivis par les différentes trières ne doivent **jamais se croiser** (voir l'illustration 1.1-2 plus bas).

Notre esclave propose alors de procéder comme il suit (voir l'illustration 1.1-1) :

1. Sur la carte, représenter les positions des trières au départ de la manoeuvre (par exemple par des petits cailloux) et à la fin de celle-ci (avec des petits bouts de bois).
2. A l'aide d'une règle ou d'une tige rectiligne, tracer une ligne droite séparant à la fois la flotte d'arrivée et celle de départ en deux parties de telle façon à ce que le nombre de bateaux situés d'un côté ou de l'autre de la droite soit le même au départ qu'à l'arrivée : par exemple s'il y a 100 bateaux d'un côté et 300 de l'autre au départ, il doit y avoir également 100 bateaux à l'arrivée du même côté de la ligne que les 100 de départ, et 300 de l'autre côté, du même côté que les 300 de départ. Une telle droite est appelée une équisécante.

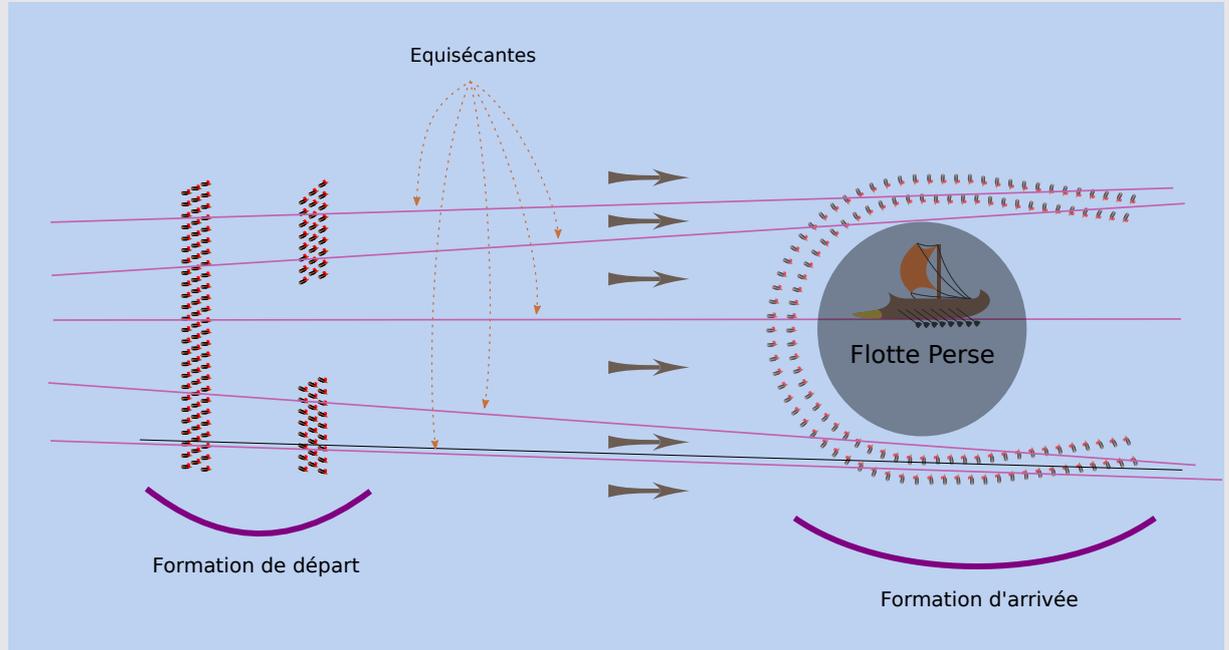
Les bateaux situés d'un côté de la ligne au départ, doivent se déplacer en restant du même côté de la ligne : en effet, si l'un d'eux changeait de côté au cours de son déplacement, un autre bateau situé de l'autre côté de la ligne devrait, en compensation d'effectifs, changer à son tour de côté, risquant ainsi de provoquer un croisement de trajectoire. Ceci aurait pour conséquence, au regard du second principe évoqué plus haut, une somme des longueurs de trajectoires plus longue que si ceux-ci ne s'étaient pas croisés.

3. Recommencer l'opération ci-dessus, et tracer autant d'équisécantes que possible. Ainsi, selon Gaspardakis Mongonos , un bateau situé entre deux équisécantes à son départ devra se déplacer en restant entre ces deux droites jusqu'à son arrivée, se prévenant ainsi des croisements de trajectoires. Heureusement pour les grecs, dans la configuration de la manoeuvre, aucune des équisécantes ne se croisent au milieu du théâtre de la bataille, ce qui aurait ici pu rendre défectueuse la méthode de l'esclave (voir l'exemple 1.1-5 p. 7).

Il fut donc décidé de procéder selon cette méthode. Le plan fut planifié, les ordres furent donnés, la bataille fut gagnée, et Athènes fut vengée !

Illustration 1.1-1. LA MANOEUVRE DES TRIÈRES

L'assignation de la trajectoire et du positionnement final de chaque trière se fait de façon à ce que chacune se déplace en ligne droite en croisant aucune des équisécantes prédéfinies (en théorie, il devrait même y avoir une infinité d'équisécantes, mais dans la pratique...).



Les équisécantes se coupent ici derrière la formation de départ (à gauche) ou derrière la formation d'arrivée (à droite). Si elles s'étaient coupées au milieu, cela aurait pu poser problème (voir l'exemple 1.1-5 p. 7).

I.1.2 Le problème des déblais et des remblais de Gaspard Monge

Nous venons, dans cette petite fiction historique (le plan de bataille ne s'est peut-être pas déroulé exactement de cette façon, mais quoi qu'il en soit la bataille fut une victoire pour les grecs!), de présenter un problème connu sous le nom de "problème des déblais et des remblais" initié par Gaspard Monge (1746-1818) à la fin du XVIIIème siècle dans [57] : l'esclave gaulois Gaspardakis Mongonos était peut-être l'un de ses lointains ancêtres. Son problème était plus exactement formulé ainsi :

LE MÉMOIRE SUR LES DÉBLAIS ET LES REMBLAIS DE GASPARD MONGE

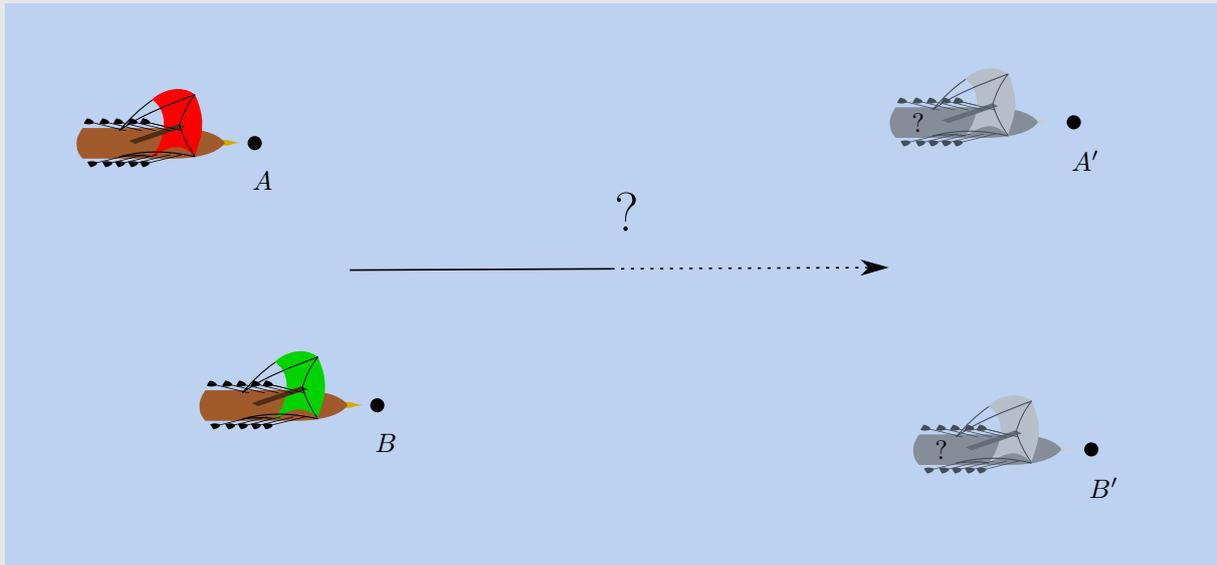
"Lorsqu'on doit transporter des terres d'un lieu dans un autre, on a coutume de donner le nom de Déblai au volume des terres que l'on doit transporter, et le nom de Remblai à l'espace qu'elles doivent occuper après le transport. Le prix du transport d'une molécule étant, toutes choses d'ailleurs égales, proportionnel à son poids et à l'espace qu'on lui fait parcourir, et par conséquent le produit du transport total devant être proportionnel à la somme des produits des molécules multipliées par l'espace parcouru, il s'ensuit que le déblai et le remblai étant donnés de figure et de position, il n'est pas indifférent que telle molécule du déblai soit transportée dans tel ou tel autre endroit du remblai, mais qu'il y a une certaine distribution à faire des molécules du premier dans le second, d'après laquelle la somme de ces produits sera la moindre possible, et le prix du transport total sera un minimum."

Ainsi, dans notre petite histoire, nous avons simplement remplacé les molécules de terre par des bateaux. On pourrait se poser la question de savoir si G. Monge considérait son problème d'un point de vue continu ou discret, au sens où on l'entend aujourd'hui : en effet, à l'époque où

G. Monge formule ce problème, ces deux notions sont encore mal définies. Dans ce problème historique, il semble implicitement supposé que l'on considère des "molécules de terre" très nombreuses ou relativement identiques du point de vue de leur taille ou de leur masse : nous sommes donc plus proche d'une formulation continue du problème. Cependant, nous continuerons dans un premier temps à discuter de notre problème de manoeuvre de trières. Après tout, G. Monge n'envisageait-t-il pas lui-même d'appliquer son problème pour des manoeuvres de fantassins ?

Illustration 1.1-2. DES TRAJECTOIRES SANS CROISEMENT

Deux trières, situées respectivement aux points A et B doivent atteindre les positions A' et B' (non-alignées aux points A et B), de façon à minimiser la somme des longueurs de leur trajectoires respectives.

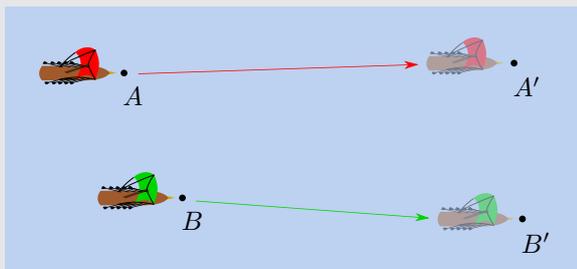


Pour minimiser cette distance parcourue, il est évident que les deux bateaux devront se déplacer en ligne droite ; la question se réduit donc à savoir à quelle position assigner chacun de ces deux bateaux : la trière située au point A doit-elle se rendre au point A' ou au point B' (même question pour celle située au point B) ?

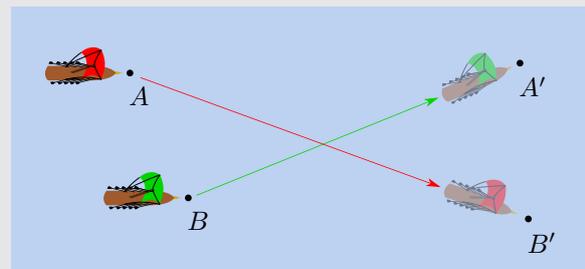
Il y a donc deux possibilités :

1. Une assignation sans croisement de trajectoire : Du point A vers le point A' et du point B vers le point B' .
2. Une assignation avec croisement de trajectoire : Du point A vers le point B' et du point B vers le point A' .

Or, dans un quadrilatère, la somme des longueurs des diagonales est toujours plus longue que celle de deux cotés opposés (i.e. $AA' + BB' < AB' + BA'$) : c'est donc la première possibilité (sans croisement) qui est la plus optimale.



Assignations sans croisements

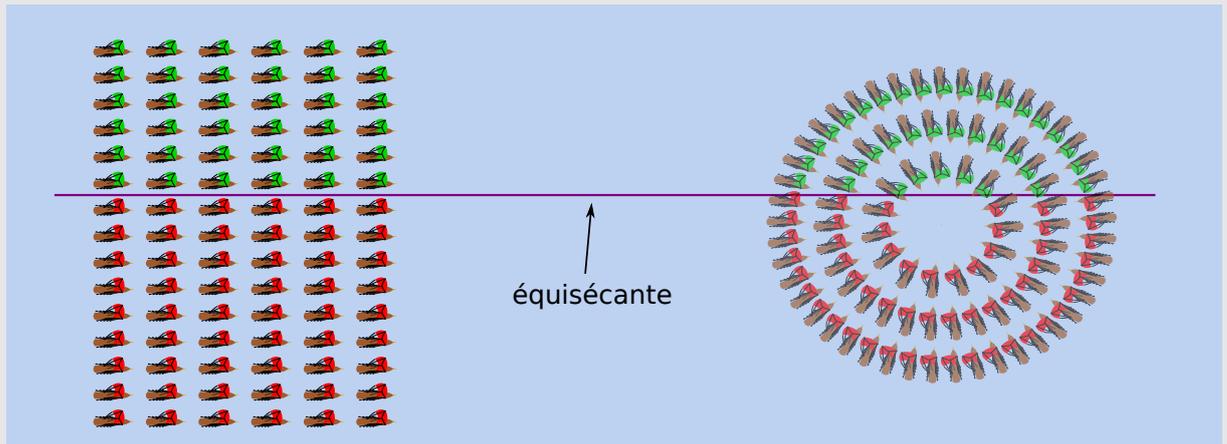


Assignations avec croisements

A la lumière de ces considérations (trajectoires en lignes droites sans croisements), expliquons un peu plus en détails le principe de la méthode des équisécantes proposée à l'époque par de G. Monge.

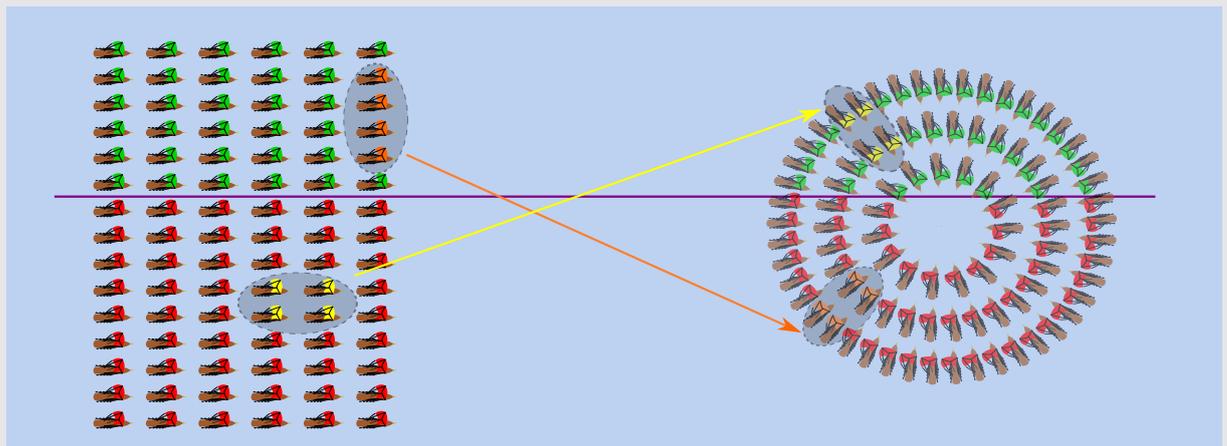
Illustration 1.1-3. LA MÉTHODE DES ÉQUISÉCANTES

Supposons que la formation de départ et d'arrivée soient coupées par une même équisécante : sur ce schéma, il y a autant de trières à voiles rouges en dessous de l'équisécante dans la formation de départ que dans celle d'arrivée, de même pour les voiles vertes.



La méthode de G. Monge prescrit que si la manoeuvre de la flotte suit un transport optimal, les trières doivent au cours de leur déplacement rester du même côté de l'équisécante : ici les voiles rouges sur les voiles rouges, et les voiles vertes sur les voiles vertes.

En effet, si par exemple certaines trières situées au dessus de l'équisécante à leur départ venaient à croiser l'équisécante (les voiles oranges sur le schéma ci-dessous), il faudrait en compensation qu'un nombre égal de trières situées en dessous de l'équisécante (les voiles jaunes) croisent également l'équisécante (car les nombres de trières situées au dessus et en dessous de l'équisécante doivent être les mêmes au départ et à l'arrivée). Une telle manoeuvre impliquerait donc un croisement de trajectoires, et la somme des longueurs de l'ensemble des trajectoires ne serait pas minimale (voir l'illustration 1.1-2).



Cette méthode ne fonctionne que si l'ensemble des équisécantes **ne se coupent pas** entre les formations de départ et d'arrivée : voir le contre-exemple 1.1-5 (p. 7).

Remarque 1.1-4. La conditions de non croisement des trajectoires n'est bien sûr valable que si les positions de départ et d'arrivée ne sont pas situées sur le même alignement. Pour que cette

condition ait toujours lieu, il faudrait chercher à minimiser la somme des **carrés des longueurs** des trajectoires plutôt que la somme des longueurs des trajectoires.

Dans le cas de notre exemple sur la bataille de Salamine, cet autre coût peut également traduire une autre signification : en effet, rappelons que nos généraux souhaitent ici que toutes les trières arrivent en même temps à destination, tout en ménageant le plus équitablement possible les forces de leurs rameurs. Les manoeuvres s'effectuant donc toutes en un même temps, la vitesse moyenne de chaque trière est donc directement proportionnelle à la longueur de la route qu'ils auront à parcourir : l'énergie dépensée par l'ensemble des marins de chaque navire est donc proportionnelle à la longueur de leur trajet.

Le coût "somme des carrés des longueurs des routes" (ou coût quadratique) est donc en théorie mieux adapté au problème que le coût "somme des longueurs des routes" (problème originel formulé par G. Monge des "déblais et des remblais"), sauf que dans le cas pratique de ce problème spécifique cela ne change rien. En effet :

- lorsqu'il n'y a pas alignement des positions de départ et d'arrivée les routes et les assignations sont les mêmes pour les deux coûts,
- dans le cas contraire, les bateaux ne peuvent en pratique (i.e. matériellement) se passer à travers. Or, l'assignation pour laquelle l'ordre des positions d'arrivée correspond à l'ordre des positions de départ de chaque bateaux est bien optimal pour le coût "quadratique" (voir l'illustration 1.1-6 p. 8).

Cependant, si les trières avaient été en mesure de se traverser sans se toucher, le problème de Monge des "déblais et des remblais" n'aurait pas eu de solution unique, contrairement au problème reformulé pour un coût quadratique.

De plus, le problème du transport optimal s'étend à des thématiques bien plus vastes et plus concrètes que des batailles navales, et en particulier à des cas continus ou discrets pour lesquels un tel coût quadratique possède un réel intérêt, autant théorique que pratique, grâce à **la formulation énergétique** du problème qu'il induit.

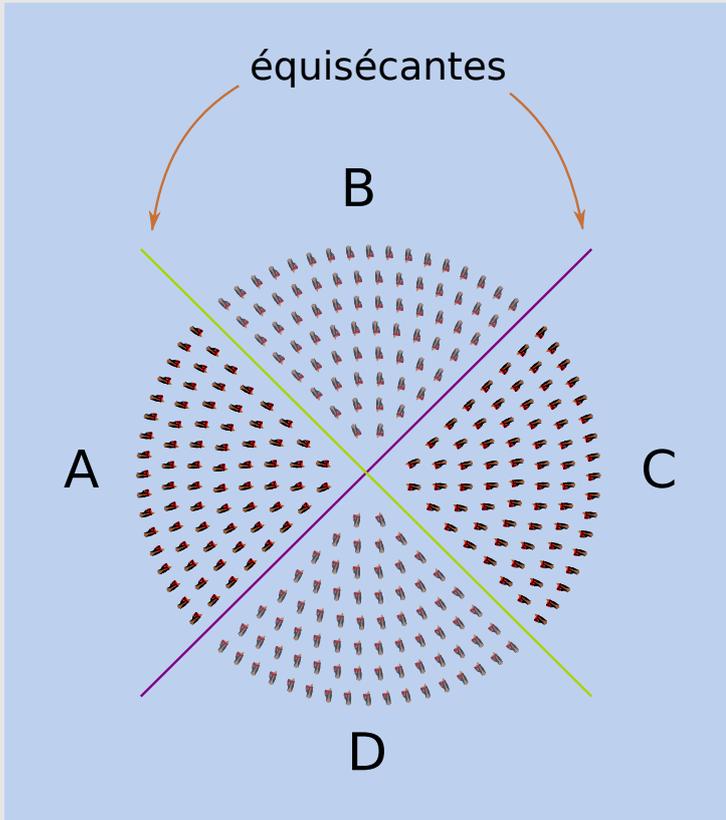
Nous verrons en effet dans les sous-sections I.1.4 (voir en particulier l'illustration 1.1-9 p. 15) et I.1.5 qu'un tel problème d'assignation optimal pour un coût quadratique correspond en fait à résoudre un problème de déplacement optimal en recherchant énergie de transport globale minimale : par exemple ici on chercherait à minimiser l'énergie globale dépensée par les rameurs, autrement dit à **minimiser l'effort global**.

Nous verrons qu'une telle formulation offre la possibilité d'aborder les choses de façon moins géométrique et moins descriptive, et permet au contraire d'envisager la possibilité d'appliquer des méthodes plus modernes d'optimisation convexe sous contraintes (méthodes entre autre plus "numériques" et plus facilement traitables par des ordinateurs).

La méthode que nous avons illustrée (méthode des équisécantes), est celle que Gaspard Monge avait proposée en son temps pour une résolution pratique du problème en **deux dimensions**. Cependant, quand bien même les différentes étapes de celle-ci auraient été facilement réalisables (découper à l'aide d'une même droite deux ensembles en deux parties égales n'est à priori pas un exercice évident), cette méthode n'aboutit pas forcément dans tous les cas à la détermination "transport optimal" selon la définition de G. Monge (et celui-ci en avait conscience) : par exemple, dans des cas où il pourrait y avoir des échanges d'unités entre les deux cotés de la ligne, sans qu'il y ait croisement de trajectoires, ou bien dans le cas des croisements d'équisécantes entre la masse de départ et d'arrivée (voir exemple 1.1-5, qui suit). G. Monge préconisait, pour que cette méthode soit valable, que le déblai et le remblai (i.e. les masses de départ et d'arrivée) ne soient pas coupés par la courbe formée par les points d'intersection de l'ensemble des équisécantes.

Exemple 1.1-5. CONTRE-EXEMPLE À LA MÉTHODE DE G. MONGE

Dans l'illustration 1.1-1 représentant la manoeuvre des trières grecques, il s'avère qu'aucune des équisécantes (celles qui y sont tracées, comme celles qui ne le sont pas) ne se croisent au milieu du théâtre des opérations : c'est à dire ici au sein de l'enveloppe convexe de la formation de départ et d'arrivée. Nous présentons ici une manoeuvre pour laquelle la méthodes de Monge ne peut fonctionner.

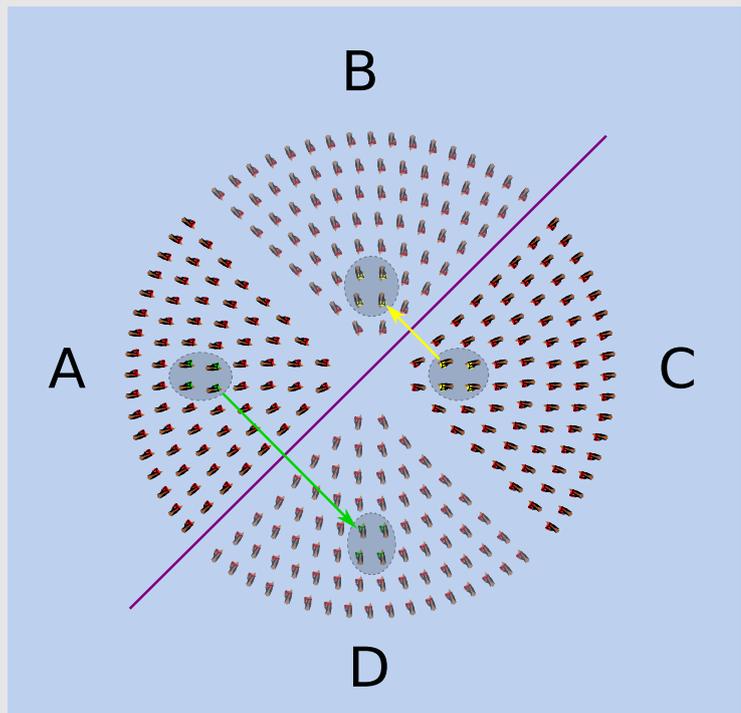


Ici la formation de départ est représentée par les trières opaques (zones A et C), et celle d'arrivée par les trières en transparence (zones B et D). Ces deux formations sont coupées par deux équisécantes (représentées en vert et en violet) qui se croisent au centre des deux formations.

La méthode de Monge se retrouve ici heurtée à une contradiction. En effet, les trières partant de la zone A doivent selon cette méthode se situer à leur arrivée du même côté de l'équiséquante verte qu'à leur départ, et doivent donc toutes se retrouver au final dans la zone D.

Cependant, elles devront se situer du même côté de l'équiséquante violette qu'à leur départ et donc toutes se retrouver en zone C (en même temps donc que se retrouver toutes en zone D). Même raisonnement pour les trières partant de la zone C.

Une explication que l'on peut donner à ce phénomène est qu'ici, contrairement au cas de l'illustration 1.1-3 (p. 5), des trière peuvent échanger leurs positions par rapport à une équisécante sans que leurs trajectoires en lignes droites ne se croisent (rappelons que la trajectoire ne représente que le segment reliant la position de départ et d'arrivée, pas la droite entière), comme l'illustre la figure suivante :

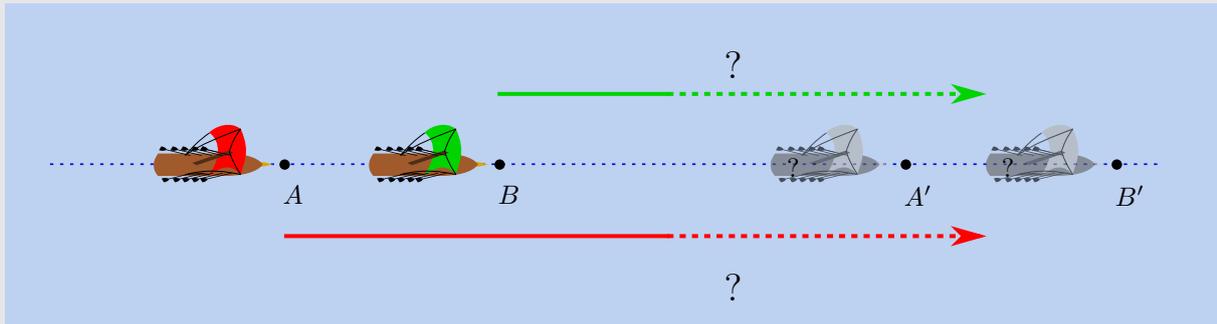


Les trières jaunes et vertes peuvent ici échanger leur position par rapport à l'équisécante violette sans que leurs trajectoires ne se croisent.

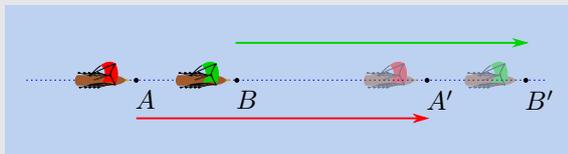
Cette méthode laisse de plus une ambiguïté, même dans la formulation "continue" (anachroniquement) de G. Monge, en ce qui concerne les déplacements en ligne droite. Cette ambiguïté ne provient en fait pas vraiment de la méthode, mais plus de la façon dont est posé le problème lui-même : le fait de considérer la simple somme des longueurs des trajectoires.

Illustration 1.1-6. AMBIGUÏTÉ DES TRAJECTOIRES EN LIGNE DROITE

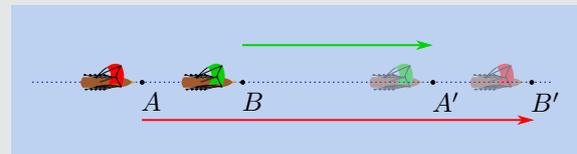
Cette illustration fait suite à la Remarque 1.1-4 (p. 5). Dans l'exemple qui suit, contrairement à l'illustration 1.1-2, ici deux trières situées aux positions A et B sont assignées à deux positions A' et B' situées dans le même alignement.



Or ici, la somme des longueurs AA' et BB' est la même que la somme de AB' et BA' . D'un point de vue strictement théorique, selon le problème posé par Monge (c'est à dire minimiser la somme des longueurs de toutes les trajectoires), les assignments $A \rightarrow A'$, $B \rightarrow B'$ et $A \rightarrow B'$, $B \rightarrow A'$ sont équivalentes pour ce problème.



Assignations sans croisement



Assignations avec croisements
(en pratique impossible en ligne droite)

Bien sûr dans la pratique, pour à la fois uniformiser la dépense d'énergie des rameurs et éviter les collisions ou toutes autres difficultés de manoeuvre, il vaut mieux dans tous les cas que les trajectoires ne se croisent pas (les trières se déplacent en ligne droite et ne peuvent donc normalement pas se croiser).

Pour faire coïncider la pratique et la théorie dans ce plan de bataille, Monge aurait par exemple pu chercher à minimiser non pas la somme des longueurs de toutes les trajectoires, mais plutôt la somme des **carrés des longueurs** de toutes les trajectoires : effet, dans cet exemple on a $AA' + BB' = AB' + BA'$, on a par contre $AA'^2 + BB'^2 < AB'^2 + BA'^2$. L'assignation sans croisement (i.e. $A \rightarrow A'$, $B \rightarrow B'$) est donc bien plus optimale pour le coût "somme des carrés des longueurs" que l'assignation avec croisement (i.e. $A \rightarrow B'$, $B \rightarrow A'$).

La méthode des équisécantes que nous avons présentée dans cette section, est celle que G. Monge avait élaborée pour résoudre son problème en deux dimensions. Cette méthode ne se généralise pas directement en trois dimensions. En effet, même en dimension trois, les trajectoires restent des lignes droites (i.e. d'une seule dimension), et acquièrent un grand nombre de degrés de liberté : quand bien même on découperait les densités de départ et d'arrivée avec un équi-plan (équivalent d'une équisécante en trois dimensions), dans la quasi-totalité des configurations, on pourrait toujours trouver une infinité de trajectoires qui coupent l'équi-plan sans se couper entre elles.

G. Monge avait cependant proposé une méthode pour tenter de résoudre le problème en trois

dimensions, valable dans un cadre "continu" (même si comme nous l'avons déjà évoqué précédemment, cette notion est quelque peu anachronique). Dans le problème en deux dimensions, lorsque la méthode des équisécantes s'applique, chaque trajectoire du transport se prolonge en une équisécante : ces trajectoires optimales s'encadrent donc mutuellement entre elles. L'idée de G. Monge pour résoudre son problème en trois dimensions est donc de découper l'espace en des sortes de "tubes" (Monge parle d'ensembles de droites congruentes [58]) dont les parois sont composées de trajectoires optimales. En effet les trajectoires intérieures à un tel "tube" ne peuvent en couper les parois (sinon elles couperaient une des trajectoires constituant le bord). L'idée de Monge est de construire ces "tubes" à partir des trajectoires dites "extrêmes", c'est à dire formés par les trajectoires optimales partant des points du bord du déblai). Les calculs menés par Monge en son temps pour tenter de déterminer et construire ces surfaces d'encadrement, selon la géométrie du déblai et du remblai, le mèneront à une équation proche de *l'équation de Monge-Ampère*, laquelle sera présentée plus bas (voir la Remarque 1.1-10 p. 17).

I.1.3 Le problème économique d'allocation optimale de ressources de Leonid Kantorovitch

Le problème de transport optimal a été redécouvert indépendamment (c'est à dire sans avoir eu connaissance des travaux de G. Monge à ce sujet) par le mathématicien russe Leonid Kantorovitch en 1942, sous un aspect plus économique : le problème économique d'allocation optimale de ressources [49].

Considérons deux types de zones : les zones de production X et les zones à approvisionner Y . Comme dans le modèle de G. Monge, on s'intéresse ici aux différents plans de transport d'une répartition de masse de la zone X (masse produite) vers une répartition de masse dans la zone Y (masse consommée). Contrairement à ce premier modèle, on considère à présent que :

- les zones X et Y ne font plus forcément partie du même "espace",
- la capacité de production ou le besoin en approvisionnement peuvent ici être variables au sein même d'une de ces zones (c'est à dire que les densités des masses de départ et d'arrivée ne sont plus uniformes),
- un élément de masse situé en un lieu x de la zone X peut, lors de son acheminement, être réparti vers différents lieux de la zone Y , et inversement un même lieu y de la zone Y peut être approvisionné par différents lieux de la zone X .

Pour reprendre l'illustration de Caffarelli telle qu'évoquée dans [77] pour un cas concret, considérons par exemple que la zone X produit une ressource \mathcal{R} (la zone X pourrait être une série de mines houillères, produisant plus ou moins de charbon, ou bien des champs de blé, etc...). Considérons une zone Y ayant des besoins plus ou moins importants d'approvisionnement en la ressource \mathcal{R} selon les différents lieux qui la constituent (exemple : zones résidentielles, zones industrielles, etc...).

Il est donc nécessaire d'acheminer la ressource \mathcal{R} depuis les lieux de production jusqu'aux lieux à approvisionner. Cependant, le transport d'une masse de \mathcal{R} d'un lieu vers un autre aura un certain coût (dépendant de la distance entre ces deux lieux, des zones à contourner, de la nature du terrain, etc...). On fait les hypothèses suivantes :

- Les coûts de production de la ressource \mathcal{R} sont partout les mêmes. Seuls comptent donc la disponibilité de cette ressource dans les différents lieux de production de la zone X , les besoins d'approvisionnement des lieux y de la zone Y , ainsi que le coût d'acheminement d'un lieu x de la zone X vers un lieu y dans la zone Y .

- La production de \mathcal{R} dans la zone X s'est organisée de façon à assurer exactement les besoins de la zone Y (pas de sur-production ni de sous-production).
- Le coût d'acheminement d'une masse m de la ressource de x vers y est proportionnel à m avec un coefficient de proportionnalité $c(x, y)$.
- Les acheminements et les coûts sont fixes sur de longues périodes : il est donc équivalent de considérer des acheminements continus et des acheminements "d'un bloc" (ce qui sera le cas de notre modèle).

A. Cas d'une répartition collective des coûts : allocation optimale

Supposons dans un premier temps que les coûts soient collectifs. L'objectif est donc de trouver un schéma d'allocation de la ressource \mathcal{R} produite dans la zone X vers la zone Y , selon la répartition des besoins, de façon à minimiser le coût de transport global.

On modélise la répartition de la production sur X par une mesure μ et la répartition des besoins sur Y par une mesure ν (X et Y étant considérés comme des espaces mesurables), le tout vérifiant $\mu(X) = \nu(Y)$ (la production de la zone X assure exactement les besoins de la zone Y). L'ensemble des schémas d'allocation de X vers Y est lui modélisé par l'ensemble $\Pi(\mu, \nu)$ des mesures π définies sur $X \times Y$ (muni d'une tribu appropriée) telles que pour tout sous-ensemble mesurable $A \subset X$ et $B \subset Y$, on ait

$$\pi(A \times Y) = \mu(A), \quad \pi(X \times B) = \nu(B). \quad (1.1-1)$$

Pour traduire qualitativement ce que représente la mesure π , disons formellement que $d\pi(x, y)$ "mesure" en tout point $(x, y) \in X \times Y$ la quantité (la masse) de ressource \mathcal{R} acheminée du lieu de production $x \in X$ au lieu approvisionné $y \in Y$ (cette masse peut bien sûr être nulle : cela signifie que rien n'est acheminé de x vers y).

Remarque 1.1-7. *La condition $\pi(A \times Y) = \mu(A)$ pour tout $A \subset X$, traduit simplement le fait que l'intégralité de la production de la zone X (traduite par la mesure μ) est livrée dans la zone Y , et la condition $\pi(X \times B) = \nu(B)$ pour tout $B \subset Y$ traduit quant à elle le fait que l'intégralité des besoins d'approvisionnement de la zone Y (traduite par la mesure ν) est assurée par la zone X .*

On peut donc formellement écrire que le coût d'acheminement de la ressource \mathcal{R} de x vers y est donné par " $c(x, y) \cdot d\pi(x, y)$ ", où $c(x, y)$ représente le coefficient proportionnel du coût de transport du point (ou lieu) $x \in X$ vers le point $y \in Y$.

Ainsi, en considérant l'ensemble des coûts de transport depuis l'ensemble des points de production $x \in X$ vers l'ensemble des points approvisionnés $y \in Y$, le coût de transport global $I(\pi)$ d'un schéma d'allocation π est dans ce modèle donné par :

$$I(\pi) = \int_{X \times Y} c(x, y) d\pi(x, y).$$

L'objectif de ce problème est donc de minimiser le coût de transport global $I(\pi)$, et dans l'idéal de déterminer un schéma d'allocation optimale $\tilde{\pi}$ pour lequel le minimum de $I(\cdot)$ est atteint. L'existence d'un tel schéma $\tilde{\pi}$ est donnée par le Théorème 1.1-2 qui sera énoncé à la fin de cette sous-section (p. 13).

B. Cas d'une répartition individualiste des coûts : dualité de Kantorovitch

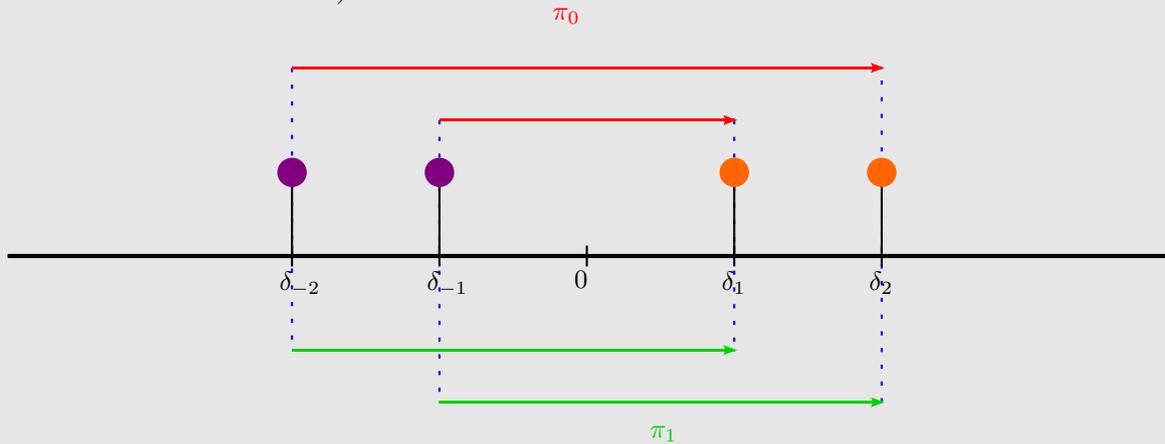
Supposons maintenant que les coûts ne soient plus collectifs mais individuels. Ici, chaque lieu y

de la zone Y va chercher à subvenir à ses besoins en la ressource \mathcal{R} selon les ressources encore disponibles de la zone de production X en cherchant à minimiser les coûts de transport nécessaires (entièrement à sa charge), et ceci de façon totalement individuelle (sans tenir comptes des autres), selon le principe : "les premiers arrivés seront les mieux servis".

Un modèle d'allocation des ressources minimisant le coût global d'acheminement, ne correspond en effet pas forcément à un coût d'acheminement minimal "individuel". Un schéma quelconque de répartition des ressources π peut de ce fait avoir un coût de transport individuellement moins cher pour certains lieux à approvisionner par rapport à un schéma $\tilde{\pi}$ minimisant le coût global (en contrepartie, d'autres lieux le paieront eux plus cher), comme décrit dans l'exemple suivant :

Exemple 1.1-8. Prenons $X = Y = \mathbb{R}$ ainsi que la fonction coût $c(x, y) = |x - y|^2$. On souhaite transporter la mesure $\mu = \delta_{-2} + \delta_{-1}$ sur la mesure $\nu = \delta_1 + \delta_2$ (les δ désignent des mesures de Dirac). On considère les deux schémas d'allocation suivants :

- $\pi_0 = \delta_{(-2,1)} + \delta_{(-1,2)}$ (qui envoie toute la masse concentrée en -2 sur 1 , et toute la masse concentrée en -1 sur 2).
- $\pi_1 = \delta_{(-2,2)} + \delta_{(-1,1)}$ (qui envoie toute la masse concentrée en -2 sur 2 , et toute la masse concentrée en -1 sur 1).



Le coût global du schéma π_0 est donc ici $I(\pi_0) = 3^2 + 3^2 = 18$, et le coût global du schéma π_1 est $I(\pi_1) = 4^2 + 2^2 = 20$. Le schéma π_1 a donc un coût global plus élevé que le schéma π_0 . Dans le schéma π_1 le coût individuel de transport de -1 sur 1 (égal à 2) est pourtant inférieur au coût de transport dans le schéma π_0 de -2 sur 1 et de -1 sur 2 (tous deux égaux à 3). Minimiser le coût global ne correspond donc pas forcément à minimiser tous les coûts individuels.

Supposons maintenant qu'une grande société de transport propose ses services en sous-traitance pour effectuer le transport de la ressource \mathcal{R} . Celle-ci possède des moyens qui lui permettent d'effectuer ces acheminements pour un coût très inférieur à celui pratiqué jusque là (c'est à dire $c(x, y) \times m$, où m est la masse à transporter de x à y).

Il est dans son intérêt d'organiser un schéma de transport fixe sur le long terme, d'avoir des revenus très stables, et d'avoir le monopole total sur les acheminements (par exemple appropriation du terrain de transit, construction d'un réseau routier ou ferroviaire, de ports ou d'aéroports, etc...).

Supposons également que pour pouvoir mettre son projet en place, cette société doit avoir l'accord unanime de tous les acheteurs.

Cette société pratique un tarif d'acheminement proportionnel à la masse transportée et à la

somme d'un coût de chargement et de déchargement. Autrement dit, pour faire livrer une masse m de \mathcal{R} d'un lieu x vers un lieu y , l'acheteur devra payer le prix $m \times [\varphi(x) + \psi(y)]$, $\varphi(x)$ étant le coût de chargement et $\psi(y)$ le coût de déchargement, les fonctions de coût φ et ψ étant fixées sur l'ensemble respectivement des zones X et Y . Autrement dit, la société pratique un prix d'achat $\psi(y)$ en chaque point $y \in Y$ et un prix $\varphi(x)$ de mise sur le marché en chaque point $x \in X$. Un tel système tarifaire lui permet d'avoir des revenus très stables.

Si l'on considère toujours que la répartition de la production sur X se modélise par une mesure μ et la répartition des besoins sur Y par une mesure ν (avec $\mu(X) = \nu(Y)$), le prix global pratiqué par la société est alors toujours le même quel que soit le schéma de répartition π , c'est à dire

$$\int_X \varphi d\mu + \int_Y \psi d\nu.$$

Ceci est illustré par le Lemme suivant :

Lemme 1.1-1.

Soient X et Y deux espaces mesurables. Soient μ une mesure positive sur X et ν une mesure positive sur Y , telles que $\mu(X) = \nu(Y)$. Définissons $\Pi(\mu, \nu)$ comme l'ensemble des mesures de Lebesgue π sur $X \times Y$ telles que pour tout sous-ensemble mesurable $A \subset X$ et $B \subset Y$,

$$\pi(A \times Y) = \mu(A), \quad \pi(X \times B) = \nu(B).$$

Alors, pour tous $(\varphi, \psi) \in L^1(d\mu) \times L^1(d\nu)$, et pour tout $\pi \in \Pi(\mu, \nu)$ on a

$$\int_{X \times Y} \varphi(x) d\pi(x, y) = \int_X \varphi d\mu, \quad \text{et} \quad \int_{X \times Y} \psi(y) d\pi(x, y) = \int_Y \psi d\nu,$$

d'où

$$\int_{X \times Y} [\varphi(x) + \psi(y)] d\pi(x, y) = \int_X \varphi d\mu + \int_Y \psi d\nu.$$

Preuve : Traitons le cas de φ . On peut, en toute généralité, supposer que φ est positive sur X . Soit $(\varphi_n)_n$ une suite de fonction étagées sur X pour la mesure μ , telle que

- $\forall n \in \mathbb{N}, \varphi_n \leq \varphi$,
- $(\varphi_n)_n$ converge simplement μ -presque partout vers φ .

Pour tout $n \in \mathbb{N}$, on définit la fonction $\bar{\varphi}_n : (x, y) \mapsto \varphi_n(x)$. La suite $(\bar{\varphi}_n)_n$ converge simplement π -presque tout vers la fonction $(x, y) \mapsto \varphi(x)$.

En effet, si E est l'ensemble des points $x \in X$ tels que $(\varphi_n(x))_n$ ne converge pas vers $\varphi(x)$, alors $E \times Y$ est l'ensemble des points $(x, y) \in X \times Y$ tels que $(\bar{\varphi}_n(x, y))_n$ ne converge pas vers $\bar{\varphi}(x, y)$. Or $\pi(E \times Y) = \mu(E) = 0$.

Pour tout $A \subset X$ mesurable, on a

$$\int_X \mathbb{1}_A d\mu = \mu(A) = \pi(A \times Y) = \int_{X \times Y} \mathbb{1}_{A \times Y} d\pi.$$

Ainsi pour tout $n \in \mathbb{N}$,

$$\int_{X \times Y} \bar{\varphi}_n(x, y) d\pi(x, y) = \int_X \varphi_n d\mu.$$

Par passage à la limite via le Théorème de Beppo-Levi, on a donc

$$\int_{X \times Y} \bar{\varphi}(x, y) d\pi(x, y) = \int_{X \times Y} \varphi(x) d\pi(x, y) = \int_X \varphi d\mu.$$

On raisonne de même pour prouver que

$$\int_{X \times Y} \psi(y) d\pi(x, y) = \int_Y \psi d\nu.$$

□

Pour pouvoir obtenir l'accord unanime des acheteurs, et ainsi pouvoir mettre en place son monopole et son projet, il est donc nécessaire pour notre société que le prix qu'elle propose pour un transport de x vers y , soit toujours inférieur au coût initial (c'est à dire sans son intervention en sous-traitance). Autrement dit, il est nécessaire que pour tout $x \in X$ et pour tout $y \in Y$, la contrainte

$$\varphi(x) + \psi(y) \leq c(x, y) \quad (1.1-2)$$

soit vérifiée, et ceci quitte à devoir pratiquer des prix négatifs en certains endroits!

Ainsi, quel que soit le schéma de répartition π , et d'après le Lemme 1.1-1, on a

$$I(\pi) = \int_{X \times Y} c(x, y) d\pi(x, y) \geq \int_{X \times Y} [\varphi(x) + \psi(y)] d\pi(x, y) = \int_X \varphi d\mu + \int_Y \psi d\nu = J(\varphi, \psi). \quad (1.1-3)$$

où $J(\varphi, \psi)$ représente donc le prix global pratiqué par la société de sous-traitance, et $I(\pi)$ le coût global qu'aurait dû assumer les acheteurs sans l'intervention de cette dernière pour le schéma π . Pour un schéma $\tilde{\pi}$ d'allocation optimal par rapport au coût global, on a donc ainsi

$$I(\pi) \geq I(\tilde{\pi}) \geq J(\varphi, \psi), \quad (1.1-4)$$

où π représente tout autre schéma quelconque de répartition.

Pour maximiser ses profits, la société de sous-traitance va chercher à fixer les fonctions de coût φ et ψ de façon à maximiser son revenu $J(\varphi, \psi)$ tout en vérifiant la contrainte (1.1-2). D'après l'inéquation (1.1-4), ce coût est majoré par $I(\tilde{\pi})$.

Il est très raisonnable de penser que cette majoration par $I(\tilde{\pi})$ se rapproche en fait d'une maximisation, ou du moins d'une "suprématie" (autrement dit que $I(\tilde{\pi})$ représente la limite supérieure des valeurs que peut atteindre $J(\varphi, \psi)$). Ce résultat est entre autre confirmé par le Théorème suivant dû à L. Kantorovitch.

Comme souvent dans la suite de ce manuscrit, nous choisirons pour simplifier les énoncés, de re-normaliser les mesures μ , ν et π de façon à obtenir des mesures de probabilité (de masse totale égale à 1). On note $P(E)$ l'ensemble des mesures de probabilité sur un espace mesurable E .

Théorème 1.1-2. DUALITÉ DE KANTOROVITCH [77]

Soient X et Y deux espaces métriques, mesurables, complets, et séparables. Soient $\mu \in P(X)$ et $\nu \in P(Y)$, et soit $c : X \times Y \rightarrow \mathbb{R}_+ \cup \{+\infty\}$ une fonction de coût semi-continue inférieurement.

Pour tous $\pi \in P(X \times Y)$ et $(\varphi, \psi) \in L^1(d\mu) \times L^1(d\nu)$, définissons

$$I(\pi) = \int_{X \times Y} c(x, y) d\pi(x, y), \quad J(\varphi, \psi) = \int_X \varphi d\mu + \int_Y \psi d\nu.$$

Définissons $\Pi(\mu, \nu)$ comme l'ensemble des mesures de probabilité de Lebesgue π sur $X \times Y$ telles que pour tout sous-ensemble mesurable $A \subset X$ et $B \subset Y$,

$$\pi(A \times Y) = \mu(A), \quad \pi(X \times B) = \nu(B),$$

et définissons Φ_c l'ensemble des fonctions $(\varphi, \psi) \in L^1(d\mu) \times L^1(d\nu)$ vérifiant

$$\varphi(x) + \psi(y) \leq c(x, y) \tag{1.1-5}$$

pour μ -presque tout $x \in X$ et ν -presque tout $y \in Y$. Définissons également C_b l'ensemble des couples de fonctions (f, g) continues et bornées sur $X \times Y$.

Alors on a

$$\inf_{\Pi(\mu, \nu)} I(\pi) = \sup_{\Phi_c} J(\varphi, \psi) = \sup_{\Phi_c \cap C_b} J(\varphi, \psi). \tag{1.1-6}$$

De plus l'infimum du terme $\inf_{\Pi(\mu, \nu)} I(\pi)$ de (1.1-6) est atteint. Autrement dit, il existe $\tilde{\pi} \in \Pi(\mu, \nu)$ tel que

$$I(\tilde{\pi}) = \inf_{\pi \in \Pi(\mu, \nu)} I(\pi).$$

On pourrait penser d'après ce théorème qu'un lien existe entre un schéma d'allocation optimal $\tilde{\pi}$ et d'éventuelles fonctions de coût $\tilde{\varphi} \in L^1(d\mu)$ et $\tilde{\psi} \in L^1(d\nu)$ maximisant le prix de revenu de la sous-traitance $J(\varphi, \psi)$. L'existence de telles fonctions de coût $\tilde{\varphi}$ et $\tilde{\psi}$ n'est cependant pas vérifiée dans le cas général.

La section suivante nous donnera un cas particulier où de telles fonctions de coût existent, et où le lien avec un (unique) $\tilde{\pi}$ est explicité. Ce cas particulier révélera en fait le cadre d'étude de la quasi-totalité des travaux de ce manuscrit.

I.1.4 Le Théorème de Brenier : un transport optimal euclidien liant les problème de Monge et de Kantorovitch

Nous allons ici présenter le problème de la précédente section dans un cadre plus restreint et plus "intuitif". Nous ajoutons les hypothèses suivantes :

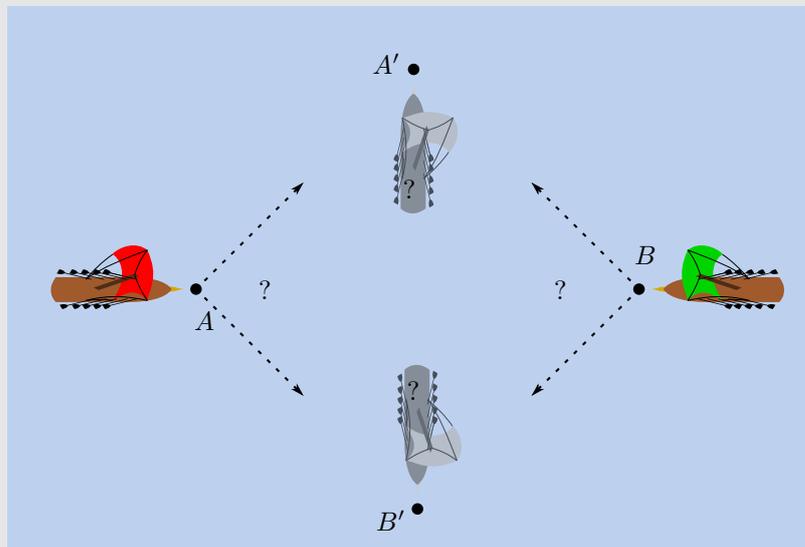
- On considère que la zone de production et la zone à approvisionner font de nouvelles parties du même espace, à savoir ici \mathbb{R}^d (i.e. $X = Y = \mathbb{R}^d$).
- On considère un coût proportionnel quadratique $c(x, y) = |x - y|^2$.
- Les mesures de production μ et d'approvisionnement ν sont deux mesures de probabilité sur la tribu de Lebesgue. De plus, la mesure de départ (de production) μ est supposée absolument continue par rapport à la mesure de Lebesgue (voir plus bas l'exemple 1.1-11 qui illustre l'importance de cette dernière hypothèse).

Autrement dit, on suppose qu'il existe une fonction $f \in L^1(\mathbb{R}^d, \mathcal{L}^d)$ à support borné telle que $\mu = f \cdot \mathcal{L}^d$. On se place donc ici dans le cadre d'une mesure de départ "diffuse" : aucun point de \mathbb{R}^d , et plus généralement aucun sous-ensemble de \mathbb{R}^d de mesure de Lebesgue nulle, ne peut concentrer de la masse.

Exemple 1.1-9. COÛT EUCLIDIEN, COÛT QUADRATIQUE

Analysons formellement les choses. Ainsi que nous l'avons déjà évoqué dans la Remarque 1.1-4 et l'illustration 1.1-6 pour des cas discrets, un transport optimal pour un tel coût quadratique (du type $c(x, y) = |x - y|^2$) traduit, par rapport au problème originel de Monge (le problème des déblais et des remblais) évoqué dans les sous-sections 1.1.1 et 1.1.2, une minimisation du coût global en énergie dépensé plutôt qu'une minimisation des distances parcourues (i.e pour un coût euclidien $c(x, y) = |x - y|$). Cependant ces deux problèmes coïncident dès lors qu'il ne s'agit pas d'un transport selon un alignement : dans le cas contraire le coût $c(x, y) = |x - y|^2$ donne une solution unique au problème, contrairement au coût $c(x, y) = |x - y|$.

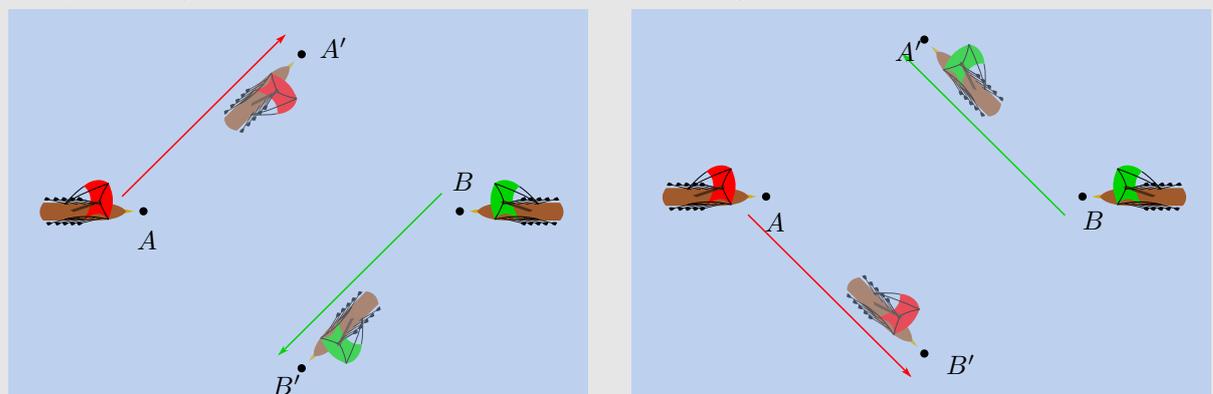
Ainsi le coût quadratique $c(x, y) = |x - y|^2$ semble plus approprié si l'on souhaite traiter un problème ayant une unique solution pour un problème proche de celui des "déblais et des remblais". Pourtant, comme le montre l'illustration ci-dessous, le problème des déblais et des remblais traité de façon discrète n'a pas toujours une solution unique même avec le coût quadratique lorsque les éléments sont tous identiques.



Reprenons nos trières pour représenter nos éléments identiques et indivisibles (sauf éperonnage par un navire perse...).

Sur l'exemple ci-contre, il y a deux possibilités d'assignation pour nos deux trières dans leurs positions de départ (la voile rouge et la voile verte) pour obtenir la formation d'arrivée (représentée comme dans les exemples précédents par les trières en gris).

Ces deux assignations ont exactement le même coût de transport, et ceci pour un coût euclidien au tant que pour un coût quadratique : en effet, quelle que soit l'assignation choisie, les deux trières parcourront exactement la même distance ; le fait d'élever ces distances à une puissance quelconque ne changera donc rien quant à l'égalité de la somme des coûts entre les deux assignations.



Ainsi, dans cet exemple, la solution au problème de Monge **n'est pas unique**. De plus, toujours dans le cadre discret, lorsque les éléments n'ont pas tous la même masse, le problème de Monge pour un coût quadratique peut même **ne pas avoir de solution**, quand bien même celui de Kantorovitch en possède une, du fait de l'indivisibilité des masses dans le problème de Monge (voir plus bas les Définitions 1.1-3 et 1.1-4, ainsi que l'exemple 1.1-11).

Nous sommes ici dans un cadre assez intuitif correspondant au problème historique de G. Monge (voir la sous-section I.1.2) pour des déblais et remblais de densités variables. Ici les géodésiques correspondent à des lignes droites. Il s'agit, comme dans de le problème de Monge de minimiser la somme de toutes les longueurs de trajectoire (pondérée de façon infinitésimale du fait que la masse de départ est ici supposée continue) : un transport optimal devrait correspondre à des déplacements en ligne droite sans aucun croisement de trajectoires (voir l'illustration 1.1-2).

Dans le problème de Monge, contrairement à celui de Kantorovitch, les "unités de masse" sont indivisibles : autant il est éventuellement possible dans le problème économique de Kantorovitch de répartir la ressource produite en un point en plusieurs points d'approvisionnement, autant il n'est pas possible dans le problème de Monge de diviser un grain de sable (ou une trière grecque) et répartir les morceaux en plusieurs points. Si le terme de **schéma d'allocation optimal** de ressources correspond mieux au problème de Kantorovitch, dans lequel on optimise sur des **mesures**, celui de **plan de transport optimal** correspond mieux à celui de Monge : on optimise sur des applications mesurables, i.e. on assigne à un point x **une seule position** $T(x)$.

Bien sûr le passage d'un cadre discret comme celui de Monge, à un cadre potentiellement continue où la masse à répartir peut avoir une densité variable rend un tel raisonnement un peu léger : il n'est a priori pas possible d'avoir un tas de sable aux éléments identiques et de densité variable (à moins de briser les grains en les écrasant, et encore...). C'est pourtant sur ce principe que nous allons formuler un autre problème de transport optimal correspondant mieux à l'idée de Monge, sans abandonner la possibilité de densités continues et variables.

Pour donner tout de même une idée plus matérielle de ce cadre, supposons un énorme tas de sable vu à une échelle où celui-ci est modélisé de façon continue. Supposons également que les différents grains de sable ne soient pas tous identiques et composés de différentes roches plus ou moins massives ? Ainsi deux mêmes volumes extraits de ce tas de sable peuvent avoir des densités différentes.

Nous présentons (à nouveau pour ce qui est du problème de Kantorovitch) ici formellement les deux problèmes :

Définition 1.1-3. LE PROBLÈME DE KANTOROVITCH (RAPPEL)

Soient X et Y deux espaces métriques mesurables. Soient également deux mesures de probabilité $\mu \in P(X)$ et $\nu \in P(Y)$, et une fonction de coût $c : X \times Y \rightarrow \mathbb{R}_+ \cup \{+\infty\}$. On définit, pour tout $\pi \in P(X \times Y)$, le coût $I(\pi)$ par :

$$I(\pi) = \int_{X \times Y} c(x, y) d\pi(x, y),$$

ainsi que l'espace $\Pi(\mu, \nu)$ comme étant l'ensemble des mesures de probabilité π sur $X \times Y$ telles que pour tout sous-ensemble mesurable $A \subset X$ et $B \subset Y$, on ait

$$\pi(A \times Y) = \mu(A), \quad \pi(X \times B) = \nu(B),$$

Le problème de Kantorovitch consiste à trouver une mesure de probabilité $\tilde{\pi} \in P(X \times Y)$ minimisant le coût global d'acheminement, c'est à dire :

$$I(\tilde{\pi}) = \inf_{\Pi(\mu, \nu)} I(\pi).$$

Définition 1.1-4. LE PROBLÈME DE MONGE

Soit X un espace métrique mesurable, et soient également deux mesures de probabilité $\mu, \nu \in P(X)$, et une fonction de coût $c : X \times X \rightarrow \mathbb{R}_+ \cup \{+\infty\}$.

Le problème de Monge consiste à trouver une fonction mesurable \widetilde{M} (aussi appelée plan de transport) allant de l'espace X dans lui-même minimisant le coût de transport global de μ sur ν , c'est à dire :

$$\int_X c(x, \widetilde{M}(x)) d\mu(x) = \inf_{M \# \mu = \nu} \int_X c(y, M(y)) d\mu(y). \quad (1.1-7)$$

où le terme $M \# \mu$ désigne la poussée de la mesure μ par le plan de transport M : cela signifie que tout sous-ensemble mesurable A de X , on a

$$(M \# \mu)(A) = \mu(M^{-1}(A)). \quad (1.1-8)$$

Remarque 1.1-10. A PROPOS DE LA NOTION DE POUSSÉE DE MESURE

L'équation (1.1-8) traduit le fait que la mesure ν **conserve la masse** mesurée par μ : ν donne à toute zone déplacée, déformée, contractée ou dilatée par l'opérateur T , la même masse que lui donnait μ avant l'application de l'opérateur. La notion de poussée de mesure traduit donc bien une propriété de **transport conservatif de masse**. Ceci est en partie à l'origine de l'idée d'une formulation dynamique du problème de Monge (sous-section 1.1.5)

Dans le cas où la mesure de départ μ et la mesure d'arrivée ν sont toutes deux définies sur la tribu de Lebesgue de \mathbb{R}^d et absolument continues par rapport à la mesure de Lebesgue (i.e. s'il existe des densités $\rho_0, \rho_1 \in L^1(\mathbb{R}^d, \mathcal{L}^d)$ telles que $\mu = \rho_0 \mathcal{L}^d$ et $\nu = \rho_1 \mathcal{L}^d$), on peut formellement décrire la poussée de la mesure μ sur la mesure ν par un plan de transport M par **l'équation de Monge-Ampère** :

$$\nu = M \# \mu \Leftrightarrow \rho_0 = (\rho_1 \circ M) \det(\nabla M), \quad (1.1-9)$$

laquelle sera de nouveau évoquée dans la sous-section suivante.

En effet, la formulation duale de la Définition $(M \# \mu)(A) = \mu(M^{-1}(A)), \forall A \subset \mathbb{R}^d$ s'écrit :

$$(\rho_1 \mathcal{L}^d) = M \# (\rho_0 \mathcal{L}^d) \Leftrightarrow \forall f \in C_c(\mathbb{R}^d), \int_{\mathbb{R}^d} f(y) \rho_1(y) dy = \int_{\mathbb{R}^d} f(M(x)) \rho_0(x) dx. \quad (1.1-10)$$

Or, pour tout $f \in C_c(\mathbb{R}^d)$, on a (formellement) par changement de variable :

$$\int_{\mathbb{R}^d} f(y) \rho_1(y) dy = \int_{\mathbb{R}^d} f(M(x)) \rho_1(M(x)) \det(\nabla M(x)) dx \quad (1.1-11)$$

d'où :

$$\rho_0 = (\rho_1 \circ M) \det(\nabla M) \quad (1.1-12)$$

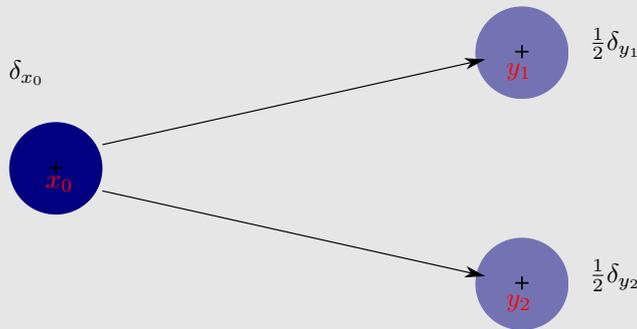
Pour plus de précision sur la notion de poussée de mesure, voir l'Annexe F, et en particulier la Proposition F.0-3 (p. 233).

Le problème de Kantorovitch, ayant des conditions plus larges, admet plus facilement une solution que celui de Monge (voir l'exemple 1.1-11 ci-dessous). Mais lorsque les deux problèmes admettent simultanément une solution, celles-ci sont en général liées. Ce sera du moins le cas dans le cadre que nous allons exposer dans cette sous-section : la solution du problème de Kantorovitch sera équivalente à celle du problème de Monge.

Exemple 1.1-11. UNE SOLUTION POUR LE PROBLÈME DE KANTOROVITCH ET NON POUR CELUI DE MONGE

Comme nous l'avons déjà vu dans l'exemple 1.1-9 (p. 15), le problème de Monge peut présenter des problèmes de consistance lorsque l'on traite le transport de mesures discrètes : en effet, contrairement au problème de Kantorovitch, il n'est pas possible dans le problème de Monge de répartir la masse contenue en un point en plusieurs lieux. Par conséquent, lorsque l'on a affaire à une répartition de masses atomiques (par exemple des mesures de Dirac), il est possible qu'il n'y ait pas de solution : par exemple dans le cas où le nombre de ces masses atomiques sont différents au départ et à l'arrivée, et que celles-ci possèdent des valeurs de masse différentes.

Prenons ici l'exemple (dans \mathbb{R}^2) dans lequel la mesure de départ μ est égale à une mesure de Dirac de masse 1 (concentrée en un point x_0 , i.e $\mu = \delta_{x_0}$) et la mesure d'arrivée ν égale à la somme de deux mesures de Dirac de masse 1/2 (concentrées aux points distincts y_1 et y_2 , i.e. $\nu = (1/2)\delta_{y_1} + (1/2)\delta_{y_2}$) : le problème de Kantorovitch possède bien une solution, qui est la mesure $\pi = (1/2)\delta_{(x_0,y_1)} + (1/2)\delta_{(x_0,y_2)}$, mais le problème de Monge lui n'en possède pas.

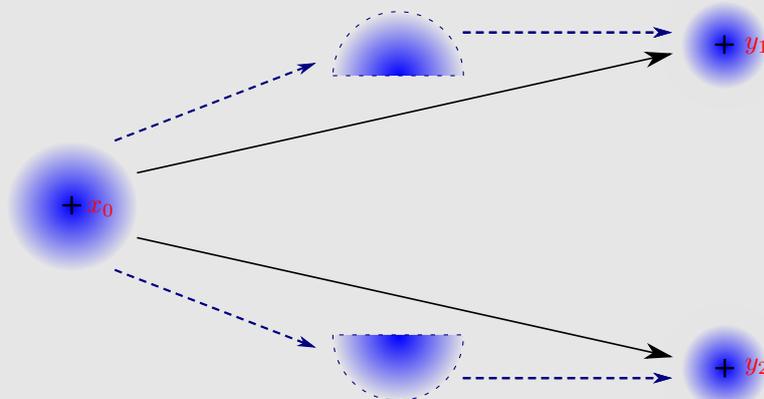


En effet, si cela avait été le cas, il existerait un plan de transport optimal T vérifiant par Définition $T\#\mu = \nu$, ce qui au sens faible s'écrit

$$\forall f \in C_c(\mathbb{R}^2), \int_{\mathbb{R}^2} f(y) d\nu(y) = \int_{\mathbb{R}^2} f(T(x)) d\mu(x) \Leftrightarrow \forall f \in C_c(\mathbb{R}^2), (1/2)[f(y_1) + f(y_2)] = f(T(x_0)).$$

On a $T(x_0) \neq y_1$ ou $T(x_0) \neq y_2$. Supposons par exemple que $T(x_0) \neq y_1$, prenons $f \in C_c(\mathbb{R}^2)$ telle que $f(T(x_0)) = f(y_2) = 1$ et $f(y_1) = 0$. On aurait donc alors $1/2 = 1 \dots$

Pour palier à ce problème on peut faire le choix de "diffuser" légèrement nos mesures de Dirac : remplaçons les par une gaussienne et deux demi-gaussiennes centrées respectivement en x_0 , y_1 et y_2 , et potentiellement très concentrées ("approchant" le Dirac). Cette formulation continue offre alors une solution au problème de Monge, qui consiste à "couper en deux" la gaussienne de départ et la répartir sur les deux demi-gaussiennes.



Avant de revenir au problème de cette sous-section, nous donnons la Définition suivante, à propos de la construction d'une métrique sur l'espace des mesures de probabilité basée sur le transport optimal. En effet, lorsque les espaces d'arrivée et de départ sont les mêmes (notons les X), il est possible de construire une distance entre deux mesures μ et ν à partir du coût du transport optimal associé à l'envoi de l'une sur l'autre, à condition que ce transport optimal soit lui-même basé sur un coût proportionnel égal à la puissance d'une distance sur X par un réel strictement positif p (c'est à dire du type $c(x, y) = d(x, y)^p$ où d est une distance sur X). Nous ne définirons ici cette distance que dans le cas où $p \in [1, +\infty[$ (l'autre cas n'étant pas nécessaire pour notre problème).

Définition 1.1-5. LA DISTANCE DE WASSERSTEIN

Soit (X, \mathcal{B}, δ) un espace métrique, mesurable, complet, et séparable (δ désigne donc ici la distance associée). Pour tout $p \in [1, +\infty[$, on définit l'ensemble $\mathcal{P}_p^\delta(X, \mathcal{B})$ des mesures de probabilité μ sur (X, \mathcal{B}) telles que, en prenant $a \in X$ on ait :

$$\int_X \delta(a, x)^p d\mu(x) < +\infty.$$

Pour tous $\mu, \nu \in \mathcal{P}_p^\delta(X, \mathcal{B})$, définissons comme dans la sous-section 1.1.3 l'ensemble $\Pi(\mu, \nu)$ des mesures de probabilité de Lebesgue π sur $X \times X$ telles que pour tout sous-ensemble mesurable $A, B \subset X$, on ait

$$\pi(A \times X) = \mu(A), \quad \pi(X \times B) = \nu(B),$$

et définissons également la grandeur :

$$\mathcal{W}_p^\delta(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \left(\int_{X \times X} \delta(x, y)^p d\pi(x, y) \right)^{\frac{1}{p}}. \quad (1.1-13)$$

L'application

$$\begin{aligned} \mathcal{W}_p^\delta : \mathcal{P}_p^\delta(X, \mathcal{B}) \times \mathcal{P}_p^\delta(X, \mathcal{B}) &\longrightarrow \mathbb{R}_+ \\ (\mu, \nu) &\longmapsto \mathcal{W}_p^\delta(\mu, \nu) \end{aligned}$$

définit une **distance** sur l'espace $\mathcal{P}_p^\delta(X, \mathcal{B})$. Cette distance entre mesures, basée sur le transport optimal, est désignée sous le nom de **distance de Wasserstein**.

Notation 1. Dans le cadre de la Définition 1.1-5, lorsque $X = \mathbb{R}^d$, muni de la tribu de Lebesgue sur \mathbb{R}^d , et que la distance associée δ_{euc} est la distance euclidienne (i.e. $\delta_{\text{euc}}(x, y) = |x - y|$), alors pour tout $p \in [1, +\infty[$, l'espace $\mathcal{P}_p^{\delta_{\text{euc}}}(\mathbb{R}^d, \mathcal{L}^d)$ sera par abus de notation simplement désigné par $\mathcal{P}_p(\mathbb{R}^d)$, de même que la distance de Wasserstein associée $\mathcal{W}_p^{\delta_{\text{euc}}}$ sera simplement désignée par \mathcal{W}_p .

Remarquons que l'espace $\mathcal{P}_p(\mathbb{R}^d)$ est par ailleurs un espace vectoriel (non normé mais métrique).

Dans toute notre étude (et en particulier dans les sous-sections qui suivent) nous prendrons en général $p = 2$.

Revenons donc maintenant au problème de cette sous-section : dans le cadre que nous avons fixé au début de celle-ci, nous obtenons les résultats suivants :

1. Il existe deux fonctions de coût (respectivement de chargement et de déchargement) $\tilde{\varphi} \in L^1(d\mu)$ et $\tilde{\psi} \in L^1(d\nu)$ maximisant le prix de revenu de la sous-traitance $J(\varphi, \psi)$.

2. Un lien explicite entre un **unique** schéma d'allocation optimal $\tilde{\pi}$ et les fonctions de coût $\tilde{\varphi} \in L^1(d\mu)$ et $\tilde{\psi} \in \times L^1(d\nu)$. Ce lien entre $\tilde{\pi}$ et $(\tilde{\varphi}, \tilde{\psi})$ sera en fait donné directement en fonction de ϕ .
3. La résolution (dans ce cadre) du problème de Monge-Kantorovitch : en effet l'unique schéma d'allocation optimal $\tilde{\pi}$ obtenu attribuera à l'unité (infinitésimal) de masse une unique destination, permettant ainsi de construire un plan de transport de μ sur ν (donné par le gradient de la fonction ϕ) optimal au sens du problème (1.1-7), et unique solution de ce problème.

Ce résultat est explicité dans le Théorème suivant, dû à M. Knott, C. Smith [73] et Y. Brenier [16] (ce Théorème est une compilation de leurs résultats combinés), cité ici à partir de [77] (p. 66). Nous précédonc ce Théorème d'une petite Définition de circonstance :

Définition 1.1-6. TRANSFORMÉE DE LEGENDRE

Soit f une fonction de \mathbb{R}^n dans $] -\infty, +\infty]$. On définit la transformée de Legendre de f , notée f^* par :

$$f^*(x) = \sup_{y \in \mathbb{R}^n} \{ \langle x, y \rangle - f(y) \}$$

Notons que f^* est convexe et semi-continue inférieurement.

Plus de propriétés des transformées de Legendre seront donnés dans la suite de ce manuscrit (voir les propriétés 3.3-5 p. 87).

Théorème 1.1-7. THÉORÈME DE BRENIER [77]

Soient $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$, c'est à dire deux mesures de probabilité sur la tribu de Lebesgue de \mathbb{R}^d vérifiant la condition

$$\int_{\mathbb{R}^d} |x|^2 d\mu(x), \int_{\mathbb{R}^d} |y|^2 d\nu(y) < +\infty. \quad (1.1-14)$$

et telle que μ soit absolument continue par rapport à la mesure de Lebesgue (i.e. telles qu'il existe une densité positive $\rho_0 \in L^1(\mathbb{R}^d, \mathcal{L}^d)$ telle que $\mu = \rho_0 \mathcal{L}^d$).

On considère le problème d'allocation optimal de Kantorovitch de μ sur ν associé à la fonction de coup quadratique $c(x, y) = |x - y|^2$, en définissant toujours $I(\pi)$, $J(\varphi, \psi)$ et $\Pi(\mu, \nu)$ comme dans le Théorème 1.1-2.

Alors il existe une fonction convexe sur \mathbb{R}^d , semi-continue inférieurement ϕ , telle que le schéma d'allocation $\tilde{\pi} \in \Pi(\mu, \nu)$ défini par :

$$\pi = (\text{id} \times \nabla \phi) \# \mu, \quad (1.1-15)$$

ou écrit de façon plus formelle

$$d\pi(x, y) = d\mu(x) \mathbb{1}_{\{y = \nabla \phi(x)\}}(x, y), \quad (1.1-16)$$

soit **l'unique** schéma d'allocation optimal de μ sur ν , c'est à dire vérifiant la relation (1.1-6).

On a de plus les propriétés suivantes :

1. Le couple (ϕ, ϕ^*) (où ϕ^* désigne ici la transformée de Legendre de ϕ) vérifie :

$$\begin{aligned} (\tilde{\varphi}, \tilde{\psi}) &= \sup_{\Phi_c} J(\varphi, \psi) \\ &= \operatorname{argmax}_{(\varphi, \psi)} \left\{ \int_{\mathbb{R}^d} \varphi d\mu + \int_{\mathbb{R}^d} \psi d\nu, \quad \forall (x, y) \in (\mathbb{R}^d)^2 \varphi(x) + \psi(y) \leq |x - y|^2 \right\}, \end{aligned} \quad (1.1-17)$$

avec $(\tilde{\varphi}, \tilde{\psi}) = \left(\frac{|\cdot|^2}{2} - \phi, \frac{|\cdot|^2}{2} - \phi^* \right)$; ou bien formulé autrement :

$$(\phi, \phi^*) = \operatorname{argmin}_{(f, g)} \left\{ \int_{\mathbb{R}^d} f d\mu + \int_{\mathbb{R}^d} g d\nu, \quad \forall (x, y) \in (\mathbb{R}^d)^2 x \cdot y \leq f(x) + g(y) \right\}. \quad (1.1-18)$$

(les couples (φ, ψ) et (f, g) des argmax et argmin étant pris dans $L^1(d\mu) \times L^1(d\nu)$).

Le couple de fonctions "coût de chargement-coût de déchargement" $(\tilde{\varphi}, \tilde{\psi})$ maximisent donc le prix de revenu de la sous-traitance $J(\varphi, \psi)$.

2. L'opérateur $\nabla\phi$ est **l'unique gradient de fonction convexe**, (unique) au sens μ -presque partout, "poussant la mesure μ sur la mesure ν ", c'est à dire telle que l'on ait $\nabla\phi\#\mu = \nu$ (voir plus haut), et représente également **l'unique solution au problème de Monge** pour le transport de μ sur ν . Autrement dit, on a :

$$\int_{\mathbb{R}^d} |x - \nabla\phi(x)|^2 d\mu(x) = \inf_{M\#\mu=\nu} \int_{\mathbb{R}^d} |x - M(x)|^2 d\mu(x) = \mathcal{W}_2(\mu, \nu)^2. \quad (1.1-19)$$

3. On a $\operatorname{supp}(\nu) = \overline{\nabla\phi(\operatorname{supp}(\mu))}$ (voir la Remarque 1.1-13 plus bas).
4. Si la mesure ν est également (de même que la mesure μ) absolument continue par rapport à la mesure de Lebesgue \mathcal{L}^d sur \mathbb{R}^d , on a pour μ -presque tout $x \in \mathbb{R}^d$ et ν -presque tout $y \in \mathbb{R}^d$:

$$\nabla\phi^* \circ \nabla\phi(x) = x, \quad \nabla\phi \circ \nabla\phi^*(y) = y. \quad (1.1-20)$$

L'opérateur $\nabla\phi^*$ représente alors le "transport optimal inverse de $\nabla\phi$ " : c'est l'unique opérateur ν -presque partout défini "poussant la mesure ν sur la mesure μ ", i.e. $\nabla\phi^*\#\nu = \mu$, et représente l'unique solution au problème de Monge pour le transport de ν sur μ . Autrement dit, on a :

$$\int_{\mathbb{R}^d} |y - \nabla\phi^*(y)|^2 d\nu(y) = \inf_{N\#\nu=\mu} \int_{\mathbb{R}^d} |y - N(y)|^2 d\nu(y) = \mathcal{W}_2(\nu, \mu)^2 = \mathcal{W}_2(\mu, \nu)^2. \quad (1.1-21)$$

Remarque 1.1-12. Dans l'étude qui sera menée tout au long de ce manuscrit, nous utiliserons généralement des mesures positives finies μ et ν à supports bornés dans \mathbb{R}^d , et dans ce cas, la condition (1.1-14) sera automatiquement vérifiée, en effet :

$$\int_{\mathbb{R}^d} |x|^2 d\mu(x) \leq \sup(\operatorname{supp}(\mu))^2 \mu(\mathbb{R}^d) < +\infty \quad \text{et} \quad \int_{\mathbb{R}^d} |y|^2 d\nu(y) \leq \sup(\operatorname{supp}(\nu))^2 \nu(\mathbb{R}^d) < +\infty,$$

où l'on écrit $\sup(A) = \sup_{z \in A} |z|$ pour tout sous-ensemble de A de \mathbb{R}^d .

Remarque 1.1-13. Le terme $\nabla\phi(\text{supp}(\mu))$ n'est pas très rigoureux, étant donné que $\nabla\phi$ n'est défini que μ -presque partout sur \mathbb{R}^d : c'est un abus de notation et il doit être plutôt lu comme $\nabla\phi(\text{supp}(\mu) \cap E)$, où E représente l'ensemble des points de \mathbb{R}^d où ϕ est différentiable (et donc où $\nabla\phi$ est bien définie).

Les sous-sections qui suivront seront consacrées à une étude plus précise du problème de Monge-Kantorovitch tel que formulé dans cette même sous-section. En particulier, dans la prochaine sous-section, nous présenteront une formulation de type mécanique des fluides de ce problème. Cette formulation permettra, par le changement de point de vue qu'elle apporte au problème, la mise en pratique de méthodes numériques plus classiques mais efficaces pour la résolution pratique de ce dernier.

I.1.5 Une formulation dynamique du problème de transport optimal

Si le problème de Monge est à l'origine bien un problème de déplacement optimal, ce même G. Monge, une fois admis le fait que la trajectoire optimale pour se rendre d'un point à un autre était la ligne droite, a lui-même réduit ce problème à un seul problème d'assignation de position parmi les éléments à déplacer.

Il en est de même pour la formulation du problème par L. Kantorovitch : son problème lui aussi se ramène à un seul problème d'assignation (ou d'allocation). Le problème du choix des déplacements des éléments à déplacer (ou des masses infinitésimal dans le cas de densités continues) est en fait implicitement traité à travers le coût proportionnel de transport $c(x, y)$. Par exemple, dans la formulation de Kantorovitch du problème originel de Monge, la notion de déplacement en ligne droite découle du coût euclidien $|x - y|$, ou bien dans le coût quadratique $|x - y|^2$ dans le cas du problème amélioré (c'est à dire cherchant à minimiser une "énergie de déplacement" plutôt que la seule somme des distances : voir la sous-section I.1.4).

La réduction du problème de transport optimal à un problème d'assignation aura permis dans un premier temps de simplifier le traitement du problème dans sa dimension théorique (voir les divers Théorèmes cités précédemment). Cependant, pour ce qui est de traiter le problème dans la pratique, c'est à dire d'explicitier de façon plus ou moins précise le plan d'assignation optimal, cette formulation en terme d'assignation ne s'avère plus forcément la façon la plus simple de traiter le problème. Certaines méthodes choisissent même d'abandonner la notion d'assignation pour revenir entièrement à la notion de déplacement : ce sera le cas de la méthode que nous présentons ici, à propos de la résolution pratique du problème formulé dans la section précédente, dont nous rappelons les conditions :

- évolution dans un espace euclidien : nous nous placerons dans \mathbb{R}^d ,
- un coût proportionnel euclidien quadratique $c(x, y) = |x - y|^2$,
- la mesure de départ μ est supposée absolument continue par rapport à la mesure de Lebesgue, i.e. $\exists \rho_0 \in L^1(\mathbb{R}^d, \mathcal{L}^d)$, $\mu = \rho_0 \cdot \mathcal{L}^d$.

Nous présentons donc ici la méthode proposée par J. D. Benamou et Y. Brenier dans [6] pour traiter cette nouvelle formulation, à savoir une réécriture problème de Monge en un problème de déplacement continue de masse, typique de la mécanique des fluides.

La première initiative de rapprochement entre le problème d'assignation optimale et celui de déplacement optimisé fut proposée par R. J. McCann. Celui-ci envisagea l'étude de déplacements

continus entre les mesures dans l'espace $(\mathcal{P}_2(\mathbb{R}^d), \mathcal{W}_2)$. Il s'agirait de déterminer une famille de mesures intermédiaires dans $\mathcal{P}_2(\mathbb{R}^d)$ entre deux mesures μ et ν , dont la somme des coûts d'assignations optimaux de proche en proche (i.e. la distance de Wasserstein de proche en proche) serait égale au coût d'assignation optimal entre μ et ν (i.e. à la distance Wasserstein entre μ et ν). L'idée étant qu'à l'échelle infinitésimal, il n'y a pas de différence entre une assignation optimale et un déplacement optimal, entre une "téléportation à une distance infiniment proche" et un "déplacement sur une distance infiniment petite".

Le problème initié par McCann était plus précisément de déterminer une famille d'interpolation continue entre les mesures μ et ν pour la distance de Wasserstein \mathcal{W}_2 : autrement dit, de déterminer une famille de mesures $(\mu_t)_{0 \leq t \leq 1}$, telle que $\mu_0 = \mu$ et $\mu_1 = \nu$ et telle que pour tout $0 \leq t_1 \leq t_2 \leq t_3 \leq 1$ on ait

$$\mathcal{W}_2(\mu_{t_1}, \mu_{t_3}) = \mathcal{W}_2(\mu_{t_1}, \mu_{t_2}) + \mathcal{W}_2(\mu_{t_2}, \mu_{t_3}), \quad (1.1-22)$$

avec continuité sur $[0, 1]$ de l'application $t \mapsto \mu_t$ pour la topologie induite par \mathcal{W}_2 sur $\mathcal{P}_2(\mathbb{R}^d)$: c'est à dire, telle que pour tout $t_0 \in [0, 1]$, on ait $\mathcal{W}_2(\mu_t, \mu_{t_0}) \rightarrow 0$ ($t \rightarrow t_0$).

McCann propose alors dans [56] de définir pour tout $t \in [0, 1]$ la mesure

$$\mu_t = [(1-t) \text{id} + t \nabla \phi] \# \mu \quad (1.1-23)$$

La famille $(\mu_t)_{0 \leq t \leq 1}$ correspond ainsi à la poussée de la mesure μ par la famille $(T_t)_{0 \leq t \leq 1}$ des interpolations "euclidiennes" (correspondant donc au coût $c(x, y) = |x - y|^2$) entre l'opérateur identité et l'opérateur de transport optimal $\nabla \phi$, c'est à dire avec $T_t = (1-t) \text{id} + t \nabla \phi$ pour tout $t \in [0, 1]$. La famille $(\mu_t)_{0 \leq t \leq 1}$ répond alors bien à la condition de continuité pour \mathcal{W}_2 et ainsi qu'à la condition d'interpolation (1.1-22). Plus précisément, la famille $(\mu_t)_{0 \leq t \leq 1}$ vérifie la propriété suivante (laquelle implique bien les conditions de continuité et d'interpolation) :

Proposition 1.1-8. INTERPOLATION DE MCCANN [1]

Soient $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$: deux mesures de probabilité sur la tribu de Lebesgue de \mathbb{R}^d vérifiant la condition

$$\int_{\mathbb{R}^d} |x|^2 d\mu(x), \int_{\mathbb{R}^d} |y|^2 d\nu(y) < +\infty. \quad (1.1-24)$$

et telles que la mesure (de départ) μ soit absolument continue par rapport à la mesure de Lebesgue.

Soit également ϕ une fonction convexe telle que définie dans le Théorème 1.1-7 (p. 20), avec en particulier l'opérateur $\nabla \phi$ solution du problème de Monge pour le transport optimal de la mesure μ (départ) sur la mesure ν (arrivée) pour le coût proportionnel $c(x, y) = |x - y|^2$ ($x, y \in \mathbb{R}^d$). Autrement dit $\nu = \phi \# \mu$ et

$$\int_{\mathbb{R}^d} |x - \nabla \phi(x)|^2 d\mu(x) = \inf_{M \# \mu = \nu} \int_{\mathbb{R}^d} |x - M(x)|^2 d\mu(x) = \mathcal{W}_2(\mu, \nu)^2,$$

où \mathcal{W}_2 désigne la distance de Wasserstein associée au coût c décrite dans la Définition 1.1-5 et la notation 1 (p. 19).

Pour tout $t \in [0, 1]$, on définit la mesure

$$\mu_t = [(1-t) \text{id} + t \nabla \phi] \# \mu.$$

- Alors le chemin $t \in [0, 1[\mapsto \mu_t$ est un **chemin géodésique** de $\mathcal{P}_2(\mathbb{R}^d)$. Plus précisément, pour tous $0 \leq t < t' < 1$ (donc $(t, t') \neq (1, 1)$), on a $\mu_t \in \mathcal{P}_2(\mathbb{R}^d)$ et

$$\mathcal{W}_2(\mu_{t'}, \mu_t) = (t' - t)\mathcal{W}_2(\mu, \nu). \quad (1.1-25)$$

- Si de plus on suppose que la mesure d'arrivée ν est également absolument continue par rapport à la mesure de Lebesgue, alors le chemin prolongé en 1, c'est à dire $t \in [0, 1] \mapsto \mu_t$, est également un chemin géodésique de $\mathcal{P}_2(\mathbb{R}^d)$. Plus précisément, l'équation (1.1-25) est vérifiée pour tout $t, t' \in [0, 1]$, $t \leq t'$ (donc également pour t ou t' égal à 1).

L'équation (1.1-25) implique immédiatement la condition d'interpolation (1.1-22).

Preuve : Pour tout $t \in [0, 1]$, on pose $\phi_t = (1 - t)|\cdot|^2/2 + t\phi$, et donc $\nabla\phi_t = (1 - t)\text{id} + t\nabla\phi$.

- On suppose tout d'abord que seule la mesure μ est supposée absolument continue par rapport à la mesure de Lebesgue. Alors, pour tout $t \in [0, 1[$, les mesures μ_t sont également absolument continues par rapport à la mesure de Lebesgue (pour la preuve, voir la Proposition F.0-6 de l'Annexe 236).

De même, on peut affirmer que pour tout $t \in [0, 1[$, on a $\mu_t \in \mathcal{P}_2(\mathbb{R}^d)$: en effet, pour tout $t \in [0, 1]$, on a

$$\begin{aligned} \int_{\mathbb{R}^d} |y|^2 d\mu_t(y) &= \int_{\mathbb{R}^d} |y|^2 d(\nabla\phi_t \# \mu)(y) \\ &= \int_{\mathbb{R}^d} |(1-t)x + t\nabla\phi(x)|^2 d\mu(x) \\ &\leq (1-t) \int_{\mathbb{R}^d} |x|^2 d\mu(x) + t \int_{\mathbb{R}^d} |\nabla\phi(x)|^2 d\mu(x) \\ &= (1-t) \int_{\mathbb{R}^d} |x|^2 d\mu(x) + t \int_{\mathbb{R}^d} |z|^2 d(\nabla\phi \# \mu)(z) \\ &\leq \int_{\mathbb{R}^d} |x|^2 d\mu(x) + \int_{\mathbb{R}^d} |z|^2 d\nu(z) < +\infty, \end{aligned}$$

d'où $\mu_t \in \mathcal{P}_2(\mathbb{R}^d)$.

Ainsi, pour tous $0 \leq t < t' < 1$, il sera possible d'appliquer le Théorème 1.1-7 (p. 20), c'est à dire que l'opérateur de transport optimal de $\mu_{t'}$ sur μ_t est le gradient d'une fonction convexe.

Nous allons donc dans un premier temps chercher à exhiber cet opérateur de poussée de $\mu_{t'}$ sur μ_t . Pour cela, nous nous baserons sur la formulation duale de la notion de poussée de mesure, à savoir que :

$$\mu_t = M_{t'}^t \# \mu_{t'} \Leftrightarrow \forall f \in C_c(\mathbb{R}^d), \int_{\mathbb{R}^d} f(y) d\mu_t(y) = \int_{\mathbb{R}^d} f(M_{t'}^t(x)) d\mu_{t'}(x). \quad (1.1-26)$$

Soient $0 \leq t < t' < 1$. Pour tout $f \in C_c(\mathbb{R}^d)$, on a

$$\begin{aligned}
\int_{\mathbb{R}^d} f(y) d\mu_t(y) &= \int_{\mathbb{R}^d} f(y) d(\nabla\phi_t \# \mu)(y) \\
&= \int_{\mathbb{R}^d} f((1-t)x + t\nabla\phi(x)) d\mu(x) \\
&= \int_{\mathbb{R}^d} f((1-t)x + t\nabla\phi(x)) d\mu(x) \\
&= \int_{\mathbb{R}^d} f\left((1-t)x + \frac{t}{t'}((1-t')x + t'\nabla\phi(x)) - \frac{t(1-t')}{t'}x\right) d\mu(x) \\
&= \int_{\mathbb{R}^d} f\left(\frac{t'-t}{t'}x + \frac{t}{t'}\nabla\phi_{t'}(x)\right) d\mu(x)
\end{aligned}$$

Il est possible, comme nous le verrons dans le chapitre III (section III.2), de choisir la fonction ϕ assez régulière pour que celle-ci soit localement lipschitzienne sur \mathbb{R}^d et ainsi différentiable \mathcal{L}^d -presque partout sur \mathbb{R}^d (voir le Théorème de Rademacher p. 98). On peut alors également prouver que la fonction $(\phi_{t'})^*$ est de classe C^1 sur \mathbb{R}^d (voir la Définition 3.3-9 p. 89), et que pour tout $y \in \mathbb{R}^d$ tel que ϕ soit différentiable en y (donc pour presque tout $y \in \mathbb{R}^d$), on a

$$(\nabla_x(\phi_t)^* \circ \nabla\phi_t)(y) = y$$

(voir la Proposition 3.3-10 p. et la Remarque 3.3-13 p. 91). On a donc alors :

$$\begin{aligned}
\int_{\mathbb{R}^d} f(y) d\mu_t(y) &= \int_{\mathbb{R}^d} f\left(\frac{t'-t}{t'}(\nabla(\phi_{t'})^* \circ \nabla\phi_{t'})(x) + \frac{t}{t'}\nabla\phi_{t'}(x)\right) d\mu(x) \\
&= \int_{\mathbb{R}^d} f\left(\frac{t'-t}{t'}\nabla(\phi_{t'})^*(z) + \frac{t}{t'}z\right) d(\nabla\phi_{t'} \# \mu)(z).
\end{aligned}$$

Ainsi, $\mu_t = \nabla\phi_{t'}^t \# (\nabla\phi_{t'} \# \mu) = \nabla\phi_{t'}^t \# \mu_{t'}$, avec

$$\phi_{t'}^t = \left(1 - \frac{t}{t'}\right) (\phi_{t'})^* + \frac{t|\cdot|^2}{2t'} \text{id}.$$

Or, du fait que $t < t'$, on a $1 - t/t' > 0$, et donc la fonction $\phi_{t'}^t$ est une fonction s.c.i. convexe sur \mathbb{R}^d .

Ainsi, d'après le Théorème 1.1-7 (p. 20), l'opérateur $\nabla\phi_{t'}^t$ est l'unique gradient de fonction convexe (unique au sens $\mu_{t'}$ -presque partout) poussant la mesure $\mu_{t'}$ sur la mesure μ_t , et

on a donc :

$$\begin{aligned}
\mathcal{W}_2(\mu_{t'}, \mu_t) &= \left(\int_{\mathbb{R}^d} |y - \nabla \phi_{t'}^t(y)|^2 d\mu_{t'}(y) \right)^{1/2} = \left(\int_{\mathbb{R}^d} |y - \nabla \phi_{t'}^t(y)|^2 d(\nabla \phi_{t'} \# \mu)(y) \right)^{1/2} \\
&= \left(\int_{\mathbb{R}^d} \left| y - \left[\left(1 - \frac{t}{t'}\right) \nabla(\phi_{t'})^*(y) + \frac{t}{t'} y \right] \right|^2 d(\nabla \phi_{t'} \# \mu)(y) \right)^{1/2} \\
&= \left(1 - \frac{t}{t'}\right) \left(\int_{\mathbb{R}^d} |y - \nabla(\phi_{t'})^*(y)|^2 d(\nabla \phi_{t'} \# \mu)(y) \right)^{1/2} \\
&= \left(1 - \frac{t}{t'}\right) \left(\int_{\mathbb{R}^d} |\nabla \phi_{t'}(x) - (\nabla(\phi_{t'})^* \circ \nabla \phi_{t'})(x)|^2 d\mu(x) \right)^{1/2} \\
&= \left(1 - \frac{t}{t'}\right) \left(\int_{\mathbb{R}^d} |(1-t')x + t' \nabla \phi(x) - x|^2 d\mu(x) \right)^{1/2} \\
&= (t' - t) \left(\int_{\mathbb{R}^d} |\nabla \phi(x) - x|^2 d\mu(x) \right)^{1/2} \\
&= (t' - t) \mathcal{W}_2(\mu, \nu).
\end{aligned} \tag{1.1-27}$$

- Supposons maintenant que la mesure d'arrivée ν est également absolument continue par rapport à la mesure de Lebesgue, et traitons maintenant le cas où $t' = 1$. Soit $t \in [0, 1]$.

D'après le dernier point du Théorème 1.1-7, on a alors $\nu = \nabla \phi \# \mu \Leftrightarrow \mu = \nabla \phi^* \# \nu$, et pour ν -presque tout $z \in \mathbb{R}^d$, on a $\nabla \phi \circ \nabla \phi^*(z) = z$. Donc, pour tout $f \in C_c(\mathbb{R}^d)$, on a

$$\begin{aligned}
\int_{\mathbb{R}^d} f(y) d\mu_t(y) &= \int_{\mathbb{R}^d} f(y) d(\nabla \phi_t \# \mu)(y) \\
&= \int_{\mathbb{R}^d} f((1-t)x + t \nabla \phi(x)) d\mu(x) \\
&= \int_{\mathbb{R}^d} f((1-t)x + t \nabla \phi(x)) d(\nabla \phi^* \# \nu)(x) \\
&= \int_{\mathbb{R}^d} f((1-t) \nabla \phi^*(z) + t(\nabla \phi \circ \nabla \phi^*)(z)) d\nu(z) \\
&= \int_{\mathbb{R}^d} f((1-t) \nabla \phi^*(z) + tz) d\nu(z),
\end{aligned}$$

c'est à dire :

$$\mu_t = [(1-t) \text{id} + t \nabla \phi] \# \mu = [t \text{id} + (1-t) \nabla \phi^*] \# \nu.$$

L'opérateur $t \text{id} + (1-t) \nabla \phi = \nabla(t|\cdot|^2/2 + (1-t)\phi^*)$ est donc également l'unique gradient de fonction convexe (unique au sens ν -presque partout) poussant la mesure ν sur la mesure μ , d'où :

$$\begin{aligned}
\mathcal{W}_2(\mu_t, \mu_1) &= \mathcal{W}_2(\mu_1, \mu_t) = \mathcal{W}_2(\nu, \mu_t) \\
&= \left(\int_{\mathbb{R}^d} |x - (t \text{id} + (1-t) \nabla \phi^*)|^2 d\nu(x) \right)^{1/2} \\
&= (1-t) \left(\int_{\mathbb{R}^d} |x - \nabla \phi^*(x)|^2 d\nu(x) \right)^{1/2} \\
&= (1-t) \mathcal{W}_2(\nu, \mu) = (1-t) \mathcal{W}_2(\mu, \nu).
\end{aligned} \tag{1.1-28}$$

□

Nous supposons pour toute la suite, que la mesure d'arrivée ν est absolument continue par rapport à la mesure de Lebesgue (i.e. $\exists \rho_1 \in L^1(\mathbb{R}^d, \mathcal{L}^d)$, $\nu = \rho_1 \cdot \mathcal{L}^d$).

Le problème de McCann nous permet donc de disposer d'une interpolation continue de mesures $(\mu_t)_{0 \leq t \leq 1}$ entre μ et ν . Rappelons que pour tous $t \in [0, 1[$, les mesures μ_t sont également absolument continues par rapport à la mesure de Lebesgue : il existe donc tout $t \in [0, 1]$ une densité $\rho_t \in L^1(\mathbb{R}^d, \mathcal{L}^d)$ tel que $\mu_t = \rho_t \mathcal{L}^d$ (en particulier $\mu = \mu_0 = \rho_0 \mathcal{L}^d$ et $\nu = \mu_1 = \rho_1 \mathcal{L}^d$). Les densités ρ_t peuvent se définir formellement par rapport à ρ_0 et ϕ via l'équation de Monge-Ampère (voir la Remarque 1.1-10 p. 17) :

pour tout $t \in [0, 1]$, en posant $\tilde{X}(t, \cdot) = (1 - t) \text{id} + t \nabla_x \phi$, on a (formellement)

$$\rho_0 = \rho_t \circ (\tilde{X}(t, \cdot)) \det \left(\nabla_x \tilde{X}(t, \cdot) \right) = (\rho_t \circ ((1 - t) \text{id} + t \nabla_x \phi)) \det \left((1 - t)I + t \text{Hess}(\phi) \right). \quad (1.1-29)$$

ou formulé autrement :

$$\rho_t = \frac{\rho_0 \circ \left(\tilde{X}^{-1}(t, \cdot) \right)}{\det \left(\left(\nabla_x \tilde{X} \right) \left(t, \tilde{X}^{-1}(t, \cdot) \right) \right)} \quad (1.1-30)$$

Les mesures μ et ν étant deux mesures de probabilité, on a

$$\forall t \in [0, 1], \quad \int_{\mathbb{R}^d} \rho_t = \int_{\mathbb{R}^d} \rho_0 = \int_{\mathbb{R}^d} \rho_1 = 1. \quad (1.1-31)$$

L'équation ci-dessus, mais surtout la propriété de transport conservatif induite par la notion de poussée de mesure (voir la Remarque 1.1-10 p. 17) laisse à entendre que les densités ρ_t vérifient une propriété de **conservation de la masse** dans leur transport de ρ_0 sur ρ_1 .

C'est en se basant sur cette idée d'interpolation continue entre les densités ρ_0 et ρ_1 par des densités ρ_t de masse constante que des algorithmes du second ordre ont par exemple été récemment proposés pour résoudre le problème de manière rapide [5]. Ce type d'approche est néanmoins limité aux densités ne s'annulant pas. La seule approche numérique permettant de gérer de telles données prend sa source dans les travaux de Y. Brenier et J.D. Benamou, grâce à leur formulation plus "physique" de ce type de problème. Benamou et Brenier ont proposé dans [6] (article datant de 1999) une formulation continue du problème de transport optimal en faisant appel aux équations de la mécanique des fluides : il s'agirait de déterminer directement l'évolution de la densité $\rho = t \mapsto \rho_t$ ainsi que le champ de vitesse v auquel cette première est advectée, celui-ci correspondant donc au plan de transport optimal $\nabla \phi$. Un tel couple (ρ, v) doit donc en toute logique vérifier l'équation de continuité

$$\partial_t \rho + \text{div}(\rho v) = 0, \quad (1.1-32)$$

ainsi que les conditions finales et initiales

$$\rho(0, \cdot) = \rho_0, \quad \text{et} \quad \rho(1, \cdot) = \rho_1. \quad (1.1-33)$$

Le problème est donc de déterminer parmi tous les couples densité-vitesse (ρ, v) vérifiant la condition ci-dessus, celui qui correspond à notre problème, c'est à dire pour lequel on a bien $\rho_t \mathcal{L}^d = [(1 - t) \text{id} + t \nabla \phi] \# (\rho_0 \mathcal{L}^d)$ pour (presque) tout $t \in [0, 1]$.

D'après l'intuition "physique" développée dans les sous-sections précédentes (voir en particulier

la Remarque 1.1-4 p. 5, l'illustration 1.1-6 p. 8, ainsi que l'exemple 15 p. 1.1-9) un transport optimal basé sur un coût quadratique (c'est à dire ici $c(x, y) = |x - y|^2$) devrait correspondre à un déplacement minimisant la "dépense d'énergie globale" nécessaire au transport qui, comme nous l'avons plusieurs fois illustré (cf les exemples cités en début de paragraphe) correspondrait à un transport où la trajectoire de chaque élément s'effectue en ligne droite, sans croiser d'autres trajectoires. Comme nous revenons ici à une notion de déplacement et plus seulement d'assignation, nous pouvons ajouter que l'intuition physique nous amène à penser que le déplacement le long de chaque trajectoire se fera à **vitesse constante**.

L'énergie "physique" globale dépensée $\mathcal{K}(\rho, v)$ correspondant à un tel déplacement conservatif de masse correspond, pour un couple (ρ, v) donné, à la valeur suivante :

$$\mathcal{K}(\rho, v) = \frac{1}{2} \int_0^1 \int_{\mathbb{R}^d} |v|^2 \rho \, dx \, dt. \quad (1.1-34)$$

Le problème de transport optimal se ramène donc selon cette nouvelle formulation à trouver un couple $(\tilde{\rho}, \tilde{v})$ minimisant l'énergie de déplacement (1.1-34) parmi l'ensemble des couples (ρ, v) vérifiant les conditions (1.1-32) et (1.1-33).

Or, il est montré (formellement) dans [6] que, sous condition d'existence d'un champ de vitesse v approprié, l'évolution de densité vérifiant $\rho_t \mathcal{L}^d = [(1-t) \text{id} + t \nabla \phi] \# (\rho_0 \mathcal{L}^d)$ associée à ce champ de vitesse est bien une solution de ce nouveau problème. Reprenons donc ici le raisonnement utilisé par Benamou et Brenier dans [6] (l'énoncé et la preuve qui suivent sont **purement formels**) :

Énoncé formel 1.1-9. FORMULATION MÉCANIQUE DES FLUIDES

Définissons l'ensemble (formel) V des couples densité-vitesse (ρ, v) (avec donc $\rho \geq 0$ en tant que densité de probabilité) vérifiant les conditions (1.1-32) et (1.1-33), et définissons le terme d'énergie optimale de transport de la densité ρ_0 sur la densité ρ_1 .

$$\mathcal{E}(\rho_0, \rho_1) = 2 \times \inf_{(\rho, v) \in V} \mathcal{K}(\rho, v) \quad (1.1-35)$$

On a alors l'égalité

$$\mathcal{E}(\rho_0, \rho_1) = \mathcal{W}_2(\rho_0 \mathcal{L}^d, \rho_1 \mathcal{L}^d)^2, \quad (1.1-36)$$

De plus, il existe un couple $(\tilde{\rho}, \tilde{v}) \in V$ tel que $\mathcal{K}(\tilde{\rho}, \tilde{v}) = \frac{1}{2} \mathcal{E}(\rho_0, \rho_1)$ avec pour (presque) tout $t \in [0, 1]$,

$$\tilde{\rho}(t, \cdot) \mathcal{L}^d = [(1-t) \text{id} + t \nabla \phi] \# (\rho_0 \mathcal{L}^d)$$

i.e. $\rho_0 = (\tilde{\rho}(t, (1-t) \text{id} + t \nabla_x \phi) \det((1-t)I + t \text{Hess}(\phi)))$.

Preuve formelle :

- Montrons dans un premier temps que l'on a bien l'inégalité $\mathcal{E}(\rho_0, \rho_1) \leq \mathcal{W}_2(\rho_0 \mathcal{L}^d, \rho_1 \mathcal{L}^d)^2$.

Soit $(\rho, v) \in V$. Soit le champ de déplacement X résolvant pour tout $x \in \mathbb{R}^d$ le problème de Cauchy-Lipschitz :

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

On veut tout d'abord montrer que pour (presque) tout $t \in [0, 1]$, la densité ρ_t est telle que $\rho(t, \cdot) \mathcal{L}^d = X(t, \cdot) \# (\rho_0 \mathcal{L}^d)$, c'est à dire que $X(t, \cdot)$ vérifie l'équation de Monge-Ampère :

$$\rho_0 = \rho(t, X(t, \cdot)) \det(\nabla_x X(t, \cdot)). \quad (1.1-37)$$

On définit pour tout $t \in [0, 1]$ le jacobien $J(t, \cdot) = \det(\nabla_x X(t, \cdot))$ (en particulier on a $J(0, \cdot) = \det(\nabla_x \text{id}) = \det(\text{I}) = 1$)

Pour montrer l'équation (1.1-37), pour tout $t \in [0, 1]$, il suffit donc de prouver que $\frac{d}{dt}(\rho(t, X(t, \cdot)) J(t, \cdot)) = 0$. En effet, on aura alors

$$\rho(t, X(t, \cdot)) \det(\nabla_x X(t, \cdot)) = \rho(t, X(t, \cdot)) J(t, \cdot) = \rho_0 J(0, \cdot) = \rho_0.$$

autrement dit l'équation (1.1-37).

Sachant que le jacobien J vérifie la relation $\partial_t J = \text{div}_x(v)J$, on a :

$$\begin{aligned} \frac{d}{dt}(\rho(t, X(t, \cdot)) J(t, \cdot)) &= \frac{d}{dt}(\rho(t, X(t, \cdot))) J(t, \cdot) + \rho(t, X(t, \cdot)) \frac{d}{dt} J(t, \cdot) \\ &= [\partial_t \rho(t, X(t, \cdot)) + \partial_t X(t, \cdot) \nabla_x \rho(t, X(t, \cdot))] J(t, \cdot) \\ &\quad + \rho(t, X(t, \cdot)) (\text{div}_x v)(t, X(t, \cdot)) J(t, \cdot) \\ &= (\partial_t \rho + \text{div}_x(\rho v))(t, X(t, \cdot)) J(t, \cdot) = 0 \end{aligned} \quad (1.1-38)$$

Ainsi, pour (presque) tout $t \in [0, 1]$, $X(t, \cdot)$ est un plan de transport de ρ_0 sur $\rho(t, \cdot)$. En particulier, on a

$$\rho(1, \cdot) \mathcal{L}^d = \rho_1 \mathcal{L}^d = X(1, \cdot) \# (\rho_0 \mathcal{L}^d). \quad (1.1-39)$$

Comme déjà évoqué dans la Remarque 1.1-10 (p. 17), l'expression duale de la poussée de mesure $\rho(t, \cdot) \mathcal{L}^d = X(t, \cdot) \# (\rho_0 \mathcal{L}^d)$ se formule de la façon suivante :

$$\forall f \in C_c(\mathbb{R}^d), \quad \int_{\mathbb{R}^d} f(y) \rho(t, y) dy = \int_{\mathbb{R}^d} f(X(t, x)) \rho_0(x) dx. \quad (1.1-40)$$

Comme nous raisonnons de façon formelle, nous nous permettons de prolonger la formulation (1.1-40) au delà de $C_c(\mathbb{R}^d)$, de façon à pouvoir ainsi utiliser le carré du champ de vitesse $|v(t, \cdot)|^2$ comme fonction test, ainsi qu'il suit :

$$\begin{aligned} \mathcal{K}(\rho, v) &= \frac{1}{2} \int_0^1 \int_{\mathbb{R}^d} |v(t, y)|^2 \rho(t, y) dy dt = \frac{1}{2} \int_0^1 \int_{\mathbb{R}^d} |v(t, X(t, x))|^2 \rho_0(x) dx dt \\ &= \frac{1}{2} \int_0^1 \int_{\mathbb{R}^d} |\partial_t X(t, x)|^2 \rho_0(x) dx dt = \frac{1}{2} \int_{\mathbb{R}^d} \left(\int_0^1 |\partial_t X(t, x)|^2 dt \right) \rho_0(x) dx \\ &\geq \frac{1}{2} \int_{\mathbb{R}^d} \left| \int_0^1 \partial_t X(t, x) dt \right|^2 \rho_0(x) dx = \frac{1}{2} \int_{\mathbb{R}^d} |X(1, x) - X(0, x)|^2 \rho_0(x) dx \\ &= \frac{1}{2} \int_{\mathbb{R}^d} |X(1, x) - x|^2 \rho_0(x) dx. \end{aligned} \quad (1.1-41)$$

Or, d'après l'équation (1.1-39), on a $\rho_1 \mathcal{L}^d = X(1, \cdot) \# (\rho_0 \mathcal{L}^d)$. Ainsi, d'après le Théorème 1.1-7 (p. 1.1-7), on a donc :

$$\begin{aligned} \mathcal{K}(\rho, v) &= \frac{1}{2} \int_{\mathbb{R}^d} |X(1, x) - x|^2 \rho_0(x) dx \\ &\geq \frac{1}{2} \inf_{M \# \mu = \nu} \int_{\mathbb{R}^d} |x - M(x)|^2 d\mu(x) \\ &= \frac{1}{2} \int_{\mathbb{R}^d} |\nabla_x \phi(x) - x|^2 \rho_0(x) dx = \frac{1}{2} \mathcal{W}_2(\rho_0 \mathcal{L}^d, \rho_1 \mathcal{L}^d)^2. \end{aligned} \quad (1.1-42)$$

- Définissons maintenant pour tout $t \in [0, 1]$ la densité $\tilde{\rho}(t, \cdot)$ de la poussée de la mesure $\rho_0 \mathcal{L}^d$ par le plan de transport $\tilde{X}(t, \cdot) = (1 - t) \text{id} + t \nabla \phi$, c'est à dire que

$$\tilde{\rho}(t, \cdot) \mathcal{L}^d = \tilde{X}(t, \cdot) \# (\rho_0 \mathcal{L}^d) = [(1 - t) \text{id} + t \nabla \phi] \# (\rho_0 \mathcal{L}^d).$$

Définissons également un champ de vitesse \tilde{v} tel que \tilde{X} vérifie pour tout $t \in [0, 1]$:

$$\partial_t \tilde{X}(t, \cdot) = \nabla \phi - \text{id} = \tilde{v}(t, \tilde{X}(t, \cdot))$$

(le terme $\partial_t \tilde{X}(t, \cdot)$ est donc indépendant du temps).

Montrons que $(\tilde{\rho}, \tilde{v}) \in V$ et que $\mathcal{K}(\tilde{\rho}, \tilde{v}) = \frac{1}{2} \mathcal{W}_2(\rho_0 \mathcal{L}^d, \rho_1 \mathcal{L}^d)^2$, ce qui conclura sur la relation (1.1-36).

Commençons donc par montrer que le couple $(\tilde{\rho}, \tilde{v})$ vérifie l'équation (1.1-32) (les conditions (1.1-33) étant quant à elles vérifiées par Définition de $\tilde{\rho}$). Il est plus simple pour démontrer cette propriété de raisonner au sens faible : pour tout $\varphi \in C_c([0, 1] \times \mathbb{R}^d)$, on a

$$\begin{aligned} & \int_0^1 \int_{\mathbb{R}^d} (\partial_t \tilde{\rho} + \text{div}_x(\tilde{\rho} \tilde{v})) \varphi \, dy \, dt \\ &= - \int_0^1 \int_{\mathbb{R}^d} (\partial_t \varphi(t, y) + \tilde{v}(t, y) \cdot \nabla_x \varphi(t, y)) \tilde{\rho}(t, y) \, dy \, dt \\ &= - \int_0^1 \int_{\mathbb{R}^d} (\partial_t \varphi(t, X(t, x)) + \tilde{v}(t, X(t, x)) \cdot \nabla_x \varphi(t, X(t, x))) \tilde{\rho}_0(x) \, dx \, dt \\ &= - \int_{\mathbb{R}^d} \left(\int_0^1 (\partial_t \varphi(t, X(t, x)) + \partial_t X(t, x) \cdot \nabla_x \varphi(t, X(t, x))) \, dt \right) \tilde{\rho}_0(x) \, dx \\ &= - \int_{\mathbb{R}^d} \left(\int_0^1 \frac{d}{dt} \varphi(t, X(t, x)) \, dt \right) \tilde{\rho}_0(x) \, dx \\ &= - \int_{\mathbb{R}^d} (\varphi(1, X(1, x)) - \varphi(0, X(0, x))) \tilde{\rho}_0(x) \, dx \\ &= 0. \end{aligned} \tag{1.1-43}$$

On a donc bien $\partial_t \tilde{\rho} + \text{div}_x(\tilde{\rho} \tilde{v}) = 0$, et ainsi $(\tilde{\rho}, \tilde{v}) \in V$.

On raisonne ensuite comme précédemment :

$$\begin{aligned} \mathcal{K}(\tilde{\rho}, \tilde{v}) &= \frac{1}{2} \int_0^1 \int_{\mathbb{R}^d} |\tilde{v}(t, x)|^2 \tilde{\rho}(t, x) \, dx \, dt = \frac{1}{2} \int_0^1 \int_{\mathbb{R}^d} |\tilde{v}(t, \tilde{X}(t, y))|^2 \rho_0(y) \, dy \, dt \\ &= \frac{1}{2} \int_0^1 \int_{\mathbb{R}^d} |\nabla_x \phi(y) - y|^2 \rho_0(y) \, dy \, dt \\ &= \frac{1}{2} \int_{\mathbb{R}^d} |\nabla_x \phi(y) - y|^2 \rho_0(y) \, dy \\ &= \frac{1}{2} \mathcal{W}_2(\rho_0 \mathcal{L}^d, \rho_1 \mathcal{L}^d)^2. \end{aligned} \tag{1.1-44}$$

Cela conclut ainsi notre preuve formelle. □

I.1.6 Géodésique convexité

Classiquement, dans un espace vectoriel E , on dit qu'un ensemble C est convexe si et seulement si, pour tous points $a, b \in C$, le segment $[a, b]$ est inclus dans C . Si l'espace vectoriel E est normé, on Remarque que ce segment $[a, b]$ correspond en fait, relativement à la distance induite par la norme, à l'unique chemin géodésique reliant le point a au point b . D'où l'idée d'essayer d'étendre la notion de convexité à d'autres espaces vectoriels métriques adaptés (c'est à dire des espaces dans lesquels tous les points sont reliés par des géodésiques) par la notion de géodésique : un sous-ensemble C d'un espace vectoriel métrique (F, d) serait dit **géodésiquement convexe** si et seulement si, pour tous points $a, b \in C$, l'ensemble des chemins géodésiques reliant le point a au point b sont inclus dans C . Par extension, il s'agirait de substituer à l'interpolation linéaire classique, la notion d'interpolation géodésique, c'est à dire dans le cas du transport optimal, l'interpolation de McCann.

On pourrait étendre également la notion de fonction convexe. En effet, classiquement, sur un espace vectoriel E , une fonction $f : E \rightarrow \mathbb{R}$ est dite convexe si et seulement pour tous points $x, y \in E$ et tout $t \in [0, 1]$, on a

$$f(x + t(y - x)) \leq (1 - t)f(x) + tf(y).$$

On peut alors, de même que précédemment, étendre cette notion à d'autres espaces vectoriels métriques adaptés (comme précédemment) (F, d) de la façon suivante : on pourra qualifier une fonction $f : F \rightarrow \mathbb{R}$ de *géodésiquement convexe* si et seulement si, pour tous points $x, y \in F$, et pour tout chemin géodésique paramétré $\gamma : [0, 1] \rightarrow F$ reliant le point x au point y , et pour tout $t \in [0, 1]$, on a

$$f(\gamma(t)) \leq \frac{l_{\gamma(t), y}}{l_{x, y}} f(x) + \frac{l_{x, \gamma(t)}}{l_{x, y}} f(y), \quad (1.1-45)$$

où la longueur de la portion de chemin géodésique $l_{\gamma(t), \gamma(t')}$ entre deux points $\gamma(t)$ et $\gamma(t')$ ($t, t' \in [0, 1]$) est définie au sens de Jordan (rappelons que les chemins géodésiques sont toujours à variations bornées) par :

$$l_{\gamma(t), \gamma(t')} = \begin{cases} \sup_{\substack{t=t_1 < \dots < t_n = t' \\ n \in \mathbb{N}}} \left(\sum_{i=1}^{n-1} d(\gamma(t_i), \gamma(t_{i+1})) \right) & \text{si } t < t' \\ l_{\gamma(t'), \gamma(t)} & \text{si } t > t' \\ 0 & \text{si } t = t'. \end{cases}$$

Nous allons appliquer cette notion de géodésique convexité au problème du transport optimal, plus précisément à l'espace vectoriel métrique $(\mathcal{P}_2(\mathbb{R}^d), \mathcal{W}_2)$ (voir la Définition 1.1-5 et la notation 1 p. 19). En effet, l'introduction d'une telle notion pourra permettre d'aborder et de formuler par exemple des problèmes d'optimisation convexe, d'équations différentielles ou d'EDP, directement dans l'espace $(\mathcal{P}_2(\mathbb{R}^d), \mathcal{W}_2)$. Un des domaines du transport optimal exploitant sont les problèmes de flots de gradients, que nous évoquerons plus bas.

Nous avons vu dans la Proposition 1.1-8 (p. 23) que dans cet espace, deux mesures μ et ν , si elles sont absolument continues par rapport à la mesure de Lebesgue \mathcal{L}^d , sont toujours reliées par un chemin géodésique $t \in [0, 1] \mapsto \mu_t \in \mathcal{P}_2(\mathbb{R}^d)$, dit d'interpolation de McCann, vérifiant :

- pour tout $t \in [0, 1]$, la mesure μ_t est absolument continue par rapport à la mesure de Lebesgue \mathcal{L}^d : on écrit alors ρ_t sa densité,
- pour tout $t, t' \in [0, 1]$,

$$\mathcal{W}_2(\mu_{t'}, \mu_t) = (t' - t)\mathcal{W}_2(\mu, \nu). \quad (1.1-46)$$

On définit la notion de *fonction géodésiquement convexe au sens de l'interpolation de McCann* de la façon suivante :

Définition 1.1-10. FONCTION GÉODÉSIQUEMENT CONVEXE

Une fonction $\mathcal{F} : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ est dite *géodésiquement convexe* (respectivement *géodésiquement strictement convexe*) si et seulement si, pour toutes mesures $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ absolument continues par rapport à la mesure de Lebesgue, avec $\mu = \rho_0 \mathcal{L}^d$ et $\nu = \rho_1 \mathcal{L}^d$, $\nabla \phi$ le plan de transport optimal de μ sur ν tel que défini dans le théorème 1.1-7 (p. 20), et en définissant le chemin géodésique $t \in [0, 1] \mapsto \mu_t \in (\mathcal{P}_2(\mathbb{R}^d), \mathcal{W}_2)$ par

$$\forall t \in [0, 1], \quad \mu_t = \tilde{\rho}_t \mathcal{L}^d = [(1-t)\text{id} + t\nabla\phi] \# (\rho_0 \mathcal{L}^d),$$

comme dans la Définition 1.1-8 (p. 1.1-8), alors la fonction $t \mapsto \mathcal{F}(\mu_t)$ est convexe (respectivement strictement convexe) sur $[0, 1]$. Autrement dit, pour tout $t \in [0, 1]$, on a

$$\mathcal{F}(\mu_t) \leq (\text{resp. } <) (1-t)\mathcal{F}(\mu) + t\mathcal{F}(\nu). \quad (1.1-47)$$

Remarque 1.1-14. Pour raccorder la relation 1.1-47 avec la Définition de la géodésique convexit  que nous avons premi rement  nonc e par la relation 1.1-45, remarquons que la longueur $l_{\mu_t, \mu_{t'}}^{\mathcal{W}}$ des g od siques de $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ s' crit au sens de Jordan, pour $0 \leq t < t' \leq 1$,

$$\begin{aligned} l_{\mu_t, \mu_{t'}}^{\mathcal{W}} &= \sup_{\substack{t=t_1 < \dots < t_n = t' \\ n \in \mathbb{N}}} \left(\sum_{i=1}^{n-1} \mathcal{W}_2(\mu_{t_i}, \mu_{t_{i+1}}) \right) \\ &= \sup_{\substack{t=t_1 < \dots < t_n = t' \\ n \in \mathbb{N}}} \left(\mathcal{W}_2(\mu, \nu) \sum_{i=1}^{n-1} (t_{i+1} - t_i) \right) \\ &= (t' - t) \mathcal{W}_2(\mu, \nu). \end{aligned}$$

On a donc $l_{\mu, \nu}^{\mathcal{W}} = \mathcal{W}_2(\mu, \nu)$ et, pour tout $t \in [0, 1]$, $l_{\mu, \mu_t}^{\mathcal{W}} = t\mathcal{W}_2(\mu, \nu)$ et $l_{\mu_t, \nu}^{\mathcal{W}} = (1-t)\mathcal{W}_2(\mu, \nu)$. On a alors bien  quivalence entre la relation 1.1-47 et la relation 1.1-45, en remarquant que $l_{\mu, \mu_t}^{\mathcal{W}} / l_{\mu, \nu}^{\mathcal{W}} = t$ et $l_{\mu_t, \nu}^{\mathcal{W}} / l_{\mu, \nu}^{\mathcal{W}} = 1 - t$.

Dans [56], McCann met en  vidence trois cat gories de fonctions g od siquement convexes sur $\mathcal{P}_2(\mathbb{R}^d)$, qu'il qualifie d' nergies de transport, en les rapprochant de types d' nergies physiques usuelles. Nous  voquerons en effet dans la sous-section I.3.2 de certaines  nergies classiques.

Ces trois cat gories d' nergie sont les suivantes :

D finition 1.1-11.  NERGIES INTERNES, POTENTIELLES, ET D'INTERACTIONS ([77] ET [1])

• * nergie interne*

Soit $U :]-\infty, +\infty[\rightarrow [0, +\infty]$ une fonction propre, convexe, semi-continue inf rieurement telle que $U(0) = 0$.

Soit la fonctionnelle d' nergie interne $\mathcal{U} : \mathcal{P}_2(\mathbb{R}^d) \rightarrow [0, +\infty]$ d finie pour toute mesure $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ par

$$\mathcal{U}(\mu) = \begin{cases} \int_{\mathbb{R}^d} U(\rho(x)) d\mathcal{L}^d(x) & \text{si } \mu = \rho \cdot \mathcal{L}^d \in \mathcal{P}_2(\mathbb{R}^d), \\ +\infty & \text{sinon.} \end{cases} \quad (1.1-48)$$

Si la fonction $s \mapsto s^d U(s^{-d})$ est convexe et non-croissante sur $[0, +\infty[$, alors la fonctionnelle \mathcal{U} est géodésiquement convexe sur $\mathcal{P}_2(\mathbb{R}^d)$.

- **Énergie potentielle**

Soit $V : \mathbb{R}^d \rightarrow]-\infty, +\infty]$ une fonction propre, convexe, semi-continue inférieurement. Soit la fonctionnelle d'énergie potentielle $\mathcal{V} : \mathcal{P}_2(\mathbb{R}^d) \rightarrow]-\infty, +\infty]$ définie pour toute mesure $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ par

$$\mathcal{V}(\mu) = \int_{\mathbb{R}^d} V(x) d\mu(x). \quad (1.1-49)$$

Si la fonction V est convexe (respectivement strictement convexe) sur \mathbb{R}^d , alors la fonctionnelle \mathcal{V} est géodésiquement convexe (respectivement géodésiquement strictement convexe) sur $\mathcal{P}_2(\mathbb{R}^d)$.

- **Énergie d'interaction**

Soit $W : \mathbb{R}^d \rightarrow]-\infty, +\infty]$ une fonction propre, convexe, semi-continue inférieurement et symétrique (i.e. $W(x) = W(-x)$) telle qu'il existe $a, b \in \mathbb{R}_+$ et :

$$\forall x \in \mathbb{R}^d, W(x) \geq -a - b|x|^2$$

Soit la fonctionnelle d'énergie d'interaction $\mathcal{H} : \mathcal{P}_2(\mathbb{R}^d) \rightarrow]-\infty, +\infty]$ définie pour toute mesure $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ par

$$\mathcal{H}(\mu) = \begin{cases} \frac{1}{2} \int_{\mathbb{R}^d} W(x-y) \rho(x) \rho(y) d\mathcal{L}^d(x) d\mathcal{L}^d(y) & \text{si } \mu = \rho \cdot \mathcal{L}^d \in \mathcal{P}_2(\mathbb{R}^d), \\ +\infty & \text{sinon.} \end{cases} \quad (1.1-50)$$

Si la fonction W est convexe sur \mathbb{R}^d , alors la fonctionnelle \mathcal{H} est géodésiquement convexe sur $\mathcal{P}_2(\mathbb{R}^d)$. Si de plus W est strictement convexe sur \mathbb{R}^d alors, pour tout $m \in \mathbb{R}^d$, la fonctionnelle \mathcal{H} est géodésiquement strictement convexe sur le sous-espace de $\mathcal{P}_2(\mathbb{R}^d)$ des mesures de moyenne m .

Dans notre étude, nous ne ferons appel qu'à la notion d'énergie interne dans la section IV.1.1 du chapitre IV, dans le but d'obtenir des propriétés d'intégrabilité sur les densités $(\rho_t)_{t \in [0,1]}$ de l'interpolation de McCann.

Dans la section suivante, nous aborderons le traitement numérique du problème de transport optimal continu dans \mathbb{R}^d pour un coût quadratique par des méthodes se basant sur la formulation dynamique du problème de transport optimal.

I.2 Présentation des algorithmes numériques existants, basés sur la formulation dynamique du problème de transport optimal

Nous allons maintenant aborder le problème de la résolution numérique du problème de transport optimal pour des densités diffuses (non discrètes) et pour un coût quadratique. Certains travaux ont élaboré des méthodes de résolution directe de l'équation de Monge-Ampère (1.1-9). Cependant de telles méthodes ne s'appliquent que dans le cas de densités strictement positives ou à supports convexes. Citons par exemple [63], ou plus récemment [51, 31, 62] et [9, 10].

D'autres méthodes, basées sur la propriété d'un transport optimal de masse associé à un coût

proportionnel quadratique de suivre des trajectoires en ligne droite [16], ont également été proposées dans le but d'accélérer les calculs [47]. Des approximations via le transport de Knothe [22, 11] ou linéarisation du problème pour des densités proches [14] ont de même été étudiées pour la conception d'algorithmes plus rapides.

Nous ne présenterons ici que des méthodes basées sur la formulation dynamique du problème de transport optimal, telle qu'introduite dans [6] : en effet, les algorithmes basés sur cette formulation sont actuellement les seuls à pouvoir traiter en pratique les cas de densités générales positives mais non strictement positives (pouvant donc s'annuler et avoir des supports non convexes). Nous resterons ici dans le cadre stricte du transport optimal entre deux densités de même masse, mais notons tout de même que la formulation dynamique de [6] a également été utilisée pour concevoir des algorithmes de transport optimal entre de mesures de masses différentes, par exemple via l'introduction d'un terme source dans la contrainte de conservation de la masse ou sur le domaine de transport [52, 26].

Revenons donc, là où nous l'avons laissé, à notre exposé sur la formulation dynamique du problème (sous-section I.1.5). Nous y avons vu que les auteurs de [6] avaient démontré que le transport optimal pouvait être obtenu par la minimisation d'une certaine énergie de déplacement sous contrainte de conservation de la masse.

Dans [6], ces mêmes auteurs avaient également introduit un algorithme permettant d'exploiter cette formulation physique du problème de transport optimal, en se basant plus précisément sur une reformulation convexe de celui-ci. Pour cela, ils y avaient fait le choix d'introduire le terme de quantité de mouvement $m = \rho v$, et de ramener le problème à la recherche d'un couple optimal (ρ^*, m^*) plutôt que (ρ^*, v^*) . Autrement dit, de rechercher le champ de vitesse optimal v^* uniquement sur le support de la densité ρ^* (le champ de vitesse v peut en effet varier en dehors du support de ρ sans que soit modifiée la valeur de l'énergie de déplacement \mathcal{K}).

L'un des principaux intérêts d'une telle reformulation, est de rendre convexe le terme d'énergie de déplacement et linéariser la contrainte de conservation de la masse (1.1-32). Le problème devient alors le suivant : il s'agit désormais de minimiser le terme d'énergie de déplacement **convexe**, propre et semi-continue inférieurement.

$$\tilde{\mathcal{K}}(\rho, m) = \int_0^1 \int_{\Omega} \mathcal{J}(\rho, m) dx dt, \quad (1.2-51)$$

où le \mathcal{J} représente le terme d'énergie cinétique ponctuel :

$$\mathcal{J} : (a, b) \in \mathbb{R} \times \mathbb{R}^d \mapsto \begin{cases} \frac{|b|^2}{2a} & \text{si } a > 0, \\ 0 & \text{si } (a, b) = (0, 0), \\ +\infty & \text{sinon,} \end{cases}$$

pour (ρ, m) vérifiant un ensemble de contraintes **affines**, c'est à dire satisfaisant toujours les conditions initiales et finales (1.1-33), ainsi que la nouvelle contrainte de conservation de la masse (linéaire en (ρ, m)) :

$$\partial_t \rho + \operatorname{div}_x m = 0. \quad (1.2-52)$$

Remarquons que la fonction \mathcal{J} est convexe et semi-continue inférieurement sur $\mathbb{R} \times \mathbb{R}^d$, avec $\mathcal{J}(0, 0) = 0$. En invoquant par exemple la Proposition 9.32 de [4], on en déduit que la fonctionnelle $\tilde{\mathcal{K}}$ est également convexe et semi-continue inférieurement (et entre autre propre) sur $L^2((0, 1) \times \Omega)^{d+1}$.

Sous cette dernière reformulation, la recherche du couple optimal (ρ^*, m^*) peut alors se ramener au classique problème d'optimisation convexe sous contraintes :

$$(\rho^*, m^*) = \underset{(\rho, m) \in \mathcal{C}}{\operatorname{argmin}} \tilde{\mathcal{K}}(\rho, m) \quad (1.2-53)$$

où \mathcal{C} représente l'espace de contraintes :

$$\mathcal{C} = \{(\rho, m), \partial_t \rho + \operatorname{div}_x m = 0, \langle m, \vec{n} \rangle = 0, \rho(0, \cdot) = \rho_0, \rho(1, \cdot) = \rho_1\}, \quad (1.2-54)$$

(en choisissant de prendre des conditions de Neumann homogènes aux bords) lequel est un espace convexe et fermé dans $L^2((0, 1) \times \Omega)$.

Dans [6], Benamou et Brenier ont proposé de reformuler ce problème d'optimisation sous contraintes en un problème de recherche de point selle d'un certain lagrangien \mathbf{L}^{ps} . Pour cela, ils décident de gérer l'espace de contrainte \mathcal{C} par un multiplicateur de Lagrange ψ . Les auteurs de [6] proposent alors d'introduire un algorithme fondé sur la méthode d'Uzawa [76], plus précisément sur la méthode des lagrangiens augmentés telle que celle-ci fut introduite par M. Fortin et R. Glowinski dans [37]. Cependant, pour simplifier l'algorithme, le multiplicateur de Lagrange et la variable primale ont été inversés dans le lagrangien effectivement traité : le terme ψ devient la variable primale, et le terme (ρ, m) devient le multiplicateur de Lagrange. Cet algorithme sera décrit précisément dans le chapitre II (voir la sous-section II.1.3), et l'étude de sa convergence et sa consistance mathématique (résultats d'existence et d'unicité sur les points selles) sera le principal objet de ce manuscrit. C'est pourquoi nous nous contenterons de présenter ici une famille d'algorithmes plus récents (voir très récents) permettant de résoudre le problème (1.2-53).

Ceux-ci se fondent sur la méthode du "splitting proximal", qui peuvent s'avérer plus efficaces que l'algorithme de Benamou-Brenier dans certains cas. Dans les résultats numériques du chapitre V, nous avons d'ailleurs fait appel à l'un de ces algorithmes plutôt qu'à celui de Benamou-Brenier. Cependant, tous ces algorithmes (Benamou-Brenier et les méthodes de splitting proximal appliquées au problème (1.2-53)) sont équivalents du point de vue de leur consistance mathématique : leur convergence peut se prouver par des méthodes de convergence d'itérations d'opérateurs non-expansif (1-lipschitziens) vers un point fixe. De plus, l'existence d'un tel point fixe pour ces algorithmes est équivalente à l'existence d'un point fixe pour l'algorithme de Benamou-Brenier, c'est à dire d'un point selle au lagrangien traité par ce dernier. De même, l'unicité de la composante d'intérêt de celui-ci (i.e. la composante (ρ, m)) est un problème commun pour tous ces algorithmes.

I.2.1 Méthodes de "splitting proximal"

Ces méthodes de splitting proximal se fondent toutes sur l'exploitation de l'**opérateur proximal**, que nous définissons ci-dessous :

Définition 1.2-12. L'OPÉRATEUR PROXIMAL ([4] chapitre 12)

Soit H un espace de Hilbert. Soit f une fonction de H dans \mathbb{R} propre, s.c.i. et convexe, et soit $x \in H$. L'opérateur proximal de f en x , noté $\operatorname{Prox}_f(x)$ est l'unique minimiseur de $f + \frac{1}{2}|x - \cdot|^2$ dans H . Autrement dit :

$$\operatorname{Prox}_f(x) = \underset{y \in H}{\operatorname{argmin}} \left(f(y) + \frac{1}{2}|y - x|^2 \right) \quad (1.2-55)$$

On peut caractériser l'opérateur proximal par la relation suivante :

$$y = \text{Prox}_f(x) \Leftrightarrow x - y \in \partial\phi(y). \quad (1.2-56)$$

Il est à noter que l'opérateur Prox_f (de même que $\text{id} - \text{Prox}_f$) est **fermement non-expansifs**, c'est à dire que

$$\forall x, y \in H, \|\text{Prox}_f(x) - \text{Prox}_f(y)\|^2 \leq \langle x - y, \text{Prox}_f(x) - \text{Prox}_f(y) \rangle \quad (1.2-57)$$

énonçons également une propriété utile, liant l'opérateur proximal de f avec celui de sa transformée de Legendre f^* : **l'identité de Moreau**,

$$\text{Prox}_{\gamma f^*} = \text{id} - \gamma \text{Prox}_{f/\gamma}(\cdot/\gamma) \quad (1.2-58)$$

Remarque 1.2-15. La relation 1.2-56 montre qu'un point a de H minimise la fonction f si et seulement si $a = \text{Prox}_f(a)$. En effet,

$$a \in \text{argmin } f \Leftrightarrow 0 \in \partial f(a) \Leftrightarrow a - a \in \partial f(a).$$

Chercher un minimiseur d'une fonction convexe f revient donc à chercher un point fixe à l'opérateur Prox_f . C'est la raison pour laquelle il est possible de développer à partir de l'opérateur proximal de nombreux algorithmes basés sur la recherche de points fixes d'opérateurs non-expansifs (voir la section II.3, chapitre II), ces derniers disposant déjà d'une littérature assez large [4].

Remarque 1.2-16. Un cas assez fréquent dans les problèmes d'optimisation convexe, est le cas d'une optimisation convexe sous contraintes (comme c'est le cas de notre problème de transport optimal), où l'espace des contraintes est un ensemble convexe fermé C . Cependant ces problèmes peuvent toujours se ramener à un problème d'optimisation convexe classique (sans contrainte), en ajoutant à la fonctionnelle convexe à minimiser d'origine, une fonctionnelle indicatrice de l'ensemble C . Autrement dit, en notant F la fonctionnelle convexe à minimiser sur C , on a

$$\text{argmin}_{x \in C} F(x) = \text{argmin}_x (F(x) + \iota_C(x)),$$

où la fonctionnelle ι_C indicatrice de l'ensemble C s'écrit :

$$\iota_C(x) = \begin{cases} 0, & \text{si } x \in C, \\ +\infty & \text{sinon.} \end{cases} \quad (1.2-59)$$

La fonctionnelle ι_C est propre, semi-continue inférieurement et convexe sur H . Or l'utilisation d'algorithmes de proximal splitting (voir juste en dessous) pour résoudre ce type de problèmes, requiert en général le calcul des opérateurs proximaux associés (séparément) à la fonctionnelle F et à la fonctionnelle ι_C : or, il s'avère qu'un calcul simple montre que l'opérateur proximal Prox_{ι_C} associé à cette dernière est en fait **un opérateur de projection orthogonale** sur C .

Cette égalité entre l'opérateur proximal associé à l'indicatrice d'un ensemble convexe fermé et l'opérateur de projection orthogonale sur ce même convexe peut également se révéler très utile pour le calcul de l'opérateur proximal associé à une fonctionnelle **homogène** : en effet, si une fonction F est homogène (i.e. si $\lambda > 0$, $F(\lambda x) = \lambda F(x)$), alors sa transformée de Legendre F^* est

une indicatrice d'un ensemble convexe fermé (indicatrice au sens de l'équation (1.2-59)). Notons K cet ensemble convexe fermé : $F^* = \iota_K$. Or, d'après l'identité de Moreau (1.2-58), on a alors

$$\text{Prox}_F = \text{id} - \text{Prox}_{F^*} = \text{id} - \text{Prox}_{\iota_K} = \text{id} - \text{Proj}_K.$$

Autrement dit le calcul de l'opérateur proximal Prox_F se ramène à celui de Proj_K la projection orthogonale sur l'ensemble K . Ce cas sera également présent dans notre problème de transport optimal : en effet, le terme d'énergie \tilde{K} définie en (1.2-51) est homogène !

Énonçons quelques algorithmes d'optimisation convexe courant faisant appel à l'opérateur proximal : dans tous ces algorithmes, H désignera un espace de Hilbert.

A. L'algorithme du point proximal

Commençons par le plus naturel de ces algorithmes : soit f une fonction de H dans \mathbb{R} propre, s.c.i. et convexe tel que $\text{argmin } f \neq \emptyset$.

Algorithme 1. POINT PROXIMAL

Soit $x_0 \in H$, et soit $\gamma > 0$. On définit la suite $(x_n)_n$ par la récurrence :

$$\forall n \in \mathbb{N}, \quad x_{n+1} = \text{Prox}_{\gamma f}(x_n).$$

Alors la suite $(x_n)_n$ est une suite minimisante de f , et converge faiblement vers un point $x \in \text{argmin } f$.

Pour la preuve, on pourra appliquer le Théorème 2.3-5 (p. 64). En effet, $\text{argmin } f \neq \emptyset$, c'est à dire $\text{Fix}(\text{Prox}_{\gamma f}) \neq \emptyset$, pour $\text{Fix}(\text{Prox}_{\gamma f})$ désignant l'ensemble des points fixes de $\text{Prox}_{\gamma f}$, pour tout $\gamma > 0$ (le paramètre γ est ici un paramètre d'ajustement permettant d'accélérer l'algorithme). Pour plus de précisions sur l'algorithme du point proximal, voir par exemple [4].

B. L'algorithme Forward-Backward

Cet algorithme traite le cas où l'on doit minimiser la somme d'une fonction f de H dans \mathbb{R} propre, s.c.i. et convexe, et d'une fonction h de H dans \mathbb{R} convexe, différentiable sur H et de gradient $1/\beta$ -lipschitzienne, pour $\beta > 0$, telles que $\text{argmin}(f + h) \neq \emptyset$.

Algorithme 2. FORWARD-BACKWARD

Soit $x_0 \in H$, et soient $\gamma \in]0, 2\beta[$ et $\lambda \in]0, \delta[$ avec $\delta = \min(1, \beta/\gamma) + 1/2$. On définit la suite $(x_n)_n$ par la récurrence :

$$\forall n \in \mathbb{N}, \quad x_{n+1} = (1 - \lambda)x_n + \lambda \text{Prox}_{\gamma f}(x_n - \gamma \nabla h(x_n)).$$

Alors la suite $(x_n)_n$ converge faiblement dans H vers un point $x \in \text{argmin}(f + h)$, et la suite $\nabla h(x_n)$ converge fortement dans H vers le point $\nabla h(x)$.

Pour la preuve, voir par exemple [4].

C. L'algorithme Forward-Backward Généralisé

Cet algorithme est une généralisation de l'algorithme 2 pour le cas où on aurait à minimiser la somme de plusieurs fonctions convexes avec une fonction convexe lisse. Pour simplifier, prenons

simplement le cas de la somme de deux fonctions f et g , propres, s.c.i. et convexes, et d'une fonction h convexe, différentiable sur H et de gradient $1/\beta$ -lipschitzienne ($\beta > 0$), telles que $\operatorname{argmin}(f + g + h) \neq \emptyset$.

Algorithme 3. FORWARD-BACKWARD GÉNÉRALISÉ

Soient $x_0, y_0 \in H$, et soient $\gamma \in]0, 2\beta[$ et $\lambda \in]0, \delta[$ avec $\delta = \min(1, \beta/\gamma) + 1/2$. On définit les suites $(x_n)_n$, $(y_n)_n$, et $(u_n)_n$ par la récurrence :

$$\forall n \in \mathbb{N}, \quad \begin{cases} u_n = \alpha x_n + (1 - \alpha)y_n, \\ x_{n+1} = x_n + \lambda (\operatorname{Prox}_{(\gamma/\alpha)f}[2u_n - x_n - \gamma \nabla h(u_n)] - u_n), \\ y_{n+1} = y_n + \lambda (\operatorname{Prox}_{(\gamma/(1-\alpha))g}[2u_n - y_n - \gamma \nabla h(u_n)] - u_n). \end{cases}$$

Alors la suite $(u_n)_n$ converge faiblement dans H vers un point $u \in \operatorname{argmin}(f + g + h)$.

Pour la preuve, voir par exemple [70].

D. L'algorithme de Douglas-Rachford

Un algorithme un peu plus complexe dans le cas où on aurait à minimiser la somme de deux fonctions f et g , toutes deux propres, s.c.i. et convexes telles que $\operatorname{argmin}(f + g) \neq \emptyset$. Ici, contrairement aux deux algorithmes précédent, aucune des deux fonctions n'est supposée plus régulière.

Algorithme 4. DOUGLAS-RACHFORD

Soient $x_0, y_0 \in H$, et soient $\gamma > 0$ et $\lambda \in]0, 2[$. On définit les suites $(x_n)_n$ et $(y_n)_n$ par la récurrence :

$$\forall n \in \mathbb{N}, \quad \begin{cases} x_{n+1} = x_n + \lambda (\operatorname{Prox}_{\gamma f}(2y_n - x_n) - y_n), \\ y_{n+1} = \operatorname{Prox}_{\gamma g}(x_{n+1}). \end{cases}$$

Alors les suites $(x_n)_n$ et $(y_n)_n$ convergent faiblement dans H , respectivement vers des points x et y vérifiant $y = \operatorname{Prox}_{\gamma g}(x) \in \operatorname{argmin}(f + g)$.

La preuve est similaire à celle de l'algorithme 1. Pour plus de précisions sur l'algorithme de Douglas-Rachford, voir par exemple [50, 28]

E. L'algorithme Primal-Dual

Soient ici H et H' deux espaces de Hilbert. Cet algorithme permet de résoudre un problème de minimisation d'une fonctionnelle du type $f \circ \varphi + g$, où $f : H' \rightarrow \mathbb{R}$ et $g : H \rightarrow \mathbb{R}$ sont deux fonctions propres, s.c.i. et convexes, et φ un isomorphisme de H dans H' , telles que $\operatorname{argmin}(f \circ \varphi + g) \neq \emptyset$. L'idée de cet algorithme est de traiter la forme primale -duale de ce problème sous la forme d'un problème de recherche de point selle, c'est à dire de déterminer

$$\operatorname{argmin}_{x \in H} [f(\varphi(x)) + g(x)] = \operatorname{argmin}_{x \in H} \max_{y \in H'} [\langle \varphi(x), y \rangle_{H'} + g(x) - f^*(y)],$$

en rappelant que la transformée de Legendre de f s'écrit :

$$f^*(y') = \max_z [\langle z, y' \rangle_{H'} - f(z)] \Leftrightarrow f(z') = \max_y [\langle z', y \rangle_{H'} - f^*(y)].$$

Algorithme 5. PRIMAL-DUAL

Soient $x_0 = u_0 \in H$, $y_0 \in H$, et soient $\gamma, \sigma > 0$, $\lambda \in]0, 1[$. On définit les suites $(x_n)_n$, $(y_n)_n$, et $(u_n)_n$ par la récurrence :

$$\forall n \in \mathbb{N}, \quad \begin{cases} y_{n+1} = \text{Prox}_{\gamma f^*}(y_n + \gamma \varphi(u_n)), \\ x_{n+1} = \text{Prox}_{\sigma g}(x_n - \sigma \varphi^*(y_{n+1})), \\ u_n = x_{n+1} + \lambda(x_{n+1} - x_n), \end{cases}$$

où $\varphi^* : H' \rightarrow H$ désigne l'adjoint de l'isomorphisme $\varphi : H \rightarrow H'$.

Alors les suites $(u_n)_n$ et $(x_n)_n$ convergent fortement dans H vers un point $x \in \text{argmin}(f \circ \varphi + g)$; la suite $(y_n)_n$ converge, quant à elle, vers faiblement dans H' vers un point $y \in H$ vérifiant $f(\varphi(x)) = \langle \varphi(x), y \rangle - f^*(y)$.

Pour la preuve, voir par exemple [24] ou [42].

Remarque 1.2-17. Le problème ci-dessus peut également être traité par l'algorithme 3 de Forward-Backward Généralisé, en le reformulant de la façon suivante :

$$\begin{aligned} \underset{x}{\text{argmin}} [f(\varphi(x)) + g(x)] &= \underset{(x,y) \in \text{Graph}(\varphi)}{\text{argmin}} [f(y) + g(x)] = \underset{x,y}{\text{argmin}} [f(y) + g(x) + \iota_{\text{Graph}(\varphi)}(x, y)] \\ &= \underset{x,y}{\text{argmin}} [F(x, y) + G(x, y)] \end{aligned}$$

avec $F(x, y) = f(y) + g(x)$ et $G(x, y) = \iota_{\text{Graph}(\varphi)}(x, y)$, où l'ensemble $\text{Graph}(\varphi)$ est le graphe de l'opérateur φ , c'est à dire que $\text{Graph}(\varphi) = \{(x, y) \in H \times H, y = \varphi(x)\}$: ensemble convexe et fermé si φ est linéaire continu.

Les fonctionnelles F et G sont propres, convexes et semi-continues inférieurement : il est donc possible d'appliquer l'algorithme 3, en prenant ici une fonction nulle pour le terme lisse (i.e. pour la fonction différentiable de gradient lipschitzien), et en remarquant que

- On a $\text{Prox}_F(x, y) = (\text{Prox}_f(x), \text{Prox}_g(y))$: calculer l'opérateur Prox_F revient donc à calculer indépendamment les opérateurs Prox_f et Prox_g .
- D'après la Remarque 1.2-16, on a $\text{Prox}_G = \text{Prox}_{\iota_{\text{Graph}(\varphi)}} = \text{Proj}_{\text{Graph}(\varphi)}$. Calculer l'opérateur Prox_G revient donc à calculer la projection orthogonale dans $H \times H$ sur l'ensemble convexe fermé $\text{Graph}(\varphi)$.

Ceci sera la méthode que nous utiliserons dans le chapitre V lorsque nous aurons à traiter ce type problème.

Revenons maintenant au problème (1.2-53) de transport optimal dynamique.

I.2.2 Application au problème de transport optimal dynamique

Nous présentons ici des exemples d'application de ces algorithmes proximaux au problème de transport optimal dynamique, selon deux approches : la première, développée dans [67], sort l'espace des contraintes via l'introduction d'une fonction identité (voir la Remarque 1.2-16) ; la seconde, développée par M. Henry dans [43], propose au contraire une reformulation du problème en changeant les variables de façon à obtenir une fonctionnelle convexe et libre, dont l'optimisation revient à résoudre le problème (1.2-53) (p. (1.2-53)) sans sortir de l'espace des contraintes \mathcal{C} .

A. Résolution du problème (1.2-53) en sortant de l'espace des contraintes

Comme déjà évoqué dans la Remarque 1.2-16, le problème (1.2-53) d'optimisation sous contrainte peut se ramener à un problème d'optimisation convexe sans contrainte, en opérant :

$$\operatorname{argmin}_{(\rho, m) \in \mathcal{C}} \tilde{\mathcal{K}}(\rho, m) = \operatorname{argmin}_{\rho, m} \left(\tilde{\mathcal{K}}(\rho, m) + \iota_{\mathcal{C}}(\rho, m) \right), \quad (1.2-60)$$

où, rappelons-le, le terme d'énergie $\tilde{\mathcal{K}}$ est propre, s.c.i. et convexe, et où l'espace des contraintes

$$\mathcal{C} = \{(\rho, m), \partial_t \rho + \operatorname{div}_x m = 0, \langle m, \vec{n} \rangle = 0, \rho(0, \cdot) = \rho_0, \rho(1, \cdot) = \rho_1\}$$

est un convexe fermé dans $L^2((0, 1) \times \Omega)^{d+1}$, ce qui implique que son indicatrice $\iota_{\mathcal{C}}$ est également propre, s.c.i. et convexe.

Le problème ainsi formulé entre donc totalement dans le cadre des algorithmes de Douglas-Rachford (algorithme 4) et de Forward-Backward Généralisé (algorithme 3), en supposant dans ce dernier cas que la fonction lisse est nulle. D'autant plus que, comme déjà évoqué dans la Remarque 1.2-16, l'opérateur \mathcal{K} étant homogène, le calcul de l'opérateur proximal lui étant associé se ramène au calcul d'un opérateur de projection orthogonale sur un ensemble convexe et fermé de $L^2((0, 1) \times \Omega)^{d+1}$. En effet, du fait de l'identité de Moreau, on a

$$\operatorname{Prox}_{\tilde{\mathcal{K}}} = \operatorname{id} - \operatorname{Prox}_{\tilde{\mathcal{K}}^*} = \operatorname{id} - \operatorname{Proj}_{\tilde{\mathcal{P}}},$$

où $\tilde{\mathcal{P}} = \{(\tilde{a}, \tilde{b}) \in L^2((0, 1) \times \Omega) \times L^2((0, 1) \times \Omega)^d, \tilde{a} + |\tilde{b}|^2/2 \leq 0\}$. Or, il s'avère que la projection orthogonale d'une fonction de $L^2((0, 1) \times \Omega)^{d+1}$ sur l'ensemble $\tilde{\mathcal{P}}$ revient à projeter point par point (presque partout) cette fonction sur le parabolôïde

$$\mathcal{P} = \left\{ (a, b) \in \mathbb{R} \times \mathbb{R}^d, a + \frac{|b|^2}{2} \leq 0 \right\},$$

ce qui correspond pour chaque point projeté à déterminer les racines d'un polynôme de degré 3, qui peuvent être explicitées en utilisant par exemple une méthode de Cardan. Cet étape se retrouve également dans l'algorithme de Benamou-Brenier, que nous traiterons dans le chapitre II.

La difficulté de ce problème vient en fait de la projection orthogonale dans $L^2((0, 1) \times \Omega)^{d+1}$ sur l'espace des contraintes \mathcal{C} , qui n'est pas du tout explicite (nous retrouvons également cette étape dans l'algorithme de Benamou-Brenier). Cette projection se ramène à l'inversion d'un laplacien espace-temps, qui peut être résolue par une méthode de Poisson (voir [6, 67]).

Ces méthodes impliquent une discrétisation très fine du problème, et peuvent être source d'une forte instabilité lors de leur traitement numérique. Il a été proposé dans [67] (et [45] pour une formulation différente du problème : voir plus bas) de traiter numériquement cette projection sur \mathcal{C} par une grille décalée par rapport à celle sur laquelle est effectuée la projection sur le parabolôïde \mathcal{P} , ceci dans le but d'affiner le calcul des différentes dérivées partielles nécessaires à la résolution pratique du problème, et ainsi de rapprocher le résultat numérique de cette projection sur \mathcal{C} par rapport à la projection théorique. Cette méthode, qui s'est avérée très efficace, fait appelle à un opérateur fixe de recentrage de grille L (opérateur linéaire inversible) dont l'adjoint L^* peut être calculé à l'avance. C'est pourquoi, dans ce cadre, le problème d'optimisation discrétisé devient un problème du type :

$$\operatorname{argmin}_{\rho, m} \left(\tilde{\mathcal{K}}(L(\rho, m)) + \iota_{\mathcal{C}}(\rho, m) \right), \quad (1.2-61)$$

Ce problème entre alors dans le cadre de l'algorithme 5 Primal-Dual ou bien dans celui de l'algorithme 3 de Forward-Backward Généralisé adapté à ce type de problème (voir la Remarque 1.2-17). Cette dernière méthode sera appliquée dans le chapitre V.

Pour d'avantage de détails sur cette méthode impliquant des grilles décalées voir [67] et [43].

B. Résolution du problème (1.2-53) sans sortir de l'espace des contraintes

Dans [43], M. Henry propose plusieurs méthodes permettant de traiter le problème (1.2-53) sans sortir de l'espace des contraintes \mathcal{C} , c'est à dire en ne se ramenant pas à un problème d'optimisation sans contrainte (comme le problème (1.2-60) de la sous-section précédente) :

- Une première méthode, efficace dans le cas d'un transport optimal en une dimension, propose d'écrire les éléments de \mathcal{C} comme le rotationnel d'une fonction courant ϕ (i.e. $(\rho, m) = \text{rot}_{t,x}(\phi)$), et de ramener le problème (1.2-53) à un problème de courbure minimale sur cette même fonction courant, c'est à dire à résoudre l'équation :

$$\text{div}_{t,x} \left(\frac{\nabla_{t,x}\phi}{|\nabla_{t,x}\phi|} \right) = 0,$$

(voir [43, 45]).

- Dans l'idée de la décomposition précédente, est considérée également la décomposition de Helmotz-Hodge des éléments de \mathcal{C} : il s'agit de décomposer les éléments de $(\rho, m) \in \mathcal{C}$ (pour un transport dans un espace de dimension 1 ou 2) sous la forme

$$(\rho, m) = \text{rot}_{t,x}(\phi) + \nabla_{t,x}h,$$

où le terme $h \in H^1((0, 1) \times \Omega)$ est le même pour tous les éléments de \mathcal{C} et contient à lui seul les conditions de bords des éléments de cet ensemble : ce terme est donc calculé une fois pour toute. Les fonctions de courant sont quant à elles des éléments de $H_0^1((0, 1) \times \Omega)^{d+1}$ à divergence nulle. Le problème se reformule alors

$$\underset{\phi \in C_0}{\text{argmin}} \tilde{\mathcal{K}}(\text{rot}_{t,x}(\phi) + \nabla_{t,x}h) = \underset{\phi \in H_0^1}{\text{argmin}} \left(\tilde{\mathcal{K}}(\text{rot}_{t,x}(\phi) + \nabla_{t,x}h) + \iota_{C_0}(\phi) \right)$$

où $C_0 = \{\phi \in H_0^1((0, 1) \times \Omega)^{d+1}, \text{div}_{t,x}(\phi) = 0\}$ (espace convexe fermé de $H_0^1((0, 1) \times \Omega)$).

On peut poser alors

$$F(\psi) = \tilde{\mathcal{K}}(\psi + \nabla_{t,x}h), \text{ et } G(\phi) = \iota_{C_0}(\phi)$$

(fonctions propres, s.c.i. convexes) et

$$L : \phi \in (H_0^1)^{d+1} \rightarrow \text{rot}_{t,x}(\phi) \in V,$$

où V est le sous-espace de $L^2((0, 1) \times \Omega)^{d+1}$ des fonctions à divergence nulle (espace fermé de L^2 de Hilbert également) : l'opérateur L est alors un isomorphisme. On se ramène alors à un problème

$$\underset{\phi \in H_0^1}{\text{argmin}} (F(\phi) + G(\phi)).$$

Ce problème est alors traité dans [43, 44, 45] par l'algorithme Primal-Dual (algorithme 5).

Ces deux premières méthodes, s'avèrent très efficaces en pratique et améliorent significativement la vitesse des algorithmes.

- Une autre méthode originale proposée dans [43] consiste à décomposer l'espace des contraintes \mathcal{C} dans une base d'ondelettes adaptée de façon à pouvoir effectuer une descente de gradient au sein même de cet espace. Cette méthode, du fait du petit nombre d'éléments de la base utilisée en pratique, permettrait de minimiser la complexité des algorithmes.

I.3 Transport Optimal, Applications et Variantes

Nous présentons dans cette section, de façon non exhaustive, quelques applications de la théorie du transport optimal. Tout d'abord nous traiterons de l'une des applications les plus communes du modèle dynamique de Benamou et Brenier [6], à savoir l'application au traitement d'image. Nous introduirons ensuite une théorie très au goût du jour dans la communauté du transport optimal, à savoir les modèles de flots de gradient dans un espace métrique de Wasserstein. Enfin, nous évoquerons le problème du transport branché, variante du problème de Monge-Kantorovitch dans laquelle est introduite une contrainte de regroupement.

Le transport optimal a bien sûr de nombreuses autres applications, par exemple en économie (après tout, tel était le sujet de L. Kantorovitch lorsqu'il avait formulé son problème d'allocation optimal de ressources en 1942), ou bien en théorie des probabilités et statistiques (la distance de Wasserstein \mathcal{W}_1 représentant une distance induisant une topologie équivalente à celle de la convergence en loi), etc...

I.3.1 Applications au traitement d'images

L'application du transport optimal au traitement d'image constitue une des utilisations les plus directes de la formulation dynamique de [6]. En effet, la résolution de ce problème ne donne qu'indirectement le plan de transport optimal entre deux densités ρ_0 et ρ_1 (il faudrait pour cela prendre la trace du champ de vitesse calculé en $t = 0$). La solution directe de ce problème se révèle en effet être les densités d'interpolation de McCann par ce transport optimal (voir (1.1-23) p. 23), ainsi que la quantité de mouvement associée à cette dynamique (permettant donc de calculer le champ de vitesse associé sur le support de ρ).

Cette densité d'interpolation peut cependant être directement utilisée, justement pour la réalisation d'interpolations d'images : les différentes méthodes de résolution numérique du problème dynamique (voir la section I.2) ont déjà été souvent exploitées par la communauté du traitement d'images [69, 66] pour réaliser des recalages [41], des estimations de flots optiques [27], du défloutage d'images en microscopie [18], etc...

En outre, la capacité des algorithmes de traitement numérique [6, 67, 44, 45] à estimer rapidement la distance de Wasserstein associée au transport optimal entre deux images peut s'avérer être d'une grande utilité en ce qui concerne la comparaison et la validation de modèles dynamiques des fluides.

En effet contrairement aux distances usuelles en imagerie comme les distances L^p , lesquelles estiment une moyenne pondérée (dans l'espace des couleurs) entre deux pixels de même coordonnées, la distance de Wasserstein mesure quant à elle un déplacement global. Ainsi, là où une faible translation entre deux densités peut induire une grande distance L^p , la distance de Wasserstein associée restera quant à elle relativement faible, et proportionnelle à l'importance du déplacement. Citons par exemple les recherches concernant son application en météorologie ou en océanographie, pour effectuer des ajustements de modèles par assimilation données en temps réel [35, 36].

Un autre grand intérêt de la formulation dynamique du problème de transport optimal est son

expression sous forme d'équations et d'énergies tirées de la mécanique des fluides : cette formulation rend le modèle très souple, et permet une généralisation facile à des contraintes physiques [52, 26], selon les données du problème. L'introduction de ces contraintes physiques (anisotropie du domaine, contraintes de divergence libre ou de rigidité sur le champ de vitesse) dans le problème dynamique fut le sujet d'une de nos recherches [46], et fera l'objet du chapitre V de ce manuscrit. Nous renvoyons donc à ce chapitre (lequel peut être lu indépendamment de tous les autres) pour plus de précisions.

I.3.2 Flots de gradient dans les espaces de mesures de probabilité

Nous présentons ici un sujet actuellement très en vogue dans la communauté du transport optimal : l'étude des flots de gradient dans l'espace $\mathcal{P}_2(\mathbb{R}^d)$.

Classiquement (dans \mathbb{R}^d), on qualifie de flot de gradient une équation différentielle du type $x'(t) = -\nabla F(x(t))$, où F est une fonction différentiable (ou éventuellement $x'(t) \in -\partial F(x(t))$ si F n'est pas différentiable en $x(t)$).

L'idée des flots de gradient dans les espaces de mesures de probabilité est d'assimiler le formalisme des flots de gradient usuels (c'est à dire évoluant dans un espace vectoriel normé) à ces espaces de mesures, munis de la métrique de Wasserstein. Plus précisément, il s'agit d'étudier des équations différentielles (sur un intervalle temporel I) du type :

$$\frac{d\mu_t}{dt} \in -\partial_{\mathcal{W}_2} \mathcal{F}(\mu_t), \quad (1.3-62)$$

où pour tout $t \in I$ la mesure $\mu_t \in \mathcal{P}_2(\mathbb{R}^d)$ (l'espace de mesures de probabilité défini à la notation 1 p. 19 sur lequel est défini la distance \mathcal{W}_2), où la dérivée temporelle de $t \mapsto \mu_t$ est ici une dérivée classique au sens des distributions, i.e.

$$\forall \varphi \in C_c^\infty(I, \mathbb{R}^d), \quad \left\langle \frac{d\mu_t}{dt}, \varphi \right\rangle = - \int_I \int_{\mathbb{R}^d} \partial_t \varphi(t, x) d\mu_t(x) dt,$$

et où \mathcal{F} est une fonction **géodésiquement convexe** sur $\mathcal{P}_2(\mathbb{R}^d)$ (voir la sous-section 1.1.6). Le *sous-gradient de Monge-Kantorovitch* $\partial_{\mathcal{W}_2} \mathcal{F}$ est défini par rapport à la métrique de Wasserstein, pour tout $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ par

$$\partial_{\mathcal{W}_2} \mathcal{F}(\mu) \in \{-\operatorname{div}_x(v\mu), \quad v \in \Lambda(\mu)\}, \quad (1.3-63)$$

où $\Lambda(\mu)$ est défini comme l'ensemble des $v \in L^2(\mu)$ vérifiant pour tout $\nu \in \mathcal{P}_2(\mathbb{R}^d)$:

$$\mathcal{F}(\nu) \geq \mathcal{F}(\mu) + \int_{\mathbb{R}^d} v(x) \cdot (T_\mu^\nu(x) - x) d\mu(x) + o(\mathcal{W}_2(\mu, \nu))$$

avec T_μ^ν le plan de transport optimal de μ sur ν .

Lorsque l'on traite de chemins de mesures $t \mapsto \mu_t = \rho_t \mathcal{L}^d$ absolument continues par rapport à la mesure de Lebesgue, et pour lesquels peuvent être associés une "dérivée en densité" $(\delta \mathcal{F} / \delta \rho)(\rho)$, définie comme l'unique fonction réelle telle que pour toutes densités $\rho, \xi \in \mathcal{P}_2(\mathbb{R}^d)$, on ait

$$\lim_{\epsilon \rightarrow 0^+} \frac{1}{\epsilon} \mathcal{F}(\rho + \epsilon \xi) = \int_{\mathbb{R}^d} \frac{\delta \mathcal{F}}{\delta \rho}(\rho) \xi d\mathcal{L}^d, \quad (1.3-64)$$

alors

$$\partial_{\mathcal{W}_2} \mathcal{F}(\rho \mathcal{L}^d) = \left\{ -\operatorname{div}_x \left(\rho \nabla_x \left(\frac{\delta \mathcal{F}}{\delta \rho}(\rho) \right) \right) \right\}. \quad (1.3-65)$$

Le sous-gradient de Monge-Kantorovitch $\partial_{\mathcal{W}_2}\mathcal{F}(\rho\mathcal{L}^d)$ peut alors être assimilé à un *gradient de Monge-Kantorovitch* $\nabla_{\mathcal{W}_2}\mathcal{F}(\rho)$, et le problème de flot de gradient dans $\mathcal{P}_2(\mathbb{R}^d)$ devient, une fois restreint aux mesures à densités,

$$\frac{d\rho}{dt} = -\nabla_{\mathcal{W}_2}\mathcal{F}(\rho) = -\operatorname{div}_x \left(\rho \nabla_x \left(\frac{\delta\mathcal{F}}{\delta\rho}(\rho) \right) \right). \quad (1.3-66)$$

Cette formulation sera en particulier valable pour des énergie \mathcal{F} formées d'une combinaison linéaires d'énergies de type interne, potentielle ou d'interaction, telles qu'introduites à la Définition 1.1-11 (p. 32) et rappelées plus bas.

Une telle Définition du gradient et du sous-gradient de Monge-Kantorovitch peut ne pas sembler très naturelle à première vue. Voici une analogie pouvant rendre cette Définition plus intuitive (voir aussi [68]).

La résolution d'une équation différentielle de flot de gradient "classique" (dans \mathbb{R}^d) à l'aide d'un schéma d'Euler implicite, pour un pas de temps τ revient à l'application itérative d'un opérateur proximal. En effet, considérons l'équation différentielle

$$\begin{cases} x'(t) \in -\partial F(x(t)), \\ x(0) = x_0. \end{cases} \quad (1.3-67)$$

Le schéma d'Euler implicite consiste, pour un pas temps $\tau > 0$, à construire une suite récursive $(x_k^\tau)_k$ (pour $x_0^\tau = x_0$), pour laquelle le passage du temps $k\tau$ au temps $k\tau + \tau$, i.e. le calcul de x_{k+1}^τ à partir de x_k^τ revient à résoudre

$$\frac{x_{k+1}^\tau - x_k^\tau}{\tau} \in -\partial F(x_{k+1}^\tau) \iff x_{k+1}^\tau = \operatorname{Prox}_{\tau F}(x_k^\tau), \quad (1.3-68)$$

d'après la propriété (1.2-56) (p. 36) de l'opérateur proximal. La suite $(x_k^\tau)_k$ converge alors vers une solution de (1.3-67) lorsque le pas temps τ tend vers 0.

Une idée naturelle serait donc d'appliquer un tel schéma pour représenter un problème de flot de gradient dans l'espace des mesures de probabilités. L'équivalent $\operatorname{Prox}_{\mathcal{F}}^{\mathcal{W}_2}$ de l'opérateur proximal associé à une fonctionnelle \mathcal{F} , géodésiquement convexe sur $(\mathcal{P}_2(\mathbb{R}^d), \mathcal{W}_2)$, se conçoit naturellement sous la forme suivante : pour toute mesure $\mu \in \mathcal{P}_2(\mathbb{R}^d)$,

$$\operatorname{Prox}_{\mathcal{F}}^{\mathcal{W}_2}(\mu) = \operatorname{argmin}_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} \left(\mathcal{F}(\nu) + \frac{\mathcal{W}_2(\mu, \nu)}{2} \right). \quad (1.3-69)$$

Pour résoudre un problème de flot de gradient du type

$$\begin{cases} \frac{d\rho}{dt} = -\partial_{\mathcal{W}_2}\mathcal{F}(\rho), \\ \rho(0, \cdot) = \rho_0, \end{cases} \quad (1.3-70)$$

le schéma discrétisé en temps (pour un pas τ) reviendrait donc, par analogie, à construire une suite récursive $(\rho_k^\tau)_k$ (pour $\rho_0^\tau = \rho_0$), pour laquelle le calcul de ρ_{k+1}^τ à partir de ρ_k^τ revient à résoudre

$$\frac{\rho_{k+1}^\tau - \rho_k^\tau}{\tau} = -\nabla_{\mathcal{W}_2}\mathcal{F}(\rho_{k+1}^\tau) \iff \rho_{k+1}^\tau = \operatorname{Prox}_{\tau\mathcal{F}}^{\mathcal{W}_2}(\rho_k^\tau) \quad (1.3-71)$$

Il est montré dans [1], que sous de bonnes hypothèses sur ρ_0 et \mathcal{F} , la suite $(\rho_k^\tau)_k$ converge bien, lorsque τ tend vers 0, vers une solution de (1.3-70), pour $\partial_{\mathcal{W}_2}\mathcal{F}$ définie comme en (1.3-63) : les gradients et sous-gradients de Monge-Kantorovitch peuvent donc être ainsi définis par analogie.

Dans la sous-section I.1.6 (Définition 1.1-11 p. 32), nous avons énoncé trois catégories de fonctions géodésiquement convexes sur $(\mathcal{P}_2(\mathbb{R}^d), \mathcal{W}_2)$, assimilables à des énergies physiques : les énergies *internes*, les énergies *potentielles* et les énergies *d'interaction*. En effet, de nombreuses équations issues de la physique, modélisant l'évolution de densités dans l'espace et le temps (nous entendons le terme de "densité" au sens générique, c'est à dire un champ donnant une valeur à chaque point de l'espace et du temps), peuvent être décrites comme dérivant d'une énergie ou d'une entropie.

Ces trois énergies s'écrivent, rappelons-le, pour toute densité ρ , sous les formes :

- Énergie interne : $\mathcal{U}(\rho\mathcal{L}^d) = \int_{\mathbb{R}^d} U(\rho) d\mathcal{L}^d$,
- Énergie potentielle : $\mathcal{V}(\rho\mathcal{L}^d) = \int_{\mathbb{R}^d} V \rho d\mathcal{L}^d$,
- Énergie d'interaction : $\mathcal{H}(\rho\mathcal{L}^d) = \int_{\mathbb{R}^d} W(x-y) \rho(x)\rho(y) d\mathcal{L}^d(x) d\mathcal{L}^d(y)$,

pour U mesurable sur \mathbb{R}_+ , V et W mesurables sur \mathbb{R}^d . Le gradient de Monge-Kantorovitch de ces trois énergies s'écrit donc, au sens de la relation (1.3-65) via des dérivées en densité sous la forme :

$$\frac{\delta\mathcal{U}}{\delta\rho}(\rho) = U'(\rho), \quad \frac{\delta\mathcal{V}}{\delta\rho}(\rho) = V, \quad \frac{\delta\mathcal{H}}{\delta\rho}(\rho) = \rho * W$$

Voici quelques exemples de ces équations tirées de la physique, pouvant se réécrire sous la forme de flots de gradients dans $\mathcal{P}_2(\mathbb{R}^d)$:

- *L'équation de la chaleur* :

$$\frac{\partial\rho}{\partial t} = \Delta_x \rho \quad = \operatorname{div}_x (\rho \nabla_x (\log(\rho) + 1)) = -\nabla_{\mathcal{W}_2} \mathcal{E}(\rho\mathcal{L}^d),$$

où l'énergie $\mathcal{E}(\rho\mathcal{L}^d) = \int_{\mathbb{R}^d} \rho \log(\rho) dx$, en entrant dans la catégorie des *énergies internes* au sens de la Définition 1.1-11, est géodésiquement convexe sur $\mathcal{P}_2(\mathbb{R}^d)$: en effet, la fonction $x \mapsto x \log(x)$ est convexe, non-décroissante sur \mathbb{R}_{++} (il en est de même pour la fonction $x \mapsto -d \times x^d \log(x) = x^d \log(x^{-d})$, croissante au voisinage de $x = 0$), et se prolonge par 0 en $x = 0$.

En effet, pour ce qui est du calcul de $\mathcal{E}(\rho\mathcal{L}^d)$ on a bien, d'après la relation (1.3-64),

$$U'(x) = \log(x) + 1 \quad \text{et donc} \quad \frac{\delta\mathcal{U}}{\delta\rho}(\rho) = U'(\rho) = \log(\rho) + 1.$$

Cette énergie est assimilable à l'entropie de Boltzmann [48].

- Une variante : *l'équation de Fokker-Planck* :

$$\begin{aligned} \frac{\partial\rho}{\partial t} &= \Delta_x \rho + \operatorname{div}_x (\rho \nabla_x V) \quad = \operatorname{div}_x (\rho \nabla_x (\log(\rho) + 1 + V)), \text{ avec } V \text{ convexe,} \\ &= -\nabla_{\mathcal{W}_2} (\mathcal{E} + \mathcal{V})(\rho\mathcal{L}^d), \end{aligned}$$

où le terme d'énergie interne $\mathcal{E}(\rho\mathcal{L}^d)$ est défini dans le point précédent, et où le terme $\mathcal{V}(\rho\mathcal{L}^d) = \int_{\mathbb{R}^d} V(x) \rho dx$ correspond à une *énergie potentielle* au sens de la Définition 1.1-11, et est donc également géodésiquement convexe sur $\mathcal{P}_2(\mathbb{R}^d)$.

- Un cas plus évident pour cette formulation, *les équations de McKean-Vlasov* :

$$\begin{aligned} \frac{\partial\rho}{\partial t} &= \operatorname{div}_x (\rho \nabla_x (\rho * W)), \text{ avec } W \text{ convexe et paire,} \\ &= -\nabla_{\mathcal{W}_2} \mathcal{H}(\rho\mathcal{L}^d), \end{aligned}$$

où l'énergie $\mathcal{H}(\rho\mathcal{L}^d) = \frac{1}{2} \int_{\mathbb{R}^d} W(x-y) \rho(x)\rho(y) dx dy$ est la Définition même d'une *énergie d'interaction* au sens de la Définition 1.1-11, et est donc géodésiquement convexe sur $\mathcal{P}_2(\mathbb{R}^d)$.

L'un des objectifs de la théorie des flots de gradient serait donc de pouvoir utiliser cette formulation pour modéliser et réunir un maximum systèmes dynamiques [64, 48], relier de façon général les équations aux énergies et entropies qu'elles induisent (telle que par exemple l'équation de la chaleur avec l'entropie de Boltzmann), et de pouvoir en tirer un maximum de propriétés mathématiques communes. La capacité de cette formulation à révéler des phénomènes d'entropie a par exemple été exploitée pour modéliser des mouvement de foule [54].

Enfin, la recherche d'algorithmes numériques efficaces, basés sur le schéma d'Euler implicite exposé plus haut, et donc sur un calcul efficace de l'opérateur $\text{Prox}_{\mathcal{F}}^{\mathcal{W}_2}$ est également sérieusement entreprise [8, 7, 21] : en effet, la grande variété de problèmes dynamiques pouvant être modélisés par un problème de flot de gradient dans $\mathcal{P}_2(\mathbb{R}^d)$ ferait de tels algorithmes des outils extrêmement utiles.

I.3.3 Le transport branché

Une variante pertinente du problème de Monge-Kantorovitch est celle du transport branché : plutôt que d'optimiser un coût de transport d'un point x à un point y , de la forme $c(x, y) \cdot m$, c'est à dire simplement proportionnel à la masse m transportée entre ces deux points, l'idée de cette approche est de pénaliser également la dispersion du transport. En effet, il est souvent juste de considérer qu'un transport groupé est plus efficace ou moins coûteux qu'une somme de transports individuels. C'est le cas par exemple du transport routier : pour un même itinéraire, le transport de dix personnes par dix voiture est en général beaucoup plus coûteux que le transport de ces dix personnes par un minibus. Si l'économie de ce regroupement est suffisamment élevée, il est souvent intéressant, pour plusieurs personnes ayant des itinéraires différents, de mettre en commun une partie de leurs trajets, quitte à devoir pour cela allonger la longueur de chaque trajet individuel.

Un tel modèle de transport, incitant au regroupement des trajectoires, favorise donc, comme on peut se l'imaginer, un transport sous forme d'un réseau (d'où le nom de transport branché) : des axes de fortes circulations se forment, et se subdivisent ensuite selon les différents départs et destinations. Un tel modèle pourrait donc, par exemple, modéliser la construction d'un réseau routier, mais également permettre de modéliser un réseau d'écoulements (ruisseaux, rivières, fleuves, une forte masse d'eau ayant une plus grande puissance d'écoulement lorsqu'elle est réunie que lorsqu'elle est diffuse), développement des branches d'un arbre, etc...

Un modèle simple permettant de rendre compte d'une telle contrainte est de remplacer le coût de transport de x à y proportionnel à une masse m (du type $c(x, y) \cdot m$), par un coût sous-additif, plus précisément par un coût proportionnel à m^α (i.e. du type $c(x, y) \cdot m^\alpha$), avec $0 < \alpha < 1$. En effet, supposons deux masses (non-nulles) m_1 et m_2 , on a alors $(m_1 + m_2)^\alpha < m_1^\alpha + m_2^\alpha$: le coût d'un transport commun de ces deux masses est donc inférieur à la somme de leur coût de transport individuel. Cette différence de coût est d'autant plus importante que l'exposant α est petit : le choix de α permet donc de plus ou moins doser cette contrainte de regroupement.

Il existe plusieurs méthodes et algorithmes permettant de résoudre numériquement ce type de transport, que ce soit dans le cas discret [79], ou bien le cas continu [78] : citons par exemple [65], basée sur l'alternance d'une régularisation par Γ -convergence de la fonctionnelle de coût globale du transport et d'une descente de gradient, dont la preuve de consistance en dimensions quelconque avec contrainte d'assignation à été démontrée par A. Monteil dans [59] et [60].

I.4 Présentation synthétique des résultats obtenus

L'algorithme de Benamou-Brenier a pour objet la résolution numérique de la formulation "mécanique des fluides" du problème de transport optimal entre deux densités ρ_0 et ρ_1 . Formellement, le problème de minimisation sous contraintes associé est le suivant. Soit \mathcal{K} le terme d'énergie de déplacement

$$\mathcal{K}(\rho, m) = \int_0^1 \int_{\Omega} \frac{|m|^2}{2\rho} dx dt, \quad (1.4-1)$$

on cherche un couple densité-quantité de mouvement $(\rho^*, m^*) \in L^2((0, 1) \times \Omega)^{1+d}$, tel que

$$(\rho^*, m^*) = \underset{\substack{\partial_t \rho + \operatorname{div}_x m = 0, \rho > 0 \\ \rho(0, \cdot) = \rho_0, \rho(1, \cdot) = \rho_1}}{\operatorname{argmin}} \mathcal{K}(\rho, m). \quad (1.4-2)$$

J.-D. Benamou et Y. Brenier ont établi dans [6] l'équivalence entre ce problème et la recherche d'un point selle dans $H^1/\mathbb{R} \times (L^2)^{1+d} \times (L^2)^{1+d}$ du lagrangien

$$L(\psi, q, \mu) = F(q) + G(\psi) + \langle \mu, \nabla_{t,x} \psi - q \rangle_{L^2((0,1) \times \Omega)}$$

où $\mu = (\rho, m)$ avec les fonctions convexes F (caractérisant la contrainte positivité sur ρ ainsi que la dépendance proportionnelle de m à ρ) et G (caractérisant les conditions initiales et finales sur la densité ρ).

Les auteurs de [6] proposent alors de traiter ce problème par un algorithme de lagrangien augmenté (plus précisément de type ADMM) tel que développé dans [37].

Cet algorithme est utilisé depuis 1999, et fonctionne bien en pratique, même si un certain nombre de questions théoriques restent en suspens :

- L'existence de solutions pour la formulation de Benamou-Brenier au problème (1.4-2), plus précisément l'existence d'un point selle pour L .
- La convergence (faible ou forte) de l'algorithme.
- L'unicité d'une solution.

L'objet principal de cette thèse est de proposer un cadre théorique permettant de répondre à ces trois points. L'étude proposée sera aussi l'occasion de caractériser rigoureusement certaines propriétés du champ de vitesse associé à un plan de transport optimal.

I.4.1 État de l'art (à notre connaissance) sur la convergence de l'algorithme de Benamou-Brenier

Une première étude de l'algorithme de Benamou-Brenier a été effectuée en 2004 par K. Guittet dans [40]. Les hypothèses et le cadre dans lesquels se plaçait cette étude étaient les suivants :

- des conditions aux limites périodiques en espace : en se plaçant dans $\Omega = \mathbb{T}^d$ (où \mathbb{T}^d désigne le tore unité en dimension d , i.e. $\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d$).
- une hypothèse sur la densité ρ^* solution du problème de minimisation (1.4-2) qui est supposée minorée par une constante strictement positive.

En notant $\nabla \phi$ le plan de transport optimal (voir théorème 1.1-7), si l'on considère la densité $\tilde{\rho}_t$ définie par :

$$\tilde{\rho}_t \mathcal{L}^d = [(1-t) \operatorname{id} + t \nabla \phi] \# (\rho_0 \mathcal{L}^d), \quad (1.4-3)$$

le fait que celle-ci soit minorée par une constante strictement positive implique en particulier une certaine régularité du plan de transport associé (sans discontinuité) : le potentiel ϕ doit être de

classe C^1 et de gradient lipschitzien. La régularité d'un plan de transport optimal sur un domaine convexe (c'est à dire plus précisément ici du potentiel ϕ) en fonction de la régularité des densités de départ et d'arrivée ρ_0 et ρ_1 , supposées de plus strictement positives, a notamment été étudiée par Caffarelli dans [19] et [20].

Un cas particulier évoqué dans [40], est celui où l'on considère des densités de départ et d'arrivée ρ_0 et ρ_1 strictement positives sur \mathbb{T}^d et dans $C^{\alpha,l}(\mathbb{T}^d)$, pour un certain $l \in]0, 1[$, et $\alpha \in \mathbb{N}^d$ (voir [17] pour une définition des espaces de Hölder). En effet, ces conditions impliquent, d'après une étude effectuée par Cordero-Erausqui dans [30], que dans ce cas le potentiel de transport optimal ϕ est de classe $C^{\alpha,l+2}$ et que, pour tout $t \in [0, 1]$, la densité $\tilde{\rho}_t$ possède également une régularité $C^{\alpha,l}$ sur \mathbb{T}^d et est minorée par une constante strictement positive indépendante de t .

Sous les hypothèses ci-dessus, il est établi dans [40] les résultats suivants :

- L'existence d'une solution (ρ^*, m^*) pour le problème (1.4-2), sous la première hypothèse. Cependant, il n'y est pas explicitement montré que cette densité ρ^* correspond bien à celle de McCann (i.e. que $\rho^* = \tilde{\rho} : t \mapsto \tilde{\rho}_t$).
- Sous les deux hypothèses, il est montré l'existence d'un point selle (ψ^*, q^*, μ^*) pour le lagrangien L , mais sans résultat d'unicité concernant le couple densité-quantité de mouvement $\mu^* = (\rho^*, m^*)$.
- Un résultat de convergence de l'algorithme de Benamou-Brenier : cependant, celui-ci ne donne pas explicitement la convergence forte ou faible de la principale composante du triplet (ψ_n, q_n, μ_n) . En effet, au regard du problème (1.4-2), la composante d'intérêt est $\mu_n = (\rho_n, m_n)$. Seules les convergences fortes dans H^1/\mathbb{R} et L^2 des composantes ψ_n et q_n y sont énoncées.

De plus, la preuve de cette convergence nous semble incomplète : en effet, celle-ci repose sur le fait que les suites $(\nabla_{t,x}\psi_n)_n$ et $(q_n)_n$ soient des suites de Cauchy de L^2 . Or la justification de cette dernière propriété repose sur les deux arguments suivant :

$$\sum_{n=1}^{+\infty} \|\nabla_{t,x}\psi_n - \nabla_{t,x}\psi_{n-1}\|_{L^2}^2 < +\infty \quad \text{et} \quad \sum_{n=1}^{+\infty} \|q_n - q_{n-1}\|_{L^2}^2 < +\infty,$$

lesquels nous semble incomplets pour conclure au fait que ces deux suites soient de Cauchy.

I.4.2 Notre contribution à l'étude de consistance de l'algorithme de Benamou-Brenier

Dans ce manuscrit, nous nous placerons dans le cadre plus général suivant :

- L'ouvert Ω de \mathbb{R}^d est convexe et borné, muni de conditions au bord de type Neumann homogènes homogène sur la quantité de mouvement m .
- La densité $\tilde{\rho}$ n'est pas supposée minorée par une constante strictement positive ni régulière. Nous supposons ici simplement que les densités ρ_0 et ρ_1 sont des éléments de $L^2(\Omega)$ (donc potentiellement non-régulières) nulles au voisinage de $\partial\Omega$.

Dans ce cadre, nous établirons les résultats suivants :

- L'existence d'un point selle pour lagrangien L , solution du problème (1.4-2) (section IV.1)
- L'unicité parmi l'ensemble des points selles (ψ^*, q^*, μ^*) du lagrangien L de $\mu^* = (\rho^*, m^*)$, dont on montre que la densité correspond au $\tilde{\rho}$ de l'interpolation de McCann (1.4-3).

Si les résultats d'unicité que nous nous proposons d'établir dans ce manuscrit concernent uniquement la composante μ c'est que, comme nous le verrons au début de la section IV.2, en général, **il n'y a pas unicité des points selles de L** : en effet, les composantes ψ et q peuvent varier en dehors du support de ρ .

Ces deux premiers points (résultats d'existence et d'unicité) seront explicitement établis après une étude préalable sur la régularité du champ de vitesse dans le chapitre IV.

- La convergence faible de l'algorithme de Benamou-Brenier, ainsi que la convergence forte d'une version relaxée de celui-ci, vers un point selle du lagrangien L . Ceci sera en fait le premier résultat que nous démontrerons (chapitre II), sous réserve de l'existence d'un tel point selle qui sera donc établi dans les deux chapitres suivants.

Pour finir, présentons les autres contributions de ce manuscrit en ce qui concerne la formulation mécanique des fluides du transport optimal.

I.4.3 Étude du champ vitesse pour un transport optimal dans L^2

Au cours de ce travail, le champ de vitesse d'un transport optimal dans L^2 a été caractérisé dans un cadre que nous avons souhaité le plus général possible, c'est à dire sans hypothèse de régularité sur les densités de départ et d'arrivée ρ_0 et ρ_1 . Plus précisément,

- Nous avons cherché des hypothèses suffisantes sur un champ de vitesse $v \in L_{loc}^\infty([0, 1], \mathbb{R}^d)$ pour qu'une densité transportée par v soit l'interpolation de McCann d'un (unique) transport optimal : celles-ci se réduiront aux caractéristiques mises en évidence dans les sections précédentes : trajectoires en ligne droite, à vitesse constante, et sans croisement. Mathématiquement un tel champ $v \in L_{loc}^\infty([0, 1], \mathbb{R}^d)$, devra vérifier les conditions :

$$\begin{cases} \partial_t v + \frac{1}{2} \nabla_x |v|^2 = 0, \\ \exists \psi \in W_{loc}^{1,\infty}([0, 1] \times \Omega), \quad v = \nabla_x \psi. \end{cases}$$

Ce résultat sera l'objet principal de la dernière section du chapitre IV (section IV.3), et sera précisément formulé dans le Théorème 4.3-9 (p. 159).

- Nous avons de plus établi des propriétés de régularité nécessaires d'un tel champ de vitesse (sur le support de la densité transportée), en particulier :
 - v localement lipschitzienne sur $]0, 1[\times \Omega$.
 - $\nabla_{t,x} v \in L^\infty((0, 1), L_{loc}^1(\Omega))$,
 - $\nabla_{t,x} v \in L^p((0, 1), L_{loc}^q(\Omega))$ avec $1/p + 1/q > 1$
(en particulier $v \in W_{loc}^{1,p}([0, 1] \times \Omega)$ avec $1 \leq p < 2$).

Ces résultats seront traité dans le chapitre III (section III.3). La Proposition 3.4-75 (p. 123) récapitule l'ensemble de ces propriétés de régularité.

I.4.4 Applications du transport optimal aux problèmes multi-physiques

Dans le chapitre V, lequel constitue une version revisitée d'un article co-publié avec Emmanuel Maître et Nicolas Papadakis [46], nous étudierons différentes manières d'introduire des contraintes physiques dans le modèle dynamique de Benamou et Brenier, dans le principal but de concevoir des modèles et des algorithmes numériques (basés sur des algorithmes proximaux) pour réaliser des interpolations entre images (au sens de l'interpolation de McCann) respectant un certain nombre de données contextuelles : anisotropie du domaine (obstacles et polarisations), contraintes

sur le champ de vitesse (rigidité, divergence libre, translation). Ces modèles et algorithmes seront en particulier appliqués à l'interpolation d'images océanographiques.

Chapitre II

Étude et convergence de l'algorithme de Benamou-Brenier sous hypothèse d'existence d'un point selle

Notes pour ce chapitre.

- Nous rappelons que $|\cdot|$ désigne la norme euclidienne sur \mathbb{R}^d ($d \in \mathbb{N}$).
- Sauf mention contraire, \mathbb{R}^d sera muni de la tribu de Lebesgue.
- Pour $f \in L^p(\omega)$ et $x \in \omega$, le terme " $f(x)$ " désignera, par abus de notation, une valeur $\tilde{f}(x)$ où la fonction \tilde{f} est une représentante de la classe d'équivalence f pour la mesure de Lebesgue (i.e. pour la relation d'équivalence basée sur l'égalité \mathcal{L}^d presque partout).
- Nous désignerons un intervalle entre a et b par la notation (a, b) lorsque l'ouverture et la fermeture de cet intervalle en a ou b est sans importance. Au contraire, lorsque ces bornes sont précisées, sauf erreur, cela signifie que ces bornes ont une importance. Par exemple lorsque l'on écrit $f \in L_{loc}^\infty([0, 1] \times \mathbb{R}^d)$, cela signifie en particulier que $f \in L^\infty([0, 1] \times K)$ pour tout compact K de \mathbb{R}^d : autrement dit que f est " L^∞ localement en espace mais pas en temps" (i.e. reste localement borné en $t = a$ et $t = b$).

Dans ce chapitre, nous introduisons tout d'abord dans la section II.1 la formulation dynamique du problème de transport optimal proposée par Benamou et Brenier dans [6]. Le problème de point selle qui en découle ainsi que sa résolution sont alors analysés. Les contributions de ce chapitre sont donc la caractérisation des points selles du problème (section II.2) puis l'étude de la convergence de l'algorithme de [6] vers un point selle (sections II.3 et II.4). Ces résultats sont obtenus sous hypothèse d'existence de points selle. L'existence sera quant à elle démontrée dans le chapitre suivant.

II.1 Formulation du problème et présentation de l'algorithme

II.1.1 Le problème de Monge-Kantorovich dans \mathbb{R}^d pour un coût quadratique

Dans tout ce qui suit, nous désignerons par $|\cdot|$ la norme euclidienne sur \mathbb{R}^d (pour tout $d \in \mathbb{N}$). Considérons deux densités (ρ_0, ρ_1) positives sur \mathbb{R}^d ($d \in \mathbb{N}^*$), à supports bornés et de même masse. Le problème de Monge-Kantorovich revient à déterminer un plan de transport optimal T entre ρ_0 et ρ_1 minimisant

$$\int_{\mathbb{R}^d} d(x, T(x))^2 \rho_0(x) dx, \quad (2.1-1)$$

où $d(x, y)$ représente une distance sur Ω . Pour le cas isotrope, on utilise habituellement la distance euclidienne $d(x, y) = |x - y|$. On écrit $T\#(\rho_0\mathcal{L}^d) = \rho_1\mathcal{L}^d$ la poussée de mesure par T de $\rho_0\mathcal{L}^d$ sur $\rho_1\mathcal{L}^d$, i.e. pour tout sous-ensemble borné A de \mathbb{R}^d ,

$$\int_A \rho_0 = \int_{T^{-1}(A)} \rho_1 = 1.$$

La distance de Wasserstein $\mathcal{W}_2(\rho_0, \rho_1)$ est ici rappelons-le définie par :

$$\mathcal{W}_2(\rho_0\mathcal{L}^d, \rho_1\mathcal{L}^d)^2 = \inf_{T\#(\rho_0\mathcal{L}^d)=\rho_1\mathcal{L}^d} \int_{\Omega} d(x, T(x))^2 \rho_0(x) dx.$$

Dans le cas isotrope (en prenant $d(x, y) = |x - y|$), il existe un unique plan de transport T entre ρ_0 et ρ_1 pouvant être écrit comme le gradient d'une fonction convexe s.c.i. ϕ (Théorème 1.1-7 p. 20) i.e.

$$\mathcal{W}_2(\rho_0\mathcal{L}^d, \rho_1\mathcal{L}^d)^2 = \int_{\mathbb{R}^d} |\nabla\phi(x) - x|^2 \rho_0(x) dx = \inf_{T\#(\rho_0\mathcal{L}^d)=\rho_1\mathcal{L}^d} \int_{\mathbb{R}^d} |T(x) - x|^2 \rho_0(x) dx.$$

Ce problème, dans la formulation introduite par J.D. Benamou and Y. Brenier [6], peut se ramener à un problème de minimisation d'une certaine énergie cinétique \mathcal{K} , sur le déplacement continu d'une masse ρ depuis ρ_0 jusqu'à ρ_1 , et porté par un champ de vitesse v .

Commençons donc par formuler précisément ce nouveau problème d'optimisation dans un cadre qui permettra sa résolution par l'algorithme de lagrangien augmenté également introduit dans [6], et qui sera l'objet principal de notre étude.

II.1.2 Formulation convexe et lagrangien augmenté

On se propose donc d'étudier le problème suivant : soient $\rho_0, \rho_1 \in L^2(\mathbb{R}^d)$ deux densités de probabilité, à supports bornés et de même masse, i.e. $\rho_0, \rho_1 \geq 0$ et

$$\int_{\mathbb{R}^d} \rho_0 = \int_{\mathbb{R}^d} \rho_1 = 1.$$

La formulation dynamique du transport optimal consiste à augmenter la dimension du problème avec une variable temporelle $t \in [0; 1]$. Formellement, on cherche un couple (ρ, v) , où ρ désigne une densité positive, et v un champ de vecteurs de \mathbb{R}^d , tous deux définis sur $]0, 1[\times \Omega$, où Ω est un ouvert convexe borné de \mathbb{R}^d contenant $\text{supp}(\rho_0)$ et $\text{supp}(\rho_1)$, satisfaisant l'équation de conservation de la masse,

$$\partial_t \rho + \text{div}(\rho v) = 0; \tag{2.1-2}$$

pour ρv vérifiant des conditions de Neumann homogènes aux bord, et pour ρ vérifiant les conditions initiales et finales

$$\rho(0, x) = \rho_0(x), \quad \rho(1, x) = \rho_1(x), \tag{2.1-3}$$

tels que ρ et v minimisent :

$$\mathcal{K}(\rho, v) = \frac{1}{2} \int_0^1 \int_{\Omega} |v(t, x)|^2 \rho(t, x) dx dt. \tag{2.1-4}$$

Afin de rendre le problème convexe et linéariser la contrainte de conservation de la masse, on pose à présent $m = \rho v$ et on cherche à minimiser le coût de transport [6] :

$$\tilde{\mathcal{K}}(\rho, m) = \int_0^1 \int_{\Omega} \frac{|m(t, x)|^2}{2\rho(t, x)} dx dt, \quad (2.1-5)$$

(ρ, m) satisfaisant toujours les conditions initiales et finales (2.1-3) ainsi que la nouvelle contrainte de conservation de la masse :

$$\partial_t \rho + \operatorname{div}_x m = 0 \quad (2.1-6)$$

sous la contrainte $\rho > 0$ ou $(\rho, m) = (0, 0)$ (ce qui caractérise la contrainte $\rho \geq 0$ et le changement de variable $m = \rho v$).

En introduisant un multiplicateur de Lagrange ψ pour gérer les contraintes linéaires (2.1-6) et (2.1-3), on peut établir une formulation point-selle du problème :

$$\inf_{(\rho, m)} \sup_{\psi} \left(\int_{]0,1[\times \Omega} \frac{|m|^2}{\rho} - \int_{]0,1[\times \Omega} (\partial_t \psi \rho - \nabla_x \psi \cdot m) + \int_{\Omega} (\psi(0, \cdot) \rho_0 - \psi(1, \cdot) \rho_1) \right) \quad (2.1-7)$$

Afin d'obtenir une formulation algorithmique simple, nous introduisons, en suivant [6], la transformée de Legendre de $(\rho, m) \mapsto \frac{|m|^2}{2\rho}$ qui s'écrit :

$$F(q) = F(a, b) = \sup_{(\rho, m)} \left(\rho a + \langle m, b \rangle - \frac{|m|^2}{2\rho} \right) \Leftrightarrow F(q) = \begin{cases} 0 & \text{si } q \in \mathcal{P} \\ +\infty & \text{sinon} \end{cases}$$

avec $q = (a, b) \in \mathbb{R} \times \mathbb{R}^d$ et $\mathcal{P} = \left\{ (a, b) \in \mathbb{R} \times \mathbb{R}^d, a \leq -\frac{|b|^2}{2} \right\}$.

Le coût de transport étant désormais convexe et s.c.i, il est égal à sa biconjuguée par la transformée de Legendre. On a par conséquent

$$\frac{|m|^2}{2\rho} = \sup_{(a, b)} (\rho a + m \cdot b - F(a, b)).$$

Le problème est ainsi en partie linéarisé par rapport aux variables (ρ, m) : la partie non-linéaire (i.e. F) se réduit à l'indicatrice du paraboloïde \mathcal{P} , dont nous verrons l'utilité plus bas.

Ainsi, par une interversion sup-intégrale et inf-sup, et en posant $q = (a, b)$ et $\mu = (\rho, m)$, on reformule le problème de point selle (2.1-7) comme :

$$\inf_{(\psi, q)} \sup_{\mu} L(\psi, p, \mu) \quad (2.1-8)$$

où

$$\mathbf{L}(\psi, p, \mu) = \tilde{F}(q) + G(\psi) + \langle \mu, \nabla_{t,x} \psi - q \rangle_{L^2} \quad (2.1-9)$$

avec

$$G(\psi) = \int_{\Omega} \psi(0, \cdot) \rho_0 dx - \int_{\Omega} \psi(1, \cdot) \rho_1 dx \quad (2.1-10)$$

et $F(q) = \begin{cases} 0 & \text{si } q \in \tilde{\mathcal{P}} \\ +\infty & \text{sinon} \end{cases}$,

où $\tilde{\mathcal{P}} = \left\{ (\tilde{a}, \tilde{b}) \in L^2((0, 1) \times \Omega) \times L^2((0, 1) \times \Omega)^d, \tilde{a} \leq -\frac{|\tilde{b}|^2}{2} \right\}$.

Par la suite on écrira souvent, par abus de notation, \mathcal{P} au lieu de $\tilde{\mathcal{P}}$. La formulation lagrangien augmenté est finalement :

$$\mathbf{L}_r(\psi, q, \mu) = F(q) + G(\psi) + \langle \mu, \nabla_{t,x}\psi - q \rangle_{L^2} + \frac{r}{2} \|\nabla_{t,x}\psi - q\|_{L^2}^2 \quad (2.1-11)$$

où $r > 0$.

II.1.3 Algorithme de Benamou-Brenier

Pour résoudre le problème de point selle (2.1-8), J.D. Benamou et Y. Brenier [6] ont proposé un algorithme basé sur une méthode de type Uzawa : il s'agit de reproduire le deuxième algorithme de lagrangien augmenté (ALG2) introduit par M. Fortin et R. Glowinski dans [37]. Celui-ci consiste à effectuer les opérations suivantes :

$(\psi^{n-1}, q^{n-1}, \mu^n)$ étant donnés :

1. Étape A : On détermine l'unique ψ^n tel que :

$$\mathbf{L}_r(\psi^n, q^{n-1}, \mu^n) \leq \mathbf{L}_r(\psi, q^{n-1}, \mu^n), \quad \forall \psi. \quad (2.1-12)$$

2. Étape B : On détermine l'unique $q^n = (a^n, b^n)$ tel que :

$$\mathbf{L}_r(\psi^n, q^n, \mu^n) \leq \mathbf{L}_r(\psi^n, q, \mu^n), \quad \forall q. \quad (2.1-13)$$

3. Étape C : La mise à jour :

$$(\rho^{n+1}, m^{n+1}) = \mu^{n+1} = \mu^n + r(\nabla_{t,x}\psi^n - q^n), \quad (2.1-14)$$

(où $r > 0$ est le paramètre du lagrangien augmenté : on fait le choix de prendre ce même paramètre pour le paramètre d'Uzawa (étape C) pour assurer la contrainte de positivité de ρ^n et d'annulation de m^n lorsque ρ^n s'annule, comme expliqué ci-dessous).

Plus précisément, l'algorithme (2.1-12)-(2.1-14) se décompose de la manière suivante :

1. Étape A (projection sur l'espace des contraintes (2.1-2) et (2.1-3) p. 52) :
On cherche l'unique $\psi^n \in H^1((0, 1) \times \Omega)/\mathbb{R}$ tel que :

$$\forall h \in H^1((0, 1) \times \Omega), \quad G(h) + \langle \mu^n, \nabla_{t,x}h \rangle + r \langle \nabla_{t,x}\psi^n - q^{n-1}, \nabla_{t,x}h \rangle = 0.$$

Plus formellement, il s'agit de trouver ψ^n solution de :

$$-r\Delta_{t,x}\psi^n = \operatorname{div}_{t,x}(\mu^n - rq^{n-1})$$

pour les conditions initiales et finales :

$$r\partial_t\psi^n(0, \cdot) = \rho_0 - \rho^n(0, \cdot) + ra^{n-1}(0, \cdot), \quad \text{et} \quad r\partial_t\psi^n(1, \cdot) = \rho_1 - \rho^n(1, \cdot) + ra^{n-1}(1, \cdot),$$

et pour des conditions de Neumann homogènes aux bords. Dans [6], cette étape est résolue par une méthode de Poisson. On observe que cette opération correspond à une projection orthogonale sur un espace affine de divergence espace-temps constante.

2. Étape B (projection orthogonale sur \mathcal{P}) :

$$q^n = P_{\mathcal{P}} \left(\frac{1}{r} \mu^n + \nabla_{t,x}\psi^n \right)$$

projection L^2 qui correspond donc à une projection point par point.

3. Étape C : On pose $\mu^{n+1} = \mu^n + r(\nabla_{t,x}\psi^n - q^n)$.

L'étape B se réécrit $\forall q' \in \mathcal{P} \langle \frac{1}{r}\mu^{n+1}, q' - q^n \rangle \leq 0$,

c'est à dire que $\mu^{n+1} = (\rho^{n+1}, m^{n+1})$ est orthogonal au parabolöide \mathcal{P} en q^n . On en déduit que, pour presque tout $(t, x) \in]0, 1[\times \Omega$,

$$\begin{cases} \rho^n(t, x) > 0 \\ \text{ou } (\rho^n(t, x), m^n(t, x)) = 0. \end{cases}$$

Nous préciserons cette propriété par la Proposition 2.2-3 (p. 58), dans la section suivante.

Plan du chapitre

L'objectif de ce chapitre sera, rappelons-le, de prouver la convergence de l'algorithme (2.1-12)-(2.1-14) vers un point selle du Lagrangien \mathbf{L} défini en (2.1-11).

1. La section II.2 se donnera pour objectif de caractériser les points selles de \mathbf{L} dans le but, à la fois de poser plus facilement les bases de formulation du problème d'existence de tels points selles (qui sera le principal objet d'étude du chapitre III), et d'amener à une reformulation du problème de convergence de notre algorithme vers un de ces points selles en un problème plus générique de convergence vers un point fixe d'opérateur non-expansif (section II.3 et II.4).
 - Nous commencerons donc par mettre en évidence les propriétés caractéristiques des points selles du Lagrangien augmenté \mathbf{L}_r dans les Propositions 2.2-1 et 2.2-3. La Proposition 2.2-3 prouve en particulier quelques propriétés "attendues" de ce point selle : le fait que la densité ρ soit positive, et que le moment cinétique m soit "absolument continu" par rapport à ρ .
 - Nous montrerons ensuite dans la Proposition 2.2-4 que la composante $\mu = (\rho, \rho v)$ d'un tel point selle minimise bien l'énergie (2.1-5), sous les contraintes (2.1-6) et (2.1-3) (autre propriété "attendue" du point selle).
 - Nous concluons cette section par la Proposition 2.2-5, qui prouvera l'équivalence entre notre problème de convergence de l'algorithme (2.1-12)-(2.1-14) vers un point selle de \mathbf{L}_r et un problème de convergence d'itérations successives d'un opérateur \mathbf{M} (défini comme résultant d'une itération de chaque étape de l'algorithme) vers un point fixe de ce dernier (autrement dit l'étude de la convergence d'une suite de type $x_{n+1} = \mathbf{M}x_n$ vers un point fixe de \mathbf{M}).
2. L'étude générique de cette nouvelle formulation du problème sera l'objet de la section II.3. Nous y énoncerons des résultats généraux de convergence de suites issues d'itérations d'un opérateur non-expansif (1-lipschitzien) \mathbf{M} vers un point fixe de ce dernier.
 - Dans la sous-section II.3.1, nous donnerons quelques définitions standards relatifs aux opérateur non-expansifs.
 - Dans la sous-section II.3.2, nous énoncerons des résultats relatifs à la convergence faible d'une suite de type $x_{n+1} = \mathbf{M}x_n$ vers un point fixe de \mathbf{M} (64 p. 64).
 - Dans la sous-section II.3.3, nous énoncerons des résultats relatifs à la convergence forte d'une suite version relaxée de l'algorithme (de type $x_{n+1} = \lambda_n a + (1 - \lambda_n)\mathbf{M}x_n$ vers un point fixe de \mathbf{M} (2.3-8 p. 66).
3. Pour conclure sur la convergence de l'algorithme, nous montrerons que l'opérateur \mathbf{M} défini formellement à la fin de la section II.2 est un **opérateur non-expansif** (plus précisément quasi-fermement non-expansif), et que les résultats de convergence de la section II.3 peuvent ainsi être appliqués.

II.2 Caractérisation des points selles

Afin de montrer la convergence de l'algorithme (2.1-12)-(2.1-14), nous étudions tout d'abord les caractéristiques des points selles du Lagrangien augmenté \mathbf{L}_r défini dans (2.1-11). Pour cela, on considère désormais

$$S = (H^1((0,1) \times \Omega)/\mathbb{R}) \times L^2((0,1) \times \Omega)^{d+1} \times L^2((0,1) \times \Omega)^{d+1}, \quad (2.2-1)$$

l'espace produit des variables d'intérêt (ψ, q, μ) de l'algorithme.

Proposition 2.2-1.

Un point-selle $(\psi^*, q^*, \mu^*) \in S$ de \mathbf{L}_r est caractérisé par les trois propriétés :

- (P1) $\forall q' \in \tilde{\mathcal{P}}, \langle \mu^*, q' - q^* \rangle \leq 0$
- (P2) $\forall h \in H^1((0,1) \times \Omega), G(h) + \langle \mu^*, \nabla_{t,x} h \rangle = 0.$
- (P3) $\nabla_{t,x} \psi^* = q^*.$

Preuve :

• Soit (ψ^*, q^*, μ^*) un point-selle de \mathbf{L}_r (dans ce cas $q^* \in \tilde{\mathcal{P}}$). Montrons qu'il vérifie les propriétés (P1), (P2) et (P3).

(P3) $\forall \mu \in L^2((0,1) \times \Omega)^3, \mathbf{L}_r(\psi^*, q^*, \mu) \leq \mathbf{L}_r(\psi^*, q^*, \mu^*)$, et $\mu \mapsto \mathbf{L}_r(\psi^*, q^*, \mu)$ différentiable en μ^* et le calcul de la différentielle en ce point donne :

$$\forall \mu \in L^2((0,1) \times \Omega)^3, \langle \mu, \nabla_{t,x} \psi^* - q^* \rangle = 0$$

i.e. $\nabla_{t,x} \psi^* - q^* = 0.$

(P2) De même $\forall (\psi, q) \in H^1((0,1) \times \Omega) \times \tilde{\mathcal{P}}, L(\psi, q, \mu^*) \geq L(\psi^*, q^*, \mu^*)$. On fixe $q = q^*$. $\psi \mapsto \mathbf{L}_r(\psi, q^*, \mu^*)$ différentiable en ψ^* et le calcul de la différentielle en ce point donne :

$$\forall h \in H^1((0,1) \times \Omega), G(h) + \langle \mu^*, \nabla_{t,x} h \rangle + r \langle \nabla_{t,x} \psi^* - q^*, \nabla_{t,x} h \rangle = 0$$

et donc comme $\nabla_{t,x} \psi^* - q^* = 0,$

$$\forall h \in H^1((0,1) \times \Omega), G(h) + \langle \mu^*, \nabla_{t,x} h \rangle = 0$$

(P1) $\forall q \in \tilde{\mathcal{P}}, \mathbf{L}_r(\psi^*, q, \mu^*) \geq \mathbf{L}_r(\psi^*, q^*, \mu^*)$. On fixe $q \in \tilde{\mathcal{P}} : \text{pour tout } \lambda \in]0, 1], q^* + \lambda(q - q^*) \in \tilde{\mathcal{P}}$ (convexité de $\tilde{\mathcal{P}}$), d'où $\tilde{F}(q^* + \lambda(q - q^*)) = 0$ (rappelons que \tilde{F} désigne l'indicatrice du paraboloïde $\tilde{\mathcal{P}}$). Ainsi :

$$\forall \lambda \in]0, 1], \mathbf{L}_r(\psi^*, q^* + \lambda(q - q^*), \mu^*) \geq \mathbf{L}_r(\psi^*, q^*, \mu^*)$$

C'est à dire, en rappelant que $\nabla_{t,x} \psi^* = q^*$ (P3) et $G(\psi^*) + \langle \mu^*, \nabla_{t,x} \psi^* \rangle = 0$ (P2),

$$\forall \lambda \in]0, 1], -\lambda \langle \mu^*, q - q^* \rangle + \lambda^2 \frac{r}{2} \|q^* - q\|^2 \geq 0$$

En divisant par λ et en faisant tendre λ vers 0, on conclut (P1) :

$$\forall q \in \tilde{\mathcal{P}}, \langle \mu^*, q - q^* \rangle \leq 0$$

• Inversement soit $(\psi^*, q^*, \mu^*) \in S$ vérifiant les propriétés (P1), (P2) et (P3). Montrons qu'il s'agit bien d'un point-selle de \mathbf{L}_r .

Soit $(\psi, q) \in H^1((0, 1) \times \Omega) \times \tilde{\mathcal{P}}$, on a $\mathbf{L}_r(\psi, q, \mu^*) = - \langle \mu^*, q \rangle + \frac{r}{2} \|\nabla_{t,x} \psi - q\|^2$ (P1), et donc $\mathbf{L}_r(\psi^*, q^*, \mu^*) = - \langle \mu^*, q^* \rangle$ (P3).

On a donc par (P2) l'inégalité :

$$\mathbf{L}_r(\psi, q, \mu^*) \geq \mathbf{L}_r(\psi^*, q^*, \mu^*)$$

De plus $\forall \mu \in L^2((0, 1) \times \Omega)^3$, $\mathbf{L}_r(\psi^*, q^*, \mu) = \mathbf{L}_r(\psi^*, q^*, \mu^*) = G(\psi^*)$. Ainsi, (ψ^*, q^*, μ^*) est un point-selle de \mathbf{L}_r . Un point-selle de \mathbf{L}_r est donc bien caractérisé par ces trois propriétés. \square

Remarque 2.2-1. Notons bien que la preuve de la Proposition 2.2-1 est indépendante de r et fonctionne en particulier pour $r = 0$: autrement dit, l'ensemble des points selles du lagrangien d'origine :

$$\mathbf{L}(\psi, q, \mu) = F(q) + G(\psi) + \langle \mu, \nabla_{t,x} \psi - q \rangle_{L^2},$$

est le même pour tous les lagrangiens augmentés

$$\mathbf{L}_r(\psi, q, \mu) = F(q) + G(\psi) + \langle \mu, \nabla_{t,x} \psi - q \rangle_{L^2} + \frac{r}{2} \|\nabla_{t,x} \psi - q\|_{L^2}^2,$$

et ceci quelque que soit $r > 0$.

Nous allons maintenant prouver que tout point-selle du lagrangien \mathbf{L} (2.1-11) correspond bien à une solution de la formulation dynamique (relations (2.1-2)-(2.1-4)) du problème de Monge.

Définition 2.2-2.

On considère l'espace des contraintes

$$\Lambda(\rho_0, \rho_1) = \left\{ \begin{array}{l} \mu = (\rho, m) \in L^2((0, 1) \times \Omega)^{d+1}, \rho \geq 0, \\ m = \rho v, v \in L^2((0, 1) \times \Omega, \rho \mathcal{L}^d), \\ \forall h \in H^1((0, 1) \times \Omega), G(h) + \langle \mu, \nabla_{t,x} h \rangle_{L^2} = 0, \end{array} \right\} \quad (2.2-2)$$

où $G(h) = \int_{\Omega} (h(0, \cdot) \rho_0 - h(1, \cdot) \rho_1)$, ainsi que le terme d'énergie

$$\begin{aligned} \tilde{\mathcal{K}} : \Lambda(\rho_0, \rho_1) &\longrightarrow \mathbb{R} \\ (\rho, m) &\longmapsto \frac{1}{2} \int_{(0,1) \times \Omega} |v|^2 \rho, \text{ avec } m = \rho v. \end{aligned} \quad (2.2-3)$$

Cette nouvelle formulation du problème de transport optimal consiste trouver à un minimum de $\tilde{\mathcal{K}}$ en imposant :

- la positivité de la densité ρ ,
- "l'absolue continuité" de m par rapport à ρ (i.e. $\exists v, m = \rho v$),
- la conservation de la masse et les conditions initiales et finales.

Commençons par étudier plus en détail la propriété (P1) de la Proposition 2.2-1 concernant les liens entre μ et q lorsque μ est orthogonal au parabolioïde $\tilde{\mathcal{P}}$ au point q .

On rappelle que

$$\tilde{\mathcal{P}} = \left\{ q = (a, b) \in L^2((0, 1) \times \Omega) \times L^2((0, 1) \times \Omega)^2, a + \frac{|b|^2}{2} \leq 0 \right\}. \quad (2.2-4)$$

Proposition 2.2-3.

Soient $\mu^* = (\rho^*, m^*) \in L^2((0, 1) \times \Omega) \times L^2((0, 1) \times \Omega)^2$ et $q^* = (a^*, b^*) \in \tilde{\mathcal{P}}$, tel que μ^* soit orthogonal à $\tilde{\mathcal{P}}$ au point q^* , i.e. Propriété (P1) : $\langle \mu^*, q - q^* \rangle_{L^2} \leq 0, \forall q \in \tilde{\mathcal{P}}$. Alors μ^* et q^* vérifient les trois propriétés suivantes presque partout sur $]0, 1[\times \Omega$:

1. $\rho^* \geq 0$
2. $m^* = \rho^* b^*$
3. $\rho^* \left(a^* + \frac{|b^*|^2}{2} \right) = 0$

La troisième propriété signifie en particulier que pour presque tout $(t, x) \in]0, 1[\times \Omega$:

$$\rho^*(t, x) \neq 0 \Rightarrow a^*(t, x) + \frac{|b^*(t, x)|^2}{2} = 0.$$

Preuve : Pour tout $q = (a, b) \in \tilde{\mathcal{P}}$, on a

$$\langle \mu^*, q - q^* \rangle_{L^2} = \int_{]0, 1[\times \Omega} \rho^*(a - a^*) + \int_{]0, 1[\times \Omega} m^* \cdot (b - b^*).$$

Dans toute la démonstration, on considérera indifféremment (par abus de notation) un représentant d'une classe d'équivalence \mathcal{L}^3 -presque partout avec la classe elle-même.

1. $\rho^* \geq 0$:

Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ définie comme il suit :

$$f : y \in \mathbb{R} \mapsto \begin{cases} \frac{y}{|y|} & \text{si } y \neq 0, \\ 0 & \text{si } y = 0. \end{cases}$$

Pour tout $y \in \mathbb{R}$, $yf(y) = |y|$, donc $\rho^* f(\rho^*) = |\rho^*|$ et $-1 \leq f(\rho^*) \leq 1$ (comme précisé en début de preuve, ρ^* désigne en fait ici sous la même notation, un représentant de la classe de ρ^*). On pose $q = (a, b)$ avec $a = f(\rho^*) - 1 + a^*$ et $b = b^*$. On a ainsi

$$a + \frac{|b|^2}{2} = f(\rho^*) - 1 + a^* + \frac{|b^*|^2}{2} \leq 0,$$

i.e. $q \in \tilde{\mathcal{P}}$, donc $\langle \mu^*, q - q^* \rangle_{L^2} \leq 0$. On a également

$$\langle \mu^*, q - q^* \rangle_{L^2} = \int_{]0, 1[\times \Omega} |\rho^*| - \rho^*.$$

Comme, $|\rho^*| - \rho^* \geq 0$, on en déduit que $|\rho^*| = \rho^*$ \mathcal{L}^d -presque partout, c'est à dire $\rho^* \geq 0$.

2. $m^* = b^* \rho^*$:

Soit $u \in L^\infty((0, 1) \times \Omega)^d$. Pour tout $\epsilon > 0$, on pose $q_\epsilon = (a_\epsilon, b_\epsilon)$, avec $b_\epsilon = b^* + \epsilon u$ et $a_\epsilon = -\frac{|b_\epsilon|^2}{2}$. On obtient ainsi :

$$q_\epsilon \in L^2((0, 1) \times \Omega)^{d+1} \quad \text{et} \quad a_\epsilon + \frac{|b_\epsilon|^2}{2} = 0,$$

ce qui implique que $q_\epsilon \in \tilde{\mathcal{P}}$. On a alors :

$$\begin{aligned} \langle \mu^*, q_\epsilon - q^* \rangle_{L^2} &= \int_{]0,1[\times \Omega} \rho^*(a_\epsilon - a^*) + \int_{]0,1[\times \Omega} m^* \cdot (b_\epsilon - b^*) \\ &= -\frac{\epsilon^2}{2} \int_{]0,1[\times \Omega} |u|^2 \rho^* + \epsilon \int_{]0,1[\times \Omega} (m^* - b^* \rho^*) \cdot u - \int_{]0,1[\times \Omega} \left(a^* + \frac{|b^*|^2}{2} \right) \rho^* \\ &\leq 0. \end{aligned}$$

Comme $-\int_{]0,1[\times \Omega} \left(a^* + \frac{|b^*|^2}{2} \right) \rho^* \geq 0$, on a alors

$$-\frac{\epsilon^2}{2} \int_{]0,1[\times \Omega} |u|^2 \rho^* + \epsilon \int_{]0,1[\times \Omega} (m^* - b^* \rho^*) \cdot u \leq 0$$

et donc, pour tout $\epsilon > 0$,

$$-\frac{\epsilon}{2} \int_{]0,1[\times \Omega} |u|^2 \rho^* + \int_{]0,1[\times \Omega} (m^* - b^* \rho^*) \cdot u \leq 0$$

et $\int_{]0,1[\times \Omega} |u|^2 \rho^* < +\infty$. En faisant tendre ϵ vers 0, on a

$$\int_{]0,1[\times \Omega} (m^* - b^* \rho^*) \cdot u \leq 0$$

On obtient, quitte à changer u en $-u$, $\int_{]0,1[\times \Omega} (m^* - b^* \rho^*) \cdot u = 0$ et ceci pour tout $u \in L^\infty((0,1) \times \Omega)^d$.

Pour tout $M > 0$, on pose $u_M = \text{sign}(m^* - b^* \rho^*) \max(|m^* - b^* \rho^*|, M) \in L^\infty((0,1) \times \Omega)^d$, avec $\text{sign}(y) = \frac{1}{|y|}y$ si $y \neq 0$, et $\text{sign}(y) = 0$, d'où $|u_M| \leq M$.

$(m^* - b^* \rho^*)u_M = |m^* - b^* \rho^*| \max(|m^* - b^* \rho^*|, M) \geq 0$ converge simplement vers $|m^* - b^* \rho^*|^2$ de façon croissante quand M tend vers l'infini. Donc, par le Théorème de Beppo-Levi,

$$0 = \int_{]0,1[\times \Omega} (m^* - b^* \rho^*) \cdot u_M \longrightarrow \int_{]0,1[\times \Omega} |m^* - b^* \rho^*|^2 \quad (M \rightarrow +\infty).$$

On a ainsi $\int_{]0,1[\times \Omega} |m^* - b^* \rho^*|^2 = 0$, i.e. $m^* = \rho^* b^*$.

$$3. \quad \rho^* \left(a^* + \frac{|b^*|^2}{2} \right) = 0 :$$

On pose $q = (a, b) \in \tilde{\mathcal{P}}$, avec $a = -\frac{1}{2}|b^*|^2$ et $b = b^*$, d'où

$$0 \geq \langle \mu^*, q - q^* \rangle_{L^2} = - \int_{]0,1[\times \Omega} \rho^* \left(a^* + \frac{1}{2}|b^*|^2 \right).$$

Or $\rho^* \left(a^* + \frac{1}{2}|b^*|^2 \right) \leq 0$, donc $\rho^* \left(a^* + \frac{|b^*|^2}{2} \right) = 0$ presque partout.

□

À partir de cette Proposition, on peut à présent montrer que la composante μ de tout point selle de \mathbf{L} minimise le coût de transport $\tilde{\mathcal{K}}$ sur l'espace des contraintes $\Lambda(\rho_0, \rho_1)$.

Proposition 2.2-4.

Si $(\psi^*, q^*, \mu^*) \in S$ est un point selle de \mathbf{L} , alors $\mu^* = (\rho^*, m^*) \in \Lambda(\rho_0, \rho_1)$ et minimise $\tilde{\mathcal{K}}$.

En posant $q^* = (a^*, b^*) \in L^2((0, 1) \times \Omega) \times L^2((0, 1) \times \Omega)^2$, on a de plus

$$\tilde{\mathcal{K}}(\mu^*) = \frac{1}{2} \int_{]0,1[\times \Omega} |b^*|^2 \rho^* = \langle \mu^*, q^* \rangle_{L^2} = \inf_{T \# \rho_0 = \rho_1} \int_{\mathbb{R}^d} |T(x) - x|^2 \rho_0(x) dx$$

Preuve :

1. $\mu^* \in \Lambda(\rho_0, \rho_1)$.

D'après la propriété (P1) de la Proposition 2.2-1 (p. 56), μ^* est orthogonal au paraboloïde au point q , i.e. $\langle \mu^*, q - q^* \rangle_{L^2} \leq 0$, $\forall q \in \tilde{\mathcal{P}}$, avec $\tilde{\mathcal{P}}$ défini par (2.2-4) p. 57. D'après la Proposition 2.2-3 (p. 58),

- $\rho^* \geq 0$
- $m^* = b^* \rho^*$
- $\rho^* \left(a^* + \frac{|b^*|^2}{2} \right) = 0$

De plus par la propriété (P2) de la Proposition 2.2-1, μ^* vérifie bien

$$\forall h \in H^1((0, 1) \times \Omega), \quad G(h) + \langle \mu^*, \nabla_{t,x} h \rangle_{L^2} = 0.$$

ce qui implique, pour $h = \psi^*$, et grâce à (P3), que :

$$-\mathbf{L}(\psi^*, q^*, \mu^*) = \langle \mu^*, q^* \rangle_{L^2} = \int_{]0,1[\times \Omega} (\rho^* a^* + m^* \cdot b^*).$$

D'après les trois propriétés de la Proposition 2.2-3 :

$$\int_{]0,1[\times \Omega} (\rho^* a^* + m^* \cdot b^*) = \int_{]0,1[\times \Omega} |b^*|^2 \rho^*,$$

et ainsi,

$$\int_{]0,1[\times \Omega} |b^*|^2 \rho^* = \langle \mu^*, q^* \rangle_{L^2} \leq \|\mu^*\|_{L^2} \cdot \|q^*\|_{L^2} < +\infty.$$

On a donc $\mu^* \in \Lambda(\rho_0, \rho_1)$, et on peut prendre $v^* = b^*$.

2. $\mu^* = (\rho^*, m^*)$ minimise $\tilde{\mathcal{K}}$.

Soit $\mu = (\rho, m) \in \Lambda(\rho_0, \rho_1)$ et v tel que $m = v\rho$, alors

$$\begin{aligned} \tilde{\mathcal{K}}(\mu) &= \frac{1}{2} \int_{]0,1[\times \Omega} |v|^2 \rho = \frac{1}{2} \int_{]0,1[\times \Omega} (|v|^2 + |b^*|^2) \rho - \frac{1}{2} \int_{]0,1[\times \Omega} |b^*|^2 \rho \\ &\geq \int_{]0,1[\times \Omega} v \cdot b^* \rho - \frac{1}{2} \int_{]0,1[\times \Omega} |b^*|^2 \rho \end{aligned}$$

Rappelons que q^* est un élément du paraboloïde \mathcal{P} et donc $a^* + \frac{|b^*|^2}{2} \leq 0$, d'où

$$-\frac{1}{2} \int_{]0,1[\times \Omega} |b^*|^2 \rho \geq \int_{]0,1[\times \Omega} a^* \rho.$$

Sachant que $m = \rho v$, on a alors

$$\tilde{\mathcal{K}}(\mu) \geq \int_{]0,1[\times \Omega} m \cdot b^* + \int_{]0,1[\times \Omega} a^* \rho \geq \langle (\rho, m), (a^*, b^*) \rangle_{L^2} = \langle \mu, q^* \rangle_{L^2}.$$

Comme le triplet (ψ^*, q^*, μ^*) est un point-selle, et toujours d'après la Proposition 2.2-1, on a $\nabla_{t,x} \psi^* = q^*$, et

$$G(\psi^*) + \langle \mu, \nabla_{t,x} \psi^* \rangle_{L^2} = 0 = G(\psi^*) + \langle \mu^*, \nabla_{t,x} \psi^* \rangle_{L^2}$$

car $\mu \in \Lambda(\rho_0, \rho_1)$. On obtient alors

$$\langle \mu, q^* \rangle_{L^2} = \langle \mu, \nabla_{t,x} \psi^* \rangle_{L^2} = \langle \mu^*, \nabla_{t,x} \psi^* \rangle_{L^2} = \langle \mu^*, q^* \rangle_{L^2} = \int_{]0,1[\times \Omega} a^* \rho^* + \int_{]0,1[\times \Omega} m^* \cdot b^*.$$

Rappelons que d'après la Proposition 2.2-3, $a^* \rho^* = -\frac{1}{2} |b^*|^2 \rho^*$ et $m^* = b^* \rho^*$, d'où

$$\langle \mu, q^* \rangle_{L^2} = \int_{]0,1[\times \Omega} a^* \rho^* + \int_{]0,1[\times \Omega} m^* \cdot b^* = \frac{1}{2} \int_{]0,1[\times \Omega} |b^*|^2 \rho^* = \tilde{\mathcal{K}}(\mu^*).$$

On déduit ainsi que $\tilde{\mathcal{K}}(\mu) \geq \tilde{\mathcal{K}}(\mu^*)$.

De ce résultat, on conclut que si $(\psi_1^*, q_1^*, \mu_1^*)$ et $(\psi_2^*, q_2^*, \mu_2^*)$ sont deux points selle de \mathbf{L} , alors $\tilde{\mathcal{K}}(\mu_1^*) = \tilde{\mathcal{K}}(\mu_2^*)$.

3. Pour conclure il suffit donc de prouver qu'il existe un point selle $(\hat{\psi}, \hat{q}, \hat{\mu})$ vérifiant :

$$\tilde{\mathcal{K}}(\hat{\mu}) = \int_{\mathbb{R}^d} |\nabla \phi(x) - x|^2 \rho_0(x) dx = \inf_{T \# \rho_0 = \rho_1} \int_{\mathbb{R}^d} |T(x) - x|^2 \rho_0(x) dx,$$

avec ϕ définie comme dans le Théorème 1.1-7 (p. 20) : c'est à dire telle que $\nabla \phi$ soit le plan de transport optimal de $\rho_0 \mathcal{L}^d$ sur $\rho_1 \mathcal{L}^d$.

Le chapitre suivant est consacré à la construction d'un tel point selle $(\hat{\psi}, \hat{q}, \hat{\mu})$, vérifiant en particulier : $\hat{\mu} = (\hat{\rho}, \hat{\rho} \hat{v})$ avec pour tout $t \in]0, 1[$ et $X(t, \cdot) = (1-t)id + t \nabla_x \phi$,

$$\hat{\rho}(t, \cdot) = X(t, \cdot) \# \rho_0$$

et $\hat{v} \in L^\infty((0, 1) \times \Omega)$ continue sur $]0, 1[\times \Omega$, tel que pour tout $(t, x) \in]0, 1[\times \Omega$:

$$\nabla \phi(x) - x = \partial_t X(t, x) = v(t, X(t, x))$$

(et en choisissant bien ϕ parmi toutes les possibilités).

Anticipons et concluons tout de suite qu'avec un tel point selle, on a la relation :

$$\tilde{\mathcal{K}}(\hat{\mu}) = \frac{1}{2} \int_{]0,1[\times \Omega} |\hat{v}|^2 \hat{\rho} = \int_0^1 \int_{\Omega} |\hat{v}(t, X(t, x))|^2 \rho_0(x) dx dt = \int_{\Omega} |\nabla \phi(x) - x|^2 \rho_0(x) dx.$$

Ainsi, on peut conclure d'après la Proposition 2.2-4 (p. 60), que pour tout point selle $(\psi^*, q^*, \mu^*) \in \mathcal{S}$ de \mathbf{L} , $\mu^* \in \Lambda(\rho_0, \rho_1)$ et pour tout $\mu \in \Lambda(\rho_0, \rho_1)$:

$$\tilde{\mathcal{K}}(\mu) \geq \tilde{\mathcal{K}}(\mu^*) = \int_{\Omega} |\nabla \phi(x) - x|^2 \rho_0(x) dx = \inf_{T \# \rho_0 = \rho_1} \int_{\Omega} |T(x) - x|^2 \rho_0(x) dx. \quad (2.2-5)$$

□

La Proposition suivante montre finalement qu'un point selle de \mathbf{L} correspond bien à une solution de la formulation dynamique (2.1-2)-(2.1-4) du problème de Monge.

Proposition 2.2-5.

Il y a équivalence entre un point-selle de \mathbf{L} (et donc de \mathbf{L}_r pour tout $r > 0$) et un point fixe de l'algorithme (défini à la sous-section II.1.3 p. 54). Autrement dit (ψ, q, μ) est un point-selle de \mathbf{L} si et seulement si il est resté invariant pour l'algorithme de Benamou-Brenier (ici on prend $r > 0$).

Preuve :

• Soit (ψ, q, μ) un point-selle de \mathbf{L} . On note (ψ', q', μ') le nouveau point obtenu après une étape de l'algorithme. Montrons que $(\psi', q', \mu') = (\psi, q, \mu)$ dans S .

1. Par l'étape A, on a $\forall h \in H^1((0, 1) \times \Omega)$, $G(h) + \langle \mu, \nabla_{t,x} h \rangle + r \langle \nabla_{t,x} \psi' - q, \nabla_{t,x} h \rangle = 0$.
Par la 1^{ère} propriété ($G(h) + \langle \mu, \nabla_{t,x} h \rangle = 0$), on a :

$$\forall h \in H^1((0, 1) \times \Omega), \quad r \langle \nabla_{t,x} \psi' - q, \nabla_{t,x} h \rangle = 0$$

et par la 3^{ème} propriété ($\nabla_{t,x} \psi = q$) et en prenant $h = \psi' - \psi$, on obtient :
 $\|\nabla_{t,x}(\psi' - \psi)\| = 0$. L'inégalité de Poincaré étant valable dans $H^1((0, 1) \times \Omega)/\mathbb{R}$, on obtient $\psi' = \psi$ dans $H^1((0, 1) \times \Omega)/\mathbb{R}$.

2. Dans l'étape B on cherche l'unique q' qui vérifie :

$$\forall p \in \tilde{\mathcal{P}}, \quad \left\langle \frac{1}{r} \mu + \nabla \psi' - q', p - q' \right\rangle \leq 0.$$

Comme $\psi = \psi'$, par les propriétés (P1) et (P3) caractérisant le point selle, q est un bon candidat et donc le seul, d'où $q' = q$. En effet,

$$\forall p \in \tilde{\mathcal{P}}, \quad \langle \mu, p - q \rangle \leq 0 \text{ et } \nabla_{t,x} \psi' = \nabla_{t,x} \psi = q,$$

donc

$$\forall p \in \tilde{\mathcal{P}}, \quad \left\langle \frac{1}{r} \mu + \nabla \psi - q, p - q \right\rangle \leq 0.$$

3. Au final $\nabla_{t,x} \psi' = \nabla_{t,x} \psi = q = q'$ et donc par l'étape C on a finalement $\mu' = \mu$.

• Pour finir, soit maintenant (ψ, q, μ) un point fixe de l'algorithme. Montrons que c'est un point-selle de \mathbf{L} . On a :

1. Etape A : $\forall h \in H^1((0, 1) \times \Omega)$, $G(h) + \langle \mu, \nabla_{t,x} h \rangle + r \langle \nabla_{t,x} \psi - q, \nabla_{t,x} h \rangle = 0$.
2. Etape B : $\forall p \in \tilde{\mathcal{P}}, \quad \left\langle \frac{1}{r} \mu + \nabla \psi - q, p - q \right\rangle \leq 0$.
3. Etape C : $\mu = \mu + r(\nabla_{t,x} \psi - q)$.

L'étape C donne alors $\nabla_{t,x} \psi = q$, et par conséquent l'étape B donne :

$$\forall p \in \tilde{\mathcal{P}}, \quad \frac{1}{r} \langle \mu, p - q \rangle \leq 0$$

et l'étape A donne :

$$\forall h \in H^1((0,1) \times \Omega), G(h) + \langle \mu, \nabla_{t,x} h \rangle = 0$$

Les trois propriétés de la Proposition 2.2-1 étant vérifiées, (ψ, q, μ) est donc un point selle de \mathbf{L} . \square

Il est maintenant possible, par la Proposition 2.2-5, de redéfinir formellement notre problème de convergence de l'algorithme (2.1-12)-(2.1-14) vers un point selle du lagrangien \mathbf{L} , en un problème de convergence d'une suite de type $x_{n+1} = \mathbf{M}x_n$ vers un point fixe de l'opérateur \mathbf{M} défini comme le produit d'une itération de chaque étape de l'algorithme (et en prenant par exemple $x_n = (\psi_n, q_n, \mu_n)$).

Nous verrons dans la section II.4 qu'en se plaçant dans un espace normé approprié, il est possible de caractériser un tel opérateur \mathbf{M} par un opérateur de type non-expansif (1-lipschitzien).

En préambule à cette reformulation, la section suivante sera consacrée à la formulation de quelques résultats généraux utiles à notre étude à propos des opérateurs non-expansifs.

II.3 Quelques résultats de convergence pour les opérateurs non-expansifs

Pour tout opérateur non-expansif \mathbf{M} , nous démontrerons dans cette section une série de résultats permettant d'obtenir :

1. la convergence **faible** vers un point fixe de \mathbf{M} de l'algorithme itératif $x_{n+1} = \mathbf{M}x_n$,
2. la convergence **forte** d'une version relaxée de cet algorithme.

II.3.1 Opérateur non-expansif \mathbf{M}

Nous commençons par rappeler quelques définitions standard utiles pour la suite. Soit $(H, \langle \cdot, \cdot \rangle)$ un espace de Hilbert et soit $\mathbf{M} : H \rightarrow H$.

Définition 2.3-1. OPÉRATEUR NON-EXPANSIF

\mathbf{M} est dit **non-expansif** si et seulement si il est 1-lipschitzien :

$$\forall x, y \in H, \|\mathbf{M}x - \mathbf{M}y\| \leq \|x - y\|.$$

Définition 2.3-2. OPÉRATEUR FERMEMENT NON-EXPANSIF

\mathbf{M} est dit **fermement non-expansif** si et seulement si :

$$\forall x, y \in H, \|\mathbf{M}x - \mathbf{M}y\|^2 \leq \langle x - y, \mathbf{M}x - \mathbf{M}y \rangle$$

ou bien de façon équivalente :

$$\forall x, y \in H, \|x - y\|^2 - \|\mathbf{M}x - \mathbf{M}y\|^2 \geq \|(x - y) - (\mathbf{M}x - \mathbf{M}y)\|^2.$$

Définition 2.3-3. OPÉRATEUR QUASI-FERMEMENT NON-EXPANSIF

\mathbf{M} est dit **quasi-fermement non-expansif** si et seulement si \mathbf{M} est non-expansif et pour tout éventuel point fixe x^* de \mathbf{M} :

$$\forall x \in H, \|x - x^*\|^2 - \|\mathbf{M}x - \mathbf{M}x^*\|^2 \geq \|(x - x^*) - (\mathbf{M}x - \mathbf{M}x^*)\|^2.$$

c'est à dire

$$\forall x \in H, \|x - x^*\|^2 - \|\mathbf{M}x - x^*\|^2 \geq \|x - \mathbf{M}x\|^2.$$

II.3.2 Convergence faible de l'algorithme itératif $x_{n+1} = \mathbf{M}x_n$

Nous considérons toujours $(H, \langle \cdot, \cdot \rangle)$ un espace de Hilbert. Nous commençons par énoncer un Lemme simple mais important pour la suite :

Lemme 2.3-4. LEMME D'OPIAL

Soit \mathbf{M} un opérateur non-expansif sur un espace de Hilbert H . Si (x_n) une suite d'éléments de H converge faiblement vers x avec $\mathbf{M}x_n - x_n$ convergeant fortement vers 0, alors x est un point fixe de \mathbf{M} .

Preuve :

$$\begin{aligned} \|x_n - x\|^2 - \|\mathbf{M}x_n - \mathbf{M}x\|^2 &\geq 0 \\ \Leftrightarrow \langle (x_n - \mathbf{M}x_n) + (\mathbf{M}x - x), \mathbf{M}x_n + x_n - \mathbf{M}x - x \rangle &\geq 0 \\ \Leftrightarrow \langle x_n - \mathbf{M}x_n, (\mathbf{M}x_n - x_n) + 2x_n - x - \mathbf{M}x \rangle - \langle x - \mathbf{M}x, \mathbf{M}x_n - x_n \rangle \\ &\quad - \langle x - \mathbf{M}x, 2x_n - x - \mathbf{M}x \rangle \geq 0. \end{aligned}$$

De plus $x_n \rightharpoonup x$ (i.e. convergence faible des x_n vers x , d'où le fait que (x_n) soit bornée), et $x_n - \mathbf{M}x_n \rightarrow 0$ (i.e. convergence **forte**).

On obtient alors par passage à la limite : $-\|x - \mathbf{M}x\|^2 \geq 0$, i.e. $\mathbf{M}x = x$. \square

Énonçons maintenant notre résultat de convergence applicable au problème de la convergence faible de l'algorithme de Benamou-Brenier, et qui sera le résultat principal de cette sous-section.

Théorème 2.3-5.

Soit \mathbf{M} un opérateur quasi-fermement non-expansif et supposons que l'ensemble $\text{Fix}(\mathbf{M})$ des points fixes de \mathbf{M} soit non vide. Soit $(x_n)_n$ une suite d'éléments de H vérifiant pour tout $n \in \mathbb{N}$, l'estimation : $\|\mathbf{M}(x_n) - x_{n+1}\| \leq \epsilon_n$, où $(\epsilon_n)_n$ est une suite réelle positive vérifiant $\sum_n \epsilon_n < +\infty$. Alors $(x_n)_n$ converge faiblement dans H vers un point fixe de \mathbf{M} .

Remarque 2.3-2. La suite $(\epsilon_n)_n$ représente ici les inévitables erreurs numériques inhérentes à l'implémentation d'un tel algorithme. On fait ici l'hypothèse que ces erreurs sont fortement contrôlées ($\sum_n \epsilon_n < +\infty$).

Avant d'aborder la démonstration de ce Théorème, nous commençons par établir le résultat suivant, démontré dans l'Annexe B.1.

Proposition 2.3-6.

Soit une suite $(x_n)_n$ définie comme dans l'énoncé du Théorème 2.3-5, et soient $(x_{\varphi_1(n)})_n$ et $(x_{\varphi_2(n)})_n$ deux sous-suites de $(x_n)_n$ convergeant faiblement respectivement vers x_1 et x_2 (points fixes de \mathbf{M} par le Lemme d'Opial). Alors on a $x_1 = x_2$.

Démonstration du Théorème 2.3-5 Par hypothèse l'ensemble $\text{Fix}(\mathbf{M})$ des points fixes de \mathbf{M} est non vide, on considère donc $x^* \in \text{Fix}(\mathbf{M})$. D'après la définition 2.3-3, on a la relation :

$$\|x_n - x^*\|^2 - \|\mathbf{M}x_n - x^*\|^2 \geq \|\mathbf{M}x_n - x_n\|^2.$$

Notons que $\|\mathbf{M}x_n - x^*\|^2 = \|\mathbf{M}x_n - x_{n+1}\|^2 + 2\langle \mathbf{M}x_n - x_{n+1}, x_{n+1} - x^* \rangle + \|x_{n+1} - x^*\|^2$,

et $|\langle \mathbf{M}x_n - x_{n+1}, x_{n+1} - x^* \rangle| \leq \|\mathbf{M}x_n - x_{n+1}\| \|x_{n+1} - x^*\| \leq C\epsilon_n$ ((x_n) bornée), d'où :

$$\|\mathbf{M}x_n - x_n\|^2 \leq \|x_n - x^*\|^2 - \|\mathbf{M}x_n - x^*\|^2 \leq \|x_n - x^*\|^2 - \|x_{n+1} - x^*\|^2 + C\epsilon_n.$$

Donc, pour tout $n \in \mathbb{N}$, on a $\sum_{k=0}^n \|\mathbf{M}x_k - x_k\|^2 \leq \|x_0 - x^*\|^2 + C \sum_{k=0}^n \epsilon_k < +\infty$. Ainsi $\mathbf{M}x_n - x_n \xrightarrow[n]{} 0$.

Par conséquent d'après le Lemme d'Opial, toute sous-suite de $(x_n)_n$ faiblement convergente a pour limite un point fixe de \mathbf{M} . Comme $(x_n)_n$ est bornée, on peut extraire de toute sous-suite de $(x_n)_n$, une sous-suite convergeant faiblement dans H . Pour conclure sur la convergence (ici faible), il suffit donc d'utiliser la Proposition 2.3-6 précédente pour obtenir que toutes les sous-suites faiblement convergentes tendent vers la même limite. En conséquent on peut conclure que $(x_n)_n$ converge faiblement vers un point fixe de \mathbf{M} , achevant ainsi la démonstration du Théorème 2.3-5. \square

II.3.3 Convergence forte d'une relaxation de l'algorithme $x_{n+1} = \mathbf{M}x_n$

Nous présentons à présent une version relaxée de l'algorithme itératif $x_{n+1} = \mathbf{M}x_n$ dont nous venons de démontrer la convergence faible. Cet version relaxée a pour intérêt de s'appliquer à tous les opérateurs non-expansifs et l'algorithme itératif ne nécessite plus l'hypothèse de quasi-ferme non-expansivité. Cette variation permet en outre d'obtenir une convergence **forte** et non plus seulement faible.

Nous commencerons par énoncer un Lemme caractérisant l'ensemble des points fixes d'un opérateur non-expansif \mathbf{M} .

Lemme 2.3-7.

Soit \mathbf{M} un opérateur non-expansif sur un espace de Hilbert $(H, \langle \cdot, \cdot \rangle)$. Soit $\text{Fix}(\mathbf{M})$ l'ensemble des points fixes de \mathbf{M} . Alors $\text{Fix}(\mathbf{M})$ est un convexe fermé.

Preuve : Remarquons que l'ensemble vide est bien un ensemble convexe fermé. Supposons donc

ici que $\text{Fix}(\mathbf{M})$ est non vide.

Soit $x_1, x_2 \in \text{Fix}(\mathbf{M})$ et soit $x \in [x_1, x_2]$. Montrons que $x \in \text{Fix}(\mathbf{M})$.

1. $\mathbf{M}x \in [x_1, x_2]$:

Soit $c \in [x_1, x_2]$ le projeté orthogonal de $\mathbf{M}x$ sur $[x_1, x_2]$. x_1 et x_2 sont des points fixes de \mathbf{M} , donc :

$$\|x_1 - x\| \geq \|\mathbf{M}x_1 - \mathbf{M}x\| = \|x_1 - \mathbf{M}x\| \text{ et } \|x_2 - x\| \geq \|\mathbf{M}x_2 - \mathbf{M}x\| = \|x_2 - \mathbf{M}x\|$$

$x, c \in [x_1, x_2]$ donc $\|x_1 - x_2\| = \|x_1 - x\| + \|x - x_2\|$ et $\|x_1 - x_2\| = \|x_1 - c\| + \|c - x_2\|$.
Si $\mathbf{M}x \in [x_1, x_2]$, alors $\|\mathbf{M}x - c\| > 0$. Alors

$$\begin{aligned} \|x_1 - x_2\| &= \|x_1 - x\| + \|x - x_2\| \\ &\geq \|x_1 - \mathbf{M}x\| + \|\mathbf{M}x - x_2\| \\ &= \sqrt{\|x_1 - c\|^2 + \|\mathbf{M}x - c\|^2} + \sqrt{\|x_1 - c\|^2 + \|\mathbf{M}x - c\|^2} \\ &> \|x_1 - c\| + \|c - x_2\| = \|x_1 - x_2\|. \end{aligned}$$

Ce qui est absurde, d'où $\mathbf{M}x \in [x_1, x_2]$.

2. $\mathbf{M}x = x$:

$\mathbf{M}x \in [x_1, x_2]$, donc comme précédemment $\|x_1 - x_2\| = \|x_1 - \mathbf{M}x\| + \|\mathbf{M}x - x_2\|$. Supposons que $\mathbf{M}x \in [x_1, x]$, ainsi :

$$\begin{aligned} \|x_1 - x_2\| &= \|x_1 - \mathbf{M}x\| + \|\mathbf{M}x - x\| + \|x - x_2\| \\ &\geq \|x_1 - \mathbf{M}x\| + \|\mathbf{M}x - x\| + \|\mathbf{M}x - x_2\| \\ &= \|x_1 - \mathbf{M}x\| + \|\mathbf{M}x - x\|. \end{aligned}$$

Donc $\|\mathbf{M}x - x\| = 0$, et il en est de même si l'on suppose que $\mathbf{M}x \in [x, x_2]$. Ainsi $\mathbf{M}x = x$ i.e. $x \in \text{Fix}(\mathbf{M})$.

$\text{Fix}(\mathbf{M})$ est donc convexe. De plus, \mathbf{M} est continue : on peut donc trivialement montrer que $\text{Fix}(\mathbf{M})$ est fermé. \square

Nous présentons maintenant l'algorithme en question :

Théorème 2.3-8. (H.Bauschke [3])

Soit \mathbf{M} un opérateur non-expansif et supposons que l'ensemble $\text{Fix}(\mathbf{M})$ des points fixes de \mathbf{M} soit non vide, et soit $(\lambda_n)_{n \geq 0}$ une suite de paramètres de $[0, 1[$ vérifiant :

1. $\lambda_n \rightarrow 0$
2. $\prod_n (1 - \lambda_n) = 0$ ou de manière équivalente $\sum_n \lambda_n = +\infty$
3. $\sum_n |\lambda_{n+1} - \lambda_n| < +\infty$.

Soit H un espace de Hilbert, et \mathbf{M} un opérateur non-expansif sur H .

Etant donnés a, x_0 dans H , on définit la suite (x_n) par récurrence :

$$\forall n \geq 0, \quad x_{n+1} = \lambda_n a + (1 - \lambda_n) \mathbf{M}_n x_n$$

où pour tout $n \in \mathbb{N}$ est vérifiée l'estimation $\|\mathbf{M}_n x_n - \mathbf{M}x_n\| \leq \epsilon_n$, avec $\sum_n \epsilon_n < +\infty$.

Alors la suite $(x_n)_n$ converge fortement vers P_{Fa} (où $F = \text{Fix}(\mathbf{M})$ est l'ensemble des points fixes de \mathbf{M} : c'est un ensemble convexe fermé d'après le Lemme 2.3-7).

La preuve de ce Théorème se trouve dans l'annexe B.2. Nous y reprendrons la démonstration de H. Bauschke dans [3], en ajoutant la prise en compte d'erreurs numériques contrôlées.

Remarque 2.3-3. *Les erreurs numériques que nous prenons en compte dans notre problème sont **fortement contrôlées**. En effet, dans les Théorèmes 2.3-5 et 2.3-8, la suite $(\epsilon_n)_n$ (où ϵ_n représente l'erreur numérique maximale à la n -ième itération de l'opérateur \mathbf{M}) doit vérifier la contrainte*

$$\sum_n \epsilon_n < +\infty,$$

ce qui est une hypothèse très forte sur le contrôle de l'erreur, et pas très réaliste en pratique.

Rappelons que l'opérateur \mathbf{M} représentera dans notre problème une itération de l'algorithme (2.1-12)-(2.1-14), comme nous le verrons dans la section suivante.

II.4 Formulation de l'algorithme de Benamou-Brenier en terme d'opérateur non-expansif

Nous allons maintenant pouvoir mettre en pratique les résultats de convergence énoncés dans la section précédente dans le cadre de l'algorithme Benamou-Brenier. Il suffit pour cela de montrer que les étapes de l'algorithme peuvent être considérées comme l'itération d'un certain opérateur non-expansif.

On en rappelle ici ces étapes :

1. Étape A : On détermine l'unique ψ^n tel que :

$$\mathbf{L}_r(\psi^n, q^{n-1}, \mu^n) \leq \mathbf{L}_r(\psi, q^{n-1}, \mu^n), \forall \psi. \quad (2.4-1)$$

2. Étape B : On détermine l'unique q^n tel que :

$$\mathbf{L}_r(\psi^n, q^n, \mu^n) \leq \mathbf{L}_r(\psi^n, q, \mu^n), \forall q. \quad (2.4-2)$$

3. Étape C : La mise à jour :

$$\mu^{n+1} = \mu^n + r(\nabla_{t,x}\psi^n - q^n). \quad (2.4-3)$$

Définition 2.4-1.

Soit $H = L^2((0,1) \times \Omega)^{d+1} \times L^2((0,1) \times \Omega)^{d+1}$. On munit H du produit scalaire $\langle (\mu_1, q_1), (\mu_2, q_2) \rangle_H = \langle \mu_1, \mu_2 \rangle_{L^2} + r^2 \langle q_1, q_2 \rangle_{L^2}$. $(H, \langle \cdot, \cdot \rangle)$ est un espace de Hilbert.

Définition 2.4-2.

Soit $\mathbf{B} : H \rightarrow H$ l'opérateur qui associe à (μ, q) les produits (μ', q') des deux dernières étapes (B et C, i.e. (2.4-2) et (2.4-3)) de l'algorithme de Benamou-Brenier.

Remarque 2.4-4.

- Ici ψ est éphémère : en effet, observons que si (μ^*, q^*, ψ^*) est un point-selle du lagrangien \mathbf{L} (et donc de \mathbf{L}_r) défini dans (2.1-11), alors $(\mu^*, q^*) = \mathbf{B}(\mu^*, q^*)$. Inversement si (μ^*, q^*) est un point fixe de \mathbf{B} alors (μ^*, q^*, ψ^*) est un point-selle du lagrangien (où ψ^* est l'unique élément de S qui vérifie $q^* = \nabla \psi^*$).

- Le Lemme 2.3-7 (p. 65) nous donne ainsi également la **convexité** de l'ensemble des points selles de \mathbf{L} . En effet, d'après cette Proposition, l'ensemble des points fixes de l'opérateur \mathbf{B} est un ensemble convexe : la linéarité de l'opérateur gradient $\nabla_{t,x}$ la transpose donc à l'ensemble des points selles de \mathbf{L} .

Autrement dit, pour tous points fixes (μ_1^*, q_1^*) et (μ_2^*, q_2^*) de \mathbf{L} et pour tout $\lambda \in [0, 1]$, le couple $(\lambda\mu_1^* + (1-\lambda)\mu_2^*, \lambda q_1^* + (1-\lambda)q_2^*)$ est également un point fixe de \mathbf{B} . Or par linéarité de $\nabla_{t,x}$, on a $\lambda q_1^* + (1-\lambda)q_2^* = \nabla_{t,x}(\lambda\psi_1^* + (1-\lambda)\psi_2^*)$; et ainsi le triplet

$$(\lambda\mu_1^* + (1-\lambda)\mu_2^*, \lambda q_1^* + (1-\lambda)q_2^*, \lambda\psi_1^* + (1-\lambda)\psi_2^*) = \lambda(\mu_1^*, q_1^*, \psi_1^*) + (1-\lambda)(\mu_2^*, q_2^*, \psi_2^*)$$

est un point selle de \mathbf{L} .

Proposition 2.4-3. NON-EXPANSIVITÉ QUASI-FERME DE \mathbf{B}

\mathbf{B} est un opérateur non-expansif, i.e. : pour tout $(\mu_1, q_1), (\mu_2, q_2) \in H$, on a

$$\|\mathbf{B}(\mu_1, q_1) - \mathbf{B}(\mu_2, q_2)\|_H \leq \|(\mu_1, q_1) - (\mu_2, q_2)\|_H. \quad (2.4-4)$$

De plus \mathbf{B} est quasi-fermement non-expansif : si $x^* = (\mu^*, q^*) \in H$ est un point fixe de \mathbf{B} , alors pour tout $x = (\mu, q) \in H$:

$$\|x - x^*\|_H^2 - \|\mathbf{B}x - x^*\|_H^2 \geq \|\mathbf{B}x - x\|_H^2. \quad (2.4-5)$$

Preuve : (μ_1, q_1) et (μ_2, q_2) étant donnés, on détermine $(\mu'_1, q'_1) = \mathbf{B}(\mu_1, q_1)$ et $(\mu'_2, q'_2) = \mathbf{B}(\mu_2, q_2)$ par l'itération (rappel) :

- Etape A : On cherche l'unique $\psi'_i \in (H^1/\mathbb{R})((0, 1) \times \Omega)$ tel que :

$$\forall h \in (H^1/\mathbb{R})((0, 1) \times \Omega) \quad G(h) + \langle \mu_i, \nabla h \rangle_{L^2} + r \langle \nabla \psi'_i - q_i, \nabla h \rangle_{L^2} = 0 \quad (2.4-6)$$

- Etape B : On cherche q'_i tel que :

$$\forall p \in \tilde{\mathcal{P}} \quad \langle \mu_i + r(\nabla \psi'_i - q'_i), p - q'_i \rangle_{L^2} \leq 0 \quad (2.4-7)$$

- Etape C : On définit μ'_i :

$$\mu'_i = \mu_i + r(\nabla_{t,x} \psi'_i - q'_i) \quad (2.4-8)$$

pour $i = 1$ ou 2 .

1. non-expansivité de \mathbf{B} :

Remarquons qu'en injectant (2.4-8) dans (2.4-6) et (2.4-7) on obtient les nouvelles équations :

$$\forall h \in H^1((0, 1) \times \Omega)/\mathbb{R}, \quad G(h) + \langle \mu'_i, \nabla h \rangle_{L^2} + r \langle q'_i - q_i, \nabla h \rangle_{L^2} = 0, \quad (2.4-9)$$

et

$$p \in \tilde{\mathcal{P}}, \quad \langle \mu'_i, p - q'_i \rangle_{L^2} \leq 0. \quad (2.4-10)$$

En prenant (2.4-9) _{$i=2$} - (2.4-9) _{$i=1$} et $h = \psi'_2 - \psi'_1$, on obtient :

$$\langle \mu'_2 - \mu'_1, \nabla(\psi'_2 - \psi'_1) \rangle_{L^2} + r \langle (q'_2 - q'_1) - (q_2 - q_1), \nabla(\psi'_2 - \psi'_1) \rangle_{L^2} = 0,$$

et en sommant (2.4-7)_{i=2} avec $p = q'_1$ et (2.4-7)_{i=1} avec $p = q'_2$, on obtient :

$$\langle \mu'_2 - \mu'_1, q'_2 - q'_1 \rangle_{L^2} \geq 0$$

En posant $\bar{\mu}^{(\prime)} = \mu_2^{(\prime)} - \mu_1^{(\prime)}$, $\bar{q}^{(\prime)} = q_2^{(\prime)} - q_1^{(\prime)}$ et $\bar{\psi}^{(\prime)} = \psi_2^{(\prime)} - \psi_1^{(\prime)}$, on obtient :

$$\langle \bar{\mu}', \nabla \bar{\psi}' \rangle + r \langle \bar{q}' - \bar{q}, \nabla \bar{\psi}' \rangle = 0, \quad (2.4-11)$$

$$\langle \bar{\mu}', \bar{q}' \rangle \geq 0. \quad (2.4-12)$$

En additionnant les relations (2.4-11) et (2.4-12), on a alors :

$$\langle \bar{\mu}', \nabla \bar{\psi}' - \bar{q}' \rangle + r \langle \bar{q}' - \bar{q}, \nabla \bar{\psi}' \rangle \leq 0. \quad (2.4-13)$$

Par ailleurs en factorisant le terme $|\bar{\mu}|^2 - |\bar{\mu}'|^2$, on obtient :

$$\begin{aligned} |\bar{\mu}|^2 - |\bar{\mu}'|^2 &= \langle \bar{\mu} - \bar{\mu}', \bar{\mu} + \bar{\mu}' \rangle \\ &= -r \langle \nabla \bar{\psi}' - \bar{q}', 2(\bar{\mu} + r(\nabla \bar{\psi}' - \bar{q}')) - r(\nabla \bar{\psi}' - \bar{q}') \rangle \\ &= -2r \langle \bar{\mu}', \nabla \bar{\psi}' - \bar{q}' \rangle + r^2 |\nabla \bar{\psi}' - \bar{q}'|^2. \end{aligned}$$

Donc par (2.4-13) :

$$|\bar{\mu}|^2 - |\bar{\mu}'|^2 \geq 2r^2 \langle \nabla \bar{\psi}', \bar{q}' - \bar{q} \rangle + r^2 |\nabla \bar{\psi}' - \bar{q}'|^2. \quad (2.4-14)$$

De plus :

$$\begin{aligned} \langle \nabla \bar{\psi}', \bar{q}' - \bar{q} \rangle &= \langle \nabla \bar{\psi}' - \bar{q}, \bar{q}' - \bar{q} \rangle + \langle \bar{q}, \bar{q}' - \bar{q} \rangle \\ &= \langle \nabla \bar{\psi}' - \bar{q}, \bar{q}' - \bar{q} \rangle + \frac{1}{2} (|\bar{q}'|^2 - |\bar{q}|^2 - |\bar{q}' - \bar{q}|^2) \\ &= \frac{1}{2} (|\bar{q}'|^2 - |\bar{q}|^2) + \langle \nabla \bar{\psi}' - \bar{q}, \bar{q}' - \bar{q} \rangle - \frac{1}{2} |\bar{q}' - \bar{q}|^2 \\ &= \frac{1}{2} (|\bar{q}'|^2 - |\bar{q}|^2 + |\bar{q}' - \bar{q}|^2) + \langle \nabla \bar{\psi}' - \bar{q}, \bar{q}' - \bar{q} \rangle - |\bar{q}' - \bar{q}|^2 \\ &= \frac{1}{2} (|\bar{q}'|^2 - |\bar{q}|^2 + |\bar{q}' - \bar{q}|^2) + \langle \nabla(\bar{\psi}' - \bar{q}) - (\bar{q}' - \bar{q}), \bar{q}' - \bar{q} \rangle \\ &= \frac{1}{2} (|\bar{q}'|^2 - |\bar{q}|^2 + |\bar{q}' - \bar{q}|^2) + \langle \nabla \bar{\psi}' - \bar{q}', \bar{q}' - \bar{q} \rangle. \end{aligned} \quad (2.4-15)$$

En réinjectant (2.4-15) dans (2.4-14), on obtient :

$$\begin{aligned} |\bar{\mu}|^2 - |\bar{\mu}'|^2 &\geq \left(r^2 (|\bar{q}'|^2 - |\bar{q}|^2) + r^2 |\bar{q}' - \bar{q}|^2 + 2r^2 \langle \nabla \bar{\psi}' - \bar{q}', \bar{q}' - \bar{q} \rangle \right) + r^2 |\nabla \bar{\psi}' - \bar{q}'|^2 \\ &\geq r^2 \left((|\bar{q}'|^2 - |\bar{q}|^2) + (|\bar{q}' - \bar{q}|^2 + 2 \langle \nabla \bar{\psi}' - \bar{q}', \bar{q}' - \bar{q} \rangle + |\nabla \bar{\psi}' - \bar{q}'|^2) \right) \\ &\geq r^2 \left((|\bar{q}'|^2 - |\bar{q}|^2) + |(\nabla \bar{\psi}' - \bar{q}') + (\bar{q}' - \bar{q})|^2 \right) \\ &\geq r^2 (|\bar{q}'|^2 - |\bar{q}|^2) + r^2 |\nabla \bar{\psi}' - \bar{q}'|^2, \end{aligned} \quad (2.4-16)$$

soit :

$$(|\bar{\mu}|^2 + r^2 |\bar{q}|^2) - (|\bar{\mu}'|^2 + r^2 |\bar{q}'|^2) \geq r^2 |\nabla \bar{\psi}' - \bar{q}'|^2. \quad (2.4-17)$$

Ceci implique en particulier :

$$|\bar{\mu}|^2 + r^2|\bar{q}|^2 \geq |\bar{\mu}'|^2 + r^2|\bar{q}'|^2$$

i.e.

$$\|\mathbf{B}(\mu_1, q_1) - \mathbf{B}(\mu_2, q_2)\|_H \leq \|(\mu_1, q_1) - (\mu_2, q_2)\|_H.$$

T est donc bien un opérateur non-expansif.

2. Démontrons maintenant que $\|x - x^*\|_H^2 - \|\mathbf{B}x - x^*\|_H^2 \geq \|\mathbf{B}x - x\|_H^2$.

Soit (μ^*, q^*) un point fixe de \mathbf{B} , et $x = (\mu, q) \in H$ défini par $(\mu', q') = \mathbf{B}(\mu, q) = \mathbf{B}x$. En reprenant l'équation (2.4-17) avec $(\bar{\psi}, \bar{q}, \bar{\mu}) = (\psi - \psi^*, q - q^*, \mu - \mu^*)$ et $(\bar{\psi}', \bar{q}', \bar{\mu}') = (\psi' - \psi^*, q' - q^*, \mu' - \mu^*)$, on obtient :

$$(|\mu - \mu^*|^2 + r^2|q - q^*|^2) - (|\mu' - \mu^*|^2 + r^2|q' - q^*|^2) \geq r^2|\nabla(\psi' - \psi^*) - (q - q^*)|^2. \quad (2.4-18)$$

Décomposons à présent le second membre :

$$\begin{aligned} r^2|\nabla(\psi' - \psi^*) - (q - q^*)|^2 &= r^2|\nabla\psi' - q|^2 \\ &= |(\mu' - \mu) + r(q' - q)|^2 \\ &= |\mu' - \mu|^2 + 2r\langle \mu' - \mu, q' - q \rangle + r^2|q' - q|^2, \end{aligned}$$

avec $\langle \mu' - \mu, q' - q \rangle \geq 0$ (d'après (2.4-12)). Ainsi, en procédant de la même manière que dans [37], on obtient :

$$(|\mu - \mu^*|^2 + r^2|q - q^*|^2) - (|\mu' - \mu^*|^2 + r^2|q' - q^*|^2) \geq |\mu' - \mu|^2 + r^2|q' - q|^2,$$

c'est à dire, de façon reformulée :

$$\|(\mu, q) - (\mu^*, q^*)\|_H^2 - \|(\mu', q') - (\mu^*, q^*)\|_H^2 \geq \|\mathbf{B}(\mu, q) - (\mu, q)\|_H^2.$$

□

Remarque 2.4-5. UNE VARIANTE AVEC UN OPÉRATEUR FERMEMENT NON-EXPANSIF

Soit $\tilde{H} = L^2((0, 1) \times \Omega)^{d+1}$ (muni de sa norme canonique) et soit $x \in \tilde{H}$, on pose $q_x = P_{\mathcal{P}}x$ (projection orthogonale L^2 i.e. point par point presque partout) sur le parabolöide \mathcal{P} , et $\mu_x = x - q_x$. Soit $\varphi_x \in H^1((0, 1) \times \Omega)/\mathbb{R}$ vérifiant

$$\forall h \in H^1((0, 1) \times \Omega), \quad G(h) + \langle \mu_x, \nabla_{t,x} h \rangle + r \langle \nabla_{t,x} \varphi_x - q_x, \nabla_{t,x} h \rangle = 0.$$

On définit alors l'opérateur $\tilde{\mathbf{B}}$:

$$\begin{aligned} \tilde{\mathbf{B}} : \tilde{H} &\longrightarrow \tilde{H} \\ x &\longmapsto \frac{1}{r}\mu_x + \nabla_{t,x}\varphi_x \end{aligned}$$

$\tilde{\mathbf{B}}$ est un opérateur **fermement non-expansif** et réalise les étapes (2.1-13), (2.1-14), puis (2.1-12) de l'algorithme de Benamou-Brenier. Les itérations successives de l'opérateur $\tilde{\mathbf{B}}$ et de l'opérateur \mathbf{B} représentent donc **exactement le même algorithme** : on a simplement ici un

décalage des étapes dans la formulation de l'algorithme.

On pourrait ainsi, de la même façon que précédemment, prouver la convergence faible des itérés de \mathbf{B} , ou une convergence forte en faisant appel à la variante du Théorème 2.3-8 (p. 66).

En résumé, il serait possible de reformuler l'algorithme (2.1-12)-(2.1-14) de la façon dont l'avons fait dans cette section, à l'aide d'un opérateur "plus fort" (fermement non-expansif) que l'opérateur \mathbf{B} de la définition 2.4-2 ("seulement" quasi-fermement non-expansif).

Nous ne sommes néanmoins pas parvenus à exploiter les propriétés supplémentaires apportées par un opérateur fermement non-expansif (par rapport à un quasi-fermement non-expansif) pour obtenir de meilleures propriétés de convergence. C'est pourquoi nous avons fait le choix reformuler le problème à l'aide de l'opérateur \mathbf{B} qui, bien que moins "puissant" que l'opérateur $\tilde{\mathbf{B}}$, est beaucoup plus simple à manipuler.

II.5 Conclusion du chapitre

En appliquant les résultats de la section II.3 à l'opérateur \mathbf{B} défini dans 2.4-2, nous sommes donc en mesure de démontrer la convergence faible- L^2 de l'algorithme de Benamou-Brenier vers un point fixe de \mathbf{B} (i.e. vers un point selle de \mathbf{L}). De même, par l'application du Théorème 2.3-8 (p. 66), nous pouvons aisément définir une version relaxée à convergence forte- L^2 de l'algorithme. Cette approche par la théorie des opérateurs non-expansifs a récemment été utilisée pour montrer la convergence, faible ou forte, d'autres algorithmes, en particulier les algorithmes de splitting-proximal [12].

Dans les deux chapitres suivants, nous nous attèlerons à prouver que l'ensemble $Fix(\mathbf{B})$ des points fixes de \mathbf{B} est **non vide**. En effet l'existence d'un point fixe pour \mathbf{B} est équivalent, d'après la Proposition 2.2-5 (p. 62), à l'existence d'un point selle au lagrangien \mathbf{L} (voir le Théorème 3.1-2 p. 76).

Parmi les problèmes encore ouverts (à notre connaissance), il reste en particulier la mise en place d'un réel critère de convergence, ou d'un moins un critère d'arrêt : autrement dit, de proposer un opérateur calculable et capable de déterminer si la distance L^2 (discrétisée ou non) entre une solution théorique et le résultat d'une étape de l'algorithme est inférieur à une erreur donnée. Dans le même ordre d'idée, il serait intéressant d'être en mesure de pouvoir établir une estimation théorique de la vitesse de convergence, et pouvoir être en mesure de prédire un temps de calcul relatif à une erreur d'approximation prédéfinie.

En ce qui concerne ces deux points, nous avons, principalement dans le cadre d'un stage effectué en collaboration avec Sébastien Riffaud, élève de deuxième année à l'ENSIMAG (dont le rapport est consultable [ici](#)), entrepris des expérimentations numériques empiriques. Nous avons par exemple cherché à comparer les vitesses de convergence (en se basant sur des solutions numériques lointaines) de l'algorithme de Benamou-Brenier d'origine (à convergence au moins faible) avec différentes versions relaxées de celui-ci (et donc à convergence forte). Ces expérimentations furent non-concluantes, et ne permettent pas à notre avis d'effectuer une spéculation à propos d'un comportement théorique, du fait de l'importance potentielle de l'erreur numérique liée à la discrétisation et à l'implémentation du problème et de l'algorithme.

Chapitre III

Préambule à l'étude de l'existence et des résultats d'unicité d'un point selle de L : régularité du champ de vitesse

Notes pour ce chapitre.

- Nous rappelons que $|\cdot|$ désigne la norme euclidienne sur \mathbb{R}^d ($d \in \mathbb{N}$).
- Sauf mention contraire, \mathbb{R}^d sera muni de la tribu de Lebesgue.
- Pour $f \in L^p(\omega)$ et $x \in \omega$, le terme " $f(x)$ " désignera, par abus de notation, une valeur $\tilde{f}(x)$ où la fonction \tilde{f} est une représentante de la classe d'équivalence f pour la mesure de Lebesgue (i.e. pour la relation d'équivalence basée sur l'égalité \mathcal{L}^d presque partout).
- Par abus de notation, pour toute fonction f définie sur \mathbb{R}^d , et tout sous-ensemble ω de \mathbb{R}^d , on désignera ici indifféremment f et $f|_\omega$ par la même notation f . Inversement, pour tout sous-ensemble ω de \mathbb{R}^d , et toute fonction f définie sur ω , on désignera indifféremment f et son prolongement par 0 sur \mathbb{R}^d par la même notation f .
- De même, pour toute mesure ν définie sur la tribu de Lebesgue \mathbb{R}^d , et tout sous-ensemble Lebesgue-mesurable ω de \mathbb{R}^d , on désignera ici indifféremment ν et $\nu|_\omega$ (où $\nu|_\omega$ est défini pour tout sous-ensemble Lebesgue-mesurable A de ω par $\nu|_\omega(A) = \nu(A)$) par la même notation ν . Inversement, pour tout sous-ensemble Lebesgue-mesurable ω de \mathbb{R}^d , et toute mesure ν définie sur la tribu de Lebesgue de ω , on désignera indifféremment ν et son prolongement ν_0 par 0 sur \mathbb{R}^d par la même notation ν : Le prolongement par 0 de ν est définie plus précisément pour tout sous-ensemble Lebesgue-mesurable A de \mathbb{R}^d par $\nu_0(A) = \nu(A \cap \omega)$.
- Par abus de notation, lorsqu'il n'y aura pas d'ambiguïté sur ϕ , nous désignerons par v le champ de vitesse v_ϕ défini par rapport à ϕ comme en (3.3-11), c'est à dire

$$v_\phi(t, x) = \frac{1}{t} \left(x - \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{x}{1-t} \right) \right),$$

pour tout $t \in]0, 1[$ et $x \in \mathbb{R}^d$.

- Dès lors que rien ne sera précisé, Ω désignera un ouvert convexe et borné de \mathbb{R}^d .
- Nous désignerons un intervalle entre a et b par la notation (a, b) lorsque l'ouverture et la fermeture de cet intervalle en a ou b est sans importance. Au contraire, lorsque ces bornes sont précisées, sauf erreur, cela signifie que ces bornes ont une importance. Par exemple lorsque l'on écrit $f \in L_{loc}^\infty([0, 1] \times \mathbb{R}^d)$, cela signifie en particulier que $f \in L^\infty([0, 1] \times K)$

pour tout compact K de \mathbb{R}^d : autrement dit que f est " L^∞ localement en espace mais pas en temps" (i.e. reste localement borné en $t = a$ et $t = b$).

Ces notes de chapitre seront également valable pour le chapitre suivant.

III.1 Introduction et notations

Le principal objectif du chapitre III et du début du chapitre IV est de construire directement un point selle de \mathbf{L} défini en (2.1-11)(p. 54). Définissons donc rigoureusement le cadre :

Notation 2. Soient ρ_0 et ρ_1 deux densités (i.e. positives et d'intégrale égale à 1) de $L^2(\mathbb{R}^d)$ à supports bornés, et soit Ω un ouvert convexe borné suffisamment régulier de \mathbb{R}^d tel que $\text{supp}(\rho_0) \cup \text{supp}(\rho_1) \subset \Omega$.

Pour tout $r > 0$, on note $\mathbf{L}^{ps}(\rho_0, \rho_1, \Omega)$ l'ensemble des points selles du Lagrangien \mathbf{L} éléments de $S = H^1((0, 1) \times \Omega) / \mathbb{R} \times L^2((0, 1) \times \Omega)^{d+1} \times L^2((0, 1) \times \Omega)^{d+1}$.

Rappelons que d'après la Proposition 2.2-1 (p. 56), un tel point selle $(\psi, q, \mu) \in S$ de \mathbf{L} est caractérisé par les trois propriétés :

Propriétés (I).

$$\left\{ \begin{array}{l} (\psi, q, \mu) \in S \text{ vérifie :} \\ (P1) \quad \forall q' \in \tilde{\mathcal{P}}, \quad \langle \mu, q' - q \rangle \leq 0, \\ (P2) \quad \forall h \in H^1((0, 1) \times \Omega), \quad G(h) + \langle \mu, \nabla_{t,x} h \rangle = 0, \\ (P3) \quad \nabla_{t,x} \psi = q. \end{array} \right.$$

où le parabolöide $\tilde{\mathcal{P}}$ est définie par

$$\tilde{\mathcal{P}} = \left\{ (a, b) \in L^2((0, 1) \times \Omega) \times L^2((0, 1) \times \Omega)^d, \quad a + \frac{|b|^2}{2} \leq 0 \right\},$$

et la forme linéaire G par $G(h) = \int h(0, \cdot) \rho_0 dx - \int h(1, \cdot) \rho_1 dx$, pour tout $h \in H^1((0, 1) \times \Omega)$.

En posant $\mu = (\rho, m)$ et $q = (a, b)$ (avec $a + \frac{|b|^2}{2} \leq 0$), on peut traduire ces propriétés de la façon suivante :

(P1) signifie que si $\mu(t, x)$ est non nul, alors celui-ci est orthogonal au parabolöide au point $q(t, x)$, c'est à dire **co-linéaire** au vecteur $(1, b(t, x))$. Ainsi que le traduit la Proposition 2.2-3 (p. 58), la propriété (P1) peut se traduire de la façon suivante :

(a) $\rho \geq 0$

(b) $m = \rho b$

(c) $\rho \left(a + \frac{|b|^2}{2} \right) = 0$

(P2) correspond à l'**équation de conservation de la masse** vérifiée par ρ et b (c'est à dire $\partial_t \rho + \text{div}_x(\rho v) = 0$, en prenant $v = b$) pour les densités initiales et finales ρ_0 et ρ_1 .

(P3) équivaut en dimension spatiale inférieure ou égale à 2 (et donc en dimension espace-temps inférieure ou égale à 3) à dire que $\text{rot}_{t,x}(q) = 0$ ([38] Théorème 2.9 p.31).

Sauf mention contraire, ρ_0 et ρ_1 désigneront usuellement dans ce chapitre, de même que dans le chapitre II, deux densités de $L^2(\mathbb{R}^d)$ à supports bornés, et Ω désignera un ouvert convexe borné suffisamment régulier de \mathbb{R}^d tel que $\text{supp}(\rho_0) \cup \text{supp}(\rho_1) \subset \Omega$. Soit le potentiel ϕ une fonction convexe vérifiant $\rho(t, \cdot)\mathcal{L}^d = \rho_1\mathcal{L}^d = \nabla\phi\#(\rho_0\mathcal{L}^d)$ (telle que décrite par le Théorème 1.1-7 p. 20).

Nous commençons par définir les grandeurs suivantes :

Définition 3.1-1.

Pour tout $t \in]0, 1[$, nous définissons :

1. Le déplacement caractéristique à l'instant t ,

$$X(t, \cdot) = (1-t)\text{id} + t\nabla_x\phi = \nabla_x\phi_t \quad \text{avec} \quad \phi_t = (1-t)\frac{|\cdot|^2}{2} + t\phi, \quad (3.1-1)$$

2. Le champ de vitesse v associé,

$$v(t, \cdot) = \frac{\text{id} - \nabla_x(\phi_t)^*}{t} \quad \text{avec} \quad \phi_t = (1-t)\frac{|\cdot|^2}{2} + t\phi, \quad (3.1-2)$$

où $(\phi_t)^* = (\phi_t)^*$ désigne la transformée de Legendre du potentiel ϕ_t (défini ci-dessus).

3. La densité transportée ρ ,

$$\rho(t, \cdot)\mathcal{L}^d = \rho_t\mathcal{L}^d = X(t, \cdot)\#(\rho_0\mathcal{L}^d). \quad (3.1-3)$$

Définissons également les deux séries de propriétés suivantes : celles-ci seront caractéristiques des champs de vitesse d'un transport optimal.

Pour les deux groupes de propriétés qui suivent ((II) et (III)) Ω désignera un ouvert de \mathbb{R}^d qui ne sera **pas forcément borné** !

On dit que v vérifie les propriétés (II) si et seulement si :

Propriétés (II).

1. Il existe $\psi \in W_{loc}^{1,\infty}([0, 1] \times \Omega)$ (i.e. $\psi \in W^{1,\infty}((0, 1) \times \omega)$, pour tout ouvert $\omega \subset \Omega$ borné) tel que $v = \nabla_x\psi$.
2. Le champ de vitesse v vérifie au sens des distributions l'équation de Burgers, à savoir la relation

$$\partial_t v + \frac{1}{2}\nabla_x|v|^2 = 0. \quad (3.1-4)$$

La construction d'un point selle de \mathbf{L} (qui sont caractérisés, rappelons-le, par les propriétés (I)) passera par la construction d'un champ de vitesse vérifiant les propriétés (II). De plus, nous montrerons dans le chapitre IV que ces propriétés sont en fait suffisantes pour caractériser un transport optimal dans L^2 .

Remarque 3.1-1. Les propriétés (II) traduisent les propriétés caractéristiques d'un transport optimal isotrope pour un coût quadratique :

— Le premier point (i.e. $v = \nabla_x\psi$) traduit la propriété de **non croisement des trajec-**

toires (rappelons qu'en dimension inférieure à 3 cette propriété équivaut à un rotationnel nul pour v).

— Le second point (l'équation de Burgers) traduit quant à lui la propriété de **déplacement en ligne droite**.

A la fin du chapitre IV (section IV.3), nous chercherons à donner un cadre dans lequel on peut rigoureusement caractériser un déplacement de masse de type transport optimal à partir de ces seules propriétés.

Remarque 3.1-2. D'après le Lemme 4.1-10 qui sera énoncé à la fin de la section IV.1 (p. 137), le groupement de propriétés (II) équivaut au groupement propriétés suivant :

1. $v \in L_{loc}^\infty([0, 1] \times \Omega)^d$ i.e. $v \in L^\infty((0, 1) \times \omega)^d$, pour tout ouvert $\omega \subset \Omega$ borné,
2. $\partial_t v + \frac{1}{2} \nabla_x |v|^2 = 0$ (au sens des distributions),
3. Il existe $\psi \in L_{loc}^1((0, 1) \times \Omega)$, tel que $v = \nabla_x \psi$,

lequel sera éventuellement plus simple à vérifier selon les cas.

On dit que v vérifie les propriétés (III) si et seulement si :

Propriétés (III).

1. v est localement lipschitzienne sur $]0, 1[\times \Omega$,
2. $\nabla_{t,x} v \in L^\infty((0, 1), L_{loc}^1(\Omega))$,
3. $\nabla_{t,x} v \in L^p((0, 1), L_{loc}^q(\Omega))$ avec $1/p + 1/q > 1$
(en particulier $v \in W_{loc}^{1,p}([0, 1] \times \Omega)$ avec $1 \leq p < 2$).

Nous montrerons que si un champ de vitesse v est associé à un transport optimal, alors celui-ci vérifiera les propriétés (III). En partant de la propriété :

Propriété (Γ_1). ϕ et sa transformée de Legendre ϕ^* sont convexes, continues et admettent un minimum sur \mathbb{R}^d ,

le principal résultat que nous chercherons à montrer dans cette première partie de chapitre sera le suivant :

Théorème 3.1-2. EXISTENCE D'UN POINT SELLE

Soient ρ_0, ρ_1 et Ω définis comme dans la notation 2 (p. 74).

Alors il existe $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ convexe vérifiant la propriété (Γ_1) tel que $\rho_1 \mathcal{L}^d = \nabla \phi \# (\rho_0 \mathcal{L}^d)$.

De plus, en définissant ρ sur $[0, 1] \times \mathbb{R}^d$ par

$$t \mapsto \rho(t, \cdot) = X(t, \cdot) \# \rho_0, \text{ avec } X(t, \cdot) = (1-t) \text{id} + t \nabla_x \phi,$$

et v pour tout $(t, x) \in [0, 1[\times \mathbb{R}^d$ par

$$v(t, x) = \frac{x - \nabla_x (\phi_t)^*(x)}{t},$$

où $(\phi_t)^*$ désigne la transformée de Legendre de $\phi_t = (1-t)|\cdot|^2 + t\phi$; alors en posant $\mu = (\rho, \rho v)$ et $q = (-(1/2)|v|^2, v)$, il existe $\psi \in W_{loc}^{1,\infty}([0, 1] \times \Omega)$ tel que $q = \nabla_{t,x} \psi$ et tel que (ψ, q, μ) (ou du moins sa restriction $]0, 1[\times \Omega$) soit un élément de $\mathbf{L}^{ps}(\rho_0, \rho_1, \Omega)$.

En outre v vérifie les propriétés (II) et (III).

Remarque 3.1-3. La dernière assertion du Théorème 3.1-2 annonce que le champ de vitesse v ainsi construit vérifiera sur \mathbb{R}^d les propriétés (II) et (III).

- Comme nous le verrons dans la section IV.3, le fait que v vérifie les propriétés (II) suffit à caractériser un transport optimal dans L^2 .
- En revanche, le fait que v vérifie les propriétés (III) est en lui-même un résultat à part de cette construction : ces propriétés, bien qu'intéressantes en soit, n'interviendront directement pas pour caractériser un champ de vitesse de transport optimal dans L^2 .

Cependant il sera fait appel à des propriétés très proches de celles-ci (et démontrées parallèlement) pour prouver les différents énoncés portant sur l'unicité de la composante (ρ, m) des points selles de \mathbf{L} , ainsi que ceux relatifs à la caractérisation d'un champ de vitesse de type transport optimal.

La preuve finale de ce résultat sera énoncée à la fin de la section IV.1, en compilant les résultats obtenus au cours de cette même section ainsi que de la section III.3.

Donnons-nous juste une idée de la démarche qui sera suivie : posons $\mu = (\rho, \rho v)$ et $q = (-\frac{1}{2}|v|^2, v)$. Pour construire un point selle il faudra donc que μ et q permettent de vérifier les propriétés (I) (p. 74).

Il nous suffira pour cela de construire ρ et v de tel façon à ce qu'ils vérifient le groupe de propriétés suivant.

Propriétés (I'_α) .

- $(P_\alpha 1')$ La densité ρ appartient à $L^2((0, 1) \times \Omega)$.
 - $(P_\alpha 2')$ Le champ de vitesse v vérifie les propriétés (II).
 - $(P_\alpha 3')$ $(\rho, \rho v)$ vérifie l'équation de conservation de la masse au sens des distributions, pour les conditions initiales et finales ρ_0 et ρ_1 et des conditions de Neumann homogènes aux bords : $\forall h \in C^\infty([0, 1] \times \Omega)$,
- $$\int_{]0,1[\times \Omega} (\partial_t h + v \cdot \nabla h) \rho + \int_{\Omega} h(0, x) \rho_0(x) dx - \int_{\Omega} h(1, x) \rho_1(x) dx = 0 \quad (3.1-5)$$

Nous verrons que le fait que ρ et v vérifient les propriétés (I'_α) suffit à construire (de la façon évoquée plus haut) un triplet (ψ, q, μ) vérifiant les propriétés (I). L'inverse n'est cependant pas vrai : un triplet (ψ, q, μ) vérifiant les propriétés (I) ne suffit à construire un couple densité-champ de vitesse (ρ, v) vérifiant les propriétés (I'_α) , et tel que $\mu = (\rho, \rho v)$ et $q = (-\frac{1}{2}|v|^2, v)$. En effet, la composante q pourrait éventuellement se "décoller" du bord du parabolöide $(a, b) \rightarrow a + (1/2)|b|^2 \leq 0$ en dehors du support de μ .

Commentaires 3.1-4 (A propos des propriétés (I'_α)).

1. Ainsi que cela est rappelé au début de ce chapitre (voir la notation 2), l'ouvert $\Omega \subset \mathbb{R}^d$ dans lequel on se place pour notre problème est convexe et borné. Ainsi, la propriété $\psi \in W_{loc}^{1,\infty}([0, 1] \times \Omega)$ (et donc $v \in L_{loc}^\infty([0, 1] \times \Omega)$) du groupement (II) (p. 75) devient plus précisément $\psi \in W^{1,\infty}((0, 1) \times \Omega)$ (et $v \in L_{loc}^\infty([0, 1] \times \Omega)$).
2. La propriété $(P_\alpha 1')$, à savoir $\rho \in L^2((0, 1) \times \Omega)$, sera établie par le Corollaire 4.1-4 (p. 131).

3. Les propriétés $(P_\alpha 1')$ et $(P_\alpha 2')$ assurent au point-selle d'être dans le bon espace, c'est à dire dans S (défini à la section II.2). En effet

- $q \in L^2((0, 1) \times \Omega)^{d+1} \Leftrightarrow v \in L^4((0, 1) \times \Omega)^d \subset L^\infty((0, 1) \times \Omega)^d$,
 - $\mu \in L^2((0, 1) \times \Omega)^{d+1} \Leftrightarrow \rho \in L^2((0, 1) \times \Omega)$ et $\rho v \in L^2((0, 1) \times \Omega)^d$,
 - pour le potentiel ψ , on a bien $W^{1,\infty}((0, 1) \times \Omega) \subset H^1((0, 1) \times \Omega)$.
- $\rho \in L^2((0, 1) \times \Omega)$: ceci sera établi par le Corollaire 4.1-4 (p. 131).

4. La propriété $(P_\alpha 2')$ implique la propriété $(P3)$ du groupement (I).

En effet, affirmer que q dérive d'un potentiel espace-temps revient à vérifier, pour une dimension $d \leq 2$, que $\text{rot}_{t,x}(q) = 0$ (avec rappelons le $q = (-\frac{1}{2}|v|^2, v)$) au sens des distributions (voir [38] Théorème 2.9 p.31), ce qui équivaut à

$$\begin{cases} \partial_t v + \frac{1}{2} \nabla |v|^2 = 0, \\ \text{rot}_x(v) = 0 \Leftrightarrow \exists \psi \in \mathcal{D}'((0, 1) \times \Omega), v = \nabla_x \psi \end{cases}$$

Or il est clair que la vitesse **dérive bien d'un potentiel en espace** au sens des distributions, à savoir :

$$v(t, \cdot) = \nabla_x \left(\frac{1}{t} \left(\frac{1}{2} |\cdot|^2 - (\phi_t)^* \right) \right)$$

(voir la définition de v p. 75), et que ce potentiel dont dérive v est bien un élément de $L^1_{loc}((0, 1) \times \Omega)$, ce qui d'après le Lemme 4.1-10 p. 137 (comme déjà évoqué dans la Remarque 3.1-2) équivaut à vérifier le groupement de propriétés (II) à condition que le champ v vérifie bien l'équation de Burgers

$$\partial_t v + \frac{1}{2} \nabla_x |v|^2 = 0$$

au sens des distributions.

Quand bien même la notion de rotationnel ne serait plus vraiment définie si $d > 2$ (i.e. dans un espace de dimension $d + 1 > 3$), le Lemme 4.1-10 nous autorise toujours l'implication de la condition $(P3)$ par la condition $(P_\alpha 2')$, et ceci quelle que soit la dimension d (à condition que l'on ait bien $v \in L^\infty((0, 1) \times \Omega)^d$). Notons que le rotationnel peut en fait se définir en dimension $d + 1 > 3$, mais celui-ci sera alors à valeur dans $d(d-1)/2$ (mais ici nous raisonnerons, par souci simplicité, directement avec la notion de dérivation d'un potentiel).

Ainsi, la principale difficulté pour montrer que v vérifie la propriété $(P_\alpha 2')$ (c'est à dire les propriétés (II)) sera de montrer que $v \in L^\infty((0, 1) \times \Omega)^d$ (ce qui sera établi par la Proposition 3.3-12 p. 93) et qu'elle vérifie l'équation de Burgers citée ci-dessus (établi par la Proposition 3.3-31 p. 115).

5. Pour finir, notons que la propriété $(P_\alpha 3')$ traduit la propriété $(P2)$ du groupement (I). En effet on peut facilement prolonger la relation pour $h \in H^1((0, 1) \times \Omega)$ une fois que celle-ci est établie pour $h \in C^\infty((0, 1) \times \Omega)$. La propriété $(P1)$ est quant à elle, comme précisé plus haut, naturellement vérifiée en posant $m = \rho v$.

Les principaux enjeux de la construction d'un point selle pourraient donc se re-reformuler de la façon suivante : nous allons chercher à construire un couple densité/champ de vitesse (ρ, v) vérifiant les propriétés :

Propriétés (I'_β).

($P_\beta 1'$) La densité ρ vérifie $\rho \in L^2((0, 1) \times \Omega)$.

($P_\beta 2'$) Le champ de vitesse v vérifie $v \in L^\infty((0, 1) \times \Omega)^d$.

($P_\beta 3'$) Le champ de vitesse v vérifie au sens des distributions l'équation de Burgers :

$$\partial_t v + \frac{1}{2} \nabla_x |v|^2 = 0$$

($P_\beta 4'$) (ρ, v) vérifie l'équation de conservation de la masse au sens des distributions, pour les conditions initiales et finales ρ_0 et ρ_1 et des conditions de Neumann homogènes aux bords : $\forall h \in C^\infty([0, 1] \times \bar{\Omega})$,

$$\int_{]0, 1[\times \Omega} (\partial_t h + v \cdot \nabla h) \rho + \int_{\Omega} h(0, x) \rho_0(x) dx - \int_{\Omega} h(1, x) \rho_1(x) dx = 0 \quad (3.1-6)$$

Note pour le lecteur. Ce chapitre est en quelques sorte un préambule au chapitre [IV](#) : il sera principalement consacré à l'étude d'un certain nombre de propriétés des champs de vitesse inhérents à un transport optimal, dont la plupart ne trouveront leur application que dans la preuve des différents résultats du chapitre [IV](#). Cette étude, assez complexe, contient des résultats très importants et très utiles de ce travail, mais dont la connaissance précise ne sera pas, nous le pensons, entièrement nécessaire pour aborder la suite de ce manuscrit. Il est donc possible, à partir d'ici, de poursuivre la lecture directement en début de chapitre [IV](#), quitte à revenir à ce chapitre ultérieurement.

Plan du chapitre

Nous détaillons à présent le plan de l'étude menée au sein de ce chapitre ainsi que les principaux résultats obtenus :

1. Nous commencerons dans la section [III.2](#) par montrer qu'il est possible de prendre ϕ telle que ϕ vérifie la propriété (Γ_1) , c'est à dire que ϕ et sa transformée de Legendre ϕ^* soient finies en tout point de \mathbb{R}^d , et admettent un minimum global. Pour toute la suite du chapitre, nous pourrons ainsi considérer que cette hypothèse (Γ_1) est toujours vérifiée par ϕ .
2. Dans la section [III.3.4](#), nous étudierons en détails les propriétés du champ de vitesse v (le champ de vitesse spécifique du transport optimal dans le cadre de notre problème).
 - Dans la sous-section [III.3](#), nous montrerons quelques propriétés essentielles par la suite sur la vitesse v . En particulier le fait que v vérifie la propriété $(P_\beta 2')$ et est localement lipschitzienne, ainsi que la relation

$$\partial_t X(t, x) = v(t, X(t, x)).$$

- Pour finir, dans les sous-sections [III.3.8](#), [III.3.9](#) et [III.3.10](#), nous effectuerons une étude plus ciblée de v et nous finirons de montrer la propriété $(P_\beta 3')$, c'est à dire que v vérifie l'équation de Burgers :

$$\partial_t v + \frac{1}{2} \nabla |v|^2 = 0$$

ainsi que des propriétés de régularité (et de régularisation) sur v qui se révéleront ensuite utiles pour démontrer l'unicité de la composante $\mu = (\rho, \rho v)$ des points selles de \mathbf{L} dans la section IV.2 (chapitre IV).

III.2 Propriétés du potentiel du plan de transport

Cette section a pour objectif de prouver qu'il est possible de se donner un potentiel de transport ϕ , tel que défini dans le Théorème 1.1-7 (p. 20), assez régulier et prolongeable à \mathbb{R}^d tout entier. Plus précisément, nous allons montrer que l'on peut considérer ϕ et sa transformée de Legendre ϕ^* finies et donc continues sur \mathbb{R}^d , et admettant un minimum global. Nous regrouperons ces dernières assertions sur ϕ sous la propriété (Γ_1) .

Un potentiel ϕ vérifie la propriété (Γ_1) si et seulement si :

Propriété (Γ_1) . ϕ et ϕ^* sont convexes, continues et admettent un minimum sur \mathbb{R}^d .

Définition 3.2-1. DOMAINE D'UNE FONCTION CONVEXE

Soit $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction convexe. On appelle domaine de f l'ensemble noté $\text{Dom}(f)$ des points en lesquels f est finie.

- $\text{Dom}(f)$ est un ensemble convexe,
- f est localement lipschitzienne et donc continue sur $\text{Dom}(f)$.

Définition 3.2-2.

Soient ρ_0 une densité de probabilité Lebesgue-mesurable sur \mathbb{R}^d et μ_1 une mesure de probabilité définie sur la tribu de Lebesgue de \mathbb{R}^d . Définissons l'ensemble $\Phi(\rho_0 \mathcal{L}^d, \mu_1)$ des fonctions $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ vérifiant la propriété (Γ_1) telles que $\nabla \phi \# (\rho_0 \mathcal{L}^d) = \mu_1$.

L'objet de cette section est la démonstration du résultat suivant, et de son Corollaire.

Proposition 3.2-3.

Soient ρ_0 une densité de probabilité Lebesgue-mesurable sur \mathbb{R}^d et μ_1 une mesure de probabilité définie sur la tribu de Lebesgue de \mathbb{R}^d . On suppose que ρ_0 et μ_1 sont à support bornés.

Soit ϕ une fonction convexe, propre et s.c.i. sur \mathbb{R}^d , vérifiant

- $\mu_1 = \nabla \phi \# (\rho_0 \mathcal{L}^d)$,
- $\overline{\nabla \phi(\text{supp}(\rho_0))} = \text{supp}(\mu_1)$,

Alors il existe une fonction $\bar{\phi}$ sur \mathbb{R}^d vérifiant la propriété (Γ_1) telle que

1. $\phi = \bar{\phi}$ et $\nabla \phi = \nabla \bar{\phi}$ ρ_0 -presque partout,
2. $\mu_1 = \nabla \phi \# (\rho_0 \mathcal{L}^d) = \nabla \bar{\phi} \# (\rho_0 \mathcal{L}^d)$,
3. $\int_{\mathbb{R}^d} |\nabla \bar{\phi}(x) - x|^2 \rho_0(x) dx = \int_{\mathbb{R}^d} |\nabla \phi(x) - x|^2 \rho_0(x) dx = \inf_{T \# \rho_0 = \mu_1} \int_{\mathbb{R}^d} |Tx - x|^2 \rho_0(x) dx$,
4. $\overline{\nabla \phi(\text{supp}(\rho_0))} = \overline{\nabla \bar{\phi}(\text{supp}(\rho_0))} = \text{supp}(\mu_1)$.

Note : Rappelons que le terme $\nabla\phi(\text{supp}(\rho_0))$ doit être lu comme $\nabla\phi(\text{supp}(\rho_0) \cap \Upsilon(\rho_0))$, avec :

$$\Upsilon(\rho_0) = \{x \in \text{supp}(\rho_0), \phi \text{ différentiable en } x\}. \quad (3.2-1)$$

Corollaire 3.2-4.

Soient ρ_0 une densité de probabilité Lebesgue-mesurable sur \mathbb{R}^d et μ_1 une mesure de probabilité définie sur la tribu de Lebesgue de Ω . On suppose que ρ_0 et μ_1 sont à support bornés dans \mathbb{R}^d .

Alors $\Phi(\rho_0\mathcal{L}^d, \mu_1)$ est *non vide*.

Preuve du Corollaire sous réserve de véracité de la Proposition 3.2-3 : D'après le Théorème 1.1-7 (p. 20), il existe une fonction convexe ϕ propre et s.c.i. sur \mathbb{R}^d , vérifiant $\mu_1 = (\nabla\phi)\#(\rho_0\mathcal{L}^d)$ et $\overline{\nabla\phi(\text{supp}(\rho_0))} = \text{supp}(\mu_1)$.

D'après la Proposition 3.2-3, il existe alors une fonction $\bar{\phi}$ sur \mathbb{R}^d vérifiant la propriété (Γ_1) telle que $\mu_1 = \nabla\bar{\phi}\#(\rho_0\mathcal{L}^d) = \nabla\bar{\phi}\#(\rho_0\mathcal{L}^d)$, d'où $\bar{\phi} \in \Phi(\rho_0\mathcal{L}^d, \mu_1)$. \square

La seule difficulté de cette section consistera donc à prouver la Proposition 3.2-3. Nous allons donc montrer que pour toute fonction ϕ convexe et s.c.i. sur \mathbb{R}^d vérifiant

$$\nabla\phi\#(\rho_0\mathcal{L}^d) = \mu_1 \quad (3.2-2)$$

et

$$\int_{\mathbb{R}^d} |\nabla\phi(x) - x|^2 \rho_0(x) dx = \inf_{T\#\rho_0=\mu_1} \int_{\mathbb{R}^d} |Tx - x|^2 \rho_0(x) dx, \quad (3.2-3)$$

il est possible de choisir $\bar{\phi}$ prolongeant la restriction de ϕ à $\text{supp}(\rho_0)$, et vérifiant de plus la propriété (Γ_1) . Pour cela, nous considérons l'une de ces fonctions convexes s.c.i. ϕ vérifiant (3.2-2) et (3.2-3) (dont l'existence est assurée par le Théorème 1.1-7) et nous construirons un prolongement en une fonction $\bar{\phi}$ – convexe, superlinéaire et finie sur \mathbb{R}^d (donc localement lipschitzienne) – de la restriction de ϕ à un sous-ensemble $\Theta(\rho_0)$ de son domaine contenant $\text{supp}(\rho_0)$, de telle façon qu'en tout point x de $\text{supp}(\rho_0)$ où ϕ est différentiable, $\bar{\phi}$ le soit aussi et $\nabla\phi(x) = \nabla\bar{\phi}(x)$.

Pour assurer cette dernière condition, il suffit que $\Theta(\rho_0)$ contienne un voisinage de chaque point de $\text{supp}(\rho_0)$ où ϕ est différentiable, i.e. que $\bar{\phi}(z) = \phi(z)$ au voisinage de chacun de ces points. Nous verrons que pour prolonger ϕ à partir d'un tel ensemble $\Theta(\rho_0)$, il suffit qu'il existe une constante M qui borne au moins un élément de la sous-différentielle de ϕ en chacun des points de $\Theta(\rho_0)$.

Commençons donc par énoncer deux Lemmes généraux afin de montrer que la sous-différentielle de ϕ peut être uniformément bornée au voisinage de chaque point de $\text{supp}(\rho_0)$ où ϕ est différentiable ainsi qu'un troisième donnant les conditions à vérifier sur ϕ pour que ϕ^* soit superlinéaire. Leurs preuves sont établies dans l'annexe C.

Lemme 3.2-5.

Soit f une fonction convexe sur \mathbb{R}^d différentiable en un point $x \in \mathbb{R}^d$. Alors il existe $\epsilon > 0$ tel que f soit lipschitzienne sur $B(x, \epsilon)$, avec pour constante de lipschitz $L = 4(|\nabla f(x)| + 1)$.

Remarque 3.2-5. Si K est un compact de \mathbb{R}^d , alors d'après le Lemme 3.2-5 (p. 81), pour tout $x \in K$ il existe $\epsilon_x > 0$ tel que f soit lipschitzienne sur $B(x, \epsilon)$. Par extraction d'un recouvrement fini de K par de telles boules, on conclut donc que f est lipschitzienne sur K , et donc localement lipschitzienne sur \mathbb{R}^d .

Lemme 3.2-6.

Soit f une fonction convexe sur \mathbb{R}^d différentiable en un point $x \in \mathbb{R}^d$. Alors il existe $\delta > 0$ tel que pour tout $y \in B(x, \delta)$, et pour tout $\alpha \in \partial f(y)$, $|\alpha| \leq 4(|\nabla f(x)| + 1)$.

Lemme 3.2-7.

Si f est convexe, continue et superlinéaire sur \mathbb{R}^d tel qu'il existe $a, b, c \geq 0$ avec $a < 1$ tels que

$$\forall x \in \mathbb{R}^d, f(x) \leq a|x|^2 + b|x| + c,$$

alors f^* est convexe, continue et superlinéaire sur \mathbb{R}^d (et admet un minimum sur \mathbb{R}^d par coercivité). Autrement dit, f vérifie l'hypothèse (Γ_1) .

En vertu du Lemme 3.2-6, pour tout $x \in \Upsilon(\rho_0)$, soit $1 \geq \delta_x > 0$ tel que pour tout $y \in B(x, \delta_x)$, et pour tout $\alpha \in \partial f(y)$, on ait $|\alpha| \leq 4(|\nabla f(x)| + 1)$. On définit alors l'ouvert

$$\Theta(\rho_0) = \bigcup_{x \in \Upsilon(\rho_0)} B(x, \delta_x) \subset \text{Dom}(\phi). \quad (3.2-4)$$

Nous allons maintenant définir $\bar{\phi}$, fonction convexe superlinéaire finie sur \mathbb{R}^d , prolongeant la restriction de ϕ à $\Theta(\rho_0)$.

Pour ce faire, nous faisons de plus l'hypothèse que $\nabla\phi$ est bornée sur $\Upsilon(\rho_0)$, ce qui est effectivement vérifié par le potentiel ϕ de notre plan de transport optimal. En effet d'après le Théorème 1.1-7 (p. 20), ϕ vérifie en particulier (3.2-2) et (3.2-3), ainsi que $\overline{\nabla\phi(\text{supp}(\rho_0))} = \text{supp}(\mu_1)$ (et donc $\sup_{z \in \Upsilon(\rho_0)} |\nabla\phi(z)| < +\infty$, comme déjà précisé plus haut).

D'après le Lemme ci-dessus, il sera donc possible de borner les sous-différentielles de ϕ au voisinage de chaque point de $\Upsilon(\rho_0)$ par $4(\sup_{z \in \Upsilon(\rho_0)} |\nabla\phi(z)| + 1)$. Montrons qu'il en est de même pour au moins un élément de la sous-différentielle de chacun des points de $\text{supp}(\rho_0)$.

Proposition 3.2-8.

Soient ρ_0 une densité de probabilité Lebesgue-mesurable sur \mathbb{R}^d et μ_1 une mesure de probabilité définie sur la tribu de Lebesgue de \mathbb{R}^d . On suppose que ρ_0 et μ_1 sont à support bornés.

Soit ϕ une fonction convexe s.c.i sur \mathbb{R}^d , vérifiant

- $\mu_1 = (\nabla\phi)\#\rho_0$,
- $\overline{\nabla\phi(\text{supp}(\rho_0))} = \text{supp}(\mu_1)$,

et $\Theta(\rho_0)$ définie comme en (3.2-4).

Alors il existe une fonction $\bar{\phi}$ convexe, superlinéaire et finie en tout point de \mathbb{R}^d , tel que $\bar{\phi}|_{\Theta(\rho_0)} = \phi|_{\Theta(\rho_0)}$, et pour laquelle il existe $a, b, c \geq 0$ avec $a < 1$ tels que

$$\forall x \in \mathbb{R}^d, \bar{\phi}(x) \leq a|x|^2 + b|x| + c. \quad (3.2-5)$$

Remarque 3.2-6. Pour une fonction f de \mathbb{R}^d propre et convexe de sous-différentielle non partout vide, la condition de supralinéarité de cette dernière est suffisante pour affirmer que f^* est finie en tout point de \mathbb{R}^d (et donc continue sur \mathbb{R}^d).

Preuve de la Proposition 3.2-8 : Tout d'abord remarquons que $\nabla\phi$ est uniformément borné sur $\text{supp}(\rho_0)$. En effet, $\overline{\nabla\phi(\text{supp}(\rho_0))} = \text{supp}(\mu_1)$ (et $\text{supp}(\mu_1)$ est borné par hypothèse). Donc $\nabla\phi(\text{supp}(\rho_0)) \subset \text{supp}(\mu_1)$ est borné, i.e. $\sup_{x \in \Upsilon(\rho_0)} |\nabla\phi(x)| < +\infty$.

On pose alors $M = 4(N + 1) < +\infty$ avec $N = \sup_{z \in \Upsilon(\rho_0)} |\nabla\phi(z)|$. Il est clair d'après le Lemme 3.2-6, que pour tout $x \in \Theta(\rho_0)$, il existe $a \in \partial\phi(x)$ tel que $|a| \leq M$.

Soit $\Gamma = \{u = \langle \alpha, \cdot \rangle + \beta, (\alpha, \beta) \in \mathbb{R}^d \times \mathbb{R}, \text{ tel que } \forall z \in \mathbb{R}^d, \phi(z) \geq u(z), \text{ et } |\alpha| \leq M\}$, et soit $g : \mathbb{R}^d \rightarrow \mathbb{R}$ définie par $g(x) = \sup_{u \in \Gamma} (u(x))$ pour tout $x \in \mathbb{R}^d$.

Montrons que g est convexe, finie en tout point de \mathbb{R}^d , et telle que $g|_{\Theta(\rho_0)} = \phi|_{\Theta(\rho_0)}$.

1. g convexe.

Soient $x, y \in \mathbb{R}^d$ et $\lambda \in [0, 1]$. Soit $u \in \Gamma$, $u(x) \leq g(x)$ et $u(y) \leq g(y)$. Ainsi

$$\lambda g(x) + (1 - \lambda)g(y) \geq \lambda u(x) + (1 - \lambda)u(y) = u(\lambda x + (1 - \lambda)y),$$

(en effet u est affine), et ceci pour tout $u \in \Gamma$, donc

$$g(\lambda x + (1 - \lambda)y) = \sup_{u \in \Gamma} (u(\lambda x + (1 - \lambda)y)) \leq \lambda g(x) + (1 - \lambda)g(y),$$

et g est donc convexe sur \mathbb{R}^d .

2. g finie sur \mathbb{R}^d .

Soit $x_0 \in \text{supp}(\rho_0)$ tel que $\phi(x_0) < +\infty$. Pour tout $x \in \mathbb{R}^d$ et pour tout $u = \langle \alpha, \cdot \rangle + \beta \in \Gamma$, on a

$$u(x) = \langle \alpha, x - x_0 \rangle + u(x_0) \leq |\alpha| |x - x_0| + \phi(x_0) \leq M |x - x_0| + \phi(x_0) < +\infty,$$

et donc

$$g(x) = \sup_{u \in \Gamma} (u(x)) \leq M |x - x_0| + \phi(x_0) < +\infty. \quad (3.2-6)$$

3. $g|_{\Theta(\rho_0)} = \phi|_{\Theta(\rho_0)}$.

Il est tout d'abord clair que pour tout $y \in \mathbb{R}^d$, $g(y) \leq \phi(y)$.

Soit maintenant $x \in \Theta(\rho_0)$. Comme déjà mentionné en début de preuve, pour tout $x \in \Theta(\rho_0)$, il existe $a \in \partial\phi(x)$ tel que $|a| \leq M$. Considérons donc un tel $a \in \partial\phi(x)$. On pose $\bar{u} = \langle a, \cdot - x \rangle + \phi(x)$. Comme $a \in \partial\phi(x)$, on a bien :

$$\forall z \in \mathbb{R}^d, \phi(z) \geq \langle a, z - x \rangle + \phi(x) = \bar{u}(z),$$

et donc $\bar{u} \in \Gamma$. Or $\bar{u}(x) = \phi(x)$, d'où $g(x) = \sup_{u \in \Gamma} (u(x)) = \phi(x)$ (car $g(x) \leq \phi(x)$).

Rappelons que $\Theta(\rho_0) = \bigcup_{x \in \Upsilon(\rho_0)} B(x, \delta_x)$, en ayant pris $1 \geq \delta_x > 0$. $\Theta(\rho_0)$ est borné du fait que $\text{supp}(\rho_0)$ soit borné. Soient $R > 0$ tel que $\Theta(\rho_0) \subset B(0, R)$ et $0 \leq a < 1$. On pose :

$$\bar{\phi} = g + \max(0, a| \cdot |^2 - R^2) \quad (3.2-7)$$

$\bar{\phi}$ est bien convexe, superlinéaire, finie en tout point de \mathbb{R}^d , et vérifie $\bar{\phi}|_{\Theta(\rho_0)} = \phi|_{\Theta(\rho_0)}$. De plus d'après l'inéquation (3.2-6), on a pour tout $x \in \mathbb{R}^d$,

$$\bar{\phi}(x) \leq a|x|^2 + M|x| + (M|x_0| + \phi(x_0)).$$

□

Nous sommes maintenant en mesure de prouver le résultat principal de cette section, à savoir, rappelons-le, la Proposition 3.2-3 (p. 80).

Retour à la preuve de la Proposition 3.2-3 : Soient ρ_0 une densité de probabilité Lebesgue-mesurable sur \mathbb{R}^d et μ_1 une mesure de probabilité définie sur la tribu de Lebesgue de \mathbb{R}^d telles que ρ_0 et μ_1 soient à support bornés, et soit ϕ une fonction convexe, propre et s.c.i. sur \mathbb{R}^d , vérifiant $\mu_1 = (\nabla\phi)\#(\rho_0\mathcal{L}^d)$ et $\overline{\nabla\phi(\text{supp}(\rho_0))} = \text{supp}(\mu_1)$. D'après la Proposition 3.2-8, il existe une fonction $\bar{\phi}$ convexe, superlinéaire et finie en tout point de \mathbb{R}^d , tel que $\bar{\phi}|_{\Theta(\rho_0)} = \phi|_{\Theta(\rho_0)}$, et pour laquelle il existe $a, b, c \geq 0$ avec $a < 1$ tels que

$$\forall x \in \mathbb{R}^d, \bar{\phi}(x) \leq a|x|^2 + b|x| + c. \quad (3.2-8)$$

1. Montrons tout d'abord que $\bar{\phi}$ vérifie bien la propriété (Γ_1) : $\bar{\phi}$ est convexe et finie en tout point de \mathbb{R}^d , i.e. $\text{Dom}(\bar{\phi}) = \mathbb{R}^d$: d'après la définition 3.2-1, $\bar{\phi}$ est alors localement lipschitzienne et donc continue sur \mathbb{R}^d .

$\bar{\phi}$ est donc convexe, continue et superlinéaire sur \mathbb{R}^d , et vérifie la relation 3.2-8. D'après le Lemme 3.2-7, $\bar{\phi}^*$ est donc également convexe, continue et superlinéaire sur \mathbb{R}^d . De plus, la continuité et la supralinéarité de $\bar{\phi}$ et $\bar{\phi}^*$ sur \mathbb{R}^d leur assure à toutes les deux l'existence d'un minimum sur \mathbb{R}^d . $\bar{\phi}$ vérifie donc bien la propriété (Γ_1) .

2. Montrons maintenant les relations entre ϕ et $\bar{\phi}$. $\bar{\phi}$ est égale à ϕ au voisinage de tout point x de $\Upsilon(\rho_0)$ ($\bar{\phi}(z) = \phi(z)$ pour $z \in B(x, \delta_x)$) : il donc est clair que $\nabla\phi(x) = \nabla\bar{\phi}(x)$ en chacun de ces points. Pour toute fonction continue à support borné f de \mathbb{R}^d dans \mathbb{R} , on a

$$\int_{\mathbb{R}^d} f d\mu_1 = \int_{\mathbb{R}^d} (f \circ \nabla\phi)\rho_0 = \int_{\mathbb{R}^d} (f \circ \nabla\bar{\phi})\rho_0 dx \quad (3.2-9)$$

(pour plus de précision sur la notion de poussée de mesure, voir l'Annexe F, plus particulièrement la Proposition F.0-3), i.e. $\mu_1 = \nabla\phi\#(\rho_0\mathcal{L}^d) = \nabla\bar{\phi}\#(\rho_0\mathcal{L}^d)$ et

$$\begin{aligned} \int_{\mathbb{R}^d} |\nabla\bar{\phi}(x) - x|^2 \rho_0(x) dx &= \int_{\Upsilon(\rho_0)} |\nabla\bar{\phi}(x) - x|^2 \rho_0(x) dx \\ &= \int_{\Upsilon(\rho_0)} |\nabla\phi(x) - x|^2 \rho_0(x) dx \\ &= \int_{\mathbb{R}^d} |\nabla\phi(x) - x|^2 \rho_0(x) dx = \inf_{T\#\rho_0=\mu_1} \int_{\mathbb{R}^d} |Tx - x|^2 \rho_0(x) dx \end{aligned} \quad (3.2-10)$$

et bien sûr $\overline{\nabla\bar{\phi}(\text{supp}(\rho_0))} = \overline{\nabla\phi(\text{supp}(\rho_0))} = \text{supp}(\mu_1)$. □

Remarque 3.2-7. Dans la preuve de la Proposition 3.2-8, nous avons construit $\bar{\phi}$ en prenant

$$\bar{\phi} = g + \max(0, a|\cdot|^2 - R^2),$$

(voir l'équation (3.2-7) p. 83), avec $0 \leq a < 1$ quelconque.

Ainsi, d'après la Proposition 3.2-3 et son Corollaire 3.2-4 (p. 80), pour tout $a \in [0, 1[$, on peut construire un $\bar{\phi}_a = g + \max(0, a| \cdot |^2 - R^2) \in \Phi(\rho_0 \mathcal{L}^d, \mu_1)$.

De plus, pour tous $0 \leq a_1 < a_2 < 1$, $\bar{\phi}_{a_1}$ et $\bar{\phi}_{a_2}$ diffèrent d'une fonction non constante : en effet

$$\bar{\phi}_{a_2} - \bar{\phi}_{a_1} = \max(0, a_2| \cdot |^2 - R^2) - \max(0, a_1| \cdot |^2 - R^2) \text{ (non constante).}$$

$\Phi(\rho_0 \mathcal{L}^d, \mu_1)$ possède donc une infinité d'éléments différents deux à deux d'une fonction non constante.

Cette Remarque, qui semble pour l'instant très anodine, trouvera son importance lorsque sera abordé les questions d'unicité des points selles au début de la section IV.2 du prochain chapitre : elle permettra en effet d'affirmer la non-unicité des points selles de \mathbf{L} en général, ce qui à défaut ramènera le problème uniquement à l'étude de l'unicité de la composante $\mu = (\rho, m)$ commune aux points selles (ψ, q, μ) de \mathbf{L} .

Remarque 3.2-8. Il faut tout de même noter que la fonction ϕ donnée par le Théorème de Brenier (Théorème 1.1-7 p. 20), nous donne par ϕ^* le transport optimal inverse (transportant μ_1 sur ρ_0) dans le cas où μ_1 est une mesure à densité. Plus particulièrement, pour ρ_0 -presque tout $x \in \mathbb{R}^d$ et μ_1 -presque tout $y \in \mathbb{R}^d$:

$$\nabla \phi^* \circ \nabla \phi(x) = x, \quad \nabla \phi \circ \nabla \phi^*(y) = y,$$

Ce n'est plus forcément le cas pour $\bar{\phi}$: en passant de ϕ à $\bar{\phi}$, on gagne la continuité sur \mathbb{R}^d pour $\bar{\phi}$ et $\bar{\phi}^*$, mais il semble qu'on perde la réciprocity du transport assurée par ϕ^* .

Dans tout ce qui suivra, ϕ désignera désormais $\bar{\phi}$; c'est à dire que désormais, nous pourrions considérer que ϕ vérifie la propriété (Γ_1) .

⚠ IMPORTANT : Dans la plupart des énoncés qui suivront, nous nous placerons dans un cadre général (éventuellement hors "transport optimal") en considérant des fonctions ϕ vérifiant la propriété (Γ_1) .

L'étude effectuée de cette section, formalisée par la Proposition 3.2-3 (p. 80) et son Corollaire, montre néanmoins que dans le cadre du transport optimal (c'est à dire sous les conditions du Théorème 1.1-7 p. 20), **il ne s'agit pas d'une hypothèse mais bien d'une propriété de ce transport !**

Nous avons dans cette section, donné un cadre plus "confortable" à notre étude, en nous permettant de prolonger systématiquement le plan de transport (défini alors comme le gradient d'un potentiel convexe vérifiant la propriété de régularité (Γ_1)) à \mathbb{R}^d tout entier, de manière à assurer un certain nombre de propriétés de régularité au champ de vitesse.

La prochaine section se donne, pour objectif d'étudier en détail ces propriétés de régularité du champ de vitesse prolongé à \mathbb{R}^d obtenues grâce à la propriété (Γ_1) .

III.3 Propriétés du champ de vitesse

Dans cette section nous définirons et caractériserons un champ de vitesse associé à un transport optimal sur la base du Théorème 1.1-7 (p. 20). Les résultats qui seront démontrés dans cette section constitueront **la base** des résultats d'existence et d'unicité concernant les points selles de \mathbf{L} , ainsi que des résultats généralisés de la section IV.3.

Les trois premières sous-sections seront consacrés à la définitions d'outils d'analyse convexe auxquels nous feront régulièrement appel tout au long de cette section.

III.3.1 Quelques propriétés utiles des sous-gradients de fonctions Lipschitziennes

Lemme 3.3-1.

Soit $f : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction L -lipschitzienne sur un ouvert ω . Alors

$$\sup_{x \in \omega} |\partial f(x)| \leq L.$$

Preuve : Soient $x \in \omega$ et $\xi \in \partial f(x)$. Soit $r > 0$ tel que $\overline{B}(x, r) \subset \omega$. On suppose que $\xi \neq 0$ (le cas $|\xi| = 0 \leq L$ est bien sûr évident). Par définition de ξ :

$$\forall z \in \omega, f(z) \geq f(x) + \langle \xi, z - x \rangle.$$

Ainsi pour tout $z \in \overline{B}(x, r)$,

$$Lr \geq L|z - x| \geq f(z) - f(x) \geq \langle \xi, z - x \rangle.$$

En posant $z = x + r\xi/|\xi| \in \overline{B}(x, r)$, on obtient : $Lr \geq r|\xi|$. D'où $|\xi| \leq L$, et ceci pour tout $x \in \omega$ et $\xi \in \partial f(x)$. \square

Lemme 3.3-2.

Soit $f : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction L -lipschitzienne sur un ouvert convexe ω . Alors

$$\sup_{x \in \mathcal{D}(\omega)} |\nabla f(x)| \leq L.$$

où $\mathcal{D}(\omega)$ est l'ensemble des points de ω en lesquels f est différentiable.

Preuve : Soit $x \in \mathcal{D}(\omega)$. Soit $r > 0$ tel que $\overline{B}(x, r) \subset \omega$. On suppose que $\nabla f(x) \neq 0$ (le cas $|\nabla f(x)| = 0 \leq L$ est évident). f est différentiable en x :

$$\forall z \in \omega, f(z) = f(x) + \langle \nabla f(x), z - x \rangle + o(z - x).$$

Ainsi pour tout $z \in \overline{B}(x, r)$,

$$Lr \geq L|z - x| \geq f(z) - f(x) \geq \langle \nabla f(x), z - x \rangle + o(z - x).$$

En posant $z = x + (r\nabla f(x))/|\nabla f(x)| \in \overline{B}(x, r)$, on obtient :

$$Lr \geq r|\nabla f(x)| + o(r)$$

d'où en divisant par r et en faisant tendre celui-ci vers 0, on obtient $|\nabla f(x)| \leq L$, et ceci pour tout $x \in \mathcal{D}(\omega)$. \square

Remarque 3.3-9. *Le Lemme 3.3-2 semble à première vue un sous-cas du Lemme 3.3-1. Cependant, si la différentielle d'une fonction convexe coïncide bien avec sa sous-différentielle, ce n'est plus forcément le cas pour une fonction non-convexe.*

Lemme 3.3-3.

Soit $f : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction convexe propre et s.c.i., et soient $(x_k)_k$ et $(\xi_k)_k$ deux suites de \mathbb{R}^d qui convergent dans \mathbb{R}^d respectivement vers x et ξ , et telles que

$$\forall k \in \mathbb{N}, \quad \xi_k \in \partial f(x_k).$$

Alors $\partial f(x)$ est non-vide et $\xi \in \partial f(x)$.

Preuve : Pour tout $z \in \mathbb{R}^d$, on a pour tout $k \in \mathbb{N}$,

$$f(z) \geq f(x_k) + \langle \xi_k, z - x_k \rangle.$$

Ainsi par passage à la limite inférieure on obtient pour tout $z \in \mathbb{R}^d$:

$$f(z) \geq \liminf_k (f(x_k) + \langle \xi_k, z - x_k \rangle) \geq f(x) + \langle \xi, z - x \rangle,$$

d'où $\xi \in \partial f(x)$. □

III.3.2 Rappels sur les conjuguées convexes

Nous rappelons ici la définition de la transformée de Legendre.

Définition 3.3-4. TRANSFORMÉE DE LEGENDRE

Soit f une fonction de \mathbb{R}^n dans $] - \infty, +\infty]$. On définit la transformée de Legendre de f , notée f^* par :

$$f^*(x) = \sup_{y \in \mathbb{R}^n} \{ \langle x, y \rangle - f(y) \}$$

Notons que f^* est convexe et semi-continue inférieurement.

Nous en profitons pour énoncer quelques propriétés de la transformée de Legendre qui seront utiles pour la suite :

Propriétés 3.3-5. ([4] chapitre 13)

1. $\forall \alpha > 0, (\alpha f)^* = \alpha f^*(\cdot/\alpha),$
2. $(\frac{1}{2} |\cdot|^2)^* = \frac{1}{2} |\cdot|^2,$
3. (Théorème de Fenchel-Moreau) si f est propre, s.c.i. et convexe, alors $f^{**} = (f^*)^* = f.$

III.3.3 Opérateur Proximal et enveloppe de Moreau :

Commençons par introduire la notion de de convolution infimum, ou **inf-convolution** :

Définition 3.3-6. INF-CONVOLUTION ([4] chapitre 12)

Soient f et g deux fonctions de \mathbb{R}^d dans $] -\infty, +\infty]$. L'inf-convolution de f et g , notée $f \square g$, est définie par

$$f \square g : \mathbb{R}^d \rightarrow] -\infty, +\infty] : x \mapsto \inf_{y \in \mathbb{R}^d} \{f(x-y) + g(y)\} \quad (3.3-1)$$

Dans la suite de notre problème, nous aurons besoins de la propriété suivante, en lien avec la transformée de Legendre ([4] chapitre 13) : pour toutes fonctions f et g de \mathbb{R}^d dans $] -\infty, +\infty]$, on a

$$(f \square g)^* = f^* + g^*. \quad (3.3-2)$$

Nous rappelons également la définition de l'opérateur proximal déjà évoqué au chapitre I :

Définition 3.3-7. L'OPÉRATEUR PROXIMAL ([4] chapitre 12)

Soit f une fonction de \mathbb{R}^n ($n \in \mathbb{N}^*$) dans \mathbb{R} propre, s.c.i. et convexe, et soit $x \in \mathbb{R}^n$. L'opérateur proximal de f en x , noté $\text{Prox}_f(x)$ est l'unique minimiseur de $f + \frac{1}{2}|x - \cdot|^2$ dans \mathbb{R}^n . Autrement dit :

$$\text{Prox}_f(x) = \underset{y \in \mathbb{R}^n}{\text{argmin}} \left(f(y) + \frac{1}{2}|y - x|^2 \right) \quad (3.3-3)$$

On peut caractériser l'opérateur proximal par la relation suivante :

$$y = \text{Prox}_f(x) \Leftrightarrow x - y \in \partial\phi(y). \quad (3.3-4)$$

Les opérateurs Prox_f et $\text{id} - \text{Prox}_f$ sont non-expansifs (1-lipschitziens).

Énonçons de nouveau l'**identité de Moreau** (reliant l'opérateur proximal de f avec celui de sa transformée de Legendre f^*) :

$$\text{Prox}_{\gamma f^*} = \text{id} - \gamma \text{Prox}_{f/\gamma}(\cdot/\gamma) \quad (3.3-5)$$

Remarque 3.3-10. On a ici défini l'opérateur sur \mathbb{R}^n , mais celui-ci peut parfaitement se définir sur des espaces plus généraux (les espaces de Hilbert par exemple), avec les mêmes propriétés. Voir par exemple la sous-section I.2.1 au chapitre I.

Définissons également l'enveloppe de Moreau :

Définition 3.3-8. L'ENVELOPPE DE MOREAU ([4] chapitre 12)

Soit $f : \mathbb{R}^d \rightarrow] -\infty, +\infty]$ convexe, s.c.i propre et soit $\gamma > 0$. L'enveloppe de Moreau de f de paramètre γ est définie par :

$$\gamma f = f \square \left(\frac{1}{2\gamma} |\cdot|^2 \right) \quad (3.3-6)$$

Par définition de l'opérateur proximal, on peut donc également caractériser γf , pour tout $x \in \mathbb{R}^d$ et $\gamma > 0$, par :

$$\gamma f(x) = f(\text{Prox}_{\gamma\phi}(x)) + \frac{1}{2\gamma} |x - \text{Prox}_{\gamma\phi}(x)|^2 \quad (3.3-7)$$

γf est convexe et Fréchet-différentiable sur \mathbb{R}^d , de gradient égal à :

$$\nabla(\gamma f) = \gamma^{-1}(\text{id} - \text{Prox}_{\gamma f}) = \text{Prox}_{\phi^*/\gamma}(\cdot/\gamma) \text{ (par (3.3-5)).} \quad (3.3-8)$$

| Cette dernière application (i.e. $\nabla(\gamma f)$) est donc γ^{-1} -lipschitzienne.

III.3.4 Définition et premières propriétés du champ de vitesse

Nous allons maintenant re-définir le champ de vitesse d'un transport optimal (précédemment défini en (3.1-2) p. 75) à partir d'un opérateur de type proximal : l'opérateur \mathbf{p} . Celui-ci nous permettra de traiter plus facilement les problèmes des "cassures" du champ de vitesse (qui ne sont pas forcément des discontinuités, voir l'exemple 3.3-11 p. 90). Celui-ci agit en effet de façon bijective sur les zones régulières, et au contraire referme les différentes "cassures".

Définition 3.3-9. L'OPÉRATEUR \mathbf{p}

Soit $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ vérifiant la propriété (Γ_1) (en particulier ϕ est convexe et continue en tout point de \mathbb{R}^d , et admet donc en chacun de ces points une sous-différentielle non-vide et compacte). On définit l'opérateur

$$\begin{aligned} \mathbf{p}_\phi : [0, 1[\times \mathbb{R}^d &\rightarrow \mathbb{R}^d \\ (t, x) &\mapsto \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{x}{1-t} \right) \quad (= \nabla_x(\phi_t)^* \text{ voir plus bas}) \end{aligned}$$

\mathbf{p}_ϕ vérifie les propriétés suivantes : pour tout $t \in [0, 1[$,

1. $\mathbf{p}_\phi(t, \cdot)$ est $1/(1-t)$ -lipschitzienne,
2. si $t \in]0, 1[$ ($t \neq 0$), en posant $\phi_t = (1-t)|\cdot|^2/2 + t\phi$, alors $(\phi_t)^*$ est de classe C^1 sur \mathbb{R}^d et on a :

$$\mathbf{p}_\phi(t, \cdot) = \nabla_x(\phi_t)^*, \quad (3.3-9)$$

3. $\mathbf{p}_\phi(t, \cdot)$ est surjective sur \mathbb{R}^d et pour tous $x, y \in \mathbb{R}^d$,

$$y = \mathbf{p}_\phi(t, x) \Leftrightarrow x \in (1-t)y + t\partial\phi(y). \quad (3.3-10)$$

4. $\mathbf{p}_\phi(0, \cdot) = \text{id}$.

5. la vitesse v introduite en (3.1-2) p. 75, peut être définie à partir de \mathbf{p}_ϕ par :

$$v_\phi(t, x) = \frac{x - \mathbf{p}_\phi(t, x)}{t}, \quad (3.3-11)$$

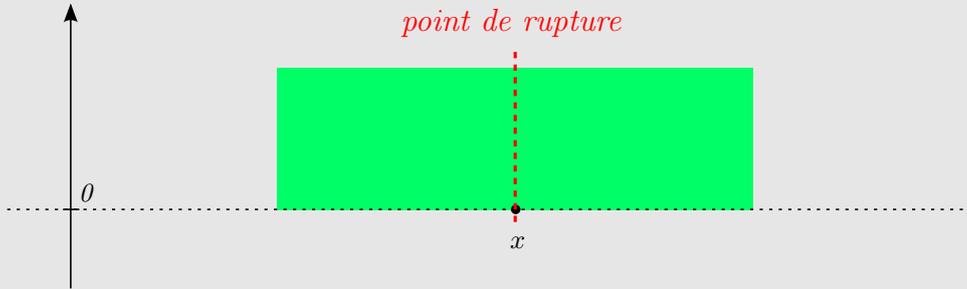
pour tous $t \in]0, 1[$ et $x \in \mathbb{R}^d$.

Note : Dans la suite de cette étude, lorsqu'il n'y aura pas d'ambiguïté sur ϕ , nous écrirons désignerons simplement par v le champ de vitesse v_ϕ .

L'illustration qui suit a pour objectif de donner une meilleure idée de ce que nous entendons par le terme de "cassure".

Exemple 3.3-11. LES "CASSURES" DU PLAN DE TRANSPORT

Nous illustrons ici deux type de "cassures" de la densité (en vert) au point x : un cas de discontinuité et un cas de non-discontinuité du plan de transport.



En effet, ce que nous appelons une "cassure" ne correspond pas forcément à une discontinuité du plan de transport, c'est à dire du fait que ϕ est une fonction convexe, à une non-différentiabilité de ϕ , comme dans la figure 1. La figure 2 illustre un cas de cassure où, bien que la fonction ϕ soit de classe C^1 , son gradient n'est pas localement lipschitzien au voisinage du point x : le plan de transport $T = \nabla\phi$ est continue en x , même si l'on y observe une "quasi-discontinuité" en ce point.

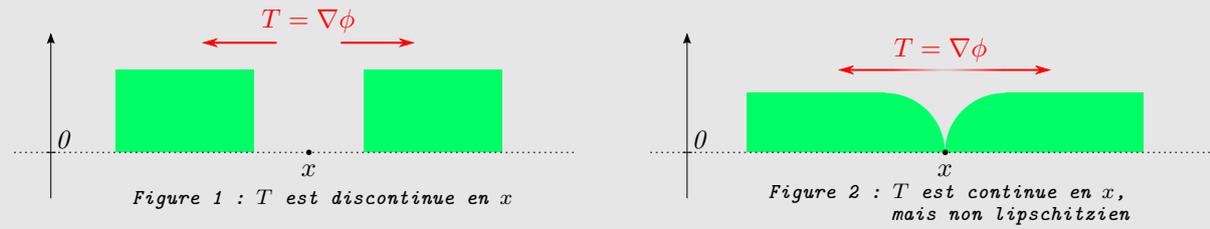


Figure 1 : T est discontinue en x

Figure 2 : T est continue en x , mais non lipschitzien

Preuve des propriétés énoncées dans la définition 3.3-9 :

1. La propriété se déduit immédiatement de la non-expansivité de l'opérateur proximal.
2. Les propriétés 3.3-5 nous permettent en particulier de montrer que $(\phi_t)^*$ peut être vu comme une enveloppe de Moreau, en effet :

$$\begin{aligned} \phi_t &= (1-t)\frac{1}{2}|\cdot|^2 + t\phi = \left((1-t)\frac{1}{2}|\cdot|^2 \right)^{**} + (t\phi)^{**} = \left(\left((1-t)\frac{1}{2}|\cdot|^2 \right)^* \square (t\phi)^* \right)^* \\ &= \left(\left((1-t)\frac{1}{2}|\cdot|^2 \right)^* \square (t\phi)^* \right)^* = \left(\frac{1}{2(1-t)}|\cdot|^2 \square (t\phi)^* \right)^* = (1-t)(t\phi)^* \end{aligned}$$

Ainsi $(\phi_t)^* = (1-t)(t\phi)^*$. En utilisant la relation (3.3-8) de la définition 3.3-8 (p. 88), on déduit donc que pour tout $t \in]0, 1[$, $(\phi_t)^*$ est de classe C^1 sur \mathbb{R}^d , et

$$\nabla_x(\phi_t)^* = \frac{\text{id} - \text{Prox}_{(1-t)(t\phi)^*}}{1-t} = \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{\cdot}{1-t} \right) = \mathbf{p}_\phi(t, \cdot). \quad (3.3-12)$$

3. La propriété se déduit immédiatement de la caractérisation (3.3-4) de l'opérateur proximal.
4. Se déduit de la précédente pour $t = 0$.
5. D'après la définition de la vitesse v (équation (3.1-2)), on a pour tout $x \in \mathbb{R}^d$:

$$v(t, x) = \frac{x - \nabla_x(\phi_t)^*(x)}{t} = \frac{x - \mathbf{p}_\phi(t, x)}{t}.$$

□

Commentaires 3.3-12 (A propos de l'opérateur \mathbf{p}_ϕ :).

1. Rappelons que le champ des trajectoires X est défini pour tout $t \in]0, 1[$ par (3.1-1) (p. 75), c'est à dire :

$$X(t, \cdot) = \nabla_x(\phi_t) = (1 - t) \text{id} + t \nabla \phi.$$

D'après l'équation (3.3-9), pour tout $t \in]0, 1[$, on a $\mathbf{p}_\phi(t, \cdot) = \nabla_x(\phi_t)^*$, c'est à dire qu'il représente formellement la "réciproque des tracés caractéristiques $X(t, \cdot) = \nabla_x(\phi_t)$ ". Cette réciproque est qualifiée de "Formelle" car, dans le cas général (i.e. sous l'hypothèse (Γ_1)), $\mathbf{p}_\phi(t, \cdot)$ n'est pas injective sur \mathbb{R}^d : au contraire l'opérateur $\mathbf{p}_\phi(t, \cdot)$ raccommode les "cassures" pouvant être générées par un plan de transport. En effet $\mathbf{p}_\phi(t, \cdot)$ re-concentre les zones générées par diffusion (par les trajectoires caractéristiques $X(t, \cdot)$) des points de cassures sur ces mêmes points.

Ainsi $\mathbf{p}_\phi(t, \cdot)$ ne peut être bijective que dans le cas où il n'y a pas de "cassures" dans le plan de transport.

Ces points de "cassures" du plan de transport correspondent aux éventuelles points où le potentiel ϕ n'est pas différentiables. Bien que d'après le Théorème 3.3-16 (p. 98) l'ensemble de ces points soit négligeable, la diffusion de ces cassures, et en particulier la torsion du champ de vitesse au niveau de ces points en $t = 0$ (ou bien $t = 1$ si l'on considère les points d'irrégularité de ϕ^*) elle ne l'est pas. En effet, les torsions du champ de vitesse au voisinage de ces points de cassures peut avoir par exemple pour conséquences **la non-régularité H^1 du champ de vitesse au niveau de ces point** (voir l'exemple 3.3-22 p. 122). Une régularité H^1 qui, si elle avait eu lieu, aurait grandement facilité l'étude abordée à la section IV.2 et traitant de résultats d'unicité dans les points selles du lagrangien \mathbf{L} . Malheureusement ce ne sera en général pas le cas...

2. On peut déduire de (3.3-11) qu'à $t \in]0, 1[$ fixé, le champ de vitesse $v(t, \cdot)$ est lipschitzien sur \mathbb{R}^d (d'après la première propriété de \mathbf{p}_ϕ énoncé ci-dessus), et qu'il est possible de lui associer une constante de Lipschitz ne dépendant que du temps t (ne dépend ni de l'espace, ni de ϕ). Par exemple, la constante de Lipschitz

$$L_t = \frac{2}{t(1-t)} \tag{3.3-13}$$

est toujours valable sur \mathbb{R}^d (pour la norme euclidienne $|\cdot|$), quelle que soit ϕ . Le champ de vitesse v est donc continue et presque partout Fréchet-différentiable en espace (par le Théorème de Rademacher [34] p.81, cité plus bas : Théorème 3.3-16 p. 98), et $\|\nabla_x v(t, x)\|$ est ainsi de plus **uniformément borné** par $L_t = 2/t(1-t)$ pour presque tout $x \in \mathbb{R}^d$ ($\|\cdot\|$ désigne ici la norme subordonnée à $|\cdot|$).

De par la nouvelle formulation de v dans le cinquième point de la définition 3.3-9, on déduit immédiatement la propriété suivante :

Proposition 3.3-10.

On suppose que ϕ vérifie la propriété (Γ_1) . Pour tout $t \in]0, 1[$, et pour tout $y \in \mathbb{R}^d$ tel que ϕ soit Fréchet-différentiable en y (c'est à dire pour presque tout $y \in \mathbb{R}^d$ étant donné que sous

$$\left| \begin{array}{l} \text{l'hypothèse } (\Gamma_1), \phi \text{ est localement lipschitzienne sur } \mathbb{R}^d, \text{ on a, en posant } X(t, \cdot) = \nabla \phi_t = \\ \nabla[(1-t)\text{id} + t\nabla\phi] : \\ \nabla\phi(y) - y = \partial_t X(t, y) = v(t, X(t, y)). \end{array} \right. \quad (3.3-14)$$

Preuve : Soit $y \in \mathbb{R}^d$ tel que ϕ soit différentiable en y (avec donc $\nabla\phi(y) = \{\nabla\phi(y)\}$). D'après (3.3-10), on a

$$y = \mathbf{p}_\phi(t, x) \Leftrightarrow x = (1-t)y + t\nabla\phi(y),$$

i.e.

$$y = \mathbf{p}_\phi(t, (1-t)y + t\nabla\phi(y)) = \mathbf{p}_\phi(t, X(t, y)). \quad (3.3-15)$$

On a donc d'après (3.3-11),

$$v(t, X(t, y)) = \frac{X(t, y) - \mathbf{p}_\phi(t, X(t, y))}{t} = \frac{X(t, y) - y}{t} = \nabla\phi(y) - y.$$

□

Remarque 3.3-13. La Proposition ci-dessus peut également se reformulé comme il suit : sachant que $X(t, \cdot) = \nabla\phi_t$ et $\mathbf{p}_\phi(t, \cdot) = \nabla_x(\phi_t)^*$ (voir le deuxième point de la définition 3.3-9) : pour tout $t \in]0, 1[$, et pour tout $x \in \mathbb{R}^d$ tel que ϕ soit Fréchet-différentiable en x , on a :

$$(\nabla_x(\phi_t)^* \circ \nabla\phi_t)(x) = x.$$

C'est la reformulation exacte de l'équation (3.3-15).

III.3.5 Contrôle du champ de vitesse

Dans cette sous-section, nous montrerons quelques propriétés de contrôle du champ v , introduit dans la relation (3.3-11) de la définition 3.3-9. En particulier, nous nous intéresserons au fait que v vérifie la propriété $(P_{\beta 2'})$ (i.e. est dans l'espace $L_{loc}^\infty([0, 1] \times \Omega)^d$).

Pour ne pas avoir à traiter indépendamment la question du contrôle de v en $t = 0$ et en $t = 1$, nous commençons par introduire une propriété de symétrie du champ v :

Lemme 3.3-11.

Soit v^* définie pour tout $t \in]0, 1[$ et $x \in \mathbb{R}^d$ par :

$$v^*(t, x) = \frac{x - \mathbf{p}_\phi^*(t, x)}{t}, \quad \text{avec} \quad \mathbf{p}_\phi^*(t, x) = \text{Prox}_{\frac{t}{1-t}\phi^*} \left(\frac{x}{1-t} \right),$$

Alors pour tout $t \in]0, 1[$ et $x \in \mathbb{R}^d$, on a

$$v(t, x) = -v^*(1-t, x). \quad (3.3-16)$$

Preuve : L'identité de Moreau (3.3-5) p. 88 nous donne :

$$\begin{aligned} \mathbf{p}_\phi(t, x) &= \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{x}{1-t} \right) = \frac{x}{1-t} - \frac{t}{1-t} \text{Prox}_{\frac{1-t}{t}\phi^*} \left(\frac{x}{t} \right) \\ &= \frac{x}{1-t} - \frac{t}{1-t} \mathbf{p}_{\phi^*}(1-t, x), \end{aligned}$$

et donc

$$v(t, x) = \frac{x - \mathbf{p}_{\phi^*}(1-t, x)}{t} = \frac{-x + \mathbf{p}_{\phi^*}(1-t, x)}{1-t} = -v^*(1-t, x). \quad (3.3-17)$$

□

On peut donc maintenant aborder le problème du contrôle du champ de vitesse, autrement dit la propriété $(P_\beta 2')$:

Proposition 3.3-12.

On suppose que ϕ vérifie la propriété (Γ_1) . Soit un champ de vitesse v défini par rapport à ϕ comme à l'équation (3.3-11) p. 89. Alors :

1. *On a $v \in L_{loc}^\infty([0, 1] \times \mathbb{R}^d)$.*

Plus précisément, pour tout $\omega \subset \mathbb{R}^d$ ouvert borné, si on pose

$$M = \sup_{x \in \omega} |\partial\phi(x)| \quad \text{et} \quad M^* = \sup_{x \in \omega} |\partial\phi^*(x)|,$$

on a

$$\sup_{(t,x) \in]0,1[\times \omega} |v(t, x)| \leq 5 (\max\{M, M^*\} + \text{sup}(\omega)). \quad (3.3-18)$$

2. *Il existe $\psi \in W_{loc}^{1,\infty}([0, 1] \times \Omega)$, tel que $v = \nabla_x \psi$.*

Par abus de notation, on désigne ici indifféremment v et $v|_\omega$ par la même notation v .

Remarque 3.3-14. *M et M^* sont bien finies : en effet ϕ est supposée vérifier la propriété (Γ_1) , et donc ϕ et ϕ^* sont supposées être finies et convexes sur \mathbb{R}^d et donc localement lipschitziennes (Lemme 3.2-5 p. 81), en particulier lipschitziennes sur ω . Ainsi, d'après le Lemme 3.3-1 (p. 86), $\partial\phi$ et $\partial\phi^*$ sont uniformément bornées sur ω .*

Remarquons également que si ϕ vérifie la propriété (Γ_1) , alors du fait que $(\phi^)^* = \phi$, ϕ^* vérifie également cette propriété.*

Preuve de la Proposition 3.3-12 : Soit $\omega \subset \mathbb{R}^d$ un ouvert borné de \mathbb{R}^d .

1. *$v(t, \cdot)$ est uniformément bornée sur ω au voisinage de $t = 0$.*

Nous allons montrer que $v(t, \cdot)$ est uniformément bornée sur ω sur $[0, 1/2]$.

Soient $t \in]0, 1/2]$ et $y \in \omega$, et soit $x \in (1-t)y + t\partial\phi(y)$. D'après (3.3-10) p. 89, on a $\mathbf{p}_\phi(t, x) = y$, donc

$$v(t, x) = \frac{x - y}{t} \in \partial\phi(y) - y. \quad (3.3-19)$$

Or $v(t, \cdot)$ est $2/t(1-t)$ -lipschitzienne (cf commentaire p. 91) sur \mathbb{R}^d , donc

$$\begin{aligned} |v(t, y) - v(t, x)| &\leq \frac{2}{t(1-t)} |x - y| \leq 4 \left| \frac{x - y}{t} \right| \quad (\text{car } t \in]0, 1/2]) \\ &= 4|v(t, x)|. \end{aligned} \quad (3.3-20)$$

En posant $M = \sup_{x \in \omega} |\partial\phi(x)|$, on a bien pour tout $t \in]0, 1/2]$ et $y \in \omega$,

$$|v(t, y)| \leq 5|v(t, x)| \leq 5(M + \sup(\omega)) \quad (\text{par (3.3-19)}).$$

2. $v(t, \cdot)$ est uniformément bornée sur ω au voisinage de $t = 1$.

Nous montrons maintenant que $v(t, \cdot)$ est uniformément bornée sur ω sur $[1/2, 1]$. Nous raisonnons par symétrie de ce qui précède.

Si on pose pour tout $t \in]0, 1[$ et $x \in \mathbb{R}^d$

$$v^*(t, x) = \frac{x - \mathbf{p}_\phi^*(t, x)}{t}, \quad \text{avec} \quad \mathbf{p}_\phi^*(t, x) = \text{Prox}_{\frac{t}{1-t}\phi^*} \left(\frac{x}{1-t} \right),$$

alors d'après le Lemme 3.3-11 (p. 92), on a

$$v(t, x) = -v^*(1-t, x).$$

En posant $M^* = \sup_{x \in \omega} |\partial\phi^*(x)|$, et en raisonnant comme précédemment, on obtient également, pour tout $t \in [1/2, 1[$ et $y \in \omega$,

$$|v(t, x)| = |v^*(1-t, x)| \leq 5(M^* + \sup(\omega)).$$

ce qui conclut la preuve du premier point.

3. Pour le second point, on rappelle que d'après le dernier point de la définition 3.3-9 (p. 89), on a :

$$v(t, \cdot) = \frac{\text{id} - \nabla_x(\phi_t)^*}{t} \quad \text{avec} \quad \phi_t = (1-t) \frac{|\cdot|^2}{2} + t\phi.$$

Ainsi, on a $v = \nabla_x \psi$ au sens des distributions, avec ψ défini pour tout $t \in]0, 1[$ et pour tout $x \in \mathbb{R}^d$ par :

$$\psi(t, x) = \frac{1}{t} \left(\frac{1}{2} |\cdot|^2 - (\phi_t)^* \right).$$

Or, on a démontré dans le premier point de cette Proposition que $v \in L^\infty([0, 1] \times \mathbb{R}^d)^d$. On en déduit que $\psi \in W_{loc}^{1, \infty}([0, 1] \times \mathbb{R}^d)$ par application du Lemme E.

□

Remarque 3.3-15. Comme nous le verrons plus loin dans la proposition 3.3-18 (p. 99), si pour tout $t \in]0, 1[$ les plans de transport interpolés $\nabla\phi_t$ avaient été inversibles, i.e. si ϕ avait été de classe C^1 de gradient lipschitzien, on aurait alors eu $\mathbf{p}_\phi(t, \cdot) = (\nabla\phi_t)^{-1} = \nabla(\phi_t)^*$, et ainsi d'après (3.3-11) $v(t, \cdot) = \nabla\phi((\nabla\phi_t)^{-1}) - (\nabla\phi_t)^{-1}$. Le champ v aurait alors été prolongeable par continuité en $t = 0$ et $t = 1$: le résultat de la proposition 3.3-12 aurait alors été évident.

Seulement, dans le cas général d'un transport pouvant induire un changement de topologie entre les supports des masses de départ et d'arrivée, c'est à dire admettant des "cassures" et donc des points de non-régularité pour le potentiel ϕ , un tel prolongement ne va pas de soit.

Pour revenir à notre problème de départ, on a bien

$$q = \left(-\frac{1}{2}|v|^2, v \right) \in L^\infty((0, 1) \times \Omega)^{d+1} \subset L^2((0, 1) \times \Omega)^{d+1}.$$

Remarque 3.3-16. Comme déjà notifié dans le commentaire 3.3-12 (p. 91, équation (3.3-13)), pour tout $t \in]0, 1[$, $v(t, \cdot)$ est continue et lipschitzienne sur \mathbb{R}^d (en munissant \mathbb{R}^d de la norme euclidienne, on peut prendre $2/t(1-t)$ comme constante de Lipschitz). Le champ $v(t, \cdot)$ est donc lipschitzien sur \mathbb{R}^d , pour une constante de Lipschitz indépendante de t sur tout intervalle $[\alpha, \beta] \subset]0, 1[$ (par exemple la constante $M_{\alpha, \beta} = \sup_{[\alpha, \beta]} 2/t(1-t)$). Il est donc possible d'appliquer le

Théorème de Cauchy-Lipschitz sur $[\alpha, \beta]$.

Donc pour tout $x \in \mathbb{R}^d$ et $t \in]0, 1[$, le problème de Cauchy

$$\begin{cases} y'_{t,x} = v(\cdot, y_{t,x}) \\ y_{t,x}(t) = x, \end{cases} \quad (3.3-21)$$

admet une unique solution maximale d'après le Théorème de Cauchy-Lipschitz sur tout intervalle $]\alpha, \beta[$, $0 < \alpha < t < \beta < 1$. On peut alors aisément établir qu'il existe une unique solution définie sur $]0, 1[$ et que celle-ci s'écrit $y_{t,x}(s) = (s-t)v(t, x) + x$ pour tout $s \in]0, 1[$. En effet une telle solution vérifie bien $y_{t,x}(t) = x$, et $y'_{t,x}(s) = v(t, x) = v(s, (s-t)v(t, x) + x) = v(s, y_{t,x}(s))$, comme démontré dans la Proposition qui suit.

Remarquons en outre que si l'on considère en particulier le problème de Cauchy :

$$\begin{cases} y' = v(\cdot, y) \\ y(t) = X(t, x) = (1-t)x + t\nabla\phi(x), \end{cases} \quad (3.3-22)$$

on a alors bien $y(s) = (s-t)v(t, X(t, x)) + X(t, x) = X(s, x)$ comme unique solution de (3.3-22) sur $]0, 1[$.

Nous finissons par une dernière propriété de contrôle sur v , de type point fixe :

Lemme 3.3-13.

On suppose que ϕ vérifie la propriété (Γ_1) . Soit $a \in \mathbb{R}^d$ tel que $\phi(a) = \inf_{\mathbb{R}^d} \phi$. Alors pour tout $t \in [0, 1[$,

1. $\mathbf{p}_\phi(t, (1-t)a) = a$,
2. $v(t, (1-t)a) = -a$.

On rappelle que

$$\mathbf{p}_\phi : (t, x) \in [0, 1[\times \mathbb{R}^d \longmapsto \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{x}{1-t} \right) = \nabla_x((1-t)|\cdot|^2/2 + t\phi)^*$$

(voir définition 3.3-9 p. 89).

Preuve :

1. Tout d'abord, ϕ étant une fonction convexe,

$$\phi(a) = \inf_{\mathbb{R}^d} \phi \Leftrightarrow 0 \in \partial\phi(a).$$

Par définition de \mathbf{p}_ϕ , en particulier par l'équivalence (3.3-10) p. 89, on a donc,

$$(1-t)a = (1-t)a + 0 \in (1-t)a + t\nabla\phi(a) \Leftrightarrow a = \mathbf{p}_\phi(t, (1-t)a).$$

2. Quant à v , on a

$$v(t, (1-t)a) = \frac{(1-t)a - \mathbf{p}_\phi(t, (1-t)a)}{t} = -a.$$

□

III.3.6 Régularité lipschitzienne en espace-temps du champ de vitesse

L'objectif de cette sous-section est de montrer que le champ v est localement lipschitzien sur $]0, 1[\times \mathbb{R}^d$, c'est à dire en **espace-temps** (et pas seulement en espace). Il s'agit du premier point des propriétés (III) sur $(0, 1) \times \mathbb{R}^d$, énoncées p. 76.

Proposition 3.3-14.

On suppose que ϕ vérifie la propriété (Γ_1) . Alors pour tous $t, s \in]0, 1[$, et pour tout $x \in \mathbb{R}^d$,

$$v(t, x) = v(s, (s-t)v(t, x) + x).$$

Preuve : On fait appel aux relations (3.3-10) et (3.3-11) p. 89 et on raisonne par équivalence : Pour tous $s, t \in]0, 1[$ et $x \in \mathbb{R}^d$,

$$\begin{aligned} v(s, (s-t)v(t, x) + x) &= v(t, x) \\ \Leftrightarrow \frac{(s-t)v(t, x) + x - \mathbf{p}_\phi(s, (s-t)v(t, x) + x)}{s} &= v(t, x) \text{ (par (3.3-11))} \\ \Leftrightarrow \mathbf{p}_\phi(s, (s-t)v(t, x) + x) &= tv(t, x) - x = \mathbf{p}_\phi(t, x) \\ \Leftrightarrow \mathbf{p}_\phi\left(s, \left(1 - \frac{s}{t}\right) \mathbf{p}_\phi(t, x) + \frac{s}{t}x\right) &= \mathbf{p}_\phi(t, x) \text{ (par (3.3-11))} \\ \Leftrightarrow \left(1 - \frac{s}{t}\right) \mathbf{p}_\phi(t, x) + \frac{s}{t}x &\in (1-s) \mathbf{p}_\phi(t, x) + s\partial\phi(\mathbf{p}_\phi(t, x)) \text{ (par (3.3-10))} \\ \Leftrightarrow \frac{s}{t}x \in s \left(\frac{1}{t} - 1\right) \mathbf{p}_\phi(t, x) + s\partial\phi(\mathbf{p}_\phi(t, x)) \\ \Leftrightarrow x \in (1-t) \mathbf{p}_\phi(t, x) + t\partial\phi(\mathbf{p}_\phi(t, x)) &\text{ (vrai par (3.3-10)).} \end{aligned}$$

□

Remarque 3.3-17. Dans la preuve qui précède, la propriété (Γ_1) n'est utile que pour les conditions sur ϕ , et non celles sur ϕ^* . En fait nous n'avons besoin ici que du fait que ϕ admette en tout point de \mathbb{R}^d une sous-différentielle non-vide et compacte.

Proposition 3.3-15.

On suppose que ϕ vérifie la propriété (Γ_1) . Alors v est localement lipschitzienne sur l'espace $]0, 1[\times \mathbb{R}^d$.

Preuve : On rappelle que pour tout $t \in]0, 1[$ (voir (3.3-13)), $v(t, \cdot)$ est lipschitzienne de constante de Lipschitz $2/t(1-t)$. Soient $0 < \alpha < \beta < 1$ et ω un ouvert borné de \mathbb{R}^d , et soient

$$M_{\alpha, \beta} = \sup_{[\alpha, \beta]} \frac{2}{t(1-t)} \quad \text{et} \quad M_v = \sup_{(t, x) \in]0, 1[\times \omega} |v(t, x)| < +\infty \quad (\text{Proposition 3.3-12 p. 93}).$$

Pour tout $(t_1, x_1), (t_2, x_2) \in [\alpha, \beta] \times \mathbb{R}^d$, on a d'après la Proposition 3.3-14,

$$\begin{aligned} |v(t_1, x_1) - v(t_2, x_2)| &= |v(t_2, (t_2 - t_1)v(t_1, x_1) + x_1) - v(t_2, x_2)| \\ &\leq \frac{2}{t_2(1-t_2)} |((t_2 - t_1)v(t_1, x_1) + x_1) - x_2| \\ &\leq M_{\alpha, \beta} (|t_2 - t_1| |v(t_1, x_1)| + |x_2 - x_1|) \\ &\leq M_{\alpha, \beta} (1 + M_v) (|t_2 - t_1| + |x_2 - x_1|). \end{aligned}$$

□

III.3.7 Préambule sur la régularité du champ de vitesse

L'une des assertions à vérifier pour pouvoir établir le Théorème 3.1-2, est de prouver que pour tout champ de vitesse v défini par

$$v(t, x) = \frac{x - \mathbf{p}_\phi(t, x)}{t} = \frac{1}{t} \left(x - \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{x}{1-t} \right) \right), \quad (3.3-23)$$

avec ϕ vérifiant (Γ_1) , s'accommode bien des propriétés (II) et (III).

On rappelle que le champ de vitesse v vérifie les propriétés (II) sur \mathbb{R}^d (définies p. 75) si et seulement si :

1. il existe $\psi \in W_{loc}^{1, \infty}([0, 1] \times \mathbb{R}^d)$ tel que $v = \nabla_x \psi$,
2. le champ de vitesse v vérifie l'équation de Burgers au sens des distributions, i.e.

$$\partial_t v + \frac{1}{2} \nabla_x |v|^2 = 0.$$

De même, v vérifie les propriétés (III) sur \mathbb{R}^d (définies p. 76) si et seulement si

1. v est localement lipschitzienne sur $]0, 1[\times \mathbb{R}^d$.
2. $\nabla_{t, x} v \in L^\infty((0, 1), L_{loc}^1(\mathbb{R}^d))$,
3. $\nabla_{t, x} v \in L^p((0, 1), L_{loc}^q(\mathbb{R}^d))$ avec $1/p + 1/q > 1$.

Le premier point des propriétés (II), a été en parti établi par la Proposition 3.3-12 (à savoir que $v \in L_{loc}^\infty([0, 1] \times \mathbb{R}^d)$), et sera totalement établie par le Lemme 4.1-10 (section IV.1).

Le premier des propriétés (III), a quant à lui été établi par la Proposition 3.3-15.

Nous allons en conséquence chercher à établir que v vérifie également la suite de ces propriétés, à savoir le second point des propriétés (II), ainsi que les deux derniers points des propriétés (III) (sous-section III.3.11).

Comme nous l'avons déjà évoqué dans la Remarque 3.1-3 (p. 77), les propriétés (III) se révèlent un peu "à part", et ne seront plus réellement utilisées dans le reste de notre étude. Cependant leur démonstration, ainsi que celle de l'équation de Burgers pour v , se feront en parallèle d'autres résultats concernant **le contrôle du gradient de champs de vitesses régularisés**, et permettront le contrôle des solutions du problème de transport engendré par le champ v . Ces résultats parallèles se révéleront alors importants dans les preuves (basées sur la méthode des caractéristiques) des principaux résultats de la sous-section IV.2.3 et de la section IV.3, relatifs à l'unicité du transport de la densité ρ et de la caractérisation du champ de vitesse v dans le cadre d'un transport optimal.

En outre les propriétés (III) illustrent bien une des difficultés relatives au champ v que nous aurons à contourner, à savoir que si celui-ci est infiniment proche d'une régularité H^1 sur $[0, 1] \times \mathbb{R}^d$, il ne l'est pour autant pas en général (voir le contre-exemple 3.3-22 p. 122), ce qui compliquera quelque peu notre étude au cours de la section IV.2 (voir le petit préambule p. 147).

Ainsi que nous l'avons déjà mentionné dans le commentaire 3.3-12, cette non-régularité H^1 est en partie dû aux éventuelles "cassures" dans le schéma de transport continu en $t = 0$ (ou au contraire à d'éventuels raccordements en $t = 1$). De telles cassures correspondent aux points de non-dérivabilité du potentiel ϕ et les raccordements en $t = 1$ sont quant à eux liés aux points de non-dérivabilité de la transformée de Legendre ϕ^* du potentiel ϕ . Cependant, du fait de la symétrie du problème, nous n'auront pas besoin de traiter les problèmes de raccordement en $t = 1$.

Nous allons donc commencer par montrer ces différentes propriétés dans un cadre "idéal", c'est à dire sans cassures : cela constituera le cadre de la sous-section III.3.8. Nous allons pour cela considérer dans un premier temps que le potentiel ϕ est assez régulier, plus précisément qu'il vérifie la propriété (Γ_2) que nous énonçons ci-dessous :

ϕ vérifie la propriété (Γ_2) si et seulement si :

Propriété (Γ_2) . ϕ vérifie la propriété (Γ_1) , est de classe C^1 , et $\nabla_x \phi$ est lipschitzien (i.e. $\phi \in C^{1,1}(\mathbb{R}^d)$).

On rappelle que ϕ vérifie la propriété (Γ_1) si et seulement si ϕ et ϕ^ sont convexes, continues et admettent un minimum sur \mathbb{R}^d .*

Une fois cela fait, nous établirons dans un deuxième temps ces propriétés dans le cadre général du transport optimal (c'est à dire sous pour ϕ vérifiant la propriété (Γ_1)), en se ramenant au cas régulier via une régularisation par enveloppe de Moreau du potentiel ϕ . Ce sera l'objet principalement de la sous-section III.3.9.

Pour cette prochaine étude, nous aurons souvent besoin de faire appel à deux résultats fondamentaux que nous énonçons ci-dessous : le Théorème de Rademacher (différentiabilité d'une fonction localement lipschitzienne) ainsi qu'un Théorème de changement de variable inspiré de la formule de co-aire (ou Théorème de changement de variable généralisé).

Théorème 3.3-16. THÉORÈME DE RADEMACHER ([34] p.81)

Soit $f : \mathbb{R}^d \rightarrow \mathbb{R}^m$ une fonction localement lipschitzienne. Alors f est \mathcal{L}^n -presque partout Fréchet-différentiable (et sa différentielle au sens de Fréchet coïncide avec sa différentielle au sens des distributions).

Donnons tout de suite un exemple d'application du Théorème de Rademacher (exemple qui sera utile à notre problème) :

Exemple 3.3-18. *On sait par la Proposition 3.3-15 valable sous la propriété (Γ_1) (et donc également sous la propriété (Γ_2)), que v est localement lipschitzienne sur $]0, 1[\times \mathbb{R}^d$. Ainsi d'après le Théorème de Rademacher 3.3-16, v est presque partout différentiable sur $]0, 1[\times \mathbb{R}^d$, et sa différentielle correspond à sa dérivée au sens des distributions. En particulier on observe que*

$$v \cdot \nabla_x v = \frac{1}{2} \nabla_x |v|^2, \quad (3.3-24)$$

au sens \mathcal{L}^{d+1} -presque partout, comme au sens des distributions.

Théorème 3.3-17. THÉORÈME DE CHANGEMENT DE VARIABLE GÉNÉRALISÉ ([34] p.117)

Soit $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$ une fonction bijective et lipschitzienne. Soit une fonction $g \in L^1(\mathbb{R}^d)$ tel que $g \circ f \in L^1(\mathbb{R}^d)$. Alors

$$\int_{\mathbb{R}^d} g(x) dx = \int_{\mathbb{R}^d} g(f(y)) J_f(y) dy$$

où est J_f est la jacobienne de f définie presque partout par le Théorème de Rademacher.

III.3.8 Equation de Burgers et résultat de régularité : cas (Γ_2)

Dans cette sous-section, nous supposons donc que ϕ vérifie l'hypothèse (Γ_2) dans le but de montrer que, dans le cas "idéal" (sans cassures), le champ de vitesse v (dont la définition est rappelée p. 97) vérifie les propriétés suivantes :

1. l'équation de Burgers au sens des distributions, i.e.

$$\partial_t v + \frac{1}{2} \nabla_x |v|^2 = 0,$$

2. Le contrôle du gradient du champ de vitesse : $\nabla_{t,x} v \in L_{loc}^\infty([0, 1[, L^1(\mathbb{R}^d))$.

(les propriétés énoncés dans les différents résultats seront en fait un peu plus précises, dans le but de mieux correspondre à nos futurs objectifs).

On rappelle que \mathbf{p}_ϕ désigne l'opérateur

$$\begin{aligned} \mathbf{p}_\phi : [0, 1[\times \mathbb{R}^d &\rightarrow \mathbb{R}^d \\ (t, x) &\mapsto \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{x}{1-t} \right) = \nabla_x(\phi_t)^* \end{aligned}$$

(voir la définition 3.3-9 ainsi que le commentaire associé 3.3-12 p. 89) et qu'il correspond en quelques sortes à une "réciproque" spatiale de l'opérateur $X(t, \cdot) = \nabla_x(\phi_t) = (1-t)\text{id} + t\nabla\phi$. "En quelques sortes" car, $X(t, \cdot)$ n'est en général pas inversible.

Proposition 3.3-18.

Sous la propriété (Γ_2) , \mathbf{p}_ϕ vérifie (c'est à dire en plus des propriétés déjà énoncés à la définition 3.3-9 p. 89), pour tout $t \in [0, 1[$, les propriétés suivantes :

1. $\mathbf{p}_\phi(t, \cdot)$ est bijective sur \mathbb{R}^d et pour tous $x, y \in \mathbb{R}^d$,

$$y = \mathbf{p}_\phi(t, x) \Leftrightarrow x = (1-t)y + t\nabla\phi(y). \quad (3.3-25)$$

2. $\mathbf{p}_\phi(0, \cdot) = \text{id}$.

3. la vitesse v définie en (3.1-2) p. 75 peut être définie à partir de \mathbf{p}_ϕ par :

$$v(t, x) = \nabla\phi(\mathbf{p}_\phi(t, x)) - \mathbf{p}_\phi(t, x) = \frac{x - \mathbf{p}_\phi(t, x)}{t}, \quad (3.3-26)$$

pour tout $(t, x) \in]0, 1[\times \mathbb{R}^d$,

4. v se prolonge de façon continue sur $[0, 1[\times \mathbb{R}^d$ (i.e. en $t = 0$) et

$$v(0, x) = \nabla\phi(x) - x, \quad (3.3-27)$$

pour tout $x \in \mathbb{R}^d$.

Preuve :

1. ϕ est supposée de classe C^1 , donc différentiable en tout point $x \in \mathbb{R}^d$, et par convexité on a $\partial\phi(x) = \{\nabla\phi(x)\}$.
2. Se déduit de la précédente pour $t = 0$.
3. On sait par la propriété précédente que :

$$x = (1-t)\mathbf{p}_\phi(t, x) + t\nabla\phi(\mathbf{p}_\phi(t, x)) \Leftrightarrow \nabla\phi(\mathbf{p}_\phi(t, x)) - \mathbf{p}_\phi(t, x) = \frac{x - \mathbf{p}_\phi(t, x)}{t} = v(t, x).$$

4. Soient $(t_n, x_n)_n \in ((0, 1) \times \mathbb{R}^d)^\mathbb{N}$ et $x \in \mathbb{R}^d$ tels que

$$(t_n, x_n) \xrightarrow{n \rightarrow +\infty} (0, x).$$

Montrons que

$$v(t_n, x_n) \xrightarrow{n \rightarrow +\infty} \nabla\phi(x) - x.$$

D'après le point précédent, pour tout $n \in \mathbb{N}$, on a

$$v(t_n, x_n) = \nabla\phi(\mathbf{p}_\phi(t_n, x_n)) - \mathbf{p}_\phi(t_n, x_n).$$

Comme ϕ vérifie la propriété (Γ_2) , $\nabla\phi$ est continue sur \mathbb{R}^d . Il suffit donc pour conclure de prouver que $\mathbf{p}_\phi(t_n, x_n)$ converge vers x lorsque $n \rightarrow +\infty$ (en d'autres termes que \mathbf{p}_ϕ est continue sur $[0, 1[\times \mathbb{R}^d$).

Soit $a \in \mathbb{R}^d$ tel que $\phi(a) = \inf_{\mathbb{R}^d} \phi$. D'après le Lemme 3.3-13 p. 95), $\mathbf{p}_\phi(t, (1-t)a) = a$ pour tout $t \in [0, 1[$. Donc pour tout $n \in \mathbb{N}$,

$$|\mathbf{p}_\phi(t_n, x_n) - a| = |\mathbf{p}_\phi(t_n, x_n) - \mathbf{p}_\phi(t_n, (1-t_n)a)| \leq \frac{|x_n - (1-t_n)a|}{1-t_n}.$$

La suite $(\mathbf{p}_\phi(t_n, x_n))_n$ est donc bornée et, $\nabla\phi$ étant continue sur \mathbb{R}^d , il en est de même pour la suite $(\nabla\phi(\mathbf{p}_\phi(t_n, x_n)))_n$.

Or pour tout $n \in \mathbb{N}$, on a d'après le deuxième point :

$$x_n = (1 - t_n) \mathbf{p}_\phi(t_n, x_n) + t \nabla\phi(\mathbf{p}_\phi(t_n, x_n)).$$

Donc $\mathbf{p}_\phi(t_n, x_n) \xrightarrow{n \rightarrow +\infty} x$, ce qui conclut la preuve. □

Remarque 3.3-19. On aurait également pu établir la propriété (3.3-25), sachant que $\mathbf{p}_\phi(t, \cdot) = \nabla_x(\phi_t)^*$, grâce au résultat :

"Si f est une fonction de classe C^1 sur \mathbb{R}^d , strictement convexe, et superlinéaire ($f(x)/\|x\|$ tend vers $+\infty$ lorsque $\|x\| \rightarrow +\infty$), alors sa transformée de Legendre f^* possède les mêmes propriétés. De plus la fonction $x \rightarrow \nabla f(x)$ est un homéomorphisme de \mathbb{R}^d sur \mathbb{R}^d , et la bijection réciproque est l'application gradient de f^* ."

En effet ici, $\phi_t = (1/2)(1-t)|\cdot|^2 + t\phi$ est de classe C^1 , strictement convexe et superlinéaire (car ϕ est convexe et $|\cdot|^2$ est strictement convexe et superlinéaire). On établit donc ainsi que pour tout $t \in]0, 1[$, $X(t, \cdot) = (1-t)\text{id} + t\nabla\phi = \nabla\phi_t$ est bijective de réciproque $\mathbf{p}_\phi(t, \cdot) = \nabla_x(\phi_t)^*$.

A. Équation de Burgers sous l'hypothèse (Γ_2)

Intéressons nous au premier résultat annoncé au début de cette sous-section sur le champ v : l'équation de Burgers.

Proposition 3.3-19.

Sous la propriété (Γ_2) , v vérifie (3.1-4) p. 75, à savoir :

$$\partial_t v + \frac{1}{2} \nabla_x |v|^2 = 0$$

au sens des distributions.

Preuve :

1. On sait par la Proposition 3.3-15 (p. 96) valable sous la propriété (Γ_1) , que v est localement lipschitzienne sur $]0, 1[\times \mathbb{R}^d$ et que d'après le Théorème de rademacher 3.3-16 (p. 98), v est presque partout différentiable (en espace et en temps) sur $]0, 1[\times \mathbb{R}^d$, et sa différentielle correspond à sa dérivée au sens des distributions.
2. Soit $(t_0, x_0) \in]0, 1[\times \mathbb{R}^d$ tel que v soit Fréchet-différentiable en (t_0, x_0) . Montrons que

$$\partial_t v(t_0, x_0) + v(t_0, x_0) \cdot \nabla_x v(t_0, x_0) = 0. \quad (3.3-28)$$

Soit $y_0 \in \mathbb{R}^d$ tel que

$$y_0 = \mathbf{p}_\phi(t_0, x_0) \Leftrightarrow x_0 = (1 - t_0)y_0 + t_0 \nabla\phi(y_0) = X(t_0, y_0) \quad (3.3-29)$$

(la fonction X définie en (3.1-1) p. 75). On pose $f(t) = (t, X(t, y_0)) \in \mathbb{R}^3$ pour tout $t \in]0, 1[$. Or on sait d'après la Proposition 3.3-10 (p. 91) que pour tout $t \in]0, 1[$,

$$(v \circ f)(t) = v(t, X(t, y_0)) = \nabla\phi(y_0) - y_0 = v(t_0, X(t_0, y_0)) = \partial_t X(t_0, y_0) = v(t_0, x_0).$$

Ainsi, v étant différentiable en $f(t_0) = (t_0, x_0)$, on obtient en dérivant la fonction constante $v \circ f$ en t_0 ,

$$\begin{aligned} 0 = (v \circ f)(t_0) &= f'(t_0) \cdot \nabla_{t,x} v(t_0, x_0) = \begin{pmatrix} 1 \\ \partial_t X(t_0, y_0) \end{pmatrix} \cdot \nabla_{t,x} v(t_0, x_0) \\ &= \partial_t v(t_0, x_0) + \partial_t X(t_0, y_0) \cdot \nabla_x v(t_0, x_0) \\ &= \partial_t v(t_0, x_0) + v(t_0, x_0) \cdot \nabla_x v(t_0, x_0) \end{aligned} \quad (3.3-30)$$

Ceci étant vrai pour presque tout $(t_0, x_0) \in]0, 1[\times \mathbb{R}^d$, et du fait que la différentielle de v définie presque partout au sens de Fréchet correspond aussi à la dérivée de v au sens des distributions, on a donc au sens des distributions l'égalité :

$$\partial_t v + v \cdot \nabla_x v = 0,$$

et donc d'après (3.3-24),

$$\partial_t v + v \cdot \nabla_x v = \partial_t v + \frac{1}{2} \nabla_x |v|^2 = 0. \quad (3.3-31)$$

□

B. Résultats intermédiaires sur l'opérateur \mathbf{p}

Afin de démontrer le second résultat de régularité sur v annoncé au début de cette sous-section, lequel sera énoncé à la Proposition 3.3-23 (p. 105), nous présentons à présent des résultats intermédiaires sur le potentiel ϕ .

Proposition 3.3-20.

On suppose que ϕ vérifie la propriété (Γ_2) . Soient $R' > R > 0$ et $a \in \mathbb{R}^d$ tel que $\phi(a) = \inf_{R^2} \phi$. Alors il existe $t_0 \in]0, 1[$ tel que pour tout $t \in [0, t_0]$,

$$\mathbf{p}_\phi(t, B(a, R)) \subset \mathbf{p}_\phi(t_0, B(a, R')). \quad (3.3-32)$$

Plus précisément, il suffit pour vérifier la propriété (3.3-32) que t_0 vérifie la condition :

$$t_0 < \min \left\{ \frac{1}{2}, \frac{R' - R}{M + 2|a|} \right\}. \quad (3.3-33)$$

avec

$$M = \sup_{x \in B(a, 2(R+|a|))} |\nabla \phi(x)|.$$

Preuve : Soient $x \in B(a, R)$ et $t, t_0 \in [0, 1[$ tels que $t_0 > 0$ et $t \in [0, t_0]$.

$\mathbf{p}_\phi(t, x) \in \mathbf{p}_\phi(t_0, B(a, R'))$ si et seulement si il existe $y \in B(a, R')$ tel que $\mathbf{p}_\phi(t_0, y) = \mathbf{p}_\phi(t, x)$. Par définition de \mathbf{p}_ϕ (par l'équivalence (3.3-25) p. 100), pour $y \in \mathbb{R}^d$:

$$\mathbf{p}_\phi(t_0, y) = \mathbf{p}_\phi(t, x) \Leftrightarrow y = (1 - t_0) \mathbf{p}_\phi(t, x) + t_0 \nabla \phi(\mathbf{p}_\phi(t, x)). \quad (3.3-34)$$

Prenons donc $y = (1 - t_0) \mathbf{p}_\phi(t, x) + t_0 \nabla \phi(\mathbf{p}_\phi(t, x))$, et cherchons une condition suffisante sur t_0 pour que $y \in B(a, R')$.

Tout d'abord rappelons que, d'après le Lemme 3.3-13 (p. 95), $\mathbf{p}_\phi(t, (1-t)a) = a$, et donc que pour tout $x \in B(a, R)$,

$$|\mathbf{p}_\phi(t, x) - a| = |\mathbf{p}_\phi(t, x) - \mathbf{p}_\phi(t, (1-t)a)| \leq \frac{1}{1-t}|x - (1-t)a| \leq \frac{1}{1-t_0}(R + |a|).$$

En prenant $t_0 \leq 1/2$, on a ainsi $\mathbf{p}_\phi(t, x) \in B(a, 2(R + |a|))$. Or sous la propriété (Γ_2) , ϕ est de classe C^1 sur \mathbb{R}^d et donc localement lipschitzienne, et ainsi lipschitzienne sur $B(a, 2(R + |a|))$. En vertu du Lemme 3.3-1 (p. 86), on peut donc poser

$$M = \sup_{x \in B(a, 2(R+|a|))} |\partial\phi(x)| < +\infty.$$

Pour tout $t \in [0, t_0]$ et $x \in B(a, R)$, et $y = (1-t_0)\mathbf{p}_\phi(t, x) + t_0\nabla\phi(\mathbf{p}_\phi(t, x))$:

$$\begin{aligned} |y - a| &\leq (1-t_0)|\mathbf{p}_\phi(t, x) - a| + t_0|\nabla\phi(\mathbf{p}_\phi(t, x)) - a| \\ &\leq (1-t_0)|\mathbf{p}_\phi(t, x) - \mathbf{p}_\phi(t, (1-t)a)| + t_0(|\nabla\phi(\mathbf{p}_\phi(t, x))| + |a|) \\ &\leq \frac{1-t_0}{1-t}|x - (1-t)a| + t_0(M + |a|) \\ &\leq |x - a| + t|a| + t_0(M + |a|) \\ &\leq R + t_0(M + 2|a|). \end{aligned}$$

Ainsi si $t_0 < \min\{1/2, (R' - R)/(M + 2|a|)\}$, alors pour tout $t \in [0, t_0]$ et $x \in B(a, R)$

1. $y = (1-t_0)\mathbf{p}_\phi(t, x) + t_0\nabla\phi(\mathbf{p}_\phi(t, x)) \in B(a, R')$,
2. $\mathbf{p}_\phi(t_0, y) = \mathbf{p}_\phi(t, x)$,

et donc $\mathbf{p}_\phi(t, B(a, R)) \subset \mathbf{p}_\phi(t_0, B(a, R'))$. □

Remarque 3.3-20. Le Lemme 3.3-13 (p. 95) et la Proposition 3.3-20 (p. 102) auraient également pu être établis avec \mathbf{p}_ϕ définie à partir seulement d'une fonction ϕ convexe continue sur \mathbb{R}^d (sans plus de régularité), par des preuves analogues en tout points, et de complexités équivalentes.

On rappelle enfin le Lemme suivant :

Lemme 3.3-21. (voir [34] p.75, ch.2)

Soit f une fonction lipschitzienne sur \mathbb{R}^d . Alors pour tout $A \subset \mathbb{R}^d$,

$$\mathcal{L}^n(A) = 0 \Rightarrow \mathcal{L}(f(A)) = 0$$

où rappelons le, \mathcal{L} désigne la mesure de Lebesgue sur \mathbb{R} .

Lemme 3.3-22.

On suppose que ϕ vérifie la propriété (Γ_2) . Pour tout $t \in [0, 1[$, $\mathbf{p}_\phi(t, \cdot)$ est presque partout différentiable sur \mathbb{R}^d (sa différentielle coïncide avec sa dérivée au sens des distributions). De plus, pour presque tout $x \in \mathbb{R}^d$, $\nabla\phi$ est différentiable en $\mathbf{p}_\phi(t, x)$ et

$$\nabla_x \mathbf{p}_\phi(t, x) = (t \text{Hess}(\phi)(\mathbf{p}_\phi(t, x)) + (1-t)I)^{-1} \quad (3.3-35)$$

où $I \in \mathcal{M}_2(\mathbb{R})$ désigne la matrice identité.

Preuve : Soit $t \in [0, 1[$. $\mathbf{p}_\phi(t, \cdot)$ est lipschitzienne et bijective, et d'après (3.3-25) p. 100 sa réciproque est

$$\mathbf{p}_\phi(t, \cdot)^{-1} = t\nabla\phi + (1-t)\text{id} = X(t, \cdot)$$

($X(t, \cdot)$ définie en (3.1-1) p. 75).

Rappelons que par hypothèse, $\nabla\phi$ est supposé lipschitzien. D'après le Théorème de Rademacher 3.3-16 (p. 98), $\nabla\phi$ et $\mathbf{p}_\phi(t, \cdot)$ sont presque partout différentiables sur \mathbb{R}^d et leurs différentielles coïncident avec leurs dérivées au sens des distributions.

Ainsi l'ensemble F des points de \mathbb{R}^d où $\nabla\phi$ n'est pas différentiable est de mesure nulle (pour la mesure de Lebesgue). Le fait que $\nabla\phi$ soit supposé lipschitzien, implique que $X(t, \cdot)$ l'est également. Par conséquent, d'après le Lemme 3.3-21, $\mathcal{L}^2(X(t, F)) = 0$. $X(t, F)$ étant l'ensemble des points $x \in \mathbb{R}^d$ pour lesquels $\nabla\phi$ n'est pas différentiable en $\mathbf{p}_\phi(t, x)$, cela signifie donc que $\nabla\phi$ est différentiable en $\mathbf{p}_\phi(t, x)$ pour presque tout $x \in \mathbb{R}^d$.

Or d'après (3.3-25), pour tout $x \in \mathbb{R}^d$,

$$x = (1-t)\mathbf{p}_\phi(t, x) + t\nabla\phi(\mathbf{p}_\phi(t, x)).$$

Ainsi, pour presque tout $x \in \mathbb{R}^d$, $\mathbf{p}_\phi(t, \cdot)$ est différentiable en x , $\nabla\phi$ est différentiable en $\mathbf{p}_\phi(t, x)$ et

$$I = (t\text{Hess}(\phi)(\mathbf{p}_\phi(t, x)) + (1-t)I)\nabla_x \mathbf{p}_\phi(t, x).$$

ϕ étant convexe, $\text{Hess}(\phi)(\mathbf{p}_\phi(t, x))$ est symétrique positive, et donc par coercivité, $t\text{Hess}(\phi)(\mathbf{p}_\phi(t, x)) + (1-t)I$ est symétrique définie positive et donc inversible dans $\mathcal{M}_2(\mathbb{R})$. En conséquence, pour presque tout $x \in \mathbb{R}^d$, $\mathbf{p}_\phi(t, \cdot)$ est différentiable en x et

$$\nabla_x \mathbf{p}_\phi(t, x) = (t\text{Hess}(\phi)(\mathbf{p}_\phi(t, x)) + (1-t)I)^{-1}.$$

□

Remarque 3.3-21. Le fait que $\nabla\phi$ soit **globalement** lipschitzien sur \mathbb{R}^d est ici très important car il assure le fait que $\mathbf{p}_\phi(t, \cdot)$ ne concentre pas des zones trop grandes de \mathbb{R}^d .

En effet, si $\nabla\phi$ ne possède pas une telle régularité, l'ensemble des x en lesquels ϕ n'est pas deux fois différentiable pourrait ne pas être de mesure nulle.

Remarquons que pour tout $t \in [0, 1[$, l'opérateur $\mathbf{p}_\phi(t, \cdot)$ s'avère être presque partout différentiable sur \mathbb{R}^d , et ceci quelle que soit la régularité de ϕ (car un opérateur proximal est toujours lipschitzien). La propriété supplémentaire apportée ici par la régularité de ϕ est en fait la bijectivité de l'opérateur $\mathbf{p}_\phi(t, \cdot)$. Le caractère globalement lipschitzien de $\nabla\phi$ nous permet cependant dans notre cas d'expliciter presque partout $\nabla_x \mathbf{p}_\phi(t, \cdot)$ en fonction de $\mathbf{p}_\phi(t, \cdot)$.

C. Régularité et contrôle du champ de vitesse dans le cas (Γ_2)

On considère les normes sur $\mathcal{M}_2(\mathbb{R})$: $|\cdot|_1$ et $|\cdot|_{Tr}$, définies pour tout $A \in \mathcal{M}_2(\mathbb{R})$ par

1. $|A|_1 = \sum_{i,j} |a_{ij}|$,

2. $|A|_{Tr} = \sqrt{\text{Tr}(A^T A)}$ (norme de Frobenius).

Par équivalence des normes en dimension finie, il existe $C_1, C_2 > 0$ tels que

$$C_1 |\cdot|_1 \leq |\cdot|_{Tr} \leq C_2 |\cdot|_1$$

Rappelons en outre que pour tous $A, P \in \mathcal{M}_2(\mathbb{R})$, avec P une matrice orthogonale ($P^T = P^{-1}$), on a $|PAP^{-1}|_{Tr} = |A|_{Tr}$, et donc

$$\frac{C_1}{C_2} |A|_1 \leq |PAP^{-1}|_1 \leq \frac{C_2}{C_1} |A|_1 \quad (3.3-36)$$

Nous avons maintenant tous les éléments pour énoncer la Proposition suivante, qui constitue un des principaux résultats de cette section : elle concerne en effet le contrôle du gradient du champ de vitesse v . Ce contrôle, ainsi que nous l'avons déjà évoqué dans la sous-section précédente, sera plus tard exploité pour contrôler les solutions du problème de transport engendré par le champ v (voir les résultats d'unicité du chapitre IV).

Cette Proposition est valable pour $v = v_\phi$ définie à partir d'un potentiel ϕ vérifiant la propriété (Γ_1) . Cependant, comme pour tous les résultats de cette sous-section, nous la démontrerons dans un premier temps en supposant que le potentiel ϕ vérifie la propriété (Γ_2) , c'est à dire, rappelons-le, que ϕ est supposé convexe, de classe C^1 , minoré sur \mathbb{R}^d et que son gradient spatiale $\nabla_x \phi$ est supposé lipschitzien. Nous généraliserons ensuite la preuve dans la sous-section suivante (c'est à dire pour un potentiel ϕ vérifiant la propriété (Γ_1)).

Proposition 3.3-23.

On suppose que ϕ vérifie la propriété (Γ_2) . Soient $R' > R > 0$ et $a \in \mathbb{R}^d$ tel que $\phi(a) = \inf_{\mathbb{R}^d} \phi$. Alors il existe des constantes C et C' (indépendantes de ϕ , a , R et R') telles que pour tout $t_0 \in]0, 1[$ vérifiant la condition 3.3-33 (p. 102), on ait la propriété :

$$\forall t \in]0, t_0], \int_{B(a, R)} |\nabla_x v(t, x)|_1 dx \leq C \int_{B(a, R')} |\nabla_x v(t_0, x)|_1 dx \leq \frac{C'}{t_0(1-t_0)} \mathcal{L}^2(B(a, R')) \quad (3.3-37)$$

et donc $\nabla_x v \in L^\infty([0, t_0], L^1(B(a, R)))$.

Preuve :

1. Soit $t \in]0, 1[$. Rappelons que par (3.3-26) p. 100, pour tout $x \in \mathbb{R}^d$,

$$v(t, x) = \nabla \phi(\mathbf{p}_\phi(t, x)) - \mathbf{p}_\phi(t, x).$$

D'après le Lemme 3.3-22 (p. 103), pour presque tout $x \in \mathbb{R}^d$, $v(t, \cdot)$ est différentiable en x et

$$\begin{aligned} \nabla_x v(t, x) &= \nabla_x \mathbf{p}_\phi(t, x) (\text{Hess}(\phi)(\mathbf{p}_\phi(t, x)) - I) \\ &= (t \text{Hess}(\phi)(\mathbf{p}_\phi(t, x)) + (1-t)I)^{-1} (\text{Hess}(\phi)(\mathbf{p}_\phi(t, x)) - I). \end{aligned} \quad (3.3-38)$$

D'après le commentaire 3.3-12 (p. 91, deuxième point), la fonction $|\nabla_x v(t, \cdot)|_1$ définie presque partout sur \mathbb{R}^d est uniformément bornée par $c/t(1-t)$ (c une constante) : en effet comme déjà mentionné à la Proposition 3.3-15 (p. 96), ou bien comme cela peut aisément se constater dans la formulation (3.3-26), $v(t, \cdot)$ est $2/t(1-t)$ -lipschitzienne (on

ne prend bien sûr pas la constante de Lipschitz la plus précise). D'après le Lemme 3.3-2 (p. 86), $|\nabla_x v(t, \cdot)|_1$ est borné par $c/t(1-t)$ (la constante c vient du fait que l'on est en dimension 2 et que l'on utilise la norme $|\cdot|_1$ qui n'est pas une norme subordonnée). Ainsi $|\nabla_x v(t, \cdot)|_1 \in L^\infty(\mathbb{R}^d) \subset L^1(\mathbb{R}^d)$. Comme la fonction $X(t, \cdot) = t\nabla\phi + (1-t)\text{id}$ est lipschitzienne et bijective (et que $\mathbf{p}_\phi(t, \cdot)$ est sa réciproque), on peut donc appliquer le Théorème 3.3-17 (p. 99) de changement de variable généralisé et on obtient :

$$\int_{B(a,R)} |\nabla_x v(t, x)|_1 dx = \int_{\mathbf{p}_\phi(t, B(a,R))} |\det(t \text{Hess}(\phi)(y) + (1-t)I)| \times |(t \text{Hess}(\phi)(y) + (1-t)I)^{-1}(\text{Hess}(\phi)(y) - I)|_1 dy. \quad (3.3-39)$$

2. Pour tout $y \in \mathbb{R}^d$ où $\nabla\phi$ est différentiable, il existe une matrice orthogonale $P(y)$ et $\lambda_1(y), \lambda_2(y) \geq 0$ tels que

$$\text{Hess}(\phi)(y) = P(y) \begin{pmatrix} \lambda_1(y) & 0 \\ 0 & \lambda_2(y) \end{pmatrix} P(y)^T. \quad (3.3-40)$$

En effet ϕ étant convexe, $\text{Hess}(\phi)$ est symétrique positive.

$P(y)$ est orthogonale donc $P(y)^T = P(y)^{-1}$, et en conséquence

$$\begin{aligned} & (t \text{Hess}(\phi)(y) + (1-t)I)^{-1}(\text{Hess}(\phi)(y) - I) \\ &= P(y) \left(t \begin{pmatrix} \lambda_1(y) & 0 \\ 0 & \lambda_2(y) \end{pmatrix} + (1-t)I \right)^{-1} \left(\begin{pmatrix} \lambda_1(y) & 0 \\ 0 & \lambda_2(y) \end{pmatrix} - I \right) P(y)^{-1} \\ &= P(y) \begin{pmatrix} t\lambda_1(y) + (1-t) & 0 \\ 0 & t\lambda_2(y) + (1-t) \end{pmatrix}^{-1} \begin{pmatrix} \lambda_1(y) - 1 & 0 \\ 0 & \lambda_2(y) - 1 \end{pmatrix} P(y)^{-1} \\ &= P(y) \begin{pmatrix} \frac{\lambda_1(y) - 1}{t\lambda_1(y) + (1-t)} & 0 \\ 0 & \frac{\lambda_2(y) - 1}{t\lambda_2(y) + (1-t)} \end{pmatrix} P(y)^{-1}. \end{aligned} \quad (3.3-41)$$

D'après les inégalités (3.3-36), il est clair qu'il existe des constantes $C_3, C_4 > 0$ indépendantes de ϕ, t et y (i.e. $C_3 = C_1/C_2$ et $C_4 = C_2/C_1$), telles que pour tout $t \in [0, 1[$ et presque tout $y \in \mathbb{R}^d$,

$$\begin{aligned} & C_3 |(t \text{Hess}(\phi)(y) + (1-t)I)^{-1}(\text{Hess}(\phi)(y) - I)|_1 \\ & \leq \sum_{i=1,2} \left| \frac{\lambda_i(y) - 1}{t\lambda_i(y) + (1-t)} \right|_1 \\ & \leq C_4 |(t \text{Hess}(\phi)(y) + (1-t)I)^{-1}(\text{Hess}(\phi)(y) - I)|_1. \end{aligned} \quad (3.3-42)$$

En outre pour tout $t \in [0, 1[$ et presque tout $y \in \mathbb{R}^d$,

$$\begin{aligned} \det(t \text{Hess}(\phi)(y) + (1-t)I) &= \det \begin{pmatrix} t\lambda_1(y) + (1-t) & 0 \\ 0 & t\lambda_2(y) + (1-t) \end{pmatrix} \\ &= (t\lambda_1(y) + (1-t))(t\lambda_2(y) + (1-t)) \end{aligned} \quad (3.3-43)$$

En injectant (3.3-42) et (3.3-43) dans (3.3-39), on obtient donc, pour tout $t \in]0, 1[$ (et tout $R > 0$) :

$$\begin{aligned} C_3 \int_{B(a,R)} |\nabla_x v(t, x)|_1 dx &\leq \sum_{(i,j)=(1,2),(2,1)} \int_{\mathbf{p}_\phi(t, B(a,R))} (t\lambda_i(y) + (1-t)) |\lambda_j(y) - 1| \\ &\leq C_4 \int_{B(a,R)} |\nabla_x v(t, x)|_1 dx \end{aligned} \quad (3.3-44)$$

en rappelant que $\lambda_1(y), \lambda_2(y) \geq 0$. C_3 et C_4 sont bien indépendantes de t et R .

3. On prend $t_0 \in]0, 1[$ vérifiant la condition (3.3-33) de la Proposition 3.3-20 (p. 102). On a donc, d'après cette Proposition, $\mathbf{p}_\phi(t, B(a, R)) \subset \mathbf{p}_\phi(t_0, B(a, R'))$.

La condition (3.3-33) donne aussi en particulier $t_0 \leq 1/2$, donc pour tout $t \in]0, t_0]$, $1-t \leq 1 \leq 2(1-t_0)$. $\lambda_i(y) \geq 0$ ($i = 1, 2$) pour presque tout $y \in \mathbb{R}^d$, et donc alors

$$0 \leq t\lambda_i(y) + (1-t) \leq 2(t_0\lambda_i(y) + (1-t_0)).$$

Ainsi, grâce à l'inégalité (3.3-44), on peut conclure :

$$\begin{aligned} \int_{B(a,R)} |\nabla_x v(t, x)|_1 dx &\leq \frac{1}{C_3} \sum_{(i,j)=(1,2),(2,1)} \int_{\mathbf{p}_\phi(t, B(a,R))} (t\lambda_i(y) + (1-t)) |\lambda_j(y) - 1| \\ &\leq \frac{2}{C_3} \sum_{(i,j)=(1,2),(2,1)} \int_{\mathbf{p}_\phi(t_0, B(a,R'))} (t_0\lambda_i(y) + (1-t_0)) |\lambda_j(y) - 1| \\ &\leq \frac{2C_4}{C_3} \int_{B(a,R')} |\nabla_x v(t_0, x)|_1 dx \end{aligned} \quad (3.3-45)$$

Pour finir, et comme déjà mentionné en début de preuve, il existe, d'après le Lemme 3.3-2 (p. 86), une constante $c > 0$ tel que $|\nabla_x v(t_0, \cdot)|_1$ soit borné par $c/t_0(1-t_0)$, ce qui achève la preuve de la Proposition. \square

La preuve de cette Proposition nous fournit également les éléments de preuve pour un autre Lemme concernant le contrôle de v qui sera utile dans la suite : ce sera en fait la combinaison des différents résultats de cette section concernant le contrôle du gradient du champ v qui nous permettra par la suite (chapitre IV) de contrôler les solutions du problème de transport, et ainsi d'obtenir nos futurs résultats d'unicité.

Lemme 3.3-24.

On suppose que ϕ vérifie la propriété (Γ_2) . Alors il existe une constante $C > 0$ tel que pour tout $t \in]0, 1[$ et presque tout $x \in \mathbb{R}^d$ où $v(t, \cdot)$ est différentiable,

$$\|\nabla_x v(t, \cdot)\|_{L^\infty(\mathbb{R}^d)} \leq \frac{C}{1-t} \left(\|\text{Hess}(\phi)\|_{L^\infty(\mathbb{R}^d)} + 2 \right)$$

(où la norme L^∞ est ici prise sur $\|\cdot\|$: la norme subordonnée à la norme euclidienne $|\cdot|$ de \mathbb{R}^d).

Preuve : On rappelle que d'après le Lemme 3.3-22 (p. 103), pour presque tout $x \in \mathbb{R}^d$, $v(t, \cdot)$ est différentiable en x et

$$\nabla_x v(t, x) = (t \text{Hess}(\phi)(\mathbf{p}_\phi(t, x)) + (1-t)I)^{-1}(\text{Hess}(\phi)(\mathbf{p}_\phi(t, x)) - I).$$

De même qu'avec l'équation (3.3-42), on peut par les relations (3.3-41) et (3.3-36), obtenir l'inégalité suivante pour tout $y \in \mathbb{R}^d$ où $\nabla\phi$ est différentiable :

$$\begin{aligned} C_3 |(t \text{Hess}(\phi)(y) + (1-t)I)^{-1}(\text{Hess}(\phi)(y) - I)|_1 &\leq \sum_{i=1,2} \left| \frac{\lambda_i(y) - 1}{t\lambda_i(y) + (1-t)} \right|_1 \leq \frac{\lambda_1(y) + \lambda_2(y) + 2}{1-t} \\ &\leq \frac{C_4}{1-t} (|\text{Hess}(\phi)(y)|_1 + 2), \end{aligned} \tag{3.3-46}$$

avec λ_1 et λ_2 toujours définies comme en (3.3-40). Par équivalence des normes il existe donc une constante $C > 0$ telle que

$$|(t \text{Hess}(\phi)(y) + (1-t)I)^{-1}(\text{Hess}(\phi)(y) - I)| \leq \frac{C}{1-t} \left(\|\text{Hess}(\phi)\|_{L^\infty(\mathbb{R}^d)} + 2 \right).$$

Ainsi pour tout $t \in]0, 1[$ et presque tout $x \in \mathbb{R}^d$ où $v(t, \cdot)$ est différentiable, on a

$$\|\nabla_x v(t, \cdot)\|_{L^\infty(\mathbb{R}^d)} \leq \frac{C}{1-t} \left(\|\text{Hess}(\phi)\|_{L^\infty(\mathbb{R}^d)} + 2 \right).$$

□

III.3.9 Equation de Burgers et résultat de régularité : cas (Γ_1)

Nous allons maintenant nous placer dans un cadre plus général, et étendre les résultats de la sous-section précédente pour des potentiels ϕ ne vérifiant plus que la propriété (Γ_1) , c'est à dire que dans cette sous-section ϕ possédera d'éventuelles points de non-dérivabilité **à l'origine de cassures dans le plan de transport**.

Nous allons dans un premier temps montrer que le résultat de la Proposition 3.3-19 (p. 101), prouvant que v vérifie l'équation de Burgers (3.1-4), à savoir rappelons le :

$$\partial_t v + \frac{1}{2} \nabla_x |v|^2 = 0,$$

reste valable pour ϕ vérifiant uniquement la propriété (Γ_1) .

Nous montrerons dans un deuxième temps, en invoquant en particulier la Proposition 3.3-23 (p. 105), que pour tous $0 < R < R'$ et $a \in \mathbb{R}^d$ tel que $\phi(a) = \inf_{\mathbb{R}^d} \phi$, et pour tout

$$0 < t_0 < \min\{1/2, (R' - R)/(M + 2|a|)\}$$

on a $\nabla_x v \in L^\infty([0, t_0], L^1(B(a, R)))$, et plus particulièrement qu'il existe une constante $C' > 0$ tel que pour tout $t \in [0, t_0]$,

$$\int_{B(a, R)} |\nabla_x v(t, x)|_1 dx \leq \frac{C'}{t_0(1-t_0)} \mathcal{L}^2(B(a, R')).$$

Dans le cadre de notre problème de transport optimal, nous avons montré à la section III.2 (p. 80) que l'on pouvait effectivement se placer sous l'hypothèse (Γ_1) . Nous allons montrer en régularisant ϕ que ∇_x est uniformément intégrable au voisinage de $t = 0$, puis nous raisonnerons par symétrie pour montrer qu'il en est de même au voisinage de $t = 1$.

Pour cette régularisation, nous utiliserons l'enveloppe de Moreau (introduite à la définition 3.3-8 p. 88) $\gamma\phi$ de ϕ pour tout $\gamma > 0$ dont nous rappelons la définition :

pour tout $x \in \mathbb{R}^d$,

$$\gamma\phi(x) = \inf_{y \in \mathbb{R}^d} \frac{1}{2\gamma}|x - y|^2 + \phi(y) = \frac{1}{2\gamma}|x - \text{Prox}_{\gamma\phi}(x)|^2 + \phi(\text{Prox}_{\gamma\phi}(x)), \quad (3.3-47)$$

Pour tout $\gamma > 0$, nous définissons également le champ de vitesse v_γ pour tout $t \in]0, 1[$ et $x \in \mathbb{R}^d$ par

$$v_\gamma(t, x) = v_{\gamma\phi}(t, x) = \nabla\phi(\mathbf{p}_{\gamma\phi}(t, x)) - \mathbf{p}_{\gamma\phi}(t, x) = \frac{x - \mathbf{p}_{\gamma\phi}(t, x)}{t}, \quad (3.3-48)$$

avec, rappelons le,

$$\mathbf{p}_{\gamma\phi}(t, x) = \text{Prox}_{\frac{t}{1-t}\gamma\phi}\left(\frac{x}{1-t}\right).$$

Nous rappelons également que $\gamma\phi$ est de classe C^1 et que pour tout $x \in \mathbb{R}^d$,

$$\nabla(\gamma\phi)(x) = \frac{x - \text{Prox}_{\gamma\phi}(x)}{\gamma} \in \partial\phi(\text{Prox}_{\gamma\phi}(x)) \quad (\text{par 3.3-4, p. 88}). \quad (3.3-49)$$

$\gamma\phi$ est donc γ^{-1} -lipschitzien.

Il faut tout d'abord montrer que si ϕ vérifie la propriété (Γ_1) , alors $\gamma\phi$ vérifie la propriété (Γ_2) , et ainsi pouvoir se ramener par l'intermédiaire des ϕ_γ , aux cas traités dans la sous-section III.3.8 (i.e. pour des ϕ réguliers et donc de plans de transport sans cassures).

Lemme 3.3-25.

Soit f une fonction de \mathbb{R}^d dans $] -\infty, +\infty]$. On a

$$(\gamma f)^* = f^* + \frac{\gamma}{2}|\cdot|^2.$$

Preuve : Rappelons les deux premières propriétés énoncées à la suite de la définition 3.3-4 p. 87 (à propos de la transformée de Legendre), à savoir

$$\forall \alpha > 0, (\alpha f)^* = \alpha f^*(\cdot/\alpha), \quad (3.3-50)$$

$$\text{et} \quad \left(\frac{1}{2}|\cdot|^2\right)^* = \frac{1}{2}|\cdot|^2. \quad (3.3-51)$$

En se référant également à (3.3-2), on a

$$\begin{aligned} (\gamma f)^* &= \left(f \square \left(\frac{1}{2\gamma}|\cdot|^2\right)\right)^* = f^* + \left(\frac{1}{2\gamma}|\cdot|^2\right)^* \quad (\text{par (3.3-2)}) \\ &= f^* + \frac{1}{\gamma} \left(\frac{1}{2}|\cdot|^2\right)^* (\gamma \cdot) \quad (\text{par (3.3-50)}) \\ &= f^* + \frac{1}{2\gamma}|\gamma \cdot|^2 = f^* + \frac{\gamma}{2}|\cdot|^2 \quad (\text{par (3.3-51)}). \end{aligned}$$

□

Lemme 3.3-26.

On suppose que ϕ vérifie la propriété (Γ_1) . Pour tout $\gamma > 0$ et $a \in \mathbb{R}^d$ minimise $\gamma\phi$ si et seulement si a minimise ϕ . Autrement dit

$$\gamma\phi(a) = \inf_{\mathbb{R}^d} \gamma\phi \Leftrightarrow \phi(a) = \inf_{\mathbb{R}^d} \phi \Leftrightarrow \text{Prox}_{\gamma\phi}(a) = a.$$

Preuve : Soit $a \in \mathbb{R}^d$ et $\gamma > 0$. Alors

$$\begin{aligned} \phi(a) = \inf_{\mathbb{R}^d} \phi &\Leftrightarrow 0 \in \partial\phi(a) \Leftrightarrow 0 \in \partial(\gamma\phi)(a) \Leftrightarrow a - a = 0 \in \partial(\gamma\phi)(a) \\ &\Leftrightarrow \text{Prox}_{\gamma\phi}(a) = a \Leftrightarrow \frac{a - \text{Prox}_{\gamma\phi}(a)}{\gamma} = \nabla\gamma\phi(a) = 0 \Leftrightarrow \gamma\phi(a) = \inf_{\mathbb{R}^d} \gamma\phi. \end{aligned}$$

□

Nous pouvons donc conclure :

Lemme 3.3-27.

Si ϕ vérifie la propriété (Γ_1) , alors $\gamma\phi$ vérifie la propriété (Γ_2) pour tout $\gamma > 0$.

Preuve :

1. Comme cela a été énoncé dans la définition 3.3-8, et comme rappelé en (3.3-49), $\gamma\phi$ est convexe de classe C^1 sur \mathbb{R}^d et $\gamma\phi$ est γ^{-1} -lipschitzien. Ainsi pour conclure il suffit maintenant de prouver que $\gamma\phi$ vérifie bien également la propriété (Γ_1) , à savoir que $\gamma\phi$ et $(\gamma\phi)^*$ sont convexes, continues et admettent un minimum sur \mathbb{R}^d .
2. Par le point précédent on sait déjà que $\gamma\phi$ est convexe et continue (et est même de classe C^1) sur \mathbb{R}^d . Il reste donc à prouver qu'il admet un minimum sur \mathbb{R}^d : le Lemme 3.3-26 montre que si ϕ vérifie la propriété (Γ_1) , en particulier l'hypothèse d'existence d'un minimum pour ϕ , alors $\gamma\phi$ partage les mêmes minimums.
3. Il reste à étudier $(\gamma\phi)^*$: d'après le Lemme 3.3-25, on a

$$(\gamma\phi)^* = \phi^* + \frac{\gamma}{2} |\cdot|^2.$$

Or ϕ vérifie la propriété (Γ_1) , donc ϕ^* est convexe et continue sur \mathbb{R}^d , d'où le fait que $(\gamma\phi)^*$ soit également convexe et continue sur \mathbb{R}^d . Le terme $\frac{\gamma}{2} |\cdot|^2$ rend de plus $(\gamma\phi)^*$ coercitive (et même superlinéaire) : $(\gamma\phi)^*$ admet donc un minimum sur \mathbb{R}^d .

□

A. Résultats de convergence sur les variables régularisées par enveloppe de Moreau

Nous énonçons à présent des résultats de convergence pour $\nabla^\gamma\phi$ et $(v_\gamma)_\gamma$.

Lemme 3.3-28.

On suppose que ϕ vérifie la propriété (Γ_1) . Soit $a \in \mathbb{R}^d$ tel que $\phi(a) = \inf_{\mathbb{R}^d} \phi$. Alors pour tout $r > 0$ et tout $\gamma > 0$, on a

$$\sup_{x \in B(a,r)} |\nabla^\gamma\phi(x)| \leq \sup_{x \in B(a,r)} |\partial\phi(x)| < +\infty. \quad (3.3-52)$$

En particulier, pour tout $x \in \mathbb{R}^d$, $\nabla^\gamma\phi(x)$ est borné.

Preuve : D'après le Lemme 3.3-26, a minimise ${}^\gamma\phi$ pour tout $\gamma > 0$. De plus, par l'équation (3.3-49), pour tout $x \in \mathbb{R}^d$,

$$\nabla({}^\gamma\phi)(x) \in \partial\phi(\text{Prox}_{\gamma\phi}(x)).$$

Or si $x \in B(a, r)$, alors $|\text{Prox}_{\gamma\phi}(x) - a| = |\text{Prox}_{\gamma\phi}(x) - \text{Prox}_{\gamma\phi}(a)| \leq |x - a| \leq r$, d'où $\text{Prox}_{\gamma\phi}(x) \in B(a, r)$, et donc

$$\nabla({}^\gamma\phi)(B(a, r)) \subset \partial\phi(B(a, r)), \quad (3.3-53)$$

ce qui nous permet de conclure, sachant que sous l'hypothèse (Γ_1) , ϕ est supposée convexe et localement lipschitzienne et donc

$$\sup_{x \in B(a,r)} |\partial\phi(x)| < +\infty,$$

d'après le Lemme 3.3-1. □

Lemme 3.3-29.

On suppose que ϕ vérifie la propriété (Γ_1) . Alors pour tout $x \in \mathbb{R}^d$, l'ensemble des valeurs d'adhérence de $(\nabla^\gamma\phi(x))_\gamma$ lorsque $\gamma > 0$ tend vers 0 est inclus dans $\partial\phi(x)$.

En particulier, si ϕ est différentiable en x , alors $\nabla^\gamma\phi(x)$ converge vers $\nabla\phi(x)$ lorsque $\gamma > 0$ tend vers 0.

$(\nabla^\gamma\phi)_\gamma$ converge donc simplement presque partout vers $\nabla\phi$.

Preuve : Reprenons la relation (3.3-49) : pour tout $x \in \mathbb{R}^d$ et tout $\gamma > 0$,

$$\nabla({}^\gamma\phi)(x) \in \partial\phi(\text{Prox}_{\gamma\phi}(x)).$$

Soit $x \in \mathbb{R}^d$. D'après le Lemme 3.3-28, $(\nabla({}^\gamma\phi)(x))_\gamma$ est bornée. Or, toujours d'après la relation (3.3-49),

$$x - \text{Prox}_{\gamma\phi}(x) = \gamma\nabla({}^\gamma\phi)(x).$$

Ainsi $\text{Prox}_{\gamma\phi}(x)$ converge vers x lorsque γ tend vers 0.

Soit $(\gamma_n)_n$ une suite de réels positifs convergeant vers 0, tel que $(\nabla({}^{\gamma_n}\phi)(x))_n$ converge vers un point $\ell \in \mathbb{R}^d$. D'après le Lemme 3.3-3 (p. 87), $\ell \in \partial\phi(x)$.

Supposons maintenant que ϕ soit différentiable en x , alors $\partial\phi(x) = \{\nabla\phi(x)\}$. Ainsi toute suite convergente $(\nabla(\gamma_n\phi)(x))_n$, avec $(\gamma_n)_n$ convergeant vers 0, converge vers $\nabla\phi(x)$.

Si $\nabla\gamma\phi(x)$ ne convergeait pas vers $\nabla\phi(x)$ lorsque $\gamma > 0$ tend vers 0, alors il existerait une suite $(\nabla(\gamma_n\phi)(x))_n$, avec $(\gamma_n)_n$ convergeant vers 0, restant à une distance fixée de $\nabla\phi(x)$. Or, $(\nabla(\gamma_n\phi)(x))_n$ étant bornée, on pourrait en extraire une sous-suite convergeant dans \mathbb{R}^d , et donc vers $\nabla\phi(x)$, ce qui aboutirait à une contradiction. \square

Proposition 3.3-30.

On suppose que ϕ vérifie la propriété (Γ_1) . La famille $(v_\gamma)_{\gamma>0}$ vérifie les propriétés de convergence suivantes :

1. Pour tout ouvert borné ω dans \mathbb{R}^d et tout $\gamma_0 > 0$, $(v_\gamma)_{\gamma_0 \geq \gamma > 0}$ est uniformément bornée sur $]0, 1[\times \omega$ par une constante indépendante de $\gamma \in]0, \gamma_0]$.
2. Pour tout ouvert borné ω dans \mathbb{R}^d et tout $\gamma_0 > 0$, il existe une constante M indépendante de $\gamma \in]0, \gamma_0]$, telle que

$$\forall \gamma \in]0, \gamma_0], \forall t \in]0, 1[, \|v(t, \cdot) - v_\gamma(t, \cdot)\|_{L^\infty(\omega)} \leq M \frac{\gamma}{t}. \quad (3.3-54)$$

On peut déduire en particulier de cette relation que v_γ converge simplement vers v sur $]0, 1[\times \mathbb{R}^d$ lorsque γ tend vers 0.

3. Pour tout $1 < p < +\infty$ et tout ouvert borné ω dans \mathbb{R}^d , v_γ converge vers v dans $L^p((0, 1) \times \omega)$ lorsque γ tend vers 0.
4. Pour tout $1 < p < +\infty$ et tout ouvert borné ω dans \mathbb{R}^d et pour tout $t \in]0, 1[$, $v_\gamma(t, \cdot)$ converge faiblement dans $W^{1,p}(\omega)$ vers $v(t, \cdot)$ lorsque γ tend vers 0.

Preuve :

1. Soient $\omega \subset \mathbb{R}^d$ un ouvert borné et $\gamma_0 > 0$, et soient $a \in \mathbb{R}^d$ tel que $\phi(a) = \inf_{\mathbb{R}^2} \phi$ et $r > 0$ tel que $\omega \subset B(a, r)$. D'après la Proposition 3.3-12 p. 93 (démontrée sous la propriété (Γ_1)), pour tout $\gamma_0 \geq \gamma > 0$, v_γ vérifie :

$$\sup_{(t,x) \in]0,1[\times B(a,r)} |v_\gamma(t, x)| \leq 5 (\max\{M_\gamma, M_\gamma^*\} + \sup(\omega)),$$

$$\text{avec } M_\gamma = \sup_{x \in B(a,r)} |\nabla\gamma\phi(x)| \quad \text{et} \quad M_\gamma^* = \sup_{x \in B(a,r)} |\partial(\gamma\phi)^*(x)|.$$

► D'après le Lemme 3.3-28 (p. 111),

$$M_\gamma = \sup_{x \in B(a,r)} |\nabla\gamma\phi(x)| \leq \sup_{x \in B(a,r)} |\partial\phi(x)| < +\infty. \quad (3.3-55)$$

► D'après le Lemme 3.3-25 (p. 109), on a

$$(\gamma\phi)^* = \phi^* + \frac{\gamma}{2} |\cdot|^2,$$

d'où pour tout $x \in \mathbb{R}^d$:

$$(\gamma\phi)^*(x) = \partial\left(\phi^* + \frac{\gamma}{2} |\cdot|^2\right)(x) = \partial\phi^*(x) + \gamma x,$$

$$\text{et donc } M_\gamma^* = \sup_{x \in B(a,r)} |\partial\phi^*(x) + \gamma x| \leq \sup_{x \in B(a,r)} |\partial\phi^*(x)| + \gamma_0 \sup(\omega) < +\infty. \quad (3.3-56)$$

On peut donc conclure en combinant (3.3-55) et (3.3-56) :

$$\begin{aligned} \sup_{(t,x) \in]0,1[\times \omega} |v(t,x)| &\leq \sup_{(t,x) \in]0,1[\times B(a,r)} |v_\gamma(t,x)| \\ &\leq 5 (\max\{M_\gamma, M_\gamma^*\} + \sup(\omega)) \\ &\leq 5 (\max\{M, M^*\} + (1 + \gamma_0) \sup(\omega)) < +\infty \end{aligned} \quad (3.3-57)$$

$$\text{avec } M = \sup_{x \in B(a,r)} |\partial\phi(x)| \quad \text{et} \quad M^* = \sup_{x \in B(a,r)} |\partial\phi^*(x)|.$$

2. On pose, pour tout $t \in]0,1[$ et $x \in \mathbb{R}^d$,

$$\mathbf{p}_\gamma^*(t,x) = \text{Prox}_{(1-t)(t\phi)^*}(x) = \text{Prox}_{(1-t)(t\phi)^* + (1-t)\frac{\gamma}{2t}|\cdot|^2}(x).$$

Par définition de l'opérateur proximal, on a pour tout $t \in]0,1[$ et $x \in \mathbb{R}^d$,

$$x - \mathbf{p}_\gamma^*(t,x) \in (1-t)\partial(t\phi)^*(\mathbf{p}_\gamma^*(t,x)) + \gamma \frac{1-t}{t} \mathbf{p}_\gamma^*(t,x),$$

$$\text{i.e.} \quad \left(x - \gamma \frac{1-t}{t} \mathbf{p}_\gamma^*(t,x) \right) - \mathbf{p}_\gamma^*(t,x) \in (1-t)\partial(t\phi)^*(\mathbf{p}_\gamma^*(t,x)),$$

$$\text{i.e.} \quad \mathbf{p}_\gamma^*(t,x) = \mathbf{p}^* \left(t, x - \gamma \frac{1-t}{t} \mathbf{p}_\gamma^*(t,x) \right).$$

Rappelons que

$$v(t,\cdot) = \frac{\text{id} - \nabla_x(\phi_t)^*}{t} \quad \text{avec } \phi_t = (1-t)\frac{|\cdot|^2}{2} + t\phi,$$

et d'après 3.3-12 (p. 90),

$$\nabla_x(\phi_t)^* = \frac{\text{id} - \text{Prox}_{(1-t)(t\phi)^*}}{1-t} = \frac{\text{id} - \mathbf{p}^*(t,\cdot)}{1-t}.$$

$$\text{Donc } v(t,\cdot) = \frac{\text{id} - \nabla_x(\phi_t)^*}{t} = \frac{\mathbf{p}^*(t,\cdot) - t \text{id}}{t(1-t)}, \quad (3.3-58)$$

$$\text{et } \forall \gamma > 0, v_\gamma(t,\cdot) = \frac{\mathbf{p}_\gamma^*(t,\cdot) - t \text{id}}{t(1-t)}. \quad (3.3-59)$$

Pour tout $t \in]0,1[$ et $x \in \mathbb{R}^d$,

$$\begin{aligned} |v_\gamma(t,x) - v(t,x)| &\leq \frac{1}{t(1-t)} |\mathbf{p}_\gamma^*(t,x) - \mathbf{p}^*(t,x)| \\ &\leq \frac{1}{t(1-t)} \left| \mathbf{p}^* \left(t, x - \gamma \frac{1-t}{t} \mathbf{p}_\gamma^*(t,x) \right) - \mathbf{p}^*(t,x) \right| \\ &\leq \frac{1}{t(1-t)} \left(\gamma \frac{1-t}{t} |\mathbf{p}_\gamma^*(t,x)| \right) = \frac{\gamma}{t^2} |\mathbf{p}_\gamma^*(t,x)|, \end{aligned}$$

la dernière inégalité provenant de la non-expansivité de l'opérateur proximal.

D'après 3.3-59 (p. 90), on a $\mathbf{p}_\gamma^*(t,x) = t(x + (1-t)v_\gamma(t,x))$, d'où

$$|v_\gamma(t,x) - v(t,x)| \leq \frac{\gamma}{t} |x + (1-t)v_\gamma(t,x)|.$$

D'après le premier point de cette Proposition, $(v_\gamma)_{\gamma_0 \geq \gamma > 0}$ est uniformément bornée sur $]0, 1[\times \omega$ par une constante indépendante de $\gamma \in]0, \gamma_0]$. Il existe donc $M > 0$ telle que pour tout $\gamma \in]0, \gamma_0]$ et tout $(t, x) \in]0, 1[\times \omega$,

$$|x + (1 - t)v_\gamma(t, x)| \leq M.$$

3. La troisième propriété se déduit immédiatement des deux précédentes en invoquant le Théorème de convergence dominée.
4. Soit $1 \leq p < \infty$. De même que pour la troisième propriété, on conclut des deux premières propriétés que $v_\gamma(t, \cdot)$ converge vers $v(t, \cdot)$ dans $L^p((0, 1) \times \omega)$ quand γ tend vers 0, via le Théorème de convergence dominée.

Remarquons que les deux premières propriétés sont valables sur **tout** $]0, 1[\times \mathbb{R}^d$ et non pas seulement **presque partout** modulo une certaine mesure. On peut donc appliquer la convergence dominée aussi bien sur $]0, 1[\times \omega$ tout entier que sur une simple "tranche" de type $\{t\} \times \omega$.

Il nous faut maintenant montrer que $\nabla_x v_\gamma(t, \cdot)$ converge faiblement dans $L^p(\omega)$ vers $\nabla_x v(t, \cdot)$ quand γ tend vers 0.

Raisonnons par l'absurde et supposons que cette assertion soit fautive. Il existerait alors $(\gamma_n)_n$ une suite de réels strictement positifs tel que $\lim_n \gamma_n = 0$ et pour laquelle cette convergence n'est vérifiée pour aucune sous-suite. Considérons donc une telle suite.

D'après le deuxième point du commentaire 3.3-12, (voir plus précisément l'équation (3.3-13) p. 91), il existe une constante c , indépendante de t et de γ (donc de γ), – ne dépendant en fait que de la norme considérée sur $\mathcal{M}_2(\mathbb{R})$, par exemple ici la norme $|\cdot|_1$ – telle que

$$|\nabla_x v_\gamma(t, \cdot)|_1 \leq \frac{c}{t(1-t)}$$

(pour tout $\gamma > 0$). $|\nabla_x v_\gamma(t, \cdot)|_1$ est donc uniformément borné dans $L^\infty(\omega)$ et donc dans $L^2(\omega)$.

On peut donc extraire une sous-suite de $(\gamma_n)_n$ (que nous appellerons également $(\gamma_n)_n$ par commodité) tel que $(\nabla_x v_{\gamma_n}(t, \cdot))_n$ converge faiblement dans $L^p(\omega)$ vers un élément w de $L^p(\omega)$, et donc en particulier, $(\nabla_x v_{\gamma_n}(t, \cdot))_n$ converge vers w au sens des distributions.

On sait de plus que $(v_{\gamma_n}(t, \cdot))_n$ converge vers $v(t, \cdot)$ dans $L^p(\omega)$, et donc que $(\nabla_x v_{\gamma_n}(t, \cdot))_n$ converge également vers $\nabla_x v(t, \cdot)$ au sens des distributions. Donc $w = \nabla_x v(t, \cdot)$ et $(\nabla_x v_{\gamma_n}(t, \cdot))_n$ converge faiblement vers $\nabla_x v(t, \cdot)$ dans $L^p(\omega)$, ce qui contredit l'hypothèse de non-convergence vers $\nabla_x v(t, \cdot)$ de toute sous-suite de $(\nabla_x v_{\gamma_n}(t, \cdot))_n$.

□

Nous avons désormais assez d'éléments pour démontrer dans le cas général, c'est à dire pour ϕ vérifiant la propriété (Γ_1) (donc pour un plan de transport admettant d'éventuelles "casures"), les résultats annoncés en début de section concernant dans un premier temps l'équation de Burgers, et dans un deuxième les résultats de contrôle et de régularité sur ∇v et les ∇v_γ .

B. Équation de Burgers sous l'hypothèse (Γ_1)

Proposition 3.3-31.

Sous la propriété (Γ_1) , v vérifie (3.1-4) p. 75, c'est à dire :

$$\partial_t v + \frac{1}{2} \nabla_x |v|^2 = 0,$$

ou de façon équivalente,

$$\partial_t v + v \cdot \nabla_x v = 0, \quad (3.3-60)$$

au sens des distributions.

Preuve : La Proposition 3.3-15 (p. 96) a montré que sous la propriété (Γ_1) , v est localement lipschitzienne sur $]0, 1[\times \mathbb{R}^d$. Ainsi d'après le Théorème 3.3-16 (p. 98), v est presque partout différentiable (en espace et en temps) sur $]0, 1[\times \mathbb{R}^d$, et sa différentielle correspond à sa dérivée au sens des distributions. En particulier on observe que $2v \cdot \nabla_x v = \nabla_x |v|^2$ (au sens \mathcal{L}^d -presque partout comme au sens des distribution). $\gamma\phi$ vérifie la propriété (Γ_2) , donc d'après la Proposition 3.3-19 (p. 101), pour tout $\gamma > 0$, v_γ vérifie

$$\partial_t v_\gamma + \frac{1}{2} \nabla_x |v_\gamma|^2 = 0$$

au sens des distributions, i.e. pour tout $\varphi \in C_c^\infty((0, 1) \times \mathbb{R}^d)$,

$$\int_0^1 \int_{\mathbb{R}^d} \partial_t \varphi \cdot v_\gamma + \frac{1}{2} \operatorname{div}_x(\varphi) |v_\gamma|^2 = 0. \quad (3.3-61)$$

Or d'après le troisième point de la Proposition 3.3-30 (p. 112), pour tout ouvert borné ω dans \mathbb{R}^d , v_γ converge vers v dans $L^2((0, 1) \times \omega)$ (on a pris $p = 2$ arbitrairement) quand γ tend vers 0. Donc pour tout $\varphi \in C_c^\infty((0, 1) \times \mathbb{R}^d)$,

1.

$$\int_0^1 \int_{\mathbb{R}^d} \partial_t \varphi \cdot v_\gamma \xrightarrow{\gamma \rightarrow 0} \int_0^1 \int_{\mathbb{R}^d} \partial_t \varphi \cdot v. \quad (3.3-62)$$

2. On pose $\Gamma =]0, 1[\times \operatorname{supp}(\varphi)$. On a la majoration :

$$\left| \int_0^1 \int_{\mathbb{R}^d} \operatorname{div}_x(\varphi) (|v_\gamma|^2 - |v|^2) \right| \leq \|v - v_\gamma\|_{L^2(\Gamma)} \cdot (\|v\|_{L^2(\Gamma)} + \|v_\gamma\|_{L^2(\Gamma)}) \sup_{\Gamma} |\operatorname{div}_x(\varphi)|,$$

avec donc

$$- \|v - v_\gamma\|_{L^2(\Gamma)} \xrightarrow{\gamma \rightarrow 0} 0,$$

— et le terme $\|v_\gamma\|_{L^2(\Gamma)} \leq \mathcal{L}^3(\Gamma) \cdot \sup_{\Gamma} |v_\gamma|$ est borné par le deuxième point de la Proposition 3.3-30.

En conséquences

$$\int_0^1 \int_{\mathbb{R}^d} |v_\gamma|^2 \operatorname{div}_x(\varphi) \xrightarrow{\gamma \rightarrow 0} \int_0^1 \int_{\mathbb{R}^d} |v|^2 \operatorname{div}_x(\varphi). \quad (3.3-63)$$

Ainsi par passage à la limite dans (3.3-61), quand γ tend vers 0, on a bien

$$\int_0^1 \int_{\mathbb{R}^d} \partial_t \varphi \cdot v + \frac{1}{2} \operatorname{div}_x(\varphi) |v|^2 = 0, \quad (3.3-64)$$

et ceci pour tout $\varphi \in C_c^\infty((0,1) \times \mathbb{R}^d)$, i.e. au sens des distributions :

$$\partial_t v + \frac{1}{2} \nabla_x |v|^2 = 0,$$

ou en appliquant l'égalité (3.3-24),

$$\partial_t v + v \cdot \nabla_x v = 0.$$

□

C. Régularité et contrôle du champ de vitesse dans le cas (Γ_1)

Proposition 3.3-32.

On suppose que ϕ vérifie la propriété (Γ_1) . Soient $R' > R > 0$ et $a \in \mathbb{R}^d$ tel que $\phi(a) = \inf_{\mathbb{R}^d} \phi$ et soit

$$M = \sup_{x \in B(a, 2(R+|a|))} |\partial \phi(x)|.$$

Alors il existe une constante C' – indépendante de ϕ , a , R et R' – telle que pour tout $t_0 \in]0, 1[$ vérifiant la condition

$$t_0 < \min \left\{ \frac{1}{2}, \frac{R' - R}{M + 2|a|} \right\}.$$

on ait la propriété :

$$\forall t \in]0, t_0], \int_{B(a, R)} |\nabla_x v(t, x)|_1 dx \leq \frac{C'}{t_0(1-t_0)} \mathcal{L}^2(B(a, R')) \quad (3.3-65)$$

et donc $\nabla_x v \in L^\infty([0, t_0], L^1(B(a, R)))$.

Preuve :

1. D'après le Lemme 3.3-26 (p. 110), a minimise ${}^\gamma \phi$ pour tout $\gamma > 0$ et, d'après le Lemme 3.3-28 (p. 111), pour tout $R > 0$ et tout $\gamma > 0$,

$$M_\gamma = \sup_{x \in B(a, 2(R+|a|))} |\nabla^\gamma \phi(x)| \leq \sup_{x \in B(a, 2(R+|a|))} |\partial \phi(x)| = M.$$

Ainsi, en prenant

$$t_0 < \min \left\{ \frac{1}{2}, \frac{R' - R}{M + 2|a|} \right\},$$

la condition (3.3-33) p. 102 est vérifiée pour tout $\gamma > 0$, et donc, du fait que ${}^\gamma \phi$ vérifie la propriété (Γ_2) , on a d'après la Proposition 3.3-23 (p. 105) : pour tout $\gamma > 0$ et $t \in]0, t_0]$,

$$\int_{B(a, R)} |\nabla_x v_\gamma(t, x)|_1 dx \leq C \int_{B(a, R')} |\nabla_x v_\gamma(t_0, x)|_1 dx \leq \frac{C'}{t_0(1-t_0)} \mathcal{L}^2(B(a, R')), \quad (3.3-66)$$

avec donc t_0 indépendant de γ .

2. D'après la Proposition 9.32 de [4] (p.138),

$$g : f \in L^2(B(a, R), \mathbb{R}^d) \longmapsto \int_\omega |f|_1 dx$$

est propre convexe et s.c.i, donc s.c.i faible sur $L^2(B(a, R), \mathbb{R}^d)$. Or d'après le quatrième point de la Proposition 3.3-30 (p. 112), pour tout $t \in]0, 1[$, $\nabla_x v_\gamma(t, \cdot)$ converge faiblement dans $L^2(B(a, R), \mathbb{R}^d)$ vers $\nabla_x v(t, \cdot)$ (le choix de prendre ici $p = 2$ est totalement arbitraire). Donc par (3.3-66),

$$\int_{B(a, R)} |\nabla_x v(t, x)|_1 dx \leq \liminf_{\gamma \rightarrow 0} \leq \int_{B(a, R)} |\nabla_x v_\gamma(t, x)|_1 dx \leq \frac{C'}{t_0(1-t_0)} \mathcal{L}^2(B(a, R')).$$

□

III.3.10 Résultats complémentaires pour les résultats d'unicité à venir

Tous les résultats de cette sous-section sont des résultats de contrôle des champs de vitesse régularisés v_γ . Ceux-ci s'avéreront utile pour l'étude des propriétés d'unicité dans les points selles (voir la section IV.2).

Pour commencer, le premier point de la preuve de la Proposition précédente, en particulier la relation (3.3-66), nous permet en particulier d'énoncer directement le Lemme suivant :

Lemme 3.3-33.

On suppose que ϕ vérifie la propriété (Γ_1) . Soient $R' > R > 0$ et $a \in \mathbb{R}^d$ tel que $\phi(a) = \inf_{R^2} \phi$, et soit

$$M = \sup_{x \in B(a, 2(R+|a|))} |\partial \phi(x)|.$$

Alors il existe une constante $C' > 0$ – indépendante de ϕ , $\gamma\phi$ (et donc de γ), a , R et R' – telle que pour tout $t_0 \in]0, 1[$ vérifiant la condition

$$t_0 < \min \left\{ \frac{1}{2}, \frac{R' - R}{M + 2|a|} \right\},$$

on ait la propriété :

$$\forall \gamma > 0, \forall t \in]0, t_0], \int_{B(a, R)} |\nabla_x v_\gamma(t, x)|_1 dx \leq \frac{C'}{t_0(1-t_0)} \mathcal{L}^2(B(a, R')). \quad (3.3-67)$$

En Corollaire nous en déduisons :

Corollaire 3.3-34.

On suppose que ϕ vérifie la propriété (Γ_1) . Soit $\Omega \subset \mathbb{R}^d$ un ouvert borné et $0 < t_m < 1$. Alors il existe une constante $K > 0$ telle que pour tout $\gamma > 0$ et tout $t \in [0, t_m]$, on ait

$$\int_{\Omega} |\nabla_x v_\gamma(t, x)|_1 dx \leq K. \quad (3.3-68)$$

Preuve : Soient $a \in \mathbb{R}^d$ tel que $\phi(a) = \inf_{\mathbb{R}^d} \phi$, et $R' > R > 0$ tel que $\Omega \subset B(a, R)$. Soient

$$M = \sup_{x \in B(a, 2(R+|a|))} |\partial\phi(x)|,$$

et $t_0 \in]0, 1[$ tel que

$$t_0 < \min \left\{ \frac{1}{2}, \frac{R' - R}{M + 2|a|} \right\}.$$

D'après le Lemme 3.3-33 (p. 117), on a l'inégalité (3.3-67) pour une constante $C' > 0$ indépendante de γ , et d'après le Lemme 3.3-2 (p. 86) et le deuxième point du commentaire 3.3-12 (p. 91), il existe une constante $c > 0$ tel que pour tout $t \in]0, 1[$ et $x \in \mathbb{R}^d$,

$$|\nabla_x v(t, x)|_1 \leq \frac{c}{t(1-t)}.$$

On conclut donc en prenant

$$K = \max \left\{ \frac{C'}{t_0(1-t_0)} \mathcal{L}^2(B(a, R')), \sup_{t \in [t_0, t_m]} \frac{c}{t(1-t)} \mathcal{L}^2(\Omega) \right\}.$$

□

Nous terminerons par le Lemme suivant :

Lemme 3.3-35.

Soient $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ convexe vérifiant la propriété (Γ_1) . Alors il existe une constante c , indépendante de γ , tel que pour tout $1 \geq \gamma > 0$, tout $t \in]0, 1[$ et presque tout $x \in \mathbb{R}^d$ où $v_\gamma(t, \cdot)$ est différentiable,

$$\|\nabla_x v_\gamma(t, \cdot)\|_{L^\infty(\mathbb{R}^d)} \leq \frac{c}{\gamma(1-t)}.$$

Preuve : Soit $1 \geq \gamma > 0$. $\nabla(\gamma\phi)(x) = (x - \text{Prox}_{\gamma\phi}(x))/\gamma$ est $1/\gamma$ -lipschitzienne sur \mathbb{R}^d , donc par le Théorème de Rademacher (Théorème 3.3-16 p. 98), $\nabla(\gamma\phi)$ est presque partout différentiable sur \mathbb{R}^d et alors

$$\|\text{Hess}(\gamma\phi)\|_{L^\infty(\mathbb{R}^d)} \leq \frac{1}{\gamma},$$

où la norme L^∞ est ici prise sur $\|\cdot\|$: la norme subordonnée à la norme euclidienne $|\cdot|$ de \mathbb{R}^d .

D'après le Lemme 3.3-27 (p. 110), $\gamma\phi$ vérifie la propriété (Γ_2) , et ainsi d'après le Lemme 3.3-24 (p. 107), il existe une constante $C > 0$ tel que pour tout $t \in]0, 1[$ et presque tout $x \in \mathbb{R}^d$ où $v_\gamma(t, \cdot)$ est différentiable,

$$\|\nabla_x v_\gamma(t, \cdot)\|_{L^\infty(\mathbb{R}^d)} \leq \frac{C}{1-t} \left(\|\text{Hess}(\gamma\phi)\|_{L^\infty(\mathbb{R}^d)} + 2 \right) \leq \frac{C}{\gamma(1-t)} (1 + 2\gamma) \leq \frac{3C}{\gamma(1-t)}.$$

□

III.3.11 Un résultat indépendant mais notable

Il est intéressant de noter que le résultat de la Proposition 3.3-32 (p. 116) peut être généralisé et que l'on peut montrer par symétrie en temps que $\nabla_{t,x}v \in L^\infty((0, 1), L^1(\Omega))$ pour tout ouvert borné Ω .

Ce résultat n'est en soit pas utile pour les résultats d'existence et d'unicité à suivre, que préparent les autres résultats de section. Il est tout de même assez notable pour que nous l'énoncions : il offre un résultat de régularité et de contrôle du champ de vitesse d'un transport optimal isotrope pour la distance L^2 . Ces propriétés complètent les propriétés (III) (voir p. 76).

Théorème 3.3-36.

On suppose que ϕ vérifie la propriété (Γ_1) . Soit Ω un ouvert borné de \mathbb{R}^d . Alors $\nabla_{t,x}v \in L^\infty((0, 1), L^1(\Omega))$, autrement dit il existe une constante $K > 0$ telle que pour tout $t \in]0, 1[$,

$$\int_{\Omega} |\nabla_{t,x}v(t, x)|_1 dx \leq K. \quad (3.3-69)$$

Preuve :

1. On pose pour tout $t \in]0, 1[$ et $x \in \mathbb{R}^d$

$$v^*(t, x) = \frac{x - \mathbf{p}_{\phi^*}(t, x)}{t}, \quad \text{avec} \quad \mathbf{p}_{\phi^*}(t, x) = \text{Prox}_{\frac{t}{1-t}\phi^*}\left(\frac{x}{1-t}\right).$$

D'après le Lemme 3.3-11 (p. 92), on a

$$v(t, x) = -v^*(1-t, x) \quad (3.3-70)$$

et donc pour tout $t \in]0, 1[$,

$$\nabla_x v(t, \cdot) = -\nabla_x v^*(1-t, \cdot).$$

2. Soient $a, a^* \in \mathbb{R}^d$ tels que $\phi(a) = \inf_{\mathbb{R}^d} \phi$ et $\phi^*(a^*) = \inf_{\mathbb{R}^d} \phi^*$, et soient $R' > R > 0$ tels que $\Omega \subset B(a, R) \cap B(a^*, R)$. On pose de plus

$$M = \sup_{x \in B(a, 2(R+|a|))} |\partial\phi(x)| \quad \text{et} \quad M^* = \sup_{x \in B(a^*, 2(R+|a^*|))} |\partial\phi^*(x)|,$$

ainsi que $0 < t_0 < t_0^* < 1$ tels que

$$t_0 < \min \left\{ \frac{1}{2}, \frac{R' - R}{M + 2|a|} \right\} \quad \text{et} \quad t_0^* > 1 - \min \left\{ \frac{1}{2}, \frac{R' - R}{M^* + 2|a^*|} \right\}.$$

D'après la Proposition 3.3-32 (p. 116), il existe des constantes $C, C^* > 0$ telles que pour tout $t \in]0, t_0]$,

$$\int_{\Omega} |\nabla_x v(t, x)|_1 dx \leq \int_{B(a, R)} |\nabla_x v(t, x)|_1 dx \leq \frac{C}{t_0(1-t_0)} \mathcal{L}^2(B(a, R'))$$

et pour tout $t \in [t_0^*, 1[$,

$$\begin{aligned} \int_{\Omega} |\nabla_x v(t, x)|_1 dx &= \int_{\Omega} |\nabla_x v^*(1-t, x)|_1 dx \leq \int_{B(a^*, R)} |\nabla_x v^*(1-t, x)|_1 dx \\ &\leq \frac{C^*}{(1-t_0^*)(1-(1-t_0^*))} \mathcal{L}^2(B(a^*, R')) \\ &= \frac{C^*}{t_0^*(1-t_0^*)} \mathcal{L}^2(B(a^*, R')). \end{aligned}$$

D'après le Lemme 3.3-2 (p. 86) et le commentaire 3.3-12 (p. 91), il existe une constante $c > 0$ tel que pour tout $t \in]0, 1[$ et $x \in \mathbb{R}^d$,

$$|\nabla_x v(t, x)|_1 \leq \frac{c}{t(1-t)}.$$

En posant

$$K_0 = \max \left\{ \frac{C}{t_0(1-t_0)} \mathcal{L}^2(B(a, R')), \frac{C^*}{t_0^*(1-t_0^*)} \mathcal{L}^2(B(a^*, R')), \sup_{t \in [t_0, t_0^*]} \frac{c}{t(1-t)} \mathcal{L}^2(\Omega) \right\},$$

on a pour tout $t \in]0, 1[$:

$$\int_{\Omega} |\nabla_x v(t, x)|_1 dx \leq K_0.$$

3. On a donc prouvé que $\nabla_x v \in L^\infty((0, 1), L^1(\Omega))$. Pour conclure il suffit donc de prouver que l'on a également $\partial_t v \in L^\infty((0, 1), L^1(\Omega))$. Or on sait d'après la Proposition 3.3-12 (p. 93) que $v \in L^\infty((0, 1) \times \Omega)$, et on sait par la Proposition 3.3-31 (p. 115) que v vérifie l'équation de Burgers

$$\partial_t v + v \cdot \nabla_x v = 0,$$

ce qui implique immédiatement le résultat voulu.

On peut donc poser $K = \left(1 + \sup_{]0, 1[\times \Omega} |v| \right) K_0$, et ainsi conclure la preuve. □

On peut étendre ce résultat :

Corollaire 3.3-37.

On suppose que ϕ vérifie la propriété (Γ_1) . Soit Ω un ouvert borné. Pour tous $p, q \geq 1$ tels que $1/p + 1/q > 1$, on a $\nabla_{t,x} v \in L^p((0, 1), L^q(\Omega))$. Et en particulier $v \in W^{1,p}((0, 1) \times \Omega)$ pour tout $1 \leq p < 2$.

Preuve :

1. D'après le commentaire 3.3-12, on sait qu'il existe une constante $c' > 0$ telle que pour tout $t \in]0, 1[$ et $x \in \mathbb{R}^d$,

$$|\nabla_{t,x} v(t, x)|_1 \leq \frac{c'}{t(1-t)}.$$

D'après le Théorème 3.3-36, il existe $K > 0$ tel que pour tout $t \in]0, 1[$,

$$\int_{\Omega} |\nabla_{t,x} v(t, x)|_1 dx \leq K.$$

Donc, pour tous $p, q \geq 1$ tels que $1/p + 1/q > 1$, on a pour tout $t \in]0, 1[$,

$$\left| \int_{\Omega} |\nabla_{t,x} v(t, x)|_1^q dx \right|^{\frac{p}{q}} \leq \frac{c'}{t^a(1-t)^a} \left| \int_{\Omega} |\nabla_{t,x} v(t, x)|_1^q dx \right|^{\frac{p}{q}} \leq \frac{K^{p/q} \times c'}{t^a(1-t)^a}.$$

avec $a = p(q-1)/q$. Or $t \mapsto 1/t^a(1-t)^a$ est intégrable sur $]0, 1[$ si et seulement si $a < 1$, i.e. $1/p + 1/q > 1$. On a donc bien

$$\int_0^1 \left| \int_{\Omega} |\nabla_{t,x} v(t, x)|_1^q dx \right|^{\frac{p}{q}} dt \leq (K^{p/q} \times c') \int_0^1 \frac{1}{t^a(1-t)^a} dt < +\infty,$$

d'où $\nabla_{t,x} v \in L^p((0, 1), L^q(\Omega))$.

2. En particulier si l'on prend $1 \leq p = q < 2$, on a bien $1/p + 1/q = 2/p > 1$, d'où $\nabla_{t,x} v \in L^p((0, 1) \times \Omega)$. Comme on sait de plus, d'après la Proposition 3.3-12 (p. 93), que $v \in L^\infty((0, 1) \times \Omega) \subset L^p((0, 1) \times \Omega)$, on a bien $v \in W^{1,p}((0, 1) \times \Omega)$.

□

Le Théorème 3.3-36 et son Corollaire 3.3-37 donnent bien les hypothèses de régularité les plus fortes que l'on puisse avoir en général pour le champ de vitesse v . En effet, d'après le Corollaire 3.3-37, un champ de vitesse v défini par rapport à un potentiel ϕ (vérifiant la propriété (Γ_1)), rappelons-le, par la relation (3.3-11) (p. 89), c'est à dire pour tout $t \in]0, 1[$ par

$$v(t, \cdot) = v_\phi(t, \cdot) = \frac{\text{id} - \mathbf{p}_\phi(t, \cdot)}{t} = \frac{1}{t} (\text{id} - (t\phi + (1-t)\text{id})^*), \quad (3.3-71)$$

est un élément de $W_{loc}^{1,p}([0, 1] \times \mathbb{R}^d)$ (c'est à dire sa restriction à tout ouvert borné Ω de \mathbb{R}^d est un élément de $W^{1,p}((0, 1) \times \Omega)$) et ceci pour tout $p < 2$. La question qui se pose naturellement est de savoir si $v = v_\phi$ ne pourrait pas être en fait un élément de $H_{loc}^1([0, 1] \times \mathbb{R}^d)$. Or, comme nous l'avons déjà évoqué plusieurs fois dans ce chapitre, plus récemment au début de la sous-section III.3.7, ce n'est en général pas le cas, et ceci en particulier lorsque le champ de vitesse se retrouve confronté à des "cassures". Voici un contre-exemple simple :

Exemple 3.3-22. EXEMPLE D'UN CHAMP DE VITESSE QUI N'EST PAS DE RÉGULARITÉ H^1

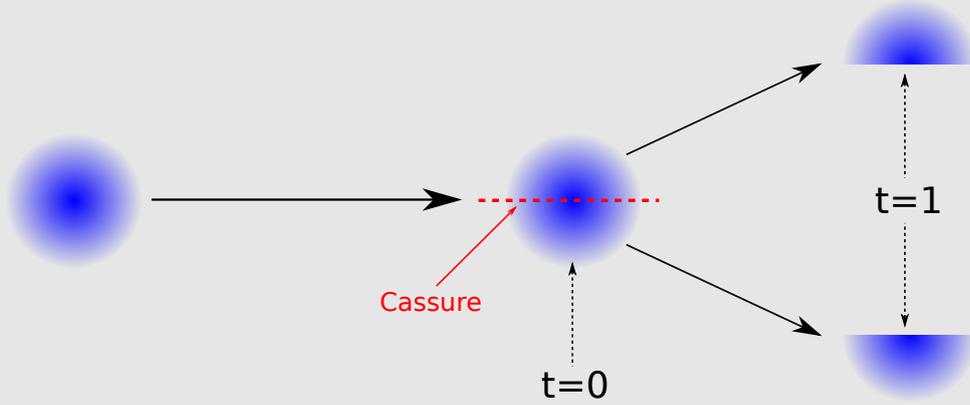
Nous reprenons ici l'exemple que Caffarelli [19, 20] avait donné pour prouver qu'en général, un plan de transport optimal n'est pas régulier (ici non-continue). Il s'agit d'un plan de transport présentant dans \mathbb{R}^2 une cassure le long de l'axe des abscisses et translatant dans les deux sens opposés de l'axe des ordonnées les densités situées de part et d'autre de cette cassure. Nous allons voir que le long de cette "cassure", le champ de vitesse n'est pas de régularité H^1 .

Pour donner un sens local à un tel plan de transport et à un tel champ de vitesse, considérons une densité dont le support supposé borné est coupé par l'axe des des abscisses (donc par la cassure). Soit le potentiel ϕ défini par :

$$\phi : (x, y) \mapsto \frac{1}{2}(|x|^2 + |y|^2) + |y|, \quad (3.3-72)$$

dont le plan de transport T associé (défini \mathcal{L}^2 -presque partout) correspond donc à

$$T = \nabla\phi : (x, y) \mapsto \begin{cases} (x, y + 1), & \text{si } y > 0, \\ (x, y - 1), & \text{si } y < 0. \end{cases}$$



D'après le Lemme 3.2-7 (p. 82), le potentiel ϕ vérifie l'hypothèse (Γ_1) sur \mathbb{R}^2 . Le champ de vitesse associé v_ϕ (c'est à dire défini par rapport à ϕ par la relation (3.3-71)) s'écrit alors, pour tout $t \in]0, 1[$ et pour tout $(x, y) \in \mathbb{R}^2$,

$$v_\phi(t, x, y) = \begin{cases} (0, 1), & \text{si } y > t, \\ (0, \frac{y}{t}), & \text{si } -t \leq y \leq t, \\ (0, -1), & \text{si } y < -t. \end{cases} \quad (3.3-73)$$

Ainsi, pour tout $t \in]0, 1[$ et \mathcal{L}^2 -presque tout $(x, y) \in \mathbb{R}^2$, on a

$$\nabla_{x,y} v_\phi(t, x, y) = \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{t} \mathbf{1}_{[-t,t]}(y) \end{pmatrix}, \quad \text{et ainsi } |\nabla_{x,y} v_\phi(t, x, y)|_1 = \frac{1}{t} \mathbf{1}_{[-t,t]}(y).$$

Si l'on se place sur le carré $] -1, 1[^2$, on a alors, pour tout $p > 1$,

$$\int_{]-1, 1[^2} |\nabla_{x,y} v_\phi(t, x, y)|_1^p dx dy = \frac{4}{t^{p-1}}.$$

Or, la fonction $t \mapsto 1/t^{p-1}$ est intégrable sur $]0, 1[$ si et seulement si $p < 2$. Ainsi, en illustration du Corollaire 3.3-37, on a bien $v_\phi \in W^{1,p}((0, 1) \times] -1, 1[^2)$ pour tout $p < 2$.

Cependant ici, on a $v_\phi \notin H^1((0, 1) \times] -1, 1[^2)$!

Remarque 3.3-23. A PROPOS DE L'EXEMPLE 3.3-22

Dans l'exemple 3.3-22, nous avons donné dans l'équation (3.3-73) l'expression du champ de

vitesse v_ϕ calculé à partir du potentiel ϕ défini en (3.3-72), et plus précisément à partir de l'opérateur \mathbf{p}_ϕ comme dans la formule (3.3-71). Pour obtenir le champ de vitesse v_ϕ de l'équation (3.3-73), cette opérateur \mathbf{p}_ϕ s'écrit nécessairement, pour tout $t \in]0, 1[$ et pour tout $(x, y) \in \mathbb{R}^2$, tel qu'il suit :

$$\mathbf{p}_\phi(t, x, y) = \begin{cases} (x, y) + (0, -t), & \text{si } y > t, \\ (x, 0), & \text{si } -t \leq y \leq t, \\ (x, y) + (0, t), & \text{si } y < -t. \end{cases} \quad (3.3-74)$$

Le détail de ce calcul fait l'objet de l'Annexe D.

III.4 Résumé et conclusion du chapitre

La section III.1 a eu pour objectif la planification du chapitre, ainsi que la définition de différents outils, notations et propriétés nécessaires à ce chapitre (et au suivant), en particulier les propriétés (I), (II), (III), (I' $_\alpha$) et (I' $_\beta$).

Dans la section III.2, nous nous sommes donné pour objectif d'affiner, pour les besoins de notre étude, les résultats du Théorème 1.1-7 (p. 20) en ce qui concerne la régularité du potentiel ϕ : le Théorème 1.1-7 prouve que, dans les conditions de notre problème, tout plan de transport optimal sur \mathbb{R}^d dérive d'un potentiel ϕ convexe, propre et s.c.i sur \mathbb{R}^d ; et nous avons prouvé dans cette section qu'il était possible de choisir un tel ϕ de façon à ce que celui-ci **vérifie la propriété** (Γ_1), que nous rappelons ici :

Propriété (Γ_1). ϕ et ϕ^* sont convexes, continues et admettent un minimum sur \mathbb{R}^d .

La section III.3 a quant à elle eu pour principal objet l'étude des différentes propriétés du champ de vitesse d'un transport optimal, que nous proposons de compiler dans la Proposition suivante :

Proposition 3.4-38. RÉSUMÉ DES PROPRIÉTÉS DU CHAMP DE VITESSE

Soit ϕ une fonction convexe sur \mathbb{R}^d vérifiant la propriété (Γ_1), et soit un champ de vitesse v défini par rapport à ϕ par (3.3-11) (p. 89), c'est à dire pour tout $t \in]0, 1[$ par

$$v(t, \cdot) = \frac{\text{id} - \mathbf{p}(t, \cdot)}{t} = \frac{1}{t} (\text{id} - \nabla_x(t\phi + (1-t)\text{id})^*), \quad (3.4-75)$$

où est $\mathbf{p}(t, \cdot)$ lui-même défini par

$$\mathbf{p}(t, \cdot) = \nabla_x(\phi_t)^* = \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{x}{1-t} \right), \text{ avec } \phi_t = t\phi + (1-t)\text{id}.$$

Alors v vérifie les propriétés (II) et (III), c'est à dire :

— **Propriétés (II) :**

1. Il existe $\psi \in W_{loc}^{1,\infty}([0, 1] \times \mathbb{R}^d)$ tel que $v = \nabla_x \psi$.
2. Le champ de vitesse v vérifie l'équation de Burgers au sens des distributions

$$\partial_t v + \frac{1}{2} \nabla_x |v|^2 = 0.$$

— **Propriétés (III) :**

1. v est localement lipschitzienne sur $]0, 1[\times \mathbb{R}^d$.

2. $\nabla_{t,x}v \in L^\infty((0, 1), L^1_{loc}(\mathbb{R}^d))$,
3. $\nabla_{t,x}v \in L^p((0, 1), L^q_{loc}(\mathbb{R}^d))$ avec $1/p + 1/q > 1$.

De plus, pour tous $t, s \in]0, 1[$, et pour tout $x \in \mathbb{R}^d$,

$$v(t, x) = v(s, (s - t)v(t, x) + x). \quad (3.4-76)$$

Preuve (par compilation des résultats de la section III.3) :

- La propriété (II)-1 est démontrée par la Proposition 3.3-12 (p. 93).
- La propriété (II)-2 est démontrée par la Proposition 3.3-31 (p. 115).
- La propriété (III)-1 est démontrée par la Proposition 3.3-15 (p. 96).
- La propriété (III)-2 est démontrée par le Théorème 3.3-36 (p. 119).
- La propriété (III)-3 découle de la précédente et est démontrée par le Corollaire 3.3-37 (p. 120).
- L'équation (3.4-76) est démontrée par la Proposition 3.3-14 (p. 96).

□

Le second point important de la section III.3 est l'ensemble des résultats énoncés dans la sous-section III.3.10, ceux-ci portant sur le contrôle des gradients des champs régularisés de v (via les enveloppe de Moreau du potentiel ϕ), à savoir les v_γ . Ces propriétés, ainsi que nous l'avons déjà mentionné, se révéleront très utiles dans les preuves des principaux résultats relatifs à l'unicité de la composante $\mu = (\rho, m)$ des points selles de \mathbf{L} , ainsi qu'à la caractérisation des champs de vitesse associés à un transport optimal dans un espace L^2 (voir le chapitre suivant).

Chapitre IV

Existence du point selle et propriétés d'unicité

Notes pour ce chapitre.

- Nous rappelons que $|\cdot|$ désigne la norme euclidienne sur \mathbb{R}^d ($d \in \mathbb{N}$).
- Sauf mention contraire, \mathbb{R}^d sera muni de la tribu de Lebesgue.
- Pour $f \in L^p(\omega)$ et $x \in \omega$, le terme " $f(x)$ " désignera, par abus de notation, une valeur $\tilde{f}(x)$ où la fonction \tilde{f} est une représentante de la classe d'équivalence f pour la mesure de Lebesgue (i.e. pour la relation d'équivalence basée sur l'égalité \mathcal{L}^d presque partout).
- Par abus de notation, pour une densité f , nous confondrons f et la mesure $f \mathcal{L}^d$.
- Toujours par abus de notation, pour toute fonction f définie sur \mathbb{R}^d , et tout sous-ensemble ω de \mathbb{R}^d , on désignera ici indifféremment f et $f|_\omega$ par la même notation f . Inversement, pour tout sous-ensemble ω de \mathbb{R}^d , et toute fonction f définie sur ω , on désignera indifféremment f et son prolongement par 0 sur \mathbb{R}^d par la même notation f .
- De même, pour toute mesure ν définie sur la tribu de Lebesgue \mathbb{R}^d , et tout sous-ensemble Lebesgue-mesurable ω de \mathbb{R}^d , on désignera ici indifféremment ν et $\nu|_\omega$ (où $\nu|_\omega$ est défini pour tout sous-ensemble Lebesgue-mesurable A de ω par $\nu|_\omega(A) = \nu(A)$) par la même notation ν . Inversement, pour tout sous-ensemble Lebesgue-mesurable ω de \mathbb{R}^d , et toute mesure ν définie sur la tribu de Lebesgue de ω , on désignera indifféremment ν et son prolongement ν_0 par 0 sur \mathbb{R}^d par la même notation ν : Le prolongement par 0 de ν est définie plus précisément pour tout sous-ensemble Lebesgue-mesurable A de \mathbb{R}^d par $\nu_0(A) = \nu(A \cap \omega)$.
- Par abus de notation, lorsqu'il n'y aura pas d'ambiguïté sur ϕ , nous désignerons par v le champ de vitesse v_ϕ défini par rapport à ϕ comme en (3.3-11) (p. 89), c'est à dire

$$v_\phi(t, x) = \frac{1}{t} \left(x - \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{x}{1-t} \right) \right),$$

pour tout $t \in]0, 1[$ et $x \in \mathbb{R}^d$.

- Dès lors que rien ne sera précisé, Ω désignera un ouvert convexe et borné de \mathbb{R}^d .
- Nous désignerons un intervalle entre a et b par la notation (a, b) lorsque l'ouverture et la fermeture de cet intervalle en a ou b est sans importance. Au contraire, lorsque ces bornes sont précisées, sauf erreur, cela signifie que ces bornes ont une importance. Par exemple lorsque l'on écrit $f \in L_{loc}^\infty([0, 1] \times \mathbb{R}^d)$, cela signifie en particulier que $f \in L^\infty([0, 1] \times K)$ pour tout compact K de \mathbb{R}^d : autrement dit que f est " L^∞ localement en espace mais pas en temps" (i.e. reste localement borné en $t = a$ et $t = b$).

Ce chapitre se donne pour objectif l'étude de résultats d'existence et d'unicité sur les points selles de \mathbf{L} par la mise en oeuvre des différents résultats du chapitre III.

Plan du chapitre

Ce chapitre se découpera en deux sections, la première étant consacrée à l'existence d'un point selle (ψ, q, μ) pour le lagrangien \mathbf{L} , la seconde à l'unicité de la composante $\mu = (\rho, \rho v)$ parmi ces points selles. Cette deuxième section sera également consacrée à un résultat de caractérisation des champs de vitesse définissant un transport optimal dans L^2 .

1. Dans la section IV.1, nous montrerons l'existence d'un point selle au lagrangien \mathbf{L} :
 - Nous montrerons tout d'abord la propriété $(P_{\beta 1'})$ dans la sous-section IV.1.1, à savoir que $\rho \in L^2((0, 1) \times \Omega)$.
 - Nous établirons ensuite la propriété $(P_{\beta 4'})$, à savoir que ρ et v vérifient l'équation de conservation de la masse : $\forall h \in C^\infty([0, 1] \times \bar{\Omega})$,

$$\int_{]0,1[\times \Omega} (\partial_t h + v \cdot \nabla h) \rho + \int_{\Omega} h(0, x) \rho_0(x) dx - \int_{\Omega} h(1, x) \rho_1(x) dx = 0$$

- Enfin, dans la sous-section IV.1.3, nous compilerons les résultats des deux-sections précédentes pour conclure finalement à l'existence d'un point selle au lagrangien \mathbf{L} : et ainsi démontrer le Théorème 3.1-2.
2. En second lieu, dans la section IV.2, nous montrerons l'unicité de la composante $\mu = (\rho, \rho v)$ parmi les points selles de \mathbf{L} , en démontrant le Théorème 4.2-1 :
 - Nous commencerons, dans la sous-section IV.2.1, par justifier cette étude d'unicité spécifique à la composante $\mu = (\rho, m)$, en montrant qu'en général l'ensemble des points selles (ψ, q, μ) de \mathbf{L} n'est pas réduite à un unique élément.
 - Dans la sous-section IV.2.2, nous montrerons l'unicité, dans le cadre de notre problème, du champ de vitesse sur l'union des support de l'ensemble des densités ρ possibles pour former un point selle (Lemme 4.2-3, p. 142). Plus précisément, cela signifie qu'il existe un champ de vitesse v telle que pour toute composante $\mu = (\tilde{\rho}, \tilde{m})$ d'un point selle de \mathbf{L} , on a $\tilde{m} = \tilde{\rho} v$.
 - Nous montrerons ensuite, dans la sous-section IV.2.3, l'unicité de la densité ρ (et donc de la composante $\mu = (\rho, \rho v)$), via le Théorème 4.2-1 et la Proposition 4.2-5 p. 146, principaux résultats de cette sous-section.
 3. Pour finir, dans la section IV.3, nous tacherons de généraliser les résultats des sous-sections IV.2.2 et IV.2.3. Nous caractériserons plus précisément un champ de vitesse déterminant un transport optimal (Théorème 4.3-9, p. 159).

IV.1 Construction d'un point selle du Lagrangien

On rappelle que Ω désigne un ouvert convexe et borné de \mathbb{R}^d et l'on s'intéresse au transport optimal entre deux densités $\rho_0, \rho_1 \in L^2(\Omega)$ à supports compacts dans Ω . Nous avons vu dans la section III.2 qu'il existait alors un potentiel ϕ sur \mathbb{R}^d vérifiant la propriété (Γ_1) , et vérifiant $\rho_1 \mathcal{L}^d = \nabla \phi \# (\rho_0 \mathcal{L}^d)$ (en considérant les prolongements par 0 à \mathbb{R}^d des densités ρ_0 et ρ_1).

Nous avons vu au chapitre II (Proposition 2.2-1 p. 56) un point selle de \mathbf{L} est caractérisé par les propriétés (I), que nous rappelons :

Propriétés (I).

$$\left. \begin{array}{l}
 (\psi, q, \mu) \in S \text{ vérifie :} \\
 (P1) \quad \forall q' \in \tilde{\mathcal{P}}, \quad \langle \mu, q' - q \rangle \leq 0, \\
 (P2) \quad \forall h \in H^1((0, 1) \times \Omega), \quad G(h) + \langle \mu, \nabla_{t,x} h \rangle = 0, \\
 (P3) \quad \nabla_{t,x} \psi = q,
 \end{array} \right|$$

où l'ensemble S est défini par :

$$S = (H^1((0, 1) \times \Omega)/\mathbb{R}) \times L^2((0, 1) \times \Omega)^{d+1} \times L^2((0, 1) \times \Omega)^{d+1},$$

le parabolöide $\tilde{\mathcal{P}}$ par :

$$\tilde{\mathcal{P}} = \left\{ (a, b) \in L^2((0, 1) \times \Omega) \times L^2((0, 1) \times \Omega)^d, \quad a + \frac{|b|^2}{2} \leq 0 \right\},$$

et la forme linéaire G par $G(h) = \int h(0, \cdot) \rho_0 dx - \int h(1, \cdot) \rho_1 dx$, pour tout $h \in H^1((0, 1) \times \Omega)$.

On rappelle également que pour construire un triplet (ψ, q, μ) vérifiant ces dernières propriétés, il suffit de construire un couple densité-champ de vitesse (ρ, v) vérifiant les propriétés (I'_β) , avec $v = \nabla_x \psi$ pour un potentiel $\psi \in L^1_{loc}([0, 1] \times \Omega)$. En effet, comme nous le démontrerons dans le Théorème IV.1.3 (p. 138), en posant $\mu = (\rho, \rho v)$ et $q = (-(1/2)|v|^2, v)$, il existe alors $\psi^* \in W^{1,\infty}((0, 1) \times \Omega)/\mathbb{R}$, tel que (ψ^*, q, μ) vérifie les propriétés (I), et soit donc un point selle de \mathbf{L} . Nous rappelons ci-dessous les propriétés (I'_β) :

Propriétés (I'_β) .

- $$\left. \begin{array}{l}
 (P_{\beta 1}') \quad \text{La densité } \rho \text{ vérifie } \rho \in L^2((0, 1) \times \Omega). \\
 (P_{\beta 2}') \quad \text{Le champ de vitesse } v \text{ vérifie } v \in L^\infty((0, 1) \times \Omega)^d. \\
 (P_{\beta 3}') \quad \text{Le champ de vitesse } v \text{ vérifie au sens des distributions l'équation de Burgers :}
 \end{array} \right|$$

$$\partial_t v + \frac{1}{2} \nabla_x |v|^2 = 0$$

- $$\left. \begin{array}{l}
 (P_{\beta 4}') \quad (\rho, \rho v) \text{ vérifie l'équation de conservation de la masse au sens des distributions, pour} \\
 \text{les conditions initiales et finales } \rho_0 \text{ et } \rho_1 \text{ et des conditions de Neumann homogènes} \\
 \text{aux bords : } \forall h \in C^\infty([0, 1] \times \bar{\Omega}),
 \end{array} \right|$$

$$\int_{]0,1[\times \Omega} (\partial_t h + v \cdot \nabla h) \rho + \int_{\Omega} h(0, x) \rho_0(x) dx - \int_{\Omega} h(1, x) \rho_1(x) dx = 0 \quad (4.1-1)$$

La densité ρ que nous choisirons comme candidate sera bien entendu la densité d'interpolation de McCann définie par $\rho : t \mapsto \rho_t$, où pour tout $t \in [0, 1]$ on a :

$$\rho_t \mathcal{L}^d = [(1-t) \text{id} + t \nabla \phi] \# (\rho_0 \mathcal{L}^d) = \nabla \phi_t \# (\rho_0 \mathcal{L}^d),$$

avec $\phi_t = (1-t)|\cdot|^2/2 + t\phi$. Quant à v , nous avons déjà choisi comme candidat au chapitre précédent le champ de vitesse définie par :

$$v = v_\phi(t, x) = \frac{x - \mathbf{p}_\phi(t, x)}{t}, \quad \text{avec} \quad \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{x}{1-t} \right) \quad (= \nabla_x(\phi_t)^*).$$

Observons qu'un tel champ v vérifie bien $v = \nabla_x \psi$ en prenant pour tout $t \in]0, 1[$:

$$\psi(t, \cdot) = \frac{1}{t} \left(\frac{1}{2} |\cdot|^2 - (\phi_t)^* \right) \in L^1_{loc}([0, 1[\times \Omega).$$

Nous avons établi dans ce même chapitre III qu'un tel champ v vérifiait bien la condition $(P_\beta 2')$ (par la Proposition 3.3-12 (p. 93)) ainsi que la condition $(P_\beta 3')$ (par la Proposition 3.3-31 p. 115).

Il ne nous reste donc plus, pour conclure à l'existence d'un point selle pour \mathbf{L} , qu'à établir la condition $(P_\beta 1')$ sur ρ , ce qui sera l'objet principal de la sous-section IV.1.1 ; ainsi que la condition $(P_\beta 4')$ sur (ρ, v) , ce qui sera l'objet de la sous-section IV.1.2.

Toutes les sections et sous-sections qui suivent feront intervenir très fréquemment la notion de poussée de mesure : nous renvoyons donc à l'Annexe F pour plus de détails à ce propos. Nous énonçons tout de même ici les deux Propositions de cette même Annexe auxquelles nous ferons le plus souvent appel (les preuves de ces dernières se trouvent dans l'Annexe en question).

La première donne une définition précise de la notion de poussée de mesure :

Proposition F.0-3.

On munit \mathbb{R}^d de la tribu de Lebesgue. Soient μ une mesure positive σ -finie sur la tribu de Lebesgue, et $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$ mesurable (toujours sur la tribu de Lebesgue). On suppose que l'une des deux hypothèses suivantes est vérifiée :

1. *La mesure μ est finie,*
2. *L'application T est coercive par rapport à μ , i.e. $T^{-1}(A)$ est borné μ -presque partout pour tout ensemble Lebesgue-mesurable borné A dans \mathbb{R}^d .*

Alors il existe une mesure positive sur la tribu de Lebesgue ν sur \mathbb{R}^d , vérifiant les propriétés (IV) p. 203, telle que

$$\forall f \in C_c^0(\mathbb{R}^d), \quad \int_{\mathbb{R}^d} f d\nu = \int_{\mathbb{R}^d} f(Tx) d\mu \quad (4.1-2)$$

De plus, pour tout ensemble Lebesgue-mesurable $A \subset \mathbb{R}^d$, on a $\nu(A) = \mu(T^{-1}(A))$.

*On dit alors que ν est la **poussée de la mesure μ** par l'opérateur T , noté*

$$\nu = T\#\mu.$$

La seconde énonce des propriétés plus précise quant à son utilisation dans le cadre de l'interpolation de McCann :

Proposition F.0-6.

Soient $\rho_0 \in L^1(\mathbb{R}^d, d\mathcal{L}^d)$ à support compact, tel que $\rho_0 \geq 0$, et $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ vérifiant la propriété (Γ_1) .

Alors, pour tout $t \in [0, 1[$, il existe une mesure positive μ_t sur la tribu de Lebesgue de \mathbb{R}^d , à support dans $(t\nabla\phi + (1-t)\text{id})(\text{supp}(\rho_0))$ (donc borné), vérifiant les propriétés (IV) p. 203, telle que $\mu_t = (t\nabla\phi + (1-t)\text{id})\#(\rho_0\mathcal{L}^d)$, i.e.

$$\forall f \in C_c^0(\mathbb{R}^d), \quad \int_{\mathbb{R}^d} f d\mu_t = \int_{\mathbb{R}^d} f(t\nabla\phi(x) + (1-t)x)\rho_0(x) d\mathcal{L}^d(x) \quad (4.1-3)$$

De plus, μ_t est absolument continue par rapport à la mesure de Lebesgue \mathcal{L}^d sur \mathbb{R}^d (i.e. $\mu_t \ll \mathcal{L}^d$), autrement dit il existe $\rho_t \in L^1(\mathbb{R}^d, \mathcal{L}^d)$ tel que $\mu_t = \rho_t\mathcal{L}^d$. Il est en outre possible, pour tout $t \in [0, 1[$, de choisir un représentant de ρ_t de tel façon à ce que $(t, x) \mapsto \rho_t(x)$ soit mesurable sur $[0, 1[\times \mathbb{R}^d$.

Pour finir, les propriétés suivantes sont vérifiées :

1. Pour tout $h \in L_{loc}^\infty(\mathbb{R}^d, \mathcal{L}^d)$, on a
 - $h \circ (t\nabla\phi + (1-t)\text{id}) \in L_{loc}^\infty(\mathbb{R}^d, \mathcal{L}^d)$,
 - $\int_{\mathbb{R}^d} h d\mu_t = \int_{\mathbb{R}^d} h(x)\rho_t(x) d\mathcal{L}^d(x) = \int_{\mathbb{R}^d} h(t\nabla\phi(x) + (1-t)x)\rho_0(x) d\mathcal{L}^d(x)$.
2. Pour tout $h \in L_{loc}^\infty([0, 1[\times \mathbb{R}^d, \mathcal{L}^{d+1})$, on a
 - $(t, x) \mapsto h(t, t\nabla\phi(x) + (1-t)x) \in L_{loc}^\infty([0, 1[\times \mathbb{R}^d)$,
 - $\int_0^1 \int_{\mathbb{R}^d} h(t, x)\rho_t(x) d\mathcal{L}^d(x) dt = \int_0^1 \int_{\mathbb{R}^d} h(t, t\nabla\phi(x) + (1-t)x)\rho_0(x) dx dt$.
3. Pour tout $h \in C_c^0([0, 1[\times \mathbb{R}^d)$, $t \mapsto \int_{\mathbb{R}^d} h(t, \cdot) d\mu_t = \int_{\mathbb{R}^d} h \rho_t d\mathcal{L}^d$ est continue sur $[0, 1[$, autrement dit $t \mapsto \mu_t$ est continue sur $[0, 1[$ dans $\mathcal{D}'(\mathbb{R}^d)$.

Désormais, lorsque nous évoquerons une densité $t \mapsto \rho_t$, cela sera toujours (implicitement si rien n'est précisé) selon la définition qui suit :

Définition 4.1-1.

Soit $\rho_0 \in L^2(\mathbb{R}^d)$, $\rho_0 \geq 0$ à support borné dans \mathbb{R}^d , et soit $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ vérifiant la propriété (Γ_1) . Pour tout $t \in [0, 1[$, on définit la mesure $\mu_t = (t\nabla\phi + (1-t)\text{id})\#(\rho_0\mathcal{L}^d)$, et on définit la densité ρ_t comme l'une des densités définies dans la Proposition IV.1 : c'est à dire telle que $\mu_t = \rho_t\mathcal{L}^d$, que $(t, x) \mapsto \rho_t(x)$ soit mesurable sur $[0, 1[\times \mathbb{R}^d$ et vérifie les différentes propriétés énoncées dans cette Proposition.

IV.1.1 Propriétés de régularité temporelle de la densité

Nous commençons par énoncer un Théorème issu de [1], qui s'avérera être une version un peu plus générale de la notion **d'énergie interne** déjà évoquée dans le chapitre I, à la définition 1.1-11 (p. 32). Celui-ci sera la base de toutes les preuves des énoncés de cette sous-section, dont les résultats principaux sont le Lemme 4.1-3 (plus précisément son Corollaire) et le Lemme 4.1-5, ceux-ci établissant la vérification de la propriété $(P_\beta 1')$ par la densité ρ .

On rappelle que pour $d \in \mathbb{N}$ et pour tout $1 \leq p < +\infty$, $\mathcal{P}_p(\mathbb{R}^d)$ désigne l'ensemble des mesures de probabilité μ sur la tribu de Lebesgue \mathbb{R}^d boréliennes, telles que $\int_{\mathbb{R}^d} |x|^p d\mu < +\infty$ (définition

1.1-5 et notation 1 p. 19).

Théorème 4.1-2. GÉODÉSIQUE CONVEXITÉ DANS L'ESPACE DE WASSERSTEIN ([1] p.212)

Soit $1 \leq p < +\infty$ et soit $F :]-\infty, +\infty[\rightarrow [0, +\infty]$ une fonction propre, convexe, semi-continue inférieurement telle que $F(0) = 0$.

Considérons la fonction $\mathcal{F} : \mathcal{P}_p(\mathbb{R}^d) \rightarrow [0, +\infty]$ définie pour toute mesure $\mu \in \mathcal{P}_p(\mathbb{R}^d)$ par

$$\mathcal{F}(\mu) = \begin{cases} \int_{\mathbb{R}^d} F(\rho(x)) d\mathcal{L}^d(x) & \text{si } \mu = \rho \cdot \mathcal{L}^d \in \mathcal{P}_p(\mathbb{R}^d), \\ +\infty & \text{sinon} \end{cases}$$

Si la fonction $s \mapsto s^d F(s^{-d})$ est convexe et non-croissante sur $[0, +\infty[$, alors \mathcal{F} est géodésiquement convexe sur $\mathcal{P}_p(\mathbb{R}^d)$ au sens de McCann. Autrement dit, pour tout $\mu_0, \mu_1 \in \mathcal{P}_p(\mathbb{R}^d)$, $t \mapsto \mathcal{F}(\mu_t)$ est convexe sur $[0, 1]$, où $\mu_t = ((1-t)\text{id} + t\mathbb{T})\#\mu_0$ avec \mathbb{T} le plan de transport optimal de μ_0 sur μ_1 (gradient d'une fonction convexe ϕ si μ_0 est assez régulière d'après le Théorème 1.1-7 p. 20).

Nous allons appliquer ce Théorème pour démontrer la continuité dans $L^p(\Omega)$ de la densité $t \mapsto \rho_t$.

Lemme 4.1-3.

Soit $1 < p < +\infty$. Soient Ω un ouvert \mathbb{R}^d , $a < b$ et $(\mu_t)_{t \in [a, b]}$ une famille de mesures positives sur Ω , définies sur la tribu de Lebesgue, telle que $t \mapsto \mu_t$ soit continue sur $[a, b]$ dans $\mathcal{D}'(\Omega)$, i.e. que pour tout $f \in C_c^\infty(\Omega)$, la fonction

$$t \mapsto \int_{\Omega} f d\mu_t$$

soit continue sur $[a, b]$. Soit la fonction $\Lambda_p : [a, b] \rightarrow [0, +\infty]$ définie pour tout $t \in [a, b]$ par

$$\Lambda_p(t) = \begin{cases} \int_{\mathbb{R}^d} |\rho_t|^p dx & \text{si } \mu_t = \rho_t \cdot \mathcal{L}^d, \text{ avec } \rho_t \in L^1(\Omega, \mathcal{L}^d) \\ +\infty & \text{sinon} \end{cases}$$

Si Λ_p est convexe et finie sur $[a, b]$, alors pour tout $t \in [a, b]$, $\mu_t \ll \mathcal{L}^d$, avec $\mu_t = \rho_t \cdot \mathcal{L}^d$, où $t \mapsto \rho_t \in C^0([a, b], L^p(\Omega))$ (continuité forte).

Preuve : Tout d'abord Λ_p étant convexe et finie (donc bornée) sur $[a, b]$, il est donc clair, par définition de Λ_p que pour tout $t \in [a, b]$, $\mu_t \ll \mathcal{L}^d$, avec $\mu_t = \rho_t \cdot \mathcal{L}^d$ et $\rho_t \in L^2(\Omega)$. De plus $t \mapsto \|\rho_t\|_{L^p(\Omega)}$ est borné sur $[a, b]$ par une constante M .

Montrons maintenant que $t \mapsto \rho_t$ est **faiblement** continue de $[a, b]$ dans $L^p(\Omega)$: soient $t_0 \in [a, b]$, $q \in]1, +\infty[$ tel que $1/p + 1/q = 1$, et soit $f \in L^q(\Omega)$.

Soit $\epsilon > 0$. L'espace $C_c^\infty(\Omega)$ étant dense dans $L^q(\Omega)$, il existe donc $g \in C_c^\infty(\Omega)$ telle que $\|f - g\|_{L^q(\Omega)} \leq \epsilon/(3M)$. $t \mapsto \mu_t$ étant continue sur $[a, b]$ dans $\mathcal{D}'(\Omega)$, il existe donc $\delta > 0$ tel que pour tout $t \in [a, b]$ vérifiant $|t - t_0| \leq \delta$, on ait :

$$\left| \int_{\Omega} g \rho_t dx - \int_{\Omega} g \rho_{t_0} dx \right| \leq \epsilon/3.$$

D'où, pour tout $t \in [a, b]$ tel que $|t - t_0| \leq \delta$,

$$\begin{aligned} \left| \int_{\Omega} f \rho_t dx - \int_{\Omega} f \rho_{t_0} dx \right| &\leq (\|\rho_{t_0}\|_{L^p} + \|\rho_t\|_{L^p}) \|f - g\|_{L^q} + \left| \int_{\Omega} g \rho_t dx - \int_{\Omega} g \rho_{t_0} dx \right| \\ &\leq 2M \|f - g\|_{L^q} + \epsilon/3 \leq \epsilon. \end{aligned}$$

On conclut donc à la continuité **faible** de $[a, b]$ dans $L^p(\Omega)$ de l'application $t \mapsto \rho_t$.

Or Λ_p étant convexe et finie sur $[a, b]$, $t \mapsto \Lambda(t) = \|\rho_t\|_{L^p(\Omega)}^p$ est **continue** sur $]a, b[$ (ouvert), admet une limite à droite en $t = a$ et une limite à gauche en $t = b$.

Ainsi, pour tout $t_0 \in]a, b[$, on a $\lim_{t \rightarrow t_0} \|\rho_t\|_{L^p(\Omega)}^p = \|\rho_{t_0}\|_{L^p(\Omega)}^p$.

De plus, en ce qui concerne la limite à droite de Λ en $t = a$ et sa limite à gauche en $t = b$, on a

$$\lim_{t \rightarrow a^+} \Lambda(t) = \lim_{t \rightarrow a^+} \|\rho_t\|_{L^p(\Omega)}^p \leq \Lambda(a) = \|\rho_a\|_{L^p(\Omega)}^p,$$

et

$$\lim_{t \rightarrow b^-} \Lambda(t) = \lim_{t \rightarrow b^-} \|\rho_t\|_{L^p(\Omega)}^p \leq \Lambda(b) = \|\rho_b\|_{L^p(\Omega)}^p.$$

Ainsi, pour tout $t_0 \in [a, b]$, on a

$$\limsup_{t \rightarrow t_0} \|\rho_t\|_{L^p(\Omega)} \leq \|\rho_{t_0}\|_{L^p(\Omega)}.$$

On peut donc conclure, en application du Théorème A.1-5 (p. A.1-5) que l'application $t \mapsto \rho_t$ est **fortement** continue de $[a, b]$ dans $L^p(\Omega)$. \square

Corollaire 4.1-4.

Soient Ω un ouvert \mathbb{R}^d , et $\rho_0, \rho_1 \in L^p(\Omega)$ deux densités sur Ω à supports bornés, et soit $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ convexe telle que $\rho_1 \mathcal{L}^d = (\nabla \phi) \# (\rho_0 \mathcal{L}^d)$. Pour tout $t \in [0, 1]$, on définit sur la tribu de Lebesgue la mesure :

$$\mu_t = (t \nabla \phi + (1 - t) \text{id}) \# (\rho_0 \mathcal{L}^d).$$

Alors pour tout $t \in [0, 1]$, $\mu_t \ll \mathcal{L}^d$, avec $\mu_t = \rho_t \cdot \mathcal{L}^d$ où $t \mapsto \rho_t \in C^0([0, 1], L^p(\Omega))$.

Preuve : Par application du Théorème 4.1-2, la fonction $\Lambda_p : [0, 1] \rightarrow [0, +\infty]$ définie pour tout $t \in [0, 1]$ par

$$\Lambda_p(t) = \begin{cases} \int_{\Omega} |\rho_t|^p dx & \text{si } \mu_t = \rho_t \cdot \mathcal{L}^d, \text{ avec } \rho_t \in L^1(\Omega, \mathcal{L}^d) \\ +\infty & \text{sinon,} \end{cases}$$

est convexe sur $[0, 1]$ et ici finie en $t = 0$ et $t = 1$ (car $\rho_0, \rho_1 \in L^p(\Omega)$, et $\rho_0, \rho_1 \in \mathcal{P}_p(\mathbb{R}^d)$), donc par convexité finie sur tout $[0, 1]$. Il est clair que $t \mapsto \mu_t$ est continue sur $[0, 1]$ dans $\mathcal{D}'(\Omega)$, on peut donc conclure par le Lemme 4.1-3. \square

Le Lemme qui suit est surtout un Lemme de rigueur. Il nous permettra d'utiliser précisément les propriétés de poussée de mesure au sens faible pour $t \mapsto \rho_t$, en particulier pour la Proposition

4.1-8 (p. 135), plus tard pour la preuve du Théorème 4.3-9 (p. 159). Il fait en particulier appel à un résultat de l'annexe F, à savoir la Proposition F.0-6, rappelée p. 128.

Lemme 4.1-5.

Soient $\rho_0 \in L^p(\mathbb{R}^d)$ et $\rho \in L^p((0, 1) \times \mathbb{R}^d)$, $\rho_0 \geq 0$ et $\rho \geq 0$, à supports bornés respectivement dans \mathbb{R}^d et $[0, 1] \times \mathbb{R}^d$, et soit $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ vérifiant la propriété (Γ_1) .

On définit $(t, x) \mapsto \rho_t(x)$ comme dans la définition 4.1-1 (p. 129).

Si pour tout $h \in C_c^0([0, 1] \times \mathbb{R}^d)$, on a

$$\int_{[0,1] \times \mathbb{R}^d} h \rho(dx \otimes dt) = \int_0^1 \int_{\mathbb{R}^d} h(t, t\nabla\phi(x) + (1-t)x) \rho_0(x) dx dt,$$

alors $t \mapsto \rho_t \in C^0([0, 1[, L^p(\mathbb{R}^d))$, et donc pour presque tout $(t, x) \in [0, 1[\times \mathbb{R}^d$, on a alors $\rho(t, x) = \rho_t(x)$.

Preuve : D'après la Proposition IV.1 (p. 128), $(t, x) \mapsto \rho_t(x)$ est mesurable sur $[0, 1[\times \mathbb{R}^d$, et pour tout $h \in C_c^0([0, 1] \times \mathbb{R}^d)$ et tout $t \in [0, 1[$,

$$\int_{\mathbb{R}^d} h(t, x) \rho_t(x) dx dt = \int_0^1 \int_{\mathbb{R}^d} h(t, t\nabla\phi(x) + (1-t)x) \rho_0(x) dx.$$

Ainsi, d'après le Théorème A.2-4 (p. 205), pour tout $h \in C_c^0([0, 1] \times \mathbb{R}^d)$,

$$\begin{aligned} \int_{[0,1] \times \mathbb{R}^d} h(t, x) \rho_t(x) (dx \otimes dt) &= \int_0^1 \int_{\mathbb{R}^d} h(t, x) \rho_t(x) dx dt \\ &= \int_0^1 \int_{\mathbb{R}^d} h(t, t\nabla\phi(x) + (1-t)x) \rho_0(x) dx dt \\ &= \int_{[0,1] \times \mathbb{R}^d} h \rho(dx \otimes dt). \end{aligned} \quad (4.1-4)$$

D'où, pour presque tout $(t, x) \in [0, 1[\times \mathbb{R}^d$, $\rho(t, x) = \rho_t(x)$, et donc $|\rho(t, x)|^p = |\rho_t(x)|^p$. Toujours d'après le Théorème A.2-4, pour tout $t \in [0, 1[$, les fonctions $x \mapsto \rho(t, x)$ et $x \mapsto \rho_t(x)$ sont mesurables, de même que les fonctions $t \mapsto \int_{\mathbb{R}^d} |\rho(t, x)|^p dx$ et $t \mapsto \int_{\mathbb{R}^d} |\rho_t(x)|^p dx$, et

$$\begin{aligned} \int_0^1 \left| \int_{\mathbb{R}^d} |\rho(t, x)|^p dx - \int_{\mathbb{R}^d} |\rho_t(x)|^p dx \right| dt &\leq \int_0^1 \int_{\mathbb{R}^d} ||\rho(t, x)|^p - |\rho_t(x)|^p| dx dt \\ &= \int_{[0,1] \times \mathbb{R}^d} ||\rho(t, x)|^p - |\rho_t(x)|^p| (dx \otimes dt) \\ &= 0. \end{aligned} \quad (4.1-5)$$

Ainsi, pour presque tout $t \in [0, 1[$, on a

$$\int_{\mathbb{R}^d} |\rho_t(x)|^p dx = \int_{\mathbb{R}^d} |\rho(t, x)|^p dx. \quad (4.1-6)$$

De plus,

$$\int_0^1 \int_{\mathbb{R}^d} |\rho(t, x)|^p dx dt = \int_{[0,1] \times \mathbb{R}^d} |\rho(t, x)|^p (dx \otimes dt) < +\infty.$$

d'où, pour presque tout $t \in [0, 1[$, $\int_{\mathbb{R}^d} |\rho(t, x)|^p dx < +\infty$. Pour presque tout $t \in [0, 1[$, on a donc

$$\int_{\mathbb{R}^d} |\rho_t(x)|^p dx = \int_{\mathbb{R}^d} |\rho(t, x)|^p dx < +\infty \quad \text{i.e.} \quad \rho_t \in L^p(\mathbb{R}^d). \quad (4.1-7)$$

Or, par application du Théorème 4.1-2 (p. 130), la fonction $\Lambda_p : [0, 1[\rightarrow [0, +\infty]$, définie pour tout $t \in [0, 1[$ par $\Lambda_p(t) = \int_{\mathbb{R}^d} |\rho_t|^p dx$, est convexe sur $[0, 1[$.

Donc pour tout $t_0 \in [0, 1[$ tel que $\rho_{t_0} \in L^p(\mathbb{R}^d)$, du fait que l'on a également $\rho_0 \in L^p(\mathbb{R}^d)$, on conclut par convexité, que Λ_p est finie sur tout $[0, t_0]$.

D'après la Proposition IV.1 (p. 128), $t \mapsto \rho_t$ est continue sur $[0, t_0]$ dans $\mathcal{D}'(\mathbb{R}^d)$, et donc par le Lemme 4.1-3 (p. 130), $t \mapsto \rho_t \in C^0([0, t_0], L^p(\mathbb{R}^d))$. Et ceci est donc vrai pour presque tout $t_0 \in [0, 1[$ d'après (4.1-7), d'où $t \mapsto \rho_t \in C^0([0, 1[, L^p(\mathbb{R}^d))$. \square

On ne peut pas, en utilisant un raisonnement identique à la preuve du Lemme 4.1-3, généraliser le résultat de celui-ci pour tout $p = +\infty$, c'est à dire prouver que $\rho \in C^0([0, 1], L^\infty(\mathbb{R}^d))$ si $\rho_0, \rho_1 \in L^\infty(\Omega)$. Cependant l'application du Théorème 4.1-2 nous permet tout de même de prouver que dans ce cas $\rho \in L^\infty([0, 1] \times \mathbb{R}^d)$.

Proposition 4.1-6.

Soient $\rho_0, \rho_1 \in L^\infty(\Omega)$ deux densités sur Ω à supports bornés, et telles qu'il existe $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ vérifiant la propriété (Γ_1) , et $\rho_1 \mathcal{L}^d = \nabla \phi \# (\rho_0 \mathcal{L}^d)$. Soit $\rho = t \mapsto \rho_t$ tel que définie dans la définition 4.1-1 (p. 129) et le Lemme 4.1-5 (p. 132).

On a alors $\rho \in L^\infty([0, 1] \times \mathbb{R}^d)$.

Preuve : Les masses ρ_0 et ρ_1 sont deux densités de probabilité à support bornées, donc $\rho_0, \rho_1 \in \mathcal{P}_2(\mathbb{R}^d)$; et $\nabla \phi$ représente le plan de transport optimal de ρ_0 sur ρ_1 d'après le Théorème 1.1-7 (p. 20). Soit $F : \mathbb{R} \rightarrow \mathbb{R} \cup +\infty$ définie pour tout $s \in \mathbb{R}$ par

$$F(s) = \begin{cases} |s| & |s| \leq M, \\ +\infty & \text{sinon.} \end{cases}$$

F est propre, convexe, semi-continue inférieurement, $F(0) = 0$ et $F(s) \geq 0$. De plus,

$$s \mapsto s^d F(s^{-d}) = \begin{cases} +\infty & \text{si } s \in]0, \frac{1}{M^{1/d}}[, \\ 1 & \text{si } s \in \left[\frac{1}{M^{1/d}}, +\infty \right[. \end{cases}$$

est convexe et décroissante sur $[0, +\infty[$. D'après le Théorème 4.1-2 (p. 130), $t \mapsto \mathcal{F}(\mu_t)$ est convexe sur $[0, 1]$ (avec \mathcal{F} définie à partir de F comme dans l'énoncé du Théorème), on a pour tout $t \in [0, 1]$:

$$\mathcal{F}(\mu_t) \leq (1-t) \int_{\mathbb{R}^d} F(\rho_0) dx + t \int_{\mathbb{R}^d} F(\rho_1) dx \leq M \left(\mathcal{L}^d(\text{supp}(\rho_0)) + \mathcal{L}^d(\text{supp}(\rho_1)) \right) < +\infty. \quad (4.1-8)$$

Ainsi, pour tout $t \in [0, 1]$, on a $\mu_t \ll \mathcal{L}^d$: il existe donc $\rho_t \in L^1(\mathbb{R}^d)$ tel que $\mu_t = \rho_t \mathcal{L}^d$. Soit

$E = \{(t, x) \in [0, 1] \times \mathbb{R}^d, |\rho_t(x)| > M\}$. On a

$$\begin{aligned} \int_{[0,1] \times \mathbb{R}^d} F(\rho_t(x)) dx \otimes dt &= \int_0^1 \int_{\mathbb{R}^d} F(\rho_t(x)) dx dt \leq M \left(\mathcal{L}^d(\text{supp}(\rho_0)) + \mathcal{L}^d(\text{supp}(\rho_1)) \right) < +\infty \\ &= \int_E F(\rho_t(x)) dx \otimes dt + \int_{[0,1] \times \mathbb{R}^d \setminus E} F(\rho_t(x)) dx \otimes dt. \end{aligned} \quad (4.1-9)$$

Si $\mathcal{L}^d(E) > 0$, alors $\int_E F(\rho_t(x)) dx \otimes dt = +\infty$, ce qui est impossible d'après (4.1-9). Ainsi $|\rho_t(x)| \leq M$ pour presque tout $(t, x) \in [0, 1] \times \mathbb{R}^d$. \square

Remarque 4.1-1.

1. Si $p = 2$, on a donc bien $\rho \in L^\infty((0, 1), L^2(\mathbb{R}^d)) \subset L^2((0, 1) \times \mathbb{R}^d)$, ce qui complète, avec la propriété $v \in L_{loc}^\infty([0, 1] \times \omega)^d$ (Proposition 3.3-12 p. 93), la condition (P1') d'existence d'un point selle au lagrangien \mathbf{L} .
2. Tout au long du raisonnement visant à montrer l'existence d'un point selle pour \mathbf{L} , l'application du Théorème 4.1-2 est le seul point qui requiert que $\nabla_x \phi$ soit le plan de transport optimal de $\rho_0 \mathcal{L}^d$ sur $\rho_1 \mathcal{L}^d$.

Pour finir, rappelons qu'au début de cette section, nous avons considéré le prolongement par 0 de Ω à \mathbb{R}^d des densités ρ_0 et ρ_1 , puis une interpolation de McCann dans \mathbb{R}^d entre ces deux masses. Pour pouvoir revenir au problème au problème d'origine, c'est à dire construire la densité ρ sur l'ouvert Ω avec conditions de Neumann homogènes aux bords, il suffit de prouver que le support de ρ est inclus dans $[0, 1] \times \Omega$ et d'en prendre à la restriction à ce dernier (en confondant dans la notation ρ et ρ_Ω comme indiqué dans les notes de début de chapitre).

Proposition 4.1-7.

Soit Ω un ouvert convexe de \mathbb{R}^d . Soient ρ_0 une densité de probabilité et μ_1 une mesure de probabilités telles que $\text{supp}(\rho_0), \text{supp}(\mu_1) \subset \Omega$, et telles qu'il existe $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ vérifiant la propriété (Γ_1) , et $\mu_1 = (\nabla \phi) \# (\rho_0 d\mathcal{L}^d)$. En définissant ρ comme dans la définition 4.1-1 et le Lemme 4.1-5, on a $\text{supp}(\rho) \subset [0, 1] \times \Omega$.

Preuve : On pose toujours $X(t, \cdot) = (1 - t) \text{id} + t \nabla_x \phi$ pour tout $t \in [0, 1]$. On sait que $\text{supp}(\rho_0) \subset \Omega$ et $\text{supp}(\mu_1) = \overline{\nabla_x \phi(\text{supp}(\rho_0))} \subset \Omega$ ([77], Théorème 2.12 p.66). Ainsi, pour tout $t \in [0, 1]$, $X(t, \cdot)(\text{supp}(\rho_0)) = ((1 - t) \text{id} + t \nabla_x \phi)(\text{supp}(\rho_0))$ est contenu dans l'enveloppe convexe de $\text{supp}(\rho_0) \cup \text{supp}(\mu_1) \subset \Omega$.

Soit $\varphi \in C_c^\infty([0, 1] \times \mathbb{R}^d)$ tel que $\text{supp}(\varphi) \cap \Omega = \emptyset$. Alors

$$\int_0^1 \int_{\mathbb{R}^d} \varphi \rho dx dt = \int_0^1 \int_{\mathbb{R}^d} \varphi(t, X(t, x)) \rho_0(x) dx dt = \int_0^1 \int_{\text{supp}(\rho_0)} \varphi(t, X(t, x)) \rho_0(x) dx dt = 0.$$

Ainsi $\text{supp}(\rho) \subset [0, 1] \times \Omega$. \square

Ainsi, les Lemme 4.1-3 (ou plutôt son Corollaire), le Lemme 4.1-5, et la Proposition 4.1-7 nous permettent de conclure que la densité d'interpolation de McCann vérifie bien la propriété $(P_\beta 1')$.

IV.1.2 Propriété de conservation de la masse

Nous allons maintenant montrer, par le Corollaire 4.1-9, que le couple (ρ, v) vérifie la propriété $(P_\beta 4')$, c'est à dire que ρ et v sont liées la relation de conservation de la masse (4.1-1) (p. 127).

Rappelons que dans cette section, la densité ρ est définie par $\rho : (t, x) \mapsto \rho_t(x)$ (voir la définition 4.1-1 p. 129 et le Lemme 4.1-5 p. 132), où pour tout $t \in [0, 1]$ on a :

$$\rho_t \mathcal{L}^d = [(1-t) \text{id} + t \nabla \phi] \# (\rho_0 \mathcal{L}^d) = \nabla \phi_t \# (\rho_0 \mathcal{L}^d), \quad (4.1-10)$$

avec $\rho_0, \rho_1 \in L^2(\Omega)$, $\phi_t = (1-t)|\cdot|^2/2 + t\phi$, et telle pour tout $t \in [0, 1]$, la densité ρ_t est choisi telle que $(t, x) \mapsto \rho_t(x)$ soit mesurable (voir la définition 4.1-1 p. 129). Cette densité, d'après le Corollaire 4.1-4 (p. 4.1-4), vérifie $t \mapsto \rho_t \in C^0([0, 1], L^2(\Omega))$ (on parle bien ici de continuité forte).

Le champ de vitesse v est quant à lui défini pour tout $t \in]0, 1[$ par :

$$v(t, \cdot) = v_\phi(t, \cdot) = \frac{\text{id} - \mathbf{p}_\phi(t, \cdot)}{t}, \quad \text{avec} \quad \mathbf{p}_\phi(t, \cdot) = \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{\cdot}{1-t} \right) = \nabla_x(\phi_t)^*. \quad (4.1-11)$$

Proposition 4.1-8.

Soit Ω un ouvert convexe de \mathbb{R}^d . Soient ρ_0 une densité de probabilité et μ_1 une mesure de probabilités telles que $\text{supp}(\rho_0), \text{supp}(\mu_1) \subset \Omega$, et telles qu'il existe $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ vérifiant la propriété (Γ_1) , et $\mu_1 = \nabla \phi \# (\rho_0 \mathcal{L}^d)$. Soient ρ et v tels que définies en (4.1-10) et (4.1-11).

Alors le couple (ρ, v) vérifie (4.1-1) p. 127 (la conservation de la masse) i.e. :

$$\int_{[0,1] \times \Omega} (\partial_t \varphi + v \cdot \nabla_x \varphi) \rho \, dx \, dt + \int_{\Omega} \varphi(0, \cdot) \rho_0 \, dx - \int_{\Omega} \varphi(1, \cdot) \, d\mu_1 = 0$$

pour tout $\varphi \in C_c^\infty([0, 1] \times \Omega)$.

Preuve : Ici, comme souvent, nous confondrons f définie sur \mathbb{R}^d et $f|_\Omega$, de même que pour g définie sur Ω avec sa prolongation par 0 à \mathbb{R}^d . On rappelle que par la Proposition 3.3-10 (p. 91), pour tout $t \in]0, 1[$, et presque tout $x \in \mathbb{R}^d$, $\partial_t X(t, x) = v(t, X(t, x))$, donc pour tout $\varphi \in C_c^\infty([0, 1] \times \Omega)$,

$$\begin{aligned} & \int_{]0,1[\times \Omega} (\partial_t \varphi + v \cdot \nabla_x \varphi) \rho \, (dx \otimes dt) \\ &= \int_0^1 \int_{\mathbb{R}^d} (\partial_t \varphi(t, x) + v(t, x) \cdot \nabla_x \varphi(t, x)) \rho(t, x) \, dx \, dt \\ & \quad (\text{En effet, par la Proposition 4.1-7 : } \text{supp}(\rho) \subset [0, 1] \times \Omega) \\ &= \int_0^1 \int_{\mathbb{R}^d} (\partial_t \varphi(t, X(t, x)) + v(t, X(t, x)) \cdot \nabla_x \varphi(t, X(t, x))) \rho_0(x) \, dx \, dt \\ &= \int_0^1 \int_{\mathbb{R}^d} (\partial_t \varphi(t, X(t, x)) + \partial_t X(t, x) \cdot \nabla_x \varphi(t, X(t, x))) \rho_0(x) \, dx \, dt \\ &= \int_0^1 \frac{d}{dt} \int_{\mathbb{R}^d} \varphi(t, X(t, x)) \rho_0(x) \, dx \, dt \end{aligned}$$

$$\begin{aligned}
&= \int_{\mathbb{R}^d} \varphi(1, \nabla_x \phi(x)) \rho_0(x) dx - \int_{\mathbb{R}^d} \varphi(0, x) \rho_0(x) dx \\
&= \int_{\mathbb{R}^d} \varphi(1, x) d\mu_1(x) - \int_{\mathbb{R}^d} \varphi(0, x) \rho_0(x) dx \\
&= \int_{\Omega} \varphi(1, x) d\mu_1(x) - \int_{\Omega} \varphi(0, x) \rho_0(x) dx
\end{aligned}$$

Le fait que $v \in L_{loc}^\infty([0, 1] \times \Omega)$ (Proposition 3.3-12 p. 93) assure que les intégrales sont bien définies. \square

Corollaire 4.1-9.

Soient $p \geq 2$ et $q \leq 2$ tels que $1/p + 1/q = 1$, et soit Ω un ouvert convexe borné de \mathbb{R}^d . Soient ρ_0 et ρ_1 deux densités de probabilité telles que $\text{supp}(\rho_0), \text{supp}(\rho_1) \subset \Omega$, et telles qu'il existe $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ vérifiant la propriété (Γ_1) , et $(\rho_1 \mathcal{L}^d) = \nabla \phi \# (\rho_0 \mathcal{L}^d)$. Soient ρ et v toujours définies par (4.1-10) et (4.1-11) (p. 135).

On suppose de plus que $\rho_0, \rho_1 \in L^p(\mathbb{R}^d)$ (d'où $\rho \in L^p((0, 1) \times \mathbb{R}^d)$).

Alors, pour tout $\varphi \in W^{1,q}((0, 1) \times \Omega)$,

$$\int_{]0,1[\times \Omega} (\partial_t \varphi + v \cdot \nabla_x \varphi) \rho dx dt + \int_{\Omega} \varphi(0, \cdot) \rho_0 dx - \int_{\Omega} \varphi(1, \cdot) \rho_1 dx = 0.$$

Preuve : D'après la Proposition 4.1-6 (p. 133), on a $\rho \in L^\infty((0, 1), L^p(\mathbb{R}^d)) \subset L^p((0, 1) \times \mathbb{R}^d)$, et d'après la Proposition 3.3-12 (p. 93), on a $v \in L_{loc}^\infty([0, 1] \times \Omega)$ (d'où $\rho v \in L^p((0, 1) \times \mathbb{R}^d)$). De plus, d'après la Proposition 4.1-8 (p. 135), on a

$$\int_{]0,1[\times \Omega} (\partial_t \varphi + v \cdot \nabla_x \varphi) \rho dx dt + \int_{\Omega} \varphi(0, \cdot) \rho_0 dx - \int_{\Omega} \varphi(1, \cdot) d\mu_1 = 0$$

pour tout $\varphi \in C_c^\infty([0, 1] \times \Omega)$. On peut donc conclure par densité de $C_c^\infty([0, 1] \times \Omega)$ dans $W^{1,q}((0, 1) \times \Omega)$ pour la norme associée. On rappelle que les opérateurs traces

$$T_0, T_1 \in C_c^\infty([0, 1] \times \Omega) \longrightarrow L^q(\Omega), f \mapsto f(0, \cdot), f(1, \cdot),$$

se prolongent à $W^{1,q}((0, 1) \times \Omega)$, ce qui permet de conclure également par densité pour les termes $\int_{\Omega} \varphi(0, \cdot) \rho_0 dx$ et $\int_{\Omega} \varphi(1, \cdot) \rho_1 dx$. \square

IV.1.3 Existence d'un point selle

Les conditions (I'_β) étant maintenant établies sur (ρ, v) (avec ρ et v toujours définies par (4.1-10) et (4.1-11)), nous sommes donc en mesure de les compiler pour prouver le Théorème 3.1-2 (p. 76), c'est à dire pour construire un point selle au lagrangien \mathbf{L} , c'est à dire construire un triplet (ψ, q, μ) vérifiant les propriétés (I).

Rappelons que pour cela, on choisit de poser $\mu = (\rho, \rho v)$ et $q = (-(1/2)|v|^2, v)$. Commençons un premier résultat visant à établir clairement la condition (P3) sur le terme q : c'est à dire que

pour un tel q , il existe alors $\psi^* \in W_{loc}^{1,\infty}([0,1] \times \Omega)$ tel que $q = \nabla_{t,x}\psi^*$. Rappelons que le champ de vitesse, en plus de vérifier les propriétés $(P_{\beta 2'})$ et $(P_{\beta 3'})$, peut s'écrire $v = \nabla_x \psi$ en prenant pour tout $t \in]0,1[$:

$$\psi(t, \cdot) = \frac{1}{t} \left(\frac{1}{2} |\cdot|^2 - (\phi_t)^* \right) \in L_{loc}^1(]0,1[\times \Omega).$$

Le triplet (ψ^*, q, μ) sera alors notre point selle, et nous montrerons dans la preuve du Théorème 3.1-2 (rappelé plus bas p. 138) qu'il vérifie également les propriétés (P1) et (P2).

Lemme 4.1-10.

Soit Ω un ouvert de \mathbb{R}^d assez régulier (tel que tout compact de Ω soit inclus dans un ouvert borné de Ω de classe C^1). Soit un champ de vitesse $v \in L_{loc}^\infty([0,1] \times \Omega)^d$ (propriété $(P_{\beta 2'})$), vérifiant au sens des distributions les propriétés :

1. $\partial_t v + \frac{1}{2} \nabla_x |v|^2 = 0$ (propriété $(P_{\beta 3'})$),

2. il existe un potentiel $\psi \in L_{loc}^1(]0,1[\times \Omega)$, tel que $v = \nabla_x \psi$.

Alors il existe un potentiel $\psi^ \in W_{loc}^{1,\infty}([0,1] \times \Omega)$, vérifiant au sens des distributions :*

$$\partial_t \psi^* + \frac{1}{2} |\nabla_x \psi^*|^2 = 0, \quad (4.1-12)$$

et

$$q = \nabla_{t,x} \psi^* \quad \text{avec} \quad q = \begin{pmatrix} -\frac{1}{2} |v|^2 \\ v \end{pmatrix} \quad (4.1-13)$$

(propriété (P3)).

La preuve de ce Lemme résulte du résultat suivant, dont la preuve fait l'objet de l'annexe E.

Lemme 4.1-11.

Soit Ω un ouvert borné de \mathbb{R}^d , et soit une distribution $T \in \mathcal{D}'((0,1) \times \Omega)$, tel que $\nabla_{t,x} T \in L_{loc}^\infty([0,1] \times \Omega)^d$ (ou du moins un de ses représentants de classe \mathcal{L}^d -mesurable que l'on notera g).

Alors $T \in W_{loc}^{1,\infty}([0,1] \times \Omega)$, ou plus précisément il existe $f \in W_{loc}^{1,\infty}([0,1] \times \Omega)$ tel que :

- $\forall h \in C_c^\infty(]0,1[\times \Omega)$, $\langle T, h \rangle = \int_0^1 \int_\Omega f h \, dx \, dt$,
- $\forall h \in C_c^\infty(]0,1[\times \Omega)^d$, $\langle T, \operatorname{div}_{t,x}(h) \rangle = \int_0^1 \int_\Omega f \cdot \operatorname{div}(h) \, dx \, dt = - \int_0^1 \int_\Omega g \cdot h \, dx \, dt$.

De plus f est lipschitzienne (ou du moins un de ses représentants) sur Ω .

Preuve du Lemme 4.1-10 : Au sens des distributions, on a

$$\begin{aligned} \partial_t v + \frac{1}{2} \nabla_x |v|^2 = 0 &\Leftrightarrow \partial_t (\nabla_x \psi) + \frac{1}{2} \nabla_x |v|^2 = 0, \\ &\Leftrightarrow \nabla_x (\partial_t \psi) + \frac{1}{2} \nabla_x |v|^2 = 0, \\ &\Leftrightarrow \nabla_x \left(\partial_t \psi + \frac{1}{2} |v|^2 \right) = 0. \end{aligned} \quad (4.1-14)$$

Il existe donc une distribution T ne dépendant que de $t \in]0, 1[$, telle que, on ait

$$\partial_t \psi + \frac{1}{2} |v|^2 = T$$

(voir par exemple le Théorème 2.16 dans [80]). On pose $\psi^* = \psi - G$, où G est une distribution primitive de T sur $]0, 1[$ (et ne dépendant toujours que de t). On a alors bien, au sens des distributions,

$$- \nabla_x \psi^* = \nabla_x \psi = v \in L_{loc}^\infty([0, 1] \times \Omega)^d,$$

$$- \partial_t \psi^* = -\frac{1}{2} |v|^2 \in L_{loc}^\infty([0, 1] \times \Omega).$$

ψ^* vérifie donc bien (4.1-12).

Pour finir, soit ω un ouvert borné et assez régulier (de classe C^1) de Ω . On a donc $\nabla_{t,x} \psi^* \in L^\infty((0, 1) \times \omega)^{d+1}$. D'après le Lemme 4.1-11 (p. 137), on a alors $\psi^* \in W^{1,\infty}((0, 1) \times \omega)$.

Ainsi, on a bien $\psi^* \in W_{loc}^{1,\infty}([0, 1] \times \Omega)$. \square

Nous sommes maintenant en mesure de démontrer le principal résultat de cette section, à savoir le Théorème 3.1-2 (p. 76) établissant l'existence d'un point selle pour le lagrangien \mathbf{L} . Nous rappelons l'énoncé du Théorème en question, dans une version réactualisée et un peu plus précise :

Théorème 3.1-2.

Soient ρ_0, ρ_1 et Ω définis comme dans la notation 2 (p. 74).

Pour tout $\phi \in \Phi(\rho_0 \mathcal{L}^d, \rho_1 \mathcal{L}^d)$ (voir la notation 3.2-2 p. 80), il existe une densité positive $\rho_\phi \in C^0([0, 1], L^2(\Omega))$, telle que pour tout $[0, 1] \times \mathbb{R}^d$, on ait

$$\rho_\phi(t, \cdot) \mathcal{L}^d = X_\phi(t, \cdot) \# (\rho_0 \mathcal{L}^d), \text{ avec } X_\phi(t, \cdot) = \nabla_x \phi_t = (1-t) \text{id} + t \nabla_x \phi,$$

et un champ de vitesse v_ϕ , définie pour tout $(t, x) \in [0, 1] \times \mathbb{R}^d$, par

$$v_\phi(t, x) = \frac{1}{t} \left(x - \text{Prox}_{\frac{t}{1-t} \phi} \left(\frac{x}{1-t} \right) \right) = \frac{x - \nabla_x (\phi_t)^*(x)}{t},$$

où $(\phi_t)^*$ désigne la transformée de Legendre de $\phi_t = (1-t)|\cdot|^2 + t\phi$.

Alors en posant $\mu_\phi = (\rho_\phi, \rho_\phi v_\phi)$ et $q_\phi = (-(1/2)|v_\phi|^2, v_\phi)$, il existe $\psi_\phi \in W_{loc}^{1,\infty}([0, 1] \times \Omega)$ tel que $q_\phi = \nabla_{t,x} \psi_\phi$ et tel que $(\psi_\phi, q_\phi, \mu_\phi)$ (ou du moins sa restriction $]0, 1[\times \Omega$) soit un élément de $\mathbf{L}^{ps}(\rho_0, \rho_1, \Omega)$.

En outre v_ϕ vérifie les propriétés (II) et (III).

Remarque 4.1-2. Rappelons que nous avons établi dans la section III.2 du chapitre précédent, que l'ensemble $\Phi(\rho_0 \mathcal{L}^d, \rho_1 \mathcal{L}^d)$ est non-vide !

Preuve : Pour la suite, rappelons qu'un élément (ψ, q, μ) de $\mathbf{L}^{ps}(\rho_0, \rho_1, \Omega)$, doit vérifier $(\psi, q, \mu) \in S$, ainsi que les propriétés (I) p. 74, que nous rappelons :

$$(P1) \quad \forall q' \in \mathcal{P}, \quad \langle \mu, q' - q \rangle \leq 0,$$

$$(P2) \quad \forall h \in H^1((0, 1) \times \Omega), \quad G(h) + \langle \mu, \nabla_{t,x} h \rangle = 0,$$

$$(P3) \quad \nabla_{t,x}\psi = q,$$

\mathcal{P} étant le paraboloïde $\{(a, b) \in L^2((0, 1) \times \Omega)^{1+d}, a + |b|^2/2 \leq 0\}$.

Tout d'abord, par la Proposition 3.3-12 (p. 93), on sait que $v_\phi \in L^\infty((0, 1) \times \Omega)^d$ (propriété $(P_\beta 2')$), et par le Corollaire 4.1-4 (p. 131) que $t \mapsto \rho_\phi(t, \cdot) = X_\phi(t, \cdot) \# \rho_0 \in C^0([0, 1], L^2(\Omega))$ (propriété $(P_\beta 1')$).

D'où $\mu_\phi = (\rho_\phi, \rho_\phi v_\phi)$, $q_\phi = (-(1/2)|v_\phi|^2, v_\phi) \in L^2((0, 1) \times \Omega)$ (ou du moins les restrictions de μ_ϕ et q_ϕ à $]0, 1[\times \Omega$). De plus, d'après la Proposition 4.1-7 (p. 134), on a $\text{supp}(\rho_\phi) \subset [0, 1] \times \Omega$, d'où $\text{supp}(\mu_\phi) \subset [0, 1] \times \Omega$: les conditions de Neumann homogènes aux bords en espace sur μ sont donc vérifiées.

1. En posant $\mu_\phi = (\rho_\phi, \rho_\phi v_\phi)$ et $q_\phi = (-(1/2)|v_\phi|^2, v_\phi)$, la condition (P1) est naturellement vérifiée. En effet, pour tout $q' = (a, b) \in \mathcal{P}$, on a

$$\begin{aligned} \langle \mu_\phi, q' - q_\phi \rangle &= \int_0^1 \int_\Omega (a\rho_\phi + b \cdot v_\phi \rho) \, dx \, dt - \int_0^1 \int_\Omega \frac{1}{2}|v_\phi|^2 \rho_\phi \, dx \, dt \\ &\leq \int_0^1 \int_\Omega \left(a + \frac{1}{2}|b|^2 + \frac{1}{2}|v|^2 \right) \rho_\phi \, dx \, dt - \int_0^1 \int_\Omega \frac{1}{2}|v_\phi|^2 \rho_\phi \, dx \, dt \\ &\leq 0. \end{aligned}$$

2. La condition (P2) résulte quant à elle du Corollaire 4.1-9 p. 136 (propriété $(P_\beta 4')$).
3. Pour finir v_ϕ vérifie la condition du Lemme 4.1-10 (p. 137) : en effet d'après la Proposition 3.3-31 (p. 115), v_ϕ vérifie au sens des distributions l'équation de Burgers (propriété $(P_\beta 3')$), à savoir

$$\partial_t v_\phi + \frac{1}{2} \nabla_x |v_\phi|^2 = 0.$$

De plus, par les propriétés énoncées à la définition 3.3-9 (p. 89), on a pour tout $t \in [0, 1[$,

$$v_\phi(t, \cdot) = \frac{\text{id} - \nabla_x(\phi_t)^*}{t} = \nabla_x \left(\frac{1}{t} \left(\frac{1}{2} |\cdot|^2 - (\phi_t)^* \right) \right),$$

avec $\phi_t = (1-t)\frac{|\cdot|^2}{2} + t\phi$. En posant donc $\tilde{\psi}_\phi$ définie pour tout $(t, x) \in [0, 1[\times \mathbb{R}^d$ par

$$\tilde{\psi}_\phi(t, x) = \frac{1}{t} \left(\frac{1}{2} |x|^2 - (\phi_t)^*(x) \right),$$

on a donc $v_\phi = \nabla_x \tilde{\psi}_\phi$ au sens des distributions.

L'application du Lemme 4.1-10 (p. 137) nous permet donc de conclure à l'existence d'un $\psi_\phi \in W_{loc}^{1,\infty}([0, 1] \times \Omega)$ tel que $q_\phi = \nabla_{t,x}\psi_\phi$ et ainsi de vérifier la condition (P3) et que le triplet $(\psi_\phi, q_\phi, \mu_\phi)$ est bien un élément de S . Ceci termine notre preuve. \square

Remarque 4.1-3. Le groupement de propriétés (I'_β) (que nous nous sommes attelés à démontrer une par une depuis le début du chapitre III) illustrent bien les caractéristiques recherchées dans un transport optimal. Cependant, un vrai point selle de \mathbf{L} est rigoureusement caractérisé par le groupement de propriétés (I) , et les propriétés (I'_β) n'en sont pas équivalentes bien que l'impliquant en grande partie (c'est à dire outre le fait que v dérive d'un potentiel et les conditions de Neumann

homogènes aux bords pour μ).

En effet, nous verrons dans la suite (sous-section IV.2.1) qu'en général le lagrangien \mathbf{L} admet une infinité de point selles, dont certains ne pouvant se construire à partir d'un couple (ρ, v) vérifiant les propriétés (V $_{\beta}$). Ces dernières propriétés resteront tout de même un "fil rouge" dans toute la suite de cette étude.

Pour finir cette étude des points selles de \mathbf{L} , nous allons montrer, dans la section qui suit, l'unicité de la composante $\mu = (\rho, \rho v)$ des points selles de \mathbf{L} , ainsi que quelques résultats de généralisation.

IV.2 Résultats d'unicité

Cette dernière section sera consacré à établir, dans le cadre de l'algorithme de Benamou-Brenier, l'unicité de la composante $\mu = (\rho, m) = (\rho, \rho v)$ des points selles de \mathbf{L} (sous-sections IV.2.2 et IV.2.3), résultat qui pourra être généralisé pour tenter de caractériser les champs de vitesse de transport optimal dans L^2 (section IV.3). Nous commençons par énoncer le principal résultat de cette section et qui sera explicitement démontré dans la dernière sous-section.

On rappelle que l'ensemble $\Phi(\rho_0 \mathcal{L}^d, \mu_1)$ désigne l'ensemble des fonctions $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ vérifiant la propriété (Γ_1) tels que $\nabla \phi \# (\rho_0 \mathcal{L}^d) = \mu_1$.

Théorème 4.2-1. UNICITÉ DE LA DENSITÉ ET DE LA QUANTITÉ DE MOUVEMENT

Si (ψ^*, q^*, μ^*) est un élément de $\mathbf{L}^{ps}(\rho_0, \rho_1, \Omega)$ (pour ρ_0, ρ_1 et Ω définis comme dans la notation 2 p. 74, définissant $\mathbf{L}^{ps}(\rho_0, \rho_1, \Omega)$), alors pour tout potentiel $\phi \in \Phi(\rho_0 \mathcal{L}^d, \rho_1 \mathcal{L}^d)$, on a $\mu^* = (\rho^*, m^*) = (\rho_\phi, \rho_\phi v_\phi)$, avec pour tout $t \in]0, 1[$:

$$v_\phi(t, x) = \frac{1}{t} \left(x - \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{x}{1-t} \right) \right) = \frac{x - \nabla_x(\phi_t)^*(x)}{t}$$

(où $(\phi_t)^*$ désigne la transformée de Legendre de $\phi_t = (1-t)|\cdot|^2 + t\phi$) et

$$\rho_\phi(t, \cdot) \mathcal{L}^d = X_\phi(t, \cdot) \# (\rho_0 \mathcal{L}^d) \in C([0, 1], L^2(\Omega)), \text{ avec } X_\phi(t, \cdot) = \nabla_x \phi_t = (1-t) \text{id} + t \nabla_x \phi.$$

Autrement dit, quel que soit le point selle $(\psi^*, q^*, \mu^*) \in \mathbf{L}^{ps}(\rho_0, \rho_1, \Omega)$, et quel que soit $\phi_1, \phi_2 \in \Phi(\rho_0 \mathcal{L}^d, \rho_1 \mathcal{L}^d)$, on a

$$\rho^* = \rho_{\phi_1} = X_{\phi_1}(t, \cdot) \# \rho_0 = \rho_{\phi_2} = X_{\phi_2}(t, \cdot) \# \rho_0,$$

et $v_{\phi_1} = v_{\phi_2}$ sur $\text{supp}(\rho^*)$. En effet, on a :

$$m^* = m_{\phi_1} = \rho_{\phi_1} v_{\phi_1} = \rho^* v_{\phi_1} = \rho^* v_{\phi_2} = \rho_{\phi_2} v_{\phi_2} = m_{\phi_2}.$$

Remarque 4.2-4. Nous ne démontrons donc pas ici l'unicité du point selle de \mathbf{L} (ce qui n'est en général pas le cas, voir la sous-section IV.2.1), mais seulement l'unicité de la composante $\mu = (\rho, m) = (\rho, \rho v)$ des points selles de \mathbf{L} . Autrement dit, l'ensemble des points selles (ψ, q, μ) de \mathbf{L} partagent la même composante μ , i.e. il y a unicité de la densité ρ et du champ de vitesse v sur le support de ρ .

Rappelons cependant que l'objectif de l'algorithme de lagrangien augmenté réside uniquement dans le calcul de $\mu = (\rho, m)$. En effet les composantes q et ψ ne sont que des éléments nécessaires au calcul, tandis que le couple (ρ, m) constitue une solution au problème initial de transport optimal (voir la section II.1) comme cela a été démontré à la Proposition 2.2-4 (p. 60).

Les composantes q et ψ ne peuvent cependant pas varier sur le support de ρ (en effet d'après la Proposition 2.2-3 p. 58, si on pose $q = (a, b)$, alors $m = \rho b$), autrement dit le champ de vitesse v est unique sur le support de ρ .

IV.2.1 Non-unicité des points selles de \mathbf{L}

Dans un premier temps, nous allons clairement montrer **la non-unicité** des points selles de \mathbf{L} en général. Nous le déduisons de la Proposition suivante :

Proposition 4.2-2.

Soient ϕ_1 et ϕ_2 deux fonctions convexes sur \mathbb{R}^d vérifiant la propriété (Γ_1) , et soient v_{ϕ_1} et v_{ϕ_2} les deux champs de vitesse sur \mathbb{R}^d définis, rappelons le, pour tout $(t, x) \in]0, 1[\times \Omega$ par

$$v_{\phi_1}(t, x) = \frac{1}{t} \left(x - \text{Prox}_{\frac{t}{1-t}\phi_1} \left(\frac{x}{1-t} \right) \right) \text{ et } v_{\phi_2}(t, x) = \frac{1}{t} \left(x - \text{Prox}_{\frac{t}{1-t}\phi_2} \left(\frac{x}{1-t} \right) \right).$$

Alors

$$v_{\phi_1} = v_{\phi_2} \implies \exists C \in \mathbb{R}, \phi_1 = \phi_2 + C. \quad (4.2-1)$$

Preuve : Supposons que $v_{\phi_1} = v_{\phi_2}$. Alors, pour tout $(t, x) \in]0, 1[\times \Omega$, on a

$$\text{Prox}_{\frac{t}{1-t}\phi_1} \left(\frac{x}{1-t} \right) = \text{Prox}_{\frac{t}{1-t}\phi_2} \left(\frac{x}{1-t} \right),$$

et donc, pour tout $\gamma \in]0, +\infty[$ et tout $y \in \mathbb{R}^d$, on a :

$$\text{Prox}_{\gamma\phi_1}(y) = \text{Prox}_{\gamma\phi_2}(y).$$

En particulier, pour tout $\gamma \in]0, +\infty[$ et tout $y \in \mathbb{R}^d$ en lequel ϕ_1 et ϕ_2 sont différentiables (ϕ_1 et ϕ_2 vérifiant la propriété (Γ_1)), ils sont différentiables presque partout sur \mathbb{R}^d , on a :

$$\gamma \nabla \phi_1(y) = y - \text{Prox}_{\gamma\phi_1}(y) = y - \text{Prox}_{\gamma\phi_2}(y) = \gamma \nabla \phi_2(y).$$

Ainsi, par le Théorème de Rademacher (p. 98), on a $\nabla \phi_1 = \nabla \phi_2$ au sens des distributions sur \mathbb{R}^d . Il existe donc une constante réelle C telle que $\phi_1 = \phi_2 + C$. \square

D'après la Remarque 3.2-7 (p. 84), il existe $\phi_1, \phi_2 \in \Phi(\rho_0 \mathcal{L}^d, \rho_1 \mathcal{L}^d)$ différant d'une fonction non constante. D'après la Proposition ci-dessus, on a alors $v_{\phi_1} \neq v_{\phi_2}$ sur \mathbb{R}^d . Or d'après le Théorème 3.1-2 (p. 76), en posant :

$$- t \mapsto \rho_1(t, \cdot) = X_1(t, \cdot) \# \rho_0, \text{ et } t \mapsto \rho_2(t, \cdot) = X_2(t, \cdot) \# \rho_0, \text{ avec}$$

$$X_1(t, \cdot) = (1-t) \text{id} + t \nabla_x \phi_1, \text{ et } X_2(t, \cdot) = (1-t) \text{id} + t \nabla_x \phi_2,$$

$$- \mu_1 = (\rho_1, \rho_1 v_{\phi_1}) \text{ et } \mu_2 = (\rho_2, \rho_2 v_{\phi_2}),$$

$$- q_1 = (-(1/2)|v_1|^2, v_1) \text{ et } q_2 = (-(1/2)|v_2|^2, v_2),$$

il existe $\psi_1, \psi_2 \in W_{loc}^{1,\infty}([0, 1] \times \Omega)$ tels que $q_1 = \nabla_{t,x} \psi_1$ et $q_2 = \nabla_{t,x} \psi_2$, et tels que les restrictions de (ψ_1, q_1, μ_1) et (ψ_2, q_2, μ_2) à $]0, 1[\times \Omega$ soit des éléments de $\mathbf{L}^{ps}(\rho_0, \rho_1, \Omega)$ (des points selles de \mathbf{L}).

Donc en prenant Ω , assez grand, on a alors $v_{\phi_1} \neq v_{\phi_2}$ sur $]0, 1[\times \Omega$, d'où $q_1 \neq q_2$ sur $]0, 1[\times \Omega$, et d'où $(\psi_1, q_1, \mu_1) \neq (\psi_2, q_2, \mu_2)$ sur $]0, 1[\times \Omega$ (on parle bien sûr ici des restrictions à $]0, 1[\times \Omega$ des différentes fonctions évoquées).

En conclusion, pour Ω assez grand, $\mathbf{L}^{ps}(\rho_0, \rho_1, \Omega)$ possède au deux éléments (et donc une infinité du fait que $\mathbf{L}^{ps}(\rho_0, \rho_1, \Omega)$ est un ensemble convexe).

Nous allons cependant montrer dans la suite de cette section que, s'il n'y a pas en général unicité des points selles (ψ, q, μ) de \mathbf{L} , il y a en revanche unicité de la composante $\mu = (\rho, m) = (\rho, \rho v)$.

Remarque 4.2-5. Nous avons dans cette Proposition construit différents points selles, mais tous construits à partir d'un couple densité vitesse (ρ, v) vérifiant les propriétés (I'_α) et (I'_β) : autrement dit le terme q ne se "décolle pas" de la paroi du paraboloïde \mathcal{P} définie rappelons par :

$$\mathcal{P} = \left\{ (a, b) \in \mathbb{R} \times \mathbb{R}^d, a \leq -\frac{|b|^2}{2} \right\}.$$

En effet on a construit ici des composantes secondaires du type $q = (a, b) = (-(1/2)|v|^2, v)$, d'où $a + (1/2)|b|^2 = 0$. Cependant en interpolant linéairement deux de ces différents points selles, on peut alors obtenir des points selles dont la composante $q = (a, b)$ vérifiant parfois (en dehors du support de μ i.e de ρ) la relation $a + (1/2)|b|^2 < 0$ (se décollant donc de la paroi de \mathcal{P}), et qui ne peuvent donc pas se construire à partir d'un couple (ρ, v) vérifiant les propriétés (I'_α) et (I'_β) .

IV.2.2 Unicité du champ de vitesse sur le support de la densité

Dans ce but, nous commencerons par démontrer que pour tous les points selles de \mathbf{L} , notés $(\psi^*, q^*, \mu^*) = (\psi^*, q^*, (\rho^*, m^*))$, les densités ρ^* se propagent selon le même champ de vitesse v .

Lemme 4.2-3.

Si (ψ^*, q^*, μ^*) est un élément de $\mathbf{L}^{ps}(\rho_0, \rho_1, \Omega)$ (pour ρ_0, ρ_1 et Ω définis dans le cadre la notation 2 p. 74), avec $\mu^* = (\rho^*, m^*)$.

Alors, pour tout $\phi \in \Phi(\rho_0 \mathcal{L}^d, \rho_1 \mathcal{L}^d)$, en posant

$$v_\phi(t, x) = \frac{1}{t} \left(x - \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{x}{1-t} \right) \right) = \frac{1}{t} (\text{id} - \mathbf{p}_\phi(t, \cdot))$$

(voir la définition 3.3-9 p. 89) pour tout $t \in]0, 1[$, on a $m^* = \rho^* v_\phi$.

Preuve : Soit (ψ^*, q^*, μ^*) un élément de $\mathbf{L}^{ps}(\rho_0, \rho_1, \Omega)$. Dans le Théorème 3.1-2 (p. 76), nous avons montré que, pour tout $\phi \in \Phi(\rho_0 \mathcal{L}^d, \rho_1 \mathcal{L}^d)$, il était possible de construire un point selle $(\psi_\phi, q_\phi, \mu_\phi)$ de \mathbf{L} , avec en particulier, $\mu_\phi = (\rho_\phi, \rho_\phi v_\phi)$ et $q_\phi = (-\frac{1}{2}|v_\phi|^2, v_\phi)$, pour tout $t \in]0, 1[$,

$$\rho_\phi(t, \cdot) = X_\phi(t, \cdot) \# \rho_0 \text{ avec } X_\phi(t, \cdot) = (1-t) \text{id} + t \nabla_x \phi,$$

et

$$v_\phi(t, \cdot) = (1/t)(\text{id} - \mathbf{p}_\phi(t, \cdot)).$$

Soit $\phi \in \Phi(\rho_0 \mathcal{L}^d, \rho_1 \mathcal{L}^d)$. Le triplet (ψ^*, q^*, μ^*) , en tant que point selle de \mathbf{L} , μ^* vérifie l'équation de conservation de la masse (voir la Proposition 2.2-1 p. 56), i.e.

$$\forall h \in H^1((0, 1) \times \Omega), \quad G(h) + \langle \mu^*, \nabla_{t,x} h \rangle = 0$$

et de plus $\nabla_{t,x}\psi^* = q^*$. Il en est de même pour μ_ϕ (i.e. conservation de la masse + $\nabla_{t,x}\psi_\phi = q_\phi$). Donc en particulier, en prenant $h = \psi_\phi$ (comme dans la preuve de la Proposition 2.2-4 p. 60),

$$\langle \mu^*, q_\phi \rangle_{L^2} = \langle \mu^*, \nabla_{t,x}\psi_\phi \rangle_{L^2} = -G(\psi_\phi) = \langle \mu_\phi, \nabla_{t,x}\psi_\phi \rangle_{L^2} = \langle \mu_\phi, q_\phi \rangle_{L^2}.$$

D'où $\langle \mu^*, q_\phi \rangle_{L^2} = \langle \mu_\phi, q_\phi \rangle_{L^2}$.

De plus, d'après la Proposition 2.2-4 (p. 60),

$$\langle \mu^*, q^* \rangle_{L^2} = \mathcal{K}(\mu^*) = \mathcal{K}(\mu_\phi) = \langle \mu_\phi, q_\phi \rangle_{L^2},$$

d'où

$$\langle \mu^*, q^* \rangle_{L^2} = \langle \mu_\phi, q_\phi \rangle_{L^2} = \langle \mu^*, q_\phi \rangle_{L^2}$$

i.e. $q^* - q_\phi \perp \mu^*$.

Pour finir, (ψ^*, q^*, μ^*) étant un point selle de \mathbf{L} , on sait que μ^* est orthogonal au paraboloidé $\tilde{\mathcal{P}}$ au point q^* , i.e. $\langle \mu^*, q' - q^* \rangle_{L^2} \leq 0$, $\forall q' \in \tilde{\mathcal{P}}$, en rappelant que

$$\tilde{\mathcal{P}} = \left\{ q = (a, b) \in L^2((0, 1) \times \Omega) \times L^2((0, 1) \times \Omega)^d, a + \frac{|b|^2}{2} \leq 0 \right\}$$

Donc pour tout $q' \in \tilde{\mathcal{P}}$, on a

$$0 \geq \langle \mu^*, q' - q^* \rangle_{L^2} = \langle \mu^*, q' \rangle_{L^2} - \langle \mu^*, q^* \rangle_{L^2} = \langle \mu^*, q' \rangle_{L^2} - \langle \mu^*, q_\phi \rangle_{L^2} = \langle \mu^*, q' - q_\phi \rangle_{L^2}$$

i.e. μ^* est orthogonal au paraboloidé $\tilde{\mathcal{P}}$ au point q . D'après la Proposition 2.2-3 (p. 58), et sachant que $q_\phi = (-\frac{1}{2}|v_\phi|^2, v_\phi)$, on peut conclure que $m^* = \rho^*v_\phi$. \square

Ainsi, pour tout $\phi \in \Phi(\rho_0\mathcal{L}^d, \rho_1\mathcal{L}^d)$ fixé, et pour tout point selle $(\psi^*, q^*, \mu^*) \in \mathbf{L}^{ps}(\rho_0, \rho_1, \Omega)$ de \mathbf{L} , on a $m^* = \rho^*v_\phi$.

Autrement dit, l'ensemble des composantes densités ρ^* de tous les points selles de \mathbf{L} sont liés au même champ de vitesse v_ϕ , c'est à dire que l'ensemble des champs de vitesses auxquels peuvent être soumis les différentes densités ρ^* des composantes de points selles sont tous égal à v_ϕ sur l'union des supports de l'ensemble des ρ^* .

On en déduit également que les différents v_ϕ , pour tout $\phi \in \Phi(\rho_0\mathcal{L}^d, \rho_1\mathcal{L}^d)$, sont égaux sur cette même union de support. Cela peut paraître un peu absurde si l'on imagine qu'il puisse y avoir des supports différents pour les différentes densités : cela revient à dire que le champ de vitesse serait imposé même sur des zones où la densité s'annule.

Cette contradiction apparente sera levé dans la prochaine sous-section, lorsque nous montrerons **l'unicité de la densité ρ^* au sein des points selles de \mathbf{L} .**

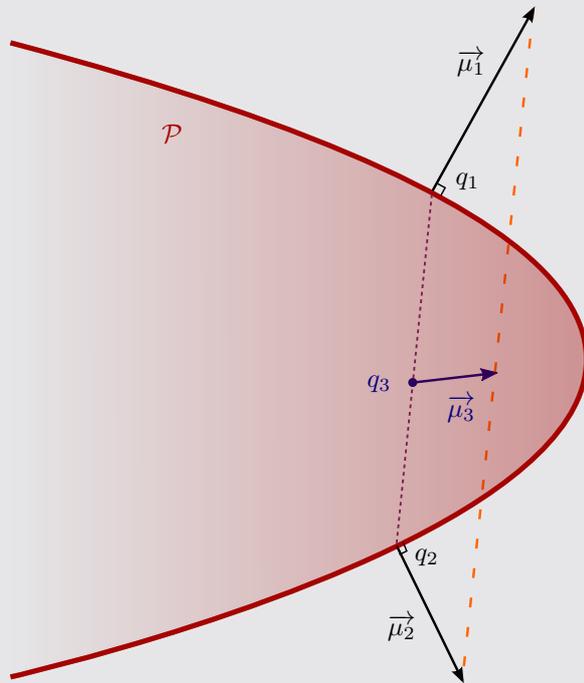
Illustration 4.2-6. UNICITÉ DU CHAMP DE VITESSE SUR LE SUPPORT DE L'ENSEMBLE DES DENSITÉS POTENTIELLES

Une vision plus "schématique" de l'unicité du champ de vitesse sur l'union des supports des densités candidates peut être abordée en invoquant la convexité de l'ensemble des points selles et la stricte convexité du parabolôïde $\mathcal{P} = \{(a, b) \in \mathbb{R} \times \mathbb{R}^d, a + |b|^2/2 \leq 0\}$.

En effet, supposons (ψ_1, q_1, μ_1) et (ψ_2, q_2, μ_2) deux points selles de \mathbf{L} . Les champs μ_1 et μ_2 sont tous deux orthogonales (au sens du produit scalaire canonique de L^2) à l'hyper-parabolôïde

$$\tilde{\mathcal{P}} = \left\{ (\tilde{a}, \tilde{b}) \in L^2((0, 1) \times \Omega) \times L^2((0, 1) \times \Omega)^d, \tilde{a} + |\tilde{b}|^2/2 \leq 0 \right\}$$

respectivement aux champs q_1 et q_2 . D'après la Remarque 2.2-1 (p. 57), l'ensemble des points selles de \mathbf{L} est convexe : $(1/2)[(\psi_1, q_1, \mu_1) + (\psi_2, q_2, \mu_2)]$ est donc également un point selle de \mathbf{L} . Le champ $(1/2)(\mu_1 + \mu_2)$ est donc également orthogonale à $\tilde{\mathcal{P}}$ au champ $(1/2)(q_1 + q_2)$.



Soit un point $(t_0, x_0) \in [0, 1] \times \Omega$ tel que les vecteurs $\mu_1(t_0, x_0)$, $\mu_2(t_0, x_0)$ ainsi que le vecteur $(1/2)(\mu_1 + \mu_2)(t_0, x_0)$ soient tous trois orthogonales au parabolôïde \mathcal{P} respectivement aux points $q_1(t_0, x_0)$, $q_2(t_0, x_0)$ et $(1/2)(q_1 + q_2)(t_0, x_0)$ (ce qui est bien le cas pour \mathcal{L}^d -presque tous point $(t, x) \in [0, 1] \times \Omega$); et tel que $\mu_1(t_0, x_0) \neq 0$ ou $\mu_2(t_0, x_0) \neq 0$: d'où $(1/2)(\mu_1 + \mu_2)(t_0, x_0) \neq 0$ (du fait l'orthogonalité des vecteurs $\mu_1(t_0, x_0)$ et $\mu_2(t_0, x_0)$ au parabolôïde \mathcal{P}).

Si l'on avait $q_1(t_0, x_0) \neq q_2(t_0, x_0)$, le point $(1/2)(q_1 + q_2)(t_0, x_0)$ serait alors intérieur au parabolôïde \mathcal{P} , du fait de la stricte convexité de celle-ci : le vecteur $(1/2)(\mu_1 + \mu_2)(t_0, x_0)$ serait alors forcément nul, ce qui est contredirait l'hypothèse faite plus haut à propos de ce vecteur.

Ainsi, on a $q_1(t_0, x_0) = q_2(t_0, x_0)$. Les vecteurs $\mu_1(t_0, x_0)$ et $\mu_2(t_0, x_0)$ sont donc tous deux orthogonales au parabolôïde \mathcal{P} au même point $q_1(t_0, x_0)$ et donc proportionnels au vecteur $(1, b_1(t_0, x_0))$.

On a alors :

$$\mu_1(t_0, x_0) = (\rho_1(t_0, x_0), m_1(t_0, x_0)) = (\rho_1(t_0, x_0), \rho_1(t_0, x_0)b_1(t_0, x_0))$$

et

$$\mu_2(t_0, x_0) = (\rho_2(t_0, x_0), m_2(t_0, x_0)) = (\rho_2(t_0, x_0), \rho_2(t_0, x_0)b_1(t_0, x_0)),$$

et ceci donc pour \mathcal{L}^d -presque tous $(t_0, x_0) \in \text{supp}(\mu_1) \cup \text{supp}(\mu_2)$.

Les champs μ_1 et μ_2 partagent donc le même champ de vitesse (i.e. le champ b_1) sur $\text{supp}(\mu_1) \cup \text{supp}(\mu_2)$.

Résumons : pour tout $\phi \in \Phi(\rho_0 \mathcal{L}^d, \rho_1 \mathcal{L}^d)$ fixé, et pour tout point selle (ψ^*, q^*, μ^*) de \mathbf{L} , ρ^* est associé au même champ de vitesse v_ϕ avec :

$$\partial_t \rho^* + \nabla \cdot (\rho^* v_\phi) = 0$$

(car $m^* = \rho^* v_\phi$) avec les conditions initiales et finales

$$\rho^*(0, \cdot) = \rho_0, \rho^*(1, \cdot) = \rho_1.$$

i.e. $\forall h \in H^1((0, 1) \times \Omega)$,

$$\int_0^1 \int_{\Omega} (\partial_t h(t, x) + v_{\phi}(t, x) \cdot \nabla_x h(t, x)) \rho^*(t, x) dx dt + \int_{\Omega} h(0, \cdot) \rho_0 - \int_{\Omega} h(1, \cdot) \rho_1 = 0.$$

Nous allons donc maintenant démontrer qu'il existe un unique ρ^* qui vérifie ces conditions, c'est à dire $\rho^*(t, \cdot) = X_{\phi}(t, \cdot) \# \rho_0$ (et ceci donc quel que soit $\phi \in \Phi(\rho_0 \mathcal{L}^d, \rho_1 \mathcal{L}^d)$). Nous nous baserons sur la méthode des caractéristiques, en nous inspirant de la Proposition 3.3-14 (p. 96).

IV.2.3 Unicité de la densité dans L^2

Dans la section précédente, nous avons montré que le champ de vitesse v associé au déplacement des densités était identique sur l'union des supports des densités candidates, et pouvait s'écrire sous une forme explicite, c'est à dire sous la forme :

$$v_{\phi}(t, \cdot) = \frac{1}{t} \left(\text{id} - \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{\cdot}{1-t} \right) \right),$$

où ϕ est un potentiel de transport optimal (vérifiant la propriété (Γ_1)) entre $\rho_0 \mathcal{L}^d$ et $\rho_1 \mathcal{L}^d$.

Ce résultat implique en particulier que, si l'on pouvait montrer l'unicité de la densité ρ , on pourrait en conclure également l'unicité des quantités de mouvement m (car $m = \rho v$). Cependant, c'est bien cette unicité même du champ de vitesse qui nous permettra, via une méthode des caractéristiques, de conclure à l'unicité de la densité et son explicitation (autrement dit qu'il s'agit bien de l'interpolation temporelle de McCann entre les densités ρ_0 et ρ_1).

L'idée de cette méthode, et de son adaptation aux difficultés du problème sera expliquée qualitativement dans le petit préambule introductif de la sous-section IV.2.4.

Dans la présente sous-section, nous énoncerons le résultat sur lequel reposera explicitement l'unicité de la densité (à savoir la Proposition 4.2-5), ainsi qu'un Lemme outil relatif à l'application de la méthode des caractéristiques (Lemme 4.2-4), laquelle sera, comme nous l'avons précisé ci-dessus, la base de notre raisonnement.

La preuve de la Proposition 4.2-5 sera finalement établie dans la sous-section IV.2.5, après le préambule introductif, lequel pourra donc guider la lecture en expliquant le raisonnement suivi. Ce résultat établi, nous pourrons enfin, dans la sous-section IV.2.6, conclure la preuve du Théorème 4.2-1 (p. 140) énoncé en début de section qui, rappelons-le, montre l'unicité de composante $\mu = (\rho, m)$ commune aux points selles du Lagrangien \mathbf{L} .

Lemme 4.2-4.

Soit $\mathcal{H} = \{f, f \in W_{loc}^{1,p}([0, 1] \times \mathbb{R}^d) \cap L^{\infty}((0, 1) \times \mathbb{R}^d), \forall p \in [1, 2]\}$. Soient de plus $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ convexe vérifiant la propriété (Γ_1) , et un champ de vitesse v définie par rapport à ϕ comme en (3.3-11) p. 89.

Pour tout $\varphi \in C_c^{\infty}((0, 1) \times \mathbb{R}^d)$, il existe $\tilde{h} \in \mathcal{H}$, définie pour tout $(t, x) \in]0, 1[\times \mathbb{R}^d$ par

$$\tilde{h}(t, x) = - \int_t^1 \varphi(s, (s-t)v(t, x) + x) ds, \quad (4.2-2)$$

et vérifiant

$$\partial_t \tilde{h} + v \cdot \nabla_x \tilde{h} = \varphi \quad (4.2-3)$$

et telle qu'il existe $0 < t_m < 1$ ne dépendant que de φ (en particulier ne dépendant pas de ϕ et donc de v), tel que $\tilde{h}(t, \cdot) = 0$ pour tout $t_m \leq t \leq 1$.

De plus, on a :

$$\nabla_{t,x} \tilde{h} = {}^t(\nabla_{t,x} v) \alpha + \beta \quad (4.2-4)$$

($\nabla_{t,x} v$ étant de taille $d \times (d+1)$), avec $\alpha \in L^\infty([0,1] \times \mathbb{R}^d)^d$, et $\beta \in L_{loc}^\infty([0,1] \times \mathbb{R}^d)^{d+1}$, tels que :

- $\|\alpha\|_{L^\infty([0,1] \times \mathbb{R}^d)} \leq \|\nabla_x \varphi\|_{L^\infty([0,1] \times \mathbb{R}^d)}$,
- pour tout ouvert borné $\omega \subset \mathbb{R}^d$,

$$\|\beta\|_{L^\infty([0,1] \times \omega)} \leq (\|v\|_{L^\infty([0,1] \times \omega)} + 1) \|\nabla_x \varphi\|_{L^\infty([0,1] \times \omega)} + \|\varphi\|_{L^\infty([0,1] \times \omega)},$$

- $\alpha(t, \cdot) = 0$ et $\beta(t, \cdot) = 0$ pour tout $t_m \leq t \leq 1$,
- $\left| \nabla_x \tilde{h}(t, x) \right| \leq (|\nabla_x v(t, x)| + 1) \|\nabla_x \varphi\|_{L^\infty([0,1] \times \mathbb{R}^d)}$, pour presque tout $(t, x) \in]0, 1[\times \mathbb{R}^d$.

La preuve de ce Lemme est établie dans l'annexe G. □

Passons donc au principal résultat de cette sous-section, qui nous permettra de conclure la preuve du Théorème 4.2-1 (p. 140).

Proposition 4.2-5.

Soient $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ convexe vérifiant la propriété (Γ_1) , $\rho_0 \in L^2(\mathbb{R}^d)$, et un champ de vitesse v défini par rapport à ϕ comme en (3.3-11) p. 89, c'est à dire rappelons-le, défini pour tout $t \in]0, 1[$ par

$$v(t, \cdot) = v_\phi(t, \cdot) = \frac{1}{t} \left(\text{id} - \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{\cdot}{1-t} \right) \right) = \frac{\text{id} - \nabla_x(\phi_t)^*}{t}$$

où $(\phi_t)^*$ désigne la transformée de Legendre de $\phi_t = (1-t)|\cdot|^2 + t\phi$. Soit $\rho \in L^2((0,1) \times \mathbb{R}^d)$, à support borné dans $[0,1] \times \mathbb{R}^d$, tel que

$$\begin{cases} \partial_t \rho + \text{div}_x(\rho v) = 0, \\ \rho(0, \cdot) = \rho_0 \end{cases}$$

(au sens des distributions), i.e.

$$\forall h \in C_c^\infty([0,1] \times \mathbb{R}^d), \int_0^1 \int_{\mathbb{R}^d} (\partial_t h + v \cdot \nabla_x h) \rho \, dx \, dt + \int_{\mathbb{R}^d} h(0, \cdot) \rho_0 \, dx = 0. \quad (4.2-5)$$

Alors $\rho(t, \cdot) = \rho_\phi(t, \cdot) = (t\nabla\phi + (1-t)\text{id})\# \rho_0$ pour presque tout $t \in [0,1]$. Autrement dit :

$$\forall \varphi \in C_c^0([0,1] \times \mathbb{R}^d), \int_0^1 \int_{\mathbb{R}^d} \varphi \rho \, dx \, dt = \int_0^1 \int_{\mathbb{R}^d} \varphi(t, t\nabla\phi(x) + (1-t)x) \rho_0(x) \, dx \, dt. \quad (4.2-6)$$

De plus $t \mapsto \rho(t, \cdot) \in C^0([0,1[, L^2(\mathbb{R}^d))$.

IV.2.4 Préambule à la démonstration de la Proposition 4.2-5

Nous allons donner un aperçu du raisonnement qui sera suivi dans cette sous-section pour prouver l'unicité de ρ (celle de v étant déjà établie par le Lemme 4.2-3 p. 142).

Soit (ψ, q, μ) est un point selle de \mathbf{L} tel que défini dans la notation 2 (p. 74), et soit $\phi \in \Phi(\rho_0 \mathcal{L}^d, \rho_1 \mathcal{L}^d)$ (vérifiant donc la propriété (Γ_1) sur \mathbb{R}^d). Celui-ci vérifie les propriétés (I), ce qui implique en particulier :

- $\forall h \in H^1((0, 1) \times \Omega)$, $G(h) + \langle \mu, \nabla_{t,x} h \rangle = 0$,
- $\mu = (\rho, m) \in L^2((0, 1) \times \Omega)$,
- $m = \rho v$, avec v définie pour tout $(t, x) \in]0, 1[\times \Omega$ par

$$v(t, x) = \frac{1}{t} \left(x - \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{x}{1-t} \right) \right), \quad (4.2-7)$$

(voir le Lemme 4.2-3) et vérifie les propriétés (II) et (III) (voir la conclusion de la section III.3), d'où $v \in \mathcal{H}$ (\mathcal{H} étant défini dans le Lemme 4.2-4).

De ces propriétés, on déduit que pour tout $h \in H^1((0, 1) \times \Omega)$:

$$\int_0^1 \int_{\Omega} (\partial_t h + v \cdot \nabla_x h) \rho \, dx \, dt = \int_{\Omega} h(1, \cdot) \rho_1 \, dx - \int_{\Omega} h(0, \cdot) \rho_0 \, dx. \quad (4.2-8)$$

Soit $\varphi \in C_c^\infty(]0, 1[\times \mathbb{R}^d)$ tel que $\text{supp}(\varphi) \subset]0, 1[\times \Omega$. D'après le Lemme 4.2-4, il existe $\tilde{h} \in \mathcal{H}$ tel que

$$\partial_t \tilde{h} + v \cdot \nabla_x \tilde{h} = \varphi$$

et

$$\nabla_{t,x} h = {}^t(\nabla_{t,x} v) \alpha + \beta \quad (4.2-9)$$

avec $\alpha \in L^\infty([0, 1] \times \mathbb{R}^d)^d$, et $\beta \in L_{loc}^\infty([0, 1] \times \mathbb{R}^d)^{d+1}$ ($\nabla_{t,x} v$ étant de taille $d \times (d+1)$).

De plus, on a

$$\tilde{h}(1, \cdot) = 0, \quad \text{et} \quad \tilde{h}(0, \cdot) = - \int_0^1 \varphi(t, X(t, \cdot)) \, dt, \quad \text{avec} \quad X(t, \cdot) = t \nabla \phi + (1-t) \text{id} \quad (4.2-10)$$

(voir la Remarque suivante).

Remarque 4.2-7 (A propos de la trace de \tilde{h} en $t = 0$). *La fonction \tilde{h} étant un élément de l'espace \mathcal{H} et donc de tout espace $W^{1,p}((0, 1) \times \Omega)$ (du moins sa restriction à Ω) pour tout $1 \leq p < 2$, elle possède une trace dans $L^2(\Omega)$ en $t = 0$ et $t = 1$, quelle que soit la dimension d de l'espace de transport : il suffit de choisir p assez proche de 2. Cependant dans les conditions (4.2-10), la relation*

$$\tilde{h}(0, \cdot) = - \int_0^1 \varphi(t, X(t, \cdot)) \, dt = - \int_0^1 \varphi(t, t \nabla \phi + (1-t) \text{id}) \, dt$$

n'est pas non plus évidente à montrer (car en général v ne se prolonge pas de façon continue en 0), même si elle est très intuitive. En effet, on a pour tout $t \in]0, 1[$, on a

$$\tilde{h}(t, X(t, x)) = - \int_t^1 \varphi(s, (s-t)v(t, X(t, x)) + X(t, x)) \, ds = - \int_t^1 \varphi(s, X(s, \cdot)) \, ds \quad (4.2-11)$$

sachant que $X(0, \cdot) = \text{id}$. Néanmoins, comme $X(t, \cdot)$ n'est en général pas inversible, et $\nabla \phi$ est en général non-continue, la relation 4.2-10 ne peut se déduire immédiatement de la relation 4.2-11.

Cependant il ne serait en fait pas nécessaire de montrer cette relation pour pouvoir conclure : nous pourrions contourner le problème de la manière suivante : le Théorème 3.1-2 donne l'existence d'un point selle $(\bar{\psi}, \bar{q}, \bar{\mu})$ vérifiant les propriétés (I), avec en particulier $\bar{\mu} = (\bar{\rho}, \bar{m})$ et $\bar{m} = \bar{\rho}v$, le champ de vitesse v étant le même que celui associé à ρ , c'est à dire tel que défini en (4.2-7), avec $\bar{\rho}$ vérifiant la relation (4.2-13). Ce qui permettrait d'écrire, pour tout $h \in H^1((0, 1) \times \Omega)$,

$$\int_0^1 \int_{\Omega} (\partial_t h + v \cdot \nabla_x h) (\rho - \bar{\rho}) dx dt = 0,$$

et en réalisant le développement (4.2-12) nous donnerait

$$\int_0^1 \int_{\Omega} \varphi(t, x) \rho(t, x) dx dt = 0,$$

d'où $\rho = \bar{\rho}$ sur $(0, 1) \times \Omega$, et nous pourrions ainsi conclure à l'unicité de la composante μ au sein des points selles de \mathbf{L} sans avoir eu à expliciter $\tilde{h}(0, \cdot)$.

Le problème ne pourra cependant pas être contourné de cette façon lorsque nous aborderons la dernière sous-section de ce chapitre. En effet le cadre du Théorème d'existence 3.1-2 est trop restreint par rapport à celui du Théorème d'unicité 4.3-9 p. 159 (qui sera énoncé dans la section IV.3), et nous ne pourrions donc pas supposer l'existence d'un point selle dans ce cadre. Nous traiterons en fait ce problème par une méthode de régularisation, ainsi que nous le verrons plus bas.

Deux cas auraient, s'ils avaient eu lieu, facilité la démonstration :

- **Cas 1** : Si l'on avait pu avoir $v \in H_{loc}^1([0, 1] \times \mathbb{R}^d)$, on aurait eu :

$$\begin{aligned} \|\nabla_{t,x} \tilde{h}\|_{L^2((0,1) \times \omega)}^2 &\leq \|\alpha\|_{L^\infty((0,1) \times \omega)}^2 \|\nabla_{t,x} v\|_{L^2((0,1) \times \omega)}^2 + \mathcal{L}^d(\omega) \|\beta\|_{L^\infty((0,1) \times \omega)}^2 \\ &< +\infty, \end{aligned}$$

pour tout ouvert borné $\omega \subset \mathbb{R}^d$. D'où l'on déduit $\tilde{h} \in H_{loc}^1([0, 1] \times \mathbb{R}^d)$ et ainsi $\tilde{h} \in H^1((0, 1) \times \Omega)$ (ou du moins le restriction de \tilde{h} à Ω).

- **Cas 2** : Si l'on avait pu avoir $\rho \in L^q((0, 1) \times \Omega)$ avec $q > 2$ (plutôt que $\rho \in L^2((0, 1) \times \Omega)$), il aurait été possible d'étendre la relation de conservation de la masse (4.2-8) à tout $h \in W^{1,p}((0, 1) \times \Omega)$ (au lieu de $h \in H^1((0, 1) \times \Omega)$) pour $1 \leq p < 2$ vérifiant $1/p + 1/q = 1$, et ainsi à tout $h \in \mathcal{H}$.

Dans ces deux cas, il aurait été possible d'injecter \tilde{h} dans la relation (4.2-8) :

$$\begin{aligned} \int_0^1 \int_{\Omega} \varphi(t, x) \rho(t, x) dx dt &= \int_0^1 \int_{\Omega} \left(\partial_t \tilde{h}(t, x) + v(t, x) \cdot \nabla_x \tilde{h}(t, x) \right) \rho(t, x) dx dt \\ &= \int_{\Omega} \tilde{h}(1, x) \rho_1(x) dx - \int_{\Omega} \tilde{h}(0, x) \rho_0(x) dx = - \int_{\Omega} \tilde{h}(0, x) \rho_0(x) dx \\ &= \int_{\Omega} \int_0^1 \varphi(t, X(t, x)) dt \rho_0(x) dx = \int_0^1 \int_{\Omega} \varphi(t, X(t, x)) \rho_0(x) dx dt. \end{aligned} \tag{4.2-12}$$

Et ceci pour tout $\varphi \in C_c^\infty([0, 1] \times \mathbb{R}^d)$ tel que $\text{supp}(\varphi) \subset]0, 1[\times \Omega$. Ceci permettrait de conclure que ρ est bien le même que celui construit dans le Théorème 3.1-2 (existence d'un point selle pour \mathbf{L}), à savoir

$$\rho(dx \otimes dt) = (\text{id}_\tau, X) \# (\rho_0 dx \otimes dt), \text{ avec } \text{id}_\tau : (t, x) \mapsto t \tag{4.2-13}$$

(quitte à prolonger à $(0, 1) \times \mathbb{R}^d$ pour bien définir la poussée de mesure en accord avec la notion de poussée de mesure telle que développée dans l'annexe F).

Malheureusement ce ne sera pas le cas : nous sommes dans l'impossibilité de montrer l'un et l'autre de ces deux cas. Pour autant nous n'abandonnons pas "l'idée", c'est à dire de traiter le problème plus ou moins par la méthode des caractéristiques :

- Pour se raccrocher au premier de ces deux cas, nous allons approcher le champ de vitesse v (associé au plan de transport $\nabla\phi$) par des champs de vitesses $v_\gamma = v_{\gamma\phi}$ associés à des plans de transport régularisés : c'est à dire des plans de transport de type $\nabla^{\gamma\phi}$, où $\gamma\phi$ désigne la γ -régularisation par enveloppe de Moreau du potentiel ϕ (voir la définition 3.3-8 p. 88). Cette régularisation a en particulier pour propriété de gommer les cassures du plan de transport, responsables du fait que v ne soit pas de régularité H^1 au voisinage de $t = 0$ (le voisinage de $t = 1$ ne posant pas de problèmes car, par construction, \tilde{h} est uniformément nulle au niveau de celui-ci).
- La convergence de la relation (4.2-8) injectée des \tilde{h}_γ (associées aux v_γ) vers une relation de type (4.2-13) fera appel aux résultats développés dans la section III.3, en particulier ceux des sous-sections III.3.9 et III.3.10. En outre, pour prouver cette convergence, nous nous raccrocherons également au second cas évoqué en faisant également appel à la densité des espaces $L^q((0, 1) \times \Omega)$ ($q > 2$) dans $L^2((0, 1) \times \Omega)$ (voir les Lemmes 4.2-6 et 4.2-7 p. 149).
- En outre, le problème de l'explicitation de la trace de \tilde{h}_γ , tel qu'abordé dans la Remarque 4.2-7 ne se posera pas étant donné que les v_γ se prolongent par continuité en $t = 0$, et ainsi de même pour les \tilde{h}_γ (voir la Proposition 3.3-18 p. 99). La preuve de cette prolongation suit d'ailleurs le même raisonnement que celui illustré par la relation (4.2-11) : en effet dans ce cas $X_\gamma(t, \cdot)$ sera inversible et $\nabla^{\gamma\phi}$ sera continue.

IV.2.5 Résultats préalables et preuve de la Proposition 4.2-5

On rappelle que pour un potentiel $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ vérifiant la propriété (Γ_1) , le champ de vitesse v_ϕ , définie en (3.3-11) (p. 89) s'écrit pour tout $t \in]0, 1[$:

$$v_\phi(t, \cdot) = \frac{1}{t} \left(\text{id} - \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{\cdot}{1-t} \right) \right) = \frac{\text{id} - \nabla_x(\phi_t)^*}{t}, \quad (4.2-14)$$

où $(\phi_t)^*$ désigne la transformée de Legendre de $\phi_t = (1-t)|\cdot|^2 + t\phi$.

Avant d'entamer la preuve de la Proposition 4.2-5, nous commencerons par démontrer les deux Lemmes suivants :

Lemme 4.2-6.

Soient $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ convexe vérifiant la propriété (Γ_1) , et un champ de vitesse v définie par rapport à ϕ comme en (3.3-11) p. 89 (c'est à dire $v = v_\phi$, voir (4.2-14)), et $0 < t_m < 1$.

Soit $\rho \in L^2((0, 1) \times \mathbb{R}^d)$, à support borné dans $[0, 1] \times \mathbb{R}^d$, et $\Omega \subset \mathbb{R}^d$ un ouvert borné tel que $\text{supp}(\rho) \subset [0, 1] \times \Omega$.

Pour tout $\gamma > 0$, on définit v_γ comme en (3.3-48) p. 109 : c'est à dire $v_\gamma = v_{\gamma\phi}$, où $\gamma\phi$ est l'enveloppe de Moreau de ϕ par γ , voir la définition 3.3-8 (p. 88).

Alors

1. Il existe une constante $M > 0$, indépendante de γ , tel que

$$\forall \gamma \in]0, 1], \int_0^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt \leq M \|\rho\|_{L^2((0,1) \times \Omega)}. \quad (4.2-15)$$

2. Si de plus il existe $p > 2$ tel que $\rho \in L^p((0, 1) \times \mathbb{R}^d)$, alors

$$\int_0^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt \xrightarrow{\gamma \rightarrow 0} 0. \quad (4.2-16)$$

Le deuxième point du Lemme 4.2-6 se révèle en fait également vrai pour le cas où l'on aurait seulement $\rho \in L^2((0, 1) \times \mathbb{R}^d)$, ainsi que le prouve Lemme suivant :

Lemme 4.2-7.

Sous les mêmes hypothèses que le Lemme 4.2-6 (p. 149), on a le résultat de convergence :

$$\int_0^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt \xrightarrow{\gamma \rightarrow 0} 0. \quad (4.2-17)$$

Preuve du Lemme 4.2-6 : Soit $\Omega \subset \mathbb{R}^d$ un ouvert borné tel que $\text{supp}(\rho) \subset [0, 1] \times \Omega$. Démontrons dans l'ordre ces deux propriétés.

1. D'après la Proposition 3.3-30 p. 112 (en particulier la relation (3.3-54)), il existe $M_0 > 0$ indépendante de $0 < \gamma \leq 1$, tel que pour tout $t \in]0, 1[$ et tout $0 < \gamma \leq 1$,

$$\|v(t, \cdot) - v_\gamma(t, \cdot)\|_{L^\infty(\Omega)} \leq M_0 \frac{\gamma}{t},$$

et d'après cette même Proposition, v_γ est uniformément bornée par une constante indépendante de $\gamma \leq 1$ que nous nommerons ici M_1 . D'où, pour tout $\gamma > 0$, $0 < \alpha < 1$ et $t \in]0, 1[$,

$$\|v(t, \cdot) - v_\gamma(t, \cdot)\|_{L^\infty(\Omega)} \leq M_0^\alpha M_1^{1-\alpha} \left(\frac{\gamma}{t}\right)^\alpha. \quad (4.2-18)$$

D'après le deuxième point du Commentaire 3.3-12 (p. 91), il existe une constante $c > 0$ indépendante de ϕ (et donc de $\gamma\phi$), tel que pour tout $t \in]0, 1[$,

$$\|\nabla_x v_\gamma(t, \cdot)\|_{L^\infty(\mathbb{R}^d)} \leq \frac{c}{t(1-t)}, \quad (4.2-19)$$

(où la norme L^∞ est ici prise sur $\|\cdot\|$: la norme subordonnée à la norme euclidienne $|\cdot|$ de \mathbb{R}^d).

De plus, d'après le Lemme 3.3-35 (p. 118), il existe $c' > 0$, tel que pour tout $t \in]0, 1[$,

$$\|\nabla_x v_\gamma(t, \cdot)\|_{L^\infty(\mathbb{R}^d)} \leq \frac{c'}{\gamma(1-t)}, \quad (4.2-20)$$

Nous allons utiliser le découpage suivant :

$$\begin{aligned} & \int_0^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt \\ &= \int_0^\gamma \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt + \int_\gamma^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt, \end{aligned}$$

et majorer chacun des deux termes : le premier en utilisant la majoration (4.2-20) et le second grâce à la majoration (4.2-19).

- Majorons donc le premier terme : pour tout $0 < \gamma < t_m$,

$$\begin{aligned} & \int_0^\gamma \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt \\ & \leq M_1 \int_0^\gamma \int_{\Omega} |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt \\ & \leq \frac{\sqrt{c'} M_1}{\sqrt{\gamma}} \int_0^\gamma \frac{1}{\sqrt{1-t}} \int_{\Omega} |\nabla_x v_\gamma(t, \cdot)|^{\frac{1}{2}} \cdot |\rho(t, \cdot)| \, dx \, dt \\ & \leq \frac{\sqrt{c'} M_1}{\sqrt{1-t_m} \sqrt{\gamma}} \int_0^\gamma \left(\int_{\Omega} |\nabla_x v_\gamma(t, \cdot)| \, dx \right)^{\frac{1}{2}} \cdot \left(\int_{\Omega} |\rho|^2 \, dx \right)^{\frac{1}{2}} \, dt. \end{aligned}$$

Or d'après le Corollaire 3.3-34 (p. 117), il existe une constante K tel que pour tout $\gamma > 0$ et pour tout $t \in]0, t_m]$, on ait

$$\int_{\Omega} |\nabla_x v_\gamma(t, x)| \, dx \leq K,$$

et ainsi

$$\begin{aligned} \int_0^\gamma \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt & \leq \frac{\sqrt{K c'} M_1}{\sqrt{1-t_m} \sqrt{\gamma}} \int_0^\gamma \left(\int_{\Omega} |\rho|^2 \, dx \right)^{\frac{1}{2}} \, dt \\ & \leq \frac{\sqrt{K c'} M_1}{\sqrt{1-t_m} \sqrt{\gamma}} \cdot \sqrt{\gamma} \left(\int_0^1 \int_{\Omega} |\rho|^2 \, dx \, dt \right)^{\frac{1}{2}} \quad (4.2-21) \\ & \leq \frac{\sqrt{K c'} M_1}{\sqrt{1-t_m}} \cdot \|\rho\|_{L^2((0,1) \times \Omega)}. \end{aligned}$$

- Passons au second terme : pour tout $0 < \gamma < t_m$,

$$\begin{aligned} & \int_\gamma^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt \\ & \leq \sqrt{c M_0 M_1} \sqrt{\gamma} \int_\gamma^{t_m} \frac{1}{t \sqrt{1-t}} \int_{\Omega} |\nabla_x v_\gamma|^{\frac{1}{2}} \cdot |\rho| \, dx \, dt \\ & \leq \sqrt{\frac{c M_0 M_1}{1-t_m}} \sqrt{\gamma} \int_\gamma^{t_m} \frac{1}{t} \left(\int_{\Omega} |\nabla_x v_\gamma(t, \cdot)| \, dx \right)^{\frac{1}{2}} \cdot \left(\int_{\Omega} |\rho|^2 \, dx \right)^{\frac{1}{2}} \, dt \quad (4.2-22) \\ & \leq \sqrt{\frac{c K M_0 M_1}{1-t_m}} \sqrt{\gamma} \left(\int_\gamma^{t_m} \frac{1}{t^2} \, dt \right)^{\frac{1}{2}} \cdot \left(\int_\gamma^{t_m} \int_{\Omega} |\rho|^2 \, dx \, dt \right)^{\frac{1}{2}} \\ & \leq \sqrt{\frac{c K M_0 M_1}{1-t_m}} \sqrt{\gamma} \sqrt{\frac{1}{\gamma} - \frac{1}{t_m}} \cdot \|\rho\|_{L^2((0,1) \times \Omega)} \\ & \leq \sqrt{\frac{c K M_0 M_1}{1-t_m}} \|\rho\|_{L^2((0,1) \times \Omega)}. \end{aligned}$$

Ainsi, en combinant (4.2-21) et (4.2-22), on obtient que

$$\begin{aligned} & \int_0^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt \\ &= \int_0^\gamma \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt + \int_\gamma^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt \\ &\leq M \|\rho\|_{L^2((0,1) \times \Omega)}, \end{aligned} \tag{4.2-23}$$

$$\text{avec } M = \max \left\{ \frac{\sqrt{Kc'} M_1}{\sqrt{1-t_m}}, \sqrt{\frac{cK M_0 M_1}{1-t_m}} \right\}.$$

2. Passons au deuxième point du Lemme : supposons qu'il existe $p > 2$, tel que $\rho \in L^p((0,1) \times \mathbb{R}^d)$. Pour commencer, remarquons que :

$$|\nabla_x v(t, x)| = |\nabla_x v(t, x)|^{\frac{1}{p} + \frac{1}{q}} \leq \left(\frac{c}{t(1-t)} \right)^{\frac{1}{p}} |\nabla_x v(t, x)|^{\frac{1}{q}}. \tag{4.2-24}$$

Pour tout $0 < \alpha < 1$, et en reprenant la relation (4.2-18), on a

$$\begin{aligned} \int_0^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt &\leq M_0^\alpha M_1^{1-\alpha} \int_0^{t_m} \left(\frac{\gamma}{t} \right)^\alpha \left(\frac{c}{t(1-t)} \right)^{\frac{1}{p}} \int_{\Omega} |\nabla_x v_\gamma|^{\frac{1}{q}} \cdot |\rho| \, dx \, dt \\ &\leq M_0^\alpha M_1^{1-\alpha} \left(\frac{c}{1-t_m} \right)^{\frac{1}{p}} \gamma^\alpha \int_0^{t_m} \frac{1}{t^{\alpha+1/p}} \int_{\Omega} |\nabla_x v_\gamma|^{\frac{1}{q}} \cdot |\rho| \, dx \, dt \end{aligned}$$

$$\text{On pose } M(\alpha) = M_0^\alpha M_1^{1-\alpha} \left(\frac{c}{1-t_m} \right)^{\frac{1}{p}}.$$

$$\begin{aligned} \int_0^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt &\leq M(\alpha) \gamma^\alpha \int_0^1 \frac{1}{t^{\alpha+1/p}} \left(\int_{\Omega} |\nabla_x v_\gamma| \, dx \right)^{\frac{1}{q}} \cdot \left(\int_{\Omega} |\rho|^p \, dx \right)^{\frac{1}{p}} dt \\ &\leq M(\alpha) K^{\frac{1}{q}} \gamma^\alpha \int_0^1 \frac{1}{t^{\alpha+1/p}} \left(\int_{\Omega} |\rho|^p \, dx \right)^{\frac{1}{p}} dt \\ &\leq M(\alpha) K^{\frac{1}{q}} \gamma^\alpha \left(\int_0^1 \frac{1}{t^{q\alpha+q/p}} \, dt \right)^{\frac{1}{q}} \left(\int_0^1 \int_{\Omega} |\rho|^p \, dx \, dt \right)^{\frac{1}{p}} \\ &= M(\alpha) K^{\frac{1}{q}} \gamma^\alpha \left(\int_0^1 \frac{1}{t^{q\alpha+q-1}} \, dt \right)^{\frac{1}{q}} \|\rho\|_{L^p((0,1) \times \Omega)}. \end{aligned}$$

Or $q < 2$, et il s'en suit qu'en fixant $0 < \alpha < 1$ assez petit, on peut avoir $q(\alpha+1) < 2 \Leftrightarrow q\alpha + q - 1 < 1$, et

$$\int_0^1 \frac{1}{t^{q\alpha+q-1}} \, dt = \frac{1}{2 - q(\alpha+1)}.$$

D'où au final

$$\int_0^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt \leq M(\alpha) K^{\frac{1}{q}} \gamma^\alpha \left(\frac{1}{2 - q(\alpha+1)} \right)^{\frac{1}{q}} \|\rho\|_{L^p((0,1) \times \Omega)}$$

et on obtient bien ainsi

$$\int_0^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt \xrightarrow{\gamma \rightarrow 0} 0.$$

□

Preuve du Lemme 4.2-7 : Nous allons dans cette preuve faire appel aux propriétés de convergence énoncés dans le Lemme 4.2-6 (p. 149), et utiliser la densité des espaces $L^p((0, 1) \times \omega)$ ($p > 2$) dans $L^2((0, 1) \times \omega)$ (pour $\omega \subset\subset \mathbb{R}^d$).

Soit $\epsilon > 0$. Pour tout $0 < \lambda < 1$, si $\rho \in L^{2/\lambda}((0, 1) \times \Omega)$, alors $|\rho|^\lambda \in L^2((0, 1) \times \Omega)$ et $2/\lambda > 2$, et nous obtenons à partir de (4.2-33), la majoration : pour tout $\gamma > 0$ et tout $0 < \lambda < 1$,

$$\begin{aligned} \int_0^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt &\leq \int_0^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot \left| |\rho| - |\rho|^\lambda \right| \, dx \, dt \\ &+ \int_0^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho|^\lambda \, dx \, dt. \end{aligned} \quad (4.2-25)$$

Le premier point du Lemme 4.2-6 nous permet d'affirmer qu'il existe une constante $M > 0$, indépendante de γ tel que pour tout $\lambda \in]0, 1[$:

$$\forall \gamma \in]0, 1], \int_0^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot \left| |\rho| - |\rho|^\lambda \right| \, dx \, dt \leq M \left\| |\rho| - |\rho|^\lambda \right\|_{L^2((0,1) \times \Omega)}. \quad (4.2-26)$$

Ainsi, $(|\rho|^\lambda)_{\lambda \in]0, 1[}$ converge simplement vers $|\rho|$ lorsque λ tend vers 1, et pour tout $0 < \lambda < 1$, on a la majoration :

$$\left| |\rho| - |\rho|^\lambda \right|^2 \leq 2 \left(|\rho|^2 + |\rho|^{2\lambda} \right) \leq 2 \left(|\rho|^2 + (|\rho| + 1)^2 \right) \in L^1((0, 1) \times \Omega).$$

En effet si $|\rho(t, x)| \leq 1$, alors $|\rho(t, x)|^\lambda \leq 1$, et si $|\rho(t, x)| \geq 1$, alors $|\rho(t, x)|^\lambda \leq |\rho(t, x)|$, d'où $|\rho|^\lambda \leq |\rho| + 1$. Ainsi par convergence dominée, le terme $\left\| |\rho| - |\rho|^\lambda \right\|_{L^2((0,1) \times \Omega)}$ tend vers 0 lorsque λ tend vers 1.

Il est donc possible de fixer alors $\lambda \in]0, 1[$ tel que

$$\left\| |\rho| - |\rho|^\lambda \right\|_{L^2((0,1) \times \Omega)} \leq \frac{\epsilon}{M},$$

et ainsi

$$\forall \gamma \in]0, 1], \int_0^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot \left| |\rho| - |\rho|^\lambda \right| \, dx \, dt \leq \epsilon.$$

Fixons donc un tel λ . On a donc

$$\int_0^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt \leq \epsilon + \int_0^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho|^\lambda \, dx \, dt. \quad (4.2-27)$$

Pour finir, du fait que $|\rho|^\lambda \in L^{2/\lambda}((0, 1) \times \Omega)$ et que $2/\lambda > 2$, le deuxième point du Lemme 4.2-6 nous donne la convergence :

$$\int_0^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho|^\lambda \, dx \, dt \xrightarrow[\gamma \rightarrow 0]{} 0. \quad (4.2-28)$$

Ainsi, pour $\gamma > 0$ assez petit, on a :

$$\int_0^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt \leq 2\epsilon.$$

On peut donc finalement conclure à la convergence vers 0 du terme

$$\int_0^{t_m} \int_{\mathbb{R}^d} |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| dx dt$$

lorsque γ tend vers 0. □

Revenons maintenant à la preuve de la Proposition 4.2-5 (p. 146).

Preuve de la Proposition 4.2-5 : Soit $\Omega \subset \mathbb{R}^d$ un ouvert borné tel que $\text{supp}(\rho_0) \subset \Omega$ et $\text{supp}(\rho) \subset [0, 1] \times \Omega$. Il est donc possible de restreindre (4.2-5) à Ω et, du fait que $\rho \in L^2((0, 1) \times \mathbb{R}^d)$, de prolonger les fonctions test à $H_c^1([0, 1] \times \overline{\Omega})$, i.e.

$$\forall h \in H_c^1([0, 1] \times \overline{\Omega}), \int_0^1 \int_{\Omega} (\partial_t h + v \cdot \nabla_x h) \rho dx dt + \int_{\Omega} h(0, \cdot) \rho_0 dx = 0 \quad (4.2-29)$$

($h \in H_c^1([0, 1] \times \overline{\Omega})$ admettant bien une trace dans $L^2(\overline{\Omega})$). Pour tout $\gamma > 0$, on définit v_γ comme dans le Lemme 4.2-6 plus haut : c'est à dire $v_\gamma = v_{\gamma\phi}$, où $\gamma\phi$ est l'enveloppe de Moreau de ϕ par γ .

Soit $\varphi \in C_c^\infty((0, 1) \times \mathbb{R}^d)$. Pour tout $\gamma > 0$, on définit \tilde{h}_γ comme dans le Lemme 4.2-4 (p. 145), c'est à dire par :

$$\tilde{h}_\gamma(t, x) = - \int_t^1 \varphi(s, (s-t)v_\gamma(t, x) + x) ds,$$

pour tout $(t, x) \in]0, 1[\times \mathbb{R}^d$. Celui-ci vérifie donc

$$\forall \gamma > 0, \partial_t \tilde{h}_\gamma + v_\gamma \cdot \nabla_x \tilde{h}_\gamma = \varphi, \quad (4.2-30)$$

Il existe de plus $0 < t_m < 1$, ne dépendant que φ (indépendant donc de v_γ et plus généralement de γ), et, pour tout $\gamma > 0$, $\alpha_\gamma \in L^\infty([0, 1] \times \mathbb{R}^d)^d$ et $\beta_\gamma \in L_{loc}^\infty([0, 1] \times \mathbb{R}^d)^{d+1}$, tels que :

- $\nabla_{t,x} \tilde{h}_\gamma = {}^t(\nabla_{t,x} v_\gamma) \alpha_\gamma + \beta_\gamma$,
 - $\|\alpha_\gamma\|_{L^\infty([0,1] \times \mathbb{R}^d)} \leq \|\nabla_x \varphi\|_{L^\infty([0,1] \times \mathbb{R}^d)}$,
 - $\|\beta_\gamma\|_{L^\infty([0,1] \times \omega)} \leq (\|v_\gamma\|_{L^\infty([0,1] \times \Omega)} + 1) \|\nabla_x \varphi\|_{L^\infty([0,1] \times \Omega)} + \|\varphi\|_{L^\infty([0,1] \times \Omega)}$,
 - $\tilde{h}_\gamma(t, \cdot) = \alpha(t, \cdot) = \beta(t, \cdot) = 0$ pour tout $t_m \leq t \leq 1$,
 - $\left| \nabla_x \tilde{h}_\gamma(t, x) \right| \leq (\|\nabla_x v_\gamma(t, x)\| + 1) \|\nabla_x \varphi\|_{L^\infty([0,1] \times \mathbb{R}^d)}$,
- pour presque tout $(t, x) \in]0, 1[\times \mathbb{R}^d$.

Rappelons que d'après le Lemme 3.3-35 (p. 118), il existe $c' > 0$, tel que pour tout $t \in]0, 1[$,

$$\|\nabla_x v_\gamma(t, \cdot)\|_{L^\infty(\mathbb{R}^d)} \leq \frac{c'}{\gamma(1-t)},$$

(où la norme L^∞ est ici prise sur $\|\cdot\|$: la norme subordonnée à la norme euclidienne $|\cdot|$ de \mathbb{R}^d).

Cette propriété a déjà été utilisé dans la preuve du Lemme 4.2-6 p. 149 (équation (4.2-20)). Donc, d'après les Propositions 3.3-30 (p. 112) et 3.3-31 (p. 115), on obtient le même type de majoration pour $|\partial_t v_\gamma(t, \cdot)|$, et ainsi, pour tout $\gamma > 0$, $\nabla_{t,x} \tilde{h}_\gamma \in L^\infty([0, 1] \times \mathbb{R}^d)$, d'où $\tilde{h}_\gamma \in H^1((0, 1) \times \Omega)$, et comme de plus $\tilde{h}_\gamma(t, \cdot) = 0$ pour tout $t_m \leq t \leq 1$, on a bien $\tilde{h}_\gamma \in H^1([0, 1] \times \overline{\Omega})$. Ajoutons que, $\gamma\phi$ vérifiant la propriété (Γ_1) (Lemme 3.3-27 p. 110), et d'après le dernier point de la Proposition 3.3-18 (p. 99), \tilde{h}_γ est continue sur $[0, 1] \times \overline{\Omega}$, et pour tout $x \in \mathbb{R}^d$,

$$\tilde{h}_\gamma(0, x) = - \int_0^1 \varphi(s, s v_\gamma(0, x) + x) ds = - \int_0^1 \varphi(s, s \nabla \gamma \phi(x) + (1-s)x) ds \quad (4.2-31)$$

(coïncidant avec la trace L^2 de \tilde{h}_γ en $t = 0$).

Donc en insérant \tilde{h}_γ dans (4.2-29), on obtient

$$\int_0^1 \int_\Omega \left(\partial_t \tilde{h}_\gamma + v \cdot \nabla_x \tilde{h}_\gamma \right) \rho \, dx \, dt + \int_\Omega \tilde{h}_\gamma(0, \cdot) \rho_0 \, dx = 0,$$

pour tout $\gamma > 0$. En insérant en outre dans cette dernière relation l'équation (4.2-30), on obtient la nouvelle équation

$$\forall \gamma > 0, \int_0^1 \int_\Omega \varphi \rho \, dx \, dt + \int_\Omega \tilde{h}_\gamma(0, \cdot) \rho_0 \, dx = \int_0^1 \int_\Omega (v_\gamma - v) \cdot \nabla_x \tilde{h}_\gamma \rho \, dx \, dt. \quad (4.2-32)$$

Ainsi, on a la majoration

$$\begin{aligned} \forall \gamma > 0, \left| \int_0^1 \int_\Omega \varphi \rho \, dx \, dt + \int_\Omega \tilde{h}_\gamma(0, \cdot) \rho_0 \, dx \right| &\leq \|\nabla_x \varphi\|_{L^\infty} \int_0^{t_m} \int_\Omega |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt \\ &\quad + \|\nabla_x \varphi\|_{L^\infty} \int_0^{t_m} \int_\Omega |v - v_\gamma| \cdot |\rho| \, dx \, dt. \end{aligned} \quad (4.2-33)$$

D'après le Lemme 4.2-7 (p. 150), on a

$$\int_0^{t_m} \int_\Omega |v - v_\gamma| \cdot |\nabla_x v_\gamma| \cdot |\rho| \, dx \, dt \xrightarrow{\gamma \rightarrow 0} 0. \quad (4.2-34)$$

De plus, la Proposition 3.3-30 (p. 112) nous apprend que $|v - v_\gamma|$ est uniformément borné et converge simplement vers 0 sur $]0, 1[\times \Omega$ lorsque γ tend vers 0. Ainsi, du fait que $\rho \in L^2((0, 1) \times \Omega)$, on conclut par convergence dominée que :

$$\int_0^{t_m} \int_\Omega |v - v_\gamma| \cdot |\rho| \, dx \, dt \xrightarrow{\gamma \rightarrow 0} 0. \quad (4.2-35)$$

Pour finir, d'après le Lemme 3.3-29 (p. 111), $\nabla^\gamma \phi(x)$ converge pour presque tout $x \in \Omega$ vers $\nabla \phi(x)$ (pour tout x où ϕ est différentiable). De plus $\varphi(s, s\nabla^\gamma \phi(x) + (1-s)x)$ est uniformément borné par $\|\varphi\|_{L^\infty}$ pour tout $(s, x) \in]0, 1[\times \Omega$. Ainsi par convergence dominée, on a

$$\begin{aligned} \int_\Omega \tilde{h}_\gamma(0, \cdot) \rho_0 \, dx &= - \int_0^1 \int_\Omega \varphi(s, s\nabla^\gamma \phi(x) + (1-s)x) \rho_0(x) \, dx \, ds \\ &\xrightarrow{\gamma \rightarrow 0} - \int_0^1 \int_\Omega \varphi(s, s\nabla \phi(x) + (1-s)x) \rho_0(x) \, dx \, ds. \end{aligned} \quad (4.2-36)$$

Ainsi, par les résultats de convergence (4.2-34), (4.2-35) et (4.2-36), on obtient par passage à la limite dans la relation 4.2-33 :

$$\int_0^1 \int_\Omega \varphi \rho \, dx \, dt = \int_0^1 \int_\Omega \varphi(t, t\nabla \phi(x) + (1-t)x) \rho_0(x) \, dx \, dt.$$

et donc, cette dernière équation étant finalement vrai pour tout $\varphi \in C^\infty((0, 1) \times \mathbb{R}^d)$, on conclut donc en prolongeant de nouveau à \mathbb{R}^d :

$$\forall \varphi \in C^\infty((0, 1) \times \mathbb{R}^d), \int_0^1 \int_{\mathbb{R}^d} \varphi \rho \, dx \, dt = \int_0^1 \int_{\mathbb{R}^d} \varphi(t, t\nabla \phi(x) + (1-t)x) \rho_0(x) \, dx \, dt.$$

$C_c^\infty((0, 1) \times \mathbb{R}^d)$ étant dense dans $C_c^0([0, 1] \times \mathbb{R}^d)$ pour la norme $\sup_{[0,1] \times \mathbb{R}^d} |\cdot|$, on peut donc conclure pour la relation (4.2-6).

Attention cependant : dans le second membre, le passage de C_c^∞ à C_c^0 se fait par convergence dominée et non par convergence uniforme : en effet $\nabla\phi$ n'est défini que presque partout sur \mathbb{R}^d . Pour finir, on peut conclure à l'aide du Lemme 4.1-5 (p. 132), que pour presque tout $(t, x) \in [0, 1[\times \mathbb{R}^d$,

$$\rho(t, x) = \rho_t(x) = ((t\nabla\phi + (1-t)\text{id})\# \rho_0)(x),$$

$(t, x) \mapsto \rho_t$ étant définie dans la Proposition F.0-6 (p. 236). Le Lemme 4.1-5 (p. 132) nous affirme également que $t \mapsto \rho(t, \cdot) \in C^0([0, 1[, L^2(\mathbb{R}^d))$. \square

IV.2.6 Retour à la preuve du Théorème 4.2-1

Maintenant que nous avons prouvé la Proposition 4.2-5 (p. 146), il nous reste plus, pour démontrer le Théorème 4.2-1 (p. 140) relatif à l'unicité de la composante $\mu = (\rho, m)$ commune des points selles de \mathbf{L} , qu'à à montrer qu'une densité ρ associée à l'un de ces points selle vérifie bien les conditions d'application de la Proposition 4.2-5, en particulier qu'elle peut se prolonger dans \mathbb{R}^d en une densité à support compact tout en vérifiant la même loi de conservation de la masse, ce qui est l'objet du Lemme suivant :

Lemme 4.2-8.

Soient ρ_0, ρ_1 et Ω définis comme dans la notation 2 (p. 74), et soit v un champ de vitesse définie sur \mathbb{R}^d tel que $v \in L^\infty((0, 1) \times K)$ pour tout compact K de \mathbb{R}^d (on rappelle que par abus de notation, nous désignerons indifféremment v et $v|_\Omega$ par v).

Soit $\rho \in L^2((0, 1) \times \Omega)$ tel que :

$$\int_{]0,1[\times \Omega} (\partial_t h + v \cdot \nabla_x h) \rho \, dx \, dt + \int_\Omega h(0, \cdot) \rho_0 \, dx - \int_\Omega h(1, \cdot) \rho_1 \, dx = 0, \quad (4.2-37)$$

pour tout $h \in H^1((0, 1) \times \Omega)$.

On pose $\bar{\rho} \in L^2((0, 1) \times \mathbb{R}^d)$, définie pour tout $t \in]0, 1[$ et $x \in \mathbb{R}^d$ par :

$$\bar{\rho} = \begin{cases} \rho(t, x) & \text{si } x \in \Omega \\ 0 & \text{sinon} \end{cases}$$

Alors $\bar{\rho}$ vérifie la conservation de la masse dans \mathbb{R}^d , avec ρ_0 et ρ_1 en conditions initiales et finales, plus précisément :

$$\int_{]0,1[\times \mathbb{R}^d} (\partial_t h + v \cdot \nabla_x h) \bar{\rho} \, dx \, dt + \int_{\mathbb{R}^d} h(0, \cdot) \rho_0 \, dx - \int_{\mathbb{R}^d} h(1, \cdot) \rho_1 \, dx = 0, \quad (4.2-38)$$

pour tout $h \in H_{loc}^1([0, 1] \times \mathbb{R}^d)$.

Remarque 4.2-8. En fait la condition de nullité du flux sortant de ρv au bord est implicitement incluse dans cette formulation faible de la conservation de la masse (qui est celle des points selles de l'algorithme de Benamou-Brenier), où seules les conditions initiales et finales apparaissent :

il n'y a pas de terme donnant les conditions aux bords, qui sont donc nulles.

On peut donc prolonger ρ par 0 sur \mathbb{R}^d , et ainsi obtenir une densité L^2 à support compact dans \mathbb{R}^d vérifiant toujours les conditions de conservation de la masse pour les mêmes conditions initiales et finales.

Preuve :

Pour tout $h \in H_{loc}^1((0,1) \times \mathbb{R}^d)$, on a $h|_{\bar{\Omega}} \in H^1((0,1) \times \Omega)$ et $\nabla_{t,x} h|_{\bar{\Omega}} = (\nabla_{t,x} h)|_{\bar{\Omega}}$ (en posant $\bar{\Omega} =]0,1[\times \Omega$).

Donc pour tout $h \in H_{loc}^1((0,1) \times \mathbb{R}^d)$, on a

$$\begin{aligned} & \int_{]0,1[\times \mathbb{R}^d} (\partial_t h + v \cdot \nabla_x h) \bar{\rho} dx dt + \int_{\mathbb{R}^d} h(0, \cdot) \rho_0 dx - \int_{\mathbb{R}^d} h(1, \cdot) \rho_1 dx \\ &= \int_{]0,1[\times \Omega} \left((\partial_t h)|_{\bar{\Omega}} + v \cdot (\nabla_x h)|_{\bar{\Omega}} \right) \rho dx dt + \int_{\Omega} h(0, \cdot)|_{\bar{\Omega}} \rho_0 dx - \int_{\Omega} h(1, \cdot)|_{\bar{\Omega}} \rho_1 dx \\ &= \int_{]0,1[\times \Omega} \left(\partial_t h|_{\bar{\Omega}} + v \cdot \nabla_x h|_{\bar{\Omega}} \right) \rho dx dt + \int_{\Omega} h|_{\bar{\Omega}}(0, \cdot) \rho_0 dx - \int_{\Omega} h|_{\bar{\Omega}}(1, \cdot) \rho_1 dx \\ &= 0. \end{aligned}$$

□

Il nous reste maintenant à conclure sur la preuve du Théorème 4.2-1 (p. 140), dont nous rappelons l'énoncé :

Théorème 4.2-1.

Si (ψ^*, q^*, μ^*) est un élément de $\mathbf{L}^{ps}(\rho_0, \rho_1, \Omega)$ (pour ρ_0, ρ_1 et Ω définis comme dans la notation 2 p. 74, définissant $\mathbf{L}^{ps}(\rho_0, \rho_1, \Omega)$), alors pour tout potentiel $\phi \in \Phi(\rho_0 \mathcal{L}^d, \rho_1 \mathcal{L}^d)$, on a $\mu^* = (\rho^*, m^*) = (\rho_\phi, \rho_\phi v_\phi)$, avec pour tout $t \in]0, 1[$:

$$v_\phi(t, x) = \frac{1}{t} \left(x - \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{x}{1-t} \right) \right) = \frac{x - \nabla_x(\phi_t)^*(x)}{t}$$

(où $(\phi_t)^*$ désigne la transformée de Legendre de $\phi_t = (1-t)|\cdot|^2 + t\phi$) et

$$\rho_\phi(t, \cdot) \mathcal{L}^d = X_\phi(t, \cdot) \# (\rho_0 \mathcal{L}^d) \in C([0, 1], L^2(\Omega)), \text{ avec } X_\phi(t, \cdot) = \nabla_x \phi_t = (1-t) \text{id} + t \nabla_x \phi.$$

Preuve du Théorème 4.2-1 : Soit (ψ^*, q^*, μ^*) est un élément de $\mathbf{L}^{ps}(\rho_0, \rho_1, \Omega)$ (i.e. un point selle). D'après la Proposition 2.2-1 (p. 56), on a

$$\forall h \in H^1((0,1) \times \Omega), \quad G(h) + \langle \mu^*, \nabla_{t,x} h \rangle = 0.$$

Soit $\phi \in \Phi(\rho_0 \mathcal{L}^d, \rho_1 \mathcal{L}^d)$ (vérifiant ainsi la propriété (Γ_1)). D'après le Lemme 4.2-3 (p. 142), en posant

$$v_\phi(t, x) = \frac{1}{t} \left(x - \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{x}{1-t} \right) \right)$$

pour tout $t \in]0, 1[$, on a $m^* = \rho^* v_\phi$. Autrement dit, $\forall h \in H^1((0,1) \times \Omega)$:

$$\int_{]0,1[\times \Omega} (\partial_t h + v_\phi \cdot \nabla_x h) \rho^* dx dt + \int_{\Omega} h(0, \cdot) \rho_0 dx - \int_{\Omega} h(1, \cdot) \rho_1 dx = 0.$$

On pose $\bar{\rho}^* \in L^2((0, 1) \times \mathbb{R}^d)$, défini comme dans le Lemme 4.2-8 (p. 156), pour tout $t \in]0, 1[$ et $x \in \mathbb{R}^d$ par :

$$\bar{\rho}^* = \begin{cases} \rho^*(t, x) & \text{si } x \in \Omega \\ 0 & \text{sinon} \end{cases}$$

D'après le Lemme 4.2-8, on a

$$\int_{]0, 1[\times \mathbb{R}^d} (\partial_t h + v_\phi \cdot \nabla_x h) \bar{\rho}^* dx dt + \int_{\mathbb{R}^d} h(0, \cdot) \rho_0 dx - \int_{\mathbb{R}^d} h(1, \cdot) \rho_1 dx = 0, \quad (4.2-39)$$

pour tout $h \in C_c^\infty([0, 1] \times \mathbb{R}^d)$. Ainsi, d'après la Proposition 4.2-5 (p. 146), on a donc bien

$$\rho^*(t, \cdot) = X_\phi(t, \cdot) \# \rho_0, \text{ avec } X_\phi(t, \cdot) = (1 - t) \text{id} + t \nabla_x \phi,$$

et d'après le Corollaire 4.1-4 (p. 131), on a bien $t \mapsto \rho^*(t, \cdot) \in C^0([0, 1], L^2(\Omega))$, ce qui conclut la preuve. \square

IV.3 Caractérisation d'un champ de vitesse de transport optimal

Dans cette dernière sous-section, nous allons tenter de caractériser moins formellement un champ de vitesse de transport optimal. Le résultat sera en gros le suivant : toute densité L^2 , à support borné et continûment portée par un champ de vitesse localement borné, dont **les trajectoires sont toutes des lignes droites qui jamais ne se croisent**, c'est à dire vérifiant les propriétés (II) correspond à un **transport optimal** et est **l'unique solution** pour un tel déplacement.

On rappelle qu'un champ de vitesse v vérifie les propriétés (II) si et seulement si :

- (II)₁ Il existe $\psi \in W_{loc}^{1, \infty}([0, 1] \times \Omega)$ (i.e. $\psi \in W^{1, \infty}((0, 1) \times \omega)$, pour tout ouvert $\omega \subset \Omega$ borné) tel que $v = \nabla_x \psi$: propriété traduisant le **non croisement des trajectoires** (le rotationnel spatial de v est nul).
- (II)₂ Le champ de vitesse v vérifie au sens des distributions **l'équation de Burgers**, c'est à dire

$$\partial_t v + \frac{1}{2} \nabla_x |v|^2 = 0. \quad (4.3-40)$$

Cette propriété traduit le fait que les déplacements induits par v s'effectuent **exclusivement en ligne droite**.

On rappelle également qu'un champ de vitesse v vérifie les propriétés (III) (énoncées p. 76) si et seulement si

- (III)₁ v est localement lipschitzienne sur $]0, 1[\times \Omega$,
- (III)₂ $\nabla_{t, x} v \in L^\infty((0, 1), L_{loc}^1(\Omega))$,
- (III)₃ $\nabla_{t, x} v \in L^p((0, 1), L_{loc}^q(\Omega))$ avec $1/p + 1/q > 1$
(en particulier $v \in W_{loc}^{1, p}([0, 1] \times \Omega)$ avec $1 \leq p < 2$).

Pour un ouvert convexe Ω de \mathbb{R}^d (pas forcément borné), définissons l'espace ${}^b L_+^2((0, 1) \times \Omega)$ des densités $\rho \in L^2((0, 1) \times \Omega)$ positives et à supports compacts dans $[0, 1] \times \Omega$. Autrement dit :

$${}^b L_+^2((0, 1) \times \Omega) = \{ \rho \in L^2((0, 1) \times \Omega), \rho \geq 0, \text{ supp}(\rho) \subset\subset \Omega \}. \quad (4.3-41)$$

Le résultat que nous souhaitons démontrer est le suivant :

Théorème 4.3-9.

Soit Ω un ouvert convexe de \mathbb{R}^d . Soit v^* un champ de vitesse sur Ω vérifiant l'ensemble de propriétés (II) p. 75 (rappelées ci-dessus), et soit $\rho_0 \in L^2(\Omega)$, tel que $\rho_0 \geq 0$ et que $\text{supp}(\rho_0)$ soit borné dans Ω .

Soit une densité $\rho^* \in {}^bL^2_+((0, 1) \times \Omega)$ (définie en (4.3-41)) solution, au sens des distributions, du système

$$\begin{cases} \partial_t \rho + \text{div}_x(\rho v^*) = 0, \\ \rho(0, \cdot) = \rho_0, \end{cases} \quad (4.3-42)$$

c'est à dire

$$\forall h \in C_c^\infty([0, 1] \times \Omega), \quad \int_0^1 \int_\Omega (\partial_t h + v^* \cdot \nabla_x h) \rho^* dx dt + \int_\Omega h(0, \cdot) \rho_0 dx = 0.$$

Alors :

- La densité ρ^* est l'**unique solution** du système (4.3-42) dans l'espace ${}^bL^2_+((0, 1) \times \Omega)$.
- On a $\rho^* \in C^0([0, 1[, L^2(\Omega)) \cap {}^bL^2_+((0, 1) \times \Omega)$
- ρ^* suit un déplacement continu de type transport optimal : autrement dit, il existe une unique mesure positive μ_1 définie sur la tribu de Lebesgue de Ω , à support dans $\text{supp}(\rho_0) \cup \left(\bigcup_{t \in [0, 1]} \text{supp}(\rho^*(t, \cdot)) \right)$, vérifiant les propriétés (IV) (p. 203), et une fonction ϕ convexe sur \mathbb{R}^d vérifiant la propriété (Γ_1) , telles que $\mu_1 = \nabla \phi \# (\rho_0 \mathcal{L}^d)$, et

$$\forall t \in [0, 1[, \quad (\rho^*(t, \cdot) \mathcal{L}^d) = (\rho_\phi(t, \cdot) \mathcal{L}^d) = (t \nabla \phi + (1-t) \text{id}) \# (\rho_0 \mathcal{L}^d),$$

c'est à dire telles que les $\rho^*(t, \cdot)$ soient les interpolées de McCann entre la mesure $\rho_0 \mathcal{L}^d$ et la mesure μ_1 .

- Le couple (ρ^*, v^*) vérifie alors le système :

$$\begin{cases} \partial_t \rho^* + \text{div}_x(\rho^* v^*) = 0, \\ \rho^*(0, \cdot) = \rho_0, \quad \rho^*(1, \cdot) = \mu_1. \end{cases}$$

c'est à dire, reformulé au sens faible,

$$\forall h \in C_c^\infty([0, 1] \times \Omega), \quad \int_0^1 \int_\Omega (\partial_t h + v^* \cdot \nabla_x h) \rho^* dx dt + \int_\Omega h(0, \cdot) \rho_0 dx - \int_\Omega h(1, \cdot) d\mu_1 = 0.$$

- Pour finir, le champ de vitesse v^* vérifie les propriétés (III) p. 76 (rappelées juste avant l'énoncé de ce Théorème) sur $\text{supp}(\rho^*)$ et, pour tout $(t, x) \in \text{supp}(\rho^*)$, on a :

$$v^*(t, x) = v_\phi = \frac{1}{t} \left(\text{id} - \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{\cdot}{1-t} \right) \right) = \frac{\text{id} - \nabla_x(\phi_t)^*}{t}, \quad (4.3-43)$$

où $(\phi_t)^*$ désigne la transformée de Legendre de $\phi_t = (1-t)|\cdot|^2 + t\phi$.

Avant d'entrer dans le vif du sujet, commençons par démontrer le Lemme suivant qui sera utile par la suite : il s'agit seulement de poser rigoureusement la conservation de la masse totale lors d'un déplacement continu au sens des distributions.

Lemme 4.3-10.

Soit Ω un ouvert de \mathbb{R}^d . Soit $(\rho, m) \in L^1_{loc}([0, 1] \times \Omega)^{d+1}$, et soit μ une mesure de probabilité sur Ω , tel que le système

$$\begin{cases} \partial_t \rho + \operatorname{div}_x(m) = 0, \\ \rho(0, \cdot) = \mu. \end{cases}$$

soit vérifié au sens des distributions, i.e.

$$\forall h \in C_c^\infty([0, 1] \times \Omega), \quad \int_0^1 \int_\Omega (\partial_t h \cdot \rho + \nabla_x h \cdot m) dx dt + \int_\Omega h(0, \cdot) d\mu = 0.$$

On suppose de plus qu'il existe un ouvert $\omega \subset\subset \Omega$ (i.e. dont l'adhérence est un compact de Ω) tel que $\operatorname{supp}(\mu) \subset \omega$ et $\operatorname{supp}((\rho, m)) \subset [0, 1] \times \omega$. Alors $\rho \in L^\infty((0, 1), L^1(\Omega))$ avec $\forall t \in]0, 1[, \int_\Omega \rho(t, \cdot) \equiv 1$, ou, écrit plus rigoureusement,

$$\forall f \in C^\infty([0, 1]), \quad \int_0^1 \int_\Omega f(t) \rho(t, x) dx dt = \int_0^1 f(t) dt.$$

Preuve : Soit $\alpha \in C_c^\infty(0, 1)$ telle que $\int_0^1 \alpha(t) dt = 1$, et soit une fonction plateau $\theta \in C_c^\infty(\Omega)$ telle que $\theta = 1$ sur $\bar{\omega}$ (rappelons que $\bar{\omega}$ est un compact de Ω). Pour tout $f \in C_c^\infty(0, 1)$, on définit $\Lambda(f) \in C_c^\infty([0, 1])$ par

$$\forall t \in [0, 1], \quad \Lambda(f)(t) = \int_0^t \left(f(s) - \alpha(s) \int_0^1 f(u) du \right) ds. \quad (4.3-44)$$

Soit $f \in C_c^\infty(0, 1)$. Il est clair que $\Lambda(f)$ est uniformément nulle au voisinage de 0 et 1, et donc que $\Lambda(f) \in C_c^\infty(0, 1)$, d'où $\theta \Lambda(f) \in C_c^\infty((0, 1) \times \Omega)$. On a alors

$$\forall t \in [0, 1], \quad f(t) = \alpha(t) \int_0^1 f(u) du + \Lambda(f)'(t).$$

D'où, en posant $C = \int_0^1 \int_\Omega \alpha(t) \rho(t, x) dx dt$,

$$\begin{aligned} \int_0^1 \int_\Omega f(t) \rho(t, x) dx dt &= \int_0^1 \int_\Omega \Lambda(f)'(t) \rho(t, x) dx dt + \left(\int_0^1 f(u) du \right) \int_0^1 \int_\Omega \alpha(t) \rho(t, x) dx dt \\ &= \int_0^1 \int_\Omega \theta(x) \Lambda(f)'(t) \rho(t, x) dx dt + C \int_0^1 f(u) du \\ &= - \int_0^1 \int_\Omega \nabla_x(\theta \Lambda(f)) \cdot m dx dt - \int_\Omega (\theta \Lambda(f))(0, \cdot) d\mu + C \int_0^1 f(u) du \\ &= - \int_0^1 \int_\Omega \Lambda(f) \nabla_x \theta \cdot m dx dt - \int_\Omega \Lambda(f)(0) \theta(x) d\mu + C \int_0^1 f(u) du \\ &= C \int_0^1 f(u) du. \end{aligned}$$

Pour tout $n \in \mathbb{N}$, on définit une fonction plateau $\chi_n \in C_c^\infty(0, 1)$ tel que $0 \leq \chi_n \leq 1$ et $\chi_n = 1$ sur $[1/n, 1 - (1/n)]$.

Soit $f \in C^\infty([0, 1])$. Pour tout $n \in \mathbb{N}$, $f\chi_n \in C_c^\infty(0, 1)$, la suite $(\chi_n)_n$ converge simplement vers 1. On a pour tout $n \in \mathbb{N}$,

$$\int_0^1 \int_\Omega \chi_n(t) f(t) \rho(t, x) dx dt = C \int_0^1 \chi_n(t) f(t) dt,$$

donc par convergence dominée lorsque n tend vers l'infini ($\rho \in L_{loc}^1([0, 1] \times \Omega)$ et f est bornée), on a

$$\forall f \in C^\infty([0, 1]), \int_0^1 \int_\Omega f(t) \rho(t, x) dx dt = C \int_0^1 f(t) dt.$$

Il reste donc à prouver que $C = 1$. Soit $f \in C^\infty([0, 1])$ telle que $f(0) = 1$ et $f(1) = 0$. On a alors :

$$\begin{aligned} C &= -C \int_0^1 f'(t) dt = - \int_0^1 \int_\Omega f'(t) \rho(t, x) dx dt = - \int_0^1 \int_\Omega \theta(x) f'(t) \rho(t, x) dx dt \\ &= - \int_0^1 \int_\Omega \partial_t(\theta f) \rho dx dt = \int_0^1 \int_\Omega \nabla_x(\theta f) \cdot m dx dt + \int_\Omega (\theta f)(0, \cdot) d\mu \\ &= \int_0^1 \int_\Omega f \nabla_x \theta \cdot m dx dt + \int_\Omega f(0) \theta(x) d\mu = \int_\Omega f(0) \theta(x) d\mu = \mu(\omega) = \mu(\Omega) = 1. \end{aligned}$$

□

Le premier point de notre preuve consiste à prouver l'existence au sens des distributions d'une mesure d'arrivée, pour toute densité vérifiant une équation de conservation de la masse, mais où seule la densité de départ est fixée. Ce sera l'objet du Lemme 4.3-12 (p. 162), dans la démonstration duquel le Lemme 4.3-11 (p. 161) intervient.

Lemme 4.3-11.

Soient $(E, \|\cdot\|_E)$ un espace vectoriel normé et D un sous-espace vectoriel dense de E . Soient $M > 0$ et (T_n) une suite de formes linéaires réelles continues sur E telles que :

- pour tout $n \in \mathbb{N}$, on ait $\|T_n\| \leq M$,
- Pour tout $x \in D$ il existe $Tx \in \mathbb{R}$ tel que $T_n x$ converge vers Tx .

Alors il existe une forme linéaire \bar{T} réelle continue sur E , tel que $T_n x$ converge vers $\bar{T}x$ pour tout $x \in E$ (avec de plus $\|\bar{T}\| \leq M$), et tel pour tout $x \in D$, $\bar{T}x = Tx$.

Preuve : Soit $x \in E$. Montrons que la suite $(T_n x)_n$ est une suite de Cauchy de \mathbb{R} . Soit $\epsilon > 0$ et soit $y \in D$ tel que $\|x - y\|_E \leq (\epsilon/3M)$. Pour tous $p, q \in \mathbb{N}$,

$$\begin{aligned} |T_p x - T_q x| &\leq |T_p x - T_p y| + |T_p y - T_q y| + |T_q x - T_q y| \\ &\leq 2M \|x - y\|_E + |T_p y - T_q y| \leq \frac{2}{3} \epsilon + |T_p y - T_q y|. \end{aligned}$$

$(T_n y)$ converge vers Ty et donc $(T_n y)$ est une suite de Cauchy : il existe donc $N \in \mathbb{N}$ tel que pour tout $p, q \geq N$, on ait $|T_p y - T_q y| \leq (1/3)\epsilon$. Donc, pour tout $p, q \geq N$, on a

$$|T_p x - T_q x| \leq \frac{2}{3} \epsilon + |T_p y - T_q y| \leq \epsilon.$$

$(T_n x)$ est donc une suite de Cauchy et donc il existe $\bar{T}x \in \mathbb{R}$ tel que $(T_n x)$ converge vers $\bar{T}x$.

Il est évident que pour tout $y \in D$ on ait $\bar{T}y = Ty$, et que $x \mapsto \bar{T}x$ soit linéaire du fait de la linéarité des T_n . De plus, pour tout $x \in E$, on a pour tout $n \in \mathbb{N}$,

$$|\bar{T}x| \leq |\bar{T}x - T_n x| + |T_n x| \leq |\bar{T}x - T_n x| + M\|x\|_E.$$

Donc, par passage à la limite lorsque $n \rightarrow \infty$, on a $|\bar{T}x| \leq M\|x\|_E$. La forme linéaire \bar{T} est donc continue et $\|\bar{T}\| \leq M$. \square

Lemme 4.3-12.

Soit Ω un ouvert \mathbb{R}^d , et soient $(\rho, m) \in L^1((0, 1), \Omega)^{d+1}$ avec $\rho \geq 0$, et μ_0 une mesure de probabilité définie sur la tribu de Lebesgue de Ω , tels que (ρ, m) vérifie au sens des distributions

$$\begin{cases} \partial_t \rho + \operatorname{div}_x(m) = 0, \\ \rho(0, \cdot) = \mu_0, \end{cases}$$

i.e.

$$\forall h \in C_c^\infty([0, 1] \times \Omega), \int_0^1 \int_\Omega (\partial_t h \cdot \rho + \nabla_x h \cdot m) dx dt + \int_\Omega h(0, \cdot) d\mu_0 = 0. \quad (4.3-45)$$

Alors il existe une mesure positive μ_1 définie sur la tribu de Lebesgue de Ω , vérifiant les propriétés (IV) p. 203, tel que

$$\forall h \in C_c^\infty([0, 1] \times \Omega), \int_0^1 \int_\Omega (\partial_t h \cdot \rho + \nabla_x h \cdot m) dx dt + \int_\Omega h(0, \cdot) d\mu_0 - \int_\Omega h(1, \cdot) d\mu_1 = 0. \quad (4.3-46)$$

De plus, s'il existe un ouvert $\omega \subset\subset \Omega$ tel que $\operatorname{supp}(\mu_0) \subset \omega$ et $\operatorname{supp}(\rho, m) \subset [0, 1] \times \omega$, alors μ_1 est une mesure de probabilité sur Ω à support compact tel que $\operatorname{supp}(\mu_1) \subset \omega$.

Preuve : On définit la forme linéaire Λ sur $C_c^\infty([0, 1] \times \Omega)$ par

$$\forall h \in C_c^\infty([0, 1] \times \Omega), \Lambda(h) = \int_0^1 \int_\Omega (\partial_t h \cdot \rho + \nabla_x h \cdot m) dx dt + \int_\Omega h(0, \cdot) d\mu_0. \quad (4.3-47)$$

et la forme linéaire T , définie sur $C_c^\infty(\Omega)$, par

$$\forall f \in C_c^\infty(\Omega), T(f) = \Lambda(\bar{f}), \text{ avec } \bar{f} : (t, x) \mapsto f(x), \quad (4.3-48)$$

d'où

$$T(f) = \int_0^1 \int_\Omega \nabla_x f \cdot m dx dt + \int_\Omega f d\mu_0.$$

Pour tout $n \in \mathbb{N}$, on choisit une fonction $\chi_n \in C^\infty([0, 1])$, vérifiant :

- $\chi_n(t) = 0$ si $0 \leq t \leq 1 - \frac{2}{n}$, et $\chi_n(t) = 1$ si $1 - \frac{1}{n} \leq t \leq 1$,
- $\forall t \in [0, 1]$, $0 \leq \chi_n'(t) \leq \frac{2}{n}$ (et donc $0 \leq \chi_n \leq 1$).

Donc pour tout $f \in C_c^\infty(\Omega)$ et tout $n \in \mathbb{N}$, on a alors $(1 - \chi_n)\bar{f} \in C_c^\infty([0, 1[\times\Omega)$, d'où $\Lambda((1 - \chi_n)\bar{f}) = 0$ (du fait de l'hypothèse (4.3-45)), et ainsi :

$$\begin{aligned} T(f) &= \Lambda(\bar{f}) = \Lambda(\chi_n\bar{f}) + \Lambda((1 - \chi_n)\bar{f}) = \Lambda(\chi_n\bar{f}) \\ &= \int_0^1 \int_\Omega (\partial_t(\chi_n\bar{f}) \cdot \rho + \nabla_x(\chi_n\bar{f}) \cdot m) dx dt + \int_\Omega \chi_n(0)\bar{f}(0, \cdot) d\mu_0 \\ &= \int_0^1 \int_\Omega \chi_n' f \rho dx dt + \int_0^1 \int_\Omega (\nabla_x f \cdot m) \chi_n dx dt. \end{aligned}$$

Or χ_n tend simplement vers 0 sur $[0, 1[$ lorsque n tend vers l'infini, et $|(\nabla_x f \cdot m) \chi_n|$ est borné par $\|\nabla_x f\|_\infty \cdot |m|$ pour tout $n \in \mathbb{N}$ (et $m \in L^1((0, 1) \times \Omega)^d$). Donc par convergence dominée, on a :

$$\int_0^1 \int_\Omega (\nabla_x f \cdot m) \chi_n dx dt \xrightarrow{n \rightarrow +\infty} 0 \quad \text{et} \quad \int_0^1 \int_\Omega \chi_n' f \rho dx dt \xrightarrow{n \rightarrow +\infty} T(f) \quad (4.3-49)$$

Pour tout $n \in \mathbb{N}$ et tout $f \in C_c^0(\Omega)$, on pose

$$T_n(f) = \int_0^1 \int_\Omega \chi_n' f \rho dx dt.$$

T_n est donc une forme linéaire sur $C_c^0(\Omega)$, et en particulier,

$$\forall n \in \mathbb{N}, \forall f \in C_c^0(\Omega), \quad |T_n(f)| \leq \int_0^1 \int_\Omega \chi_n' \rho dx dt \|f\|_\infty \leq \|f\|_\infty. \quad (4.3-50)$$

En effet, d'après le Lemme 4.3-10 (p. 160), on a

$$\int_0^1 \int_\Omega \chi_n' \rho dx dt = \int_0^1 \chi_n'(t) dt = \chi_n(1) - \chi_n(0) = 1.$$

En appliquant le Lemme 4.3-11 (p. 161) avec $E = C_c^0(\Omega)$ et $D = C_c^\infty(\Omega)$ munies de la norme $\|\cdot\|_\infty$ ($C_c^\infty(\Omega)$ est dense dans $C_c^0(\Omega)$ pour la norme $\|\cdot\|_\infty$), on obtient l'existence d'une forme linéaire \bar{T} continue sur $C_c^0(\Omega)$, tel que $(T_n(f))_n$ converge vers $\bar{T}(f)$ pour tout $f \in C_c^0(\Omega)$.

De plus, pour tout $f \in C_c^0(\Omega)$ avec $f \geq 0$, il est clair que $\bar{T}(f) \geq 0$ (car $T_n(f) \geq 0$ pour tout $n \in \mathbb{N}$). Ainsi, d'après le Théorème A.2-1 (p. 204), il existe une mesure positive μ_1 , définie sur les ensembles Lebesgue-mesurables de Ω , et vérifiant les propriétés (IV) p. 203, telle que :

$$\forall f \in C_c^0(\Omega), \quad \bar{T}(f) = \int_\Omega f d\mu_1. \quad (4.3-51)$$

En résumé :

$$\forall f \in C_c^0(\Omega), \quad \int_\Omega f d\mu_1 = \lim_{n \rightarrow +\infty} \int_0^1 \int_\Omega \chi_n'(t) f(x) \rho(t, x) dx dt, \quad (4.3-52)$$

$$\text{et } \forall f \in C_c^\infty(\Omega), \quad \int_\Omega f d\mu_1 = T(f) = \int_0^1 \int_\Omega \nabla_x f(x) \cdot m(t, x) dx dt + \int_\Omega f d\mu_0. \quad (4.3-53)$$

Il nous reste maintenant à montrer que pour tout $h \in C_c^\infty([0, 1] \times \Omega)$, on a bien

$$\int_\Omega h(1, \cdot) d\mu_1 = \Lambda(h) = \int_0^1 \int_\Omega (\partial_t h \cdot \rho + \nabla_x h \cdot m) dx dt + \int_\Omega h(0, \cdot) d\mu_0.$$

Soit $h \in C_c^\infty([0, 1] \times \Omega)$. Remarquons tout d'abord que $(1 - \chi_n)h \in C_c^\infty([0, 1[\times\Omega)$, d'où

$\Lambda((1 - \chi_n)h) = 0$ (toujours du fait de l'hypothèse (4.3-45)), et donc :

$$\begin{aligned} \Lambda(h) &= \Lambda(h) = \Lambda(\chi_n h) + \Lambda((1 - \chi_n)h) = \Lambda(\chi_n h) \\ &= \int_0^1 \int_{\Omega} (\partial_t(\chi_n h) \cdot \rho + \nabla_x(\chi_n h) \cdot m) dx dt + \int_{\Omega} \chi_n(0)h(0, \cdot) d\mu_0 \\ &= \int_0^1 \int_{\Omega} (\chi_n \partial_t h + \chi'_n h) \rho dx dt + \int_0^1 \int_{\Omega} (\nabla_x h \cdot m) \chi_n dx dt \\ &= \int_0^1 \int_{\Omega} \chi'_n h \rho dx dt + \int_0^1 \int_{\Omega} (\partial_t h \cdot \rho + \nabla_x h \cdot m) \chi_n dx dt. \end{aligned}$$

De même que précédemment, du fait que χ_n tende simplement vers 0 sur $[0, 1[$ lorsque n tend vers l'infini, et du fait que

$$|\partial_t h \cdot \rho + \nabla_x h \cdot m| \leq \|\nabla_{t,x} h\|_{\infty} \cdot (|\rho| + |m|)$$

(avec $(\rho, m) \in L^1((0, 1) \times \Omega)^{d+1}$), on a par convergence dominée :

$$\int_0^1 \int_{\Omega} (\partial_t h \cdot \rho + \nabla_x h \cdot m) \chi_n dx dt \xrightarrow{n \rightarrow +\infty} 0 \quad \text{et} \quad \int_0^1 \int_{\Omega} \chi'_n h \rho dx dt \xrightarrow{n \rightarrow +\infty} \Lambda(h). \quad (4.3-54)$$

D'où, en se référant à la relation de convergence (4.3-52), on a

$$\begin{aligned} \Lambda(h) - \int_{\Omega} h(1, \cdot) d\mu_1 &= \lim_{n \rightarrow +\infty} \int_0^1 \int_{\Omega} \chi'_n(t) (h(t, x) - h(1, x)) \rho(t, x) dx dt \\ &= \lim_{n \rightarrow +\infty} \int_{1-\frac{2}{n}}^{1-\frac{1}{n}} \int_t^1 \int_{\Omega} \chi'_n(t) \partial_t h(s, x) \rho(t, x) dx ds dt \end{aligned} \quad (4.3-55)$$

En effet χ'_n est nulle en dehors de $[1 - (2/n), 1 - (1/n)]$. Donc

$$\begin{aligned} \left| \int_{1-\frac{2}{n}}^{1-\frac{1}{n}} \int_t^1 \int_{\Omega} \chi'_n(t) \partial_t h(s, x) \rho(t, x) dx ds dt \right| &\leq \int_{1-\frac{2}{n}}^{1-\frac{1}{n}} \int_{1-\frac{2}{n}}^1 \int_{\Omega} \chi'_n(t) |\partial_t h(s, x)| \rho(t, x) dx ds dt \\ &\leq \frac{2}{n} \|\partial_t h\|_{\infty} \int_{1-\frac{2}{n}}^{1-\frac{1}{n}} \int_{\Omega} \chi'_n(t) \rho(t, x) dx dt = \frac{2}{n} \|\partial_t h\|_{\infty}. \end{aligned}$$

En effet d'après le Lemme 4.3-10 (p. 160), on a

$$\int_{1-\frac{2}{n}}^{1-\frac{1}{n}} \int_{\Omega} \chi'_n(t) \rho(t, x) dx dt = \int_0^1 \int_{\Omega} \chi'_n(t) \rho(t, x) dx dt = \int_0^1 \chi'_n(t) dt = 1.$$

On peut donc conclure que :

$$\Lambda(h) - \int_{\Omega} h(1, \cdot) d\mu_1 = \lim_{n \rightarrow +\infty} \int_{1-\frac{2}{n}}^{1-\frac{1}{n}} \int_t^1 \int_{\Omega} \chi'_n(t) \partial_t h(s, x) \rho(t, x) dx ds dt = 0, \quad (4.3-56)$$

et ceci pour tout $h \in C_c^{\infty}([0, 1] \times \Omega)$.

Pour finir, s'il existe un ouvert $\omega \subset\subset \Omega$ tel que $\text{supp}(\mu_0) \subset \omega$ et $\text{supp}((\rho, m)) \subset [0, 1] \times \omega$, alors pour tout $f \in C_c^{\infty}(\Omega)$ telle que $\text{supp}(f) \subset \Omega \setminus \omega$, alors en se référant à l'équation (4.3-53), on obtient

$$\int_{\Omega} f d\mu_1 = \int_0^1 \int_{\Omega} \nabla_x f(x) \cdot m(t, x) dx dt + \int_{\Omega} f d\mu_0 = 0,$$

d'où $\text{supp}(\mu_1) \subset \omega$. De plus, en prenant une fonction $f \in C_c^\infty(\Omega)$ telle que $f = 1$ sur $\bar{\omega}$ (voir le Théorème A.1-2 p. 201), on a

$$\begin{aligned} \mu_1(\Omega) &= \mu_1(\omega) = \int_{\Omega} f d\mu_1 = \int_0^1 \int_{\Omega} \nabla_x f(x) \cdot m(t, x) dx dt + \int_{\Omega} f d\mu_0 \\ &= \int_{\Omega} f d\mu_0 = \mu_0(\omega) = \mu_0(\Omega) = 1. \end{aligned} \quad (4.3-57)$$

□

Nous souhaitons maintenant relier les solutions du système (4.3-42) p. 159 à la solution du transport optimal entre ρ_0 et la mesure μ_1 définie dans le Lemme 4.3-12 (p. 162). Pour cela il sera nécessaire d'appliquer aux équations de conservation de la masse des fonctions test moins régulières, à savoir éléments de $W_c^{1,\infty}([0, 1] \times \Omega)$.

Lemme 4.3-13.

Soit Ω un ouvert de \mathbb{R}^d . Soient μ_0 et μ_1 deux mesures de probabilité sur la tribu de Lebesgue de \mathbb{R}^d , et soient $\rho, m \in L^1([0, 1] \times \Omega, \mathcal{L}^{d+1})$, telles que, pour tout $h \in C_c^\infty([0, 1] \times \Omega)$, on ait

$$\int_0^1 \int_{\Omega} (\rho \partial_t h + m \cdot \nabla_x h) dx dt + \int_{\Omega} h(0, \cdot) d\mu_0 - \int_{\Omega} h(1, \cdot) d\mu_1 = 0, \quad (4.3-58)$$

alors (4.3-58) est également vérifiée pour tout $h \in W_c^{1,\infty}([0, 1] \times \Omega)$.

Preuve : Soit $h \in W_c^{1,\infty}([0, 1] \times \Omega)$. Soit $R > 0$ tel que $\text{supp}(h) \subset Q = [0, 1] \times (B(0, R) \cap \Omega)$. On pose $\bar{h} = P(h|_{\text{Int}(Q)})$, avec P l'opérateur de prolongement de $W^{1,\infty}(\text{Int}(Q))$ à $W^{1,\infty}(\mathbb{R}^{d+1})$ (avec $\text{Int}(Q) =]0, 1[\times (B(0, R) \cap \Omega)$), tel que défini par le Théorème A.1-6 (p. 202).

Soit $(r_n)_n$ une suite régularisante sur \mathbb{R}^{d+1} définie par

$$\forall n \in \mathbb{N}, \quad r_n \in C_c^\infty, \quad \text{supp}(r_n) \subset B\left(0, \frac{1}{n}\right), \quad \int_{\mathbb{R}^{d+1}} r_n = 1, \quad r_n \geq 0.$$

Pour tout $n \in \mathbb{N}$, on pose $h_n = (r_n * \bar{h})\chi$, avec $\chi \in C_c^\infty(\mathbb{R}^{d+1})$, $0 \leq \chi \leq 1$, et $\chi = 1$ sur \bar{Q} . On a donc d'après le Lemme A.1-1 (p. 201) et le Théorème A.1-6 (p. 202), on a pour tout $n \in \mathbb{N}$,

1. $\forall n \in \mathbb{N}$, $h_n \in C_c^\infty(\mathbb{R}^{d+1})$ et $(h_n)_n$ converge uniformément vers $\bar{h}\chi$. En effet, \bar{h} est continue sur \mathbb{R}^{d+1} (car $W^{1,\infty}(\mathbb{R}^d)$ s'injecte dans $C^0(\mathbb{R}^d)$, voir par exemple [17]) et, d'après le Lemme A.1-1 (p. 201), $(r_n * \bar{h})$ converge uniformément vers \bar{h} sur tout compact de \mathbb{R}^{d+1} . En particulier les suites $(h_n(0, \cdot))_n$ et $(h_n(1, \cdot))_n$ convergent uniformément respectivement vers $\bar{h}(0, \cdot) = h(0, \cdot)$ et $\bar{h}(1, \cdot) = h(1, \cdot)$ sur \mathbb{R}^d .
2. $\forall n \in \mathbb{N}$, $\|h_n\|_\infty \leq \|r_n * \bar{h}\|_\infty \leq \|r_n\|_1 \cdot \|\bar{h}\|_\infty \leq C\|h\|_\infty$ ($C > 0$ ne dépendant que de Q), et

$$\begin{aligned} \forall n \in \mathbb{N}, \quad \|\nabla_{t,x} h_n\|_\infty &= \|(r_n * \nabla_{t,x} \bar{h})\chi + (r_n * \bar{h})\nabla_{t,x} \chi\|_\infty \\ &\leq \|r_n * \nabla_{t,x} \bar{h}\|_\infty + \|r_n * \bar{h}\|_\infty \cdot \|\nabla_{t,x} \chi\|_\infty \\ &\leq \|r_n\|_1 \cdot \|\nabla_{t,x} \bar{h}\|_\infty + \|r_n\|_1 \|\bar{h}\|_\infty \cdot \|\nabla_{t,x} \chi\|_\infty \\ &\leq C\|\nabla_{t,x} h\|_\infty + C\|h\|_\infty \cdot \|\nabla_{t,x} \chi\|_\infty, \end{aligned}$$

i.e. $\|h_n\|_{W^{1,\infty}(\mathbb{R}^{d+1})} \leq C'\|h\|_{W^{1,\infty}(Q)}$ (avec $C' = (2 + \|\nabla_{t,x} \chi\|_\infty)C$).

Ici $\|\cdot\|_\infty$ désigne indifféremment $\|\cdot\|_{L^\infty(Q)}$, $\|\cdot\|_{L^\infty([0,1] \times \mathbb{R}^d)}$ ou $\|\cdot\|_{L^\infty(\mathbb{R}^{d+1})}$.

3. La suite $(\nabla_{t,x}h_n)_n$ converge simplement \mathcal{L}^{d+1} -presque partout vers $\nabla_{t,x}\bar{h}$ sur Q .

En effet d'après le Lemme A.1-1 (p. 201), la suite $(r_n * \nabla_{t,x}\bar{h})_n$, converge vers $\nabla_{t,x}\bar{h}$ dans $L^p(\mathbb{R}^{d+1})$ pour tout $p \in [1, +\infty[$. La suite $(\nabla_{t,x}h_n)_n$ converge donc vers $\chi\nabla_{t,x}\bar{h} + \bar{h}\nabla_{t,x}\chi$ dans $L^p(\mathbb{R}^{d+1})$.

On peut donc considérer, en se restreignant à \bar{Q} et quitte à devoir extraire une sous-suite de $(r_n)_n$, $(\nabla_{t,x}h_n)_n$ converge simplement \mathcal{L}^{d+1} -presque partout vers $\chi\nabla_{t,x}\bar{h} + \bar{h}\nabla_{t,x}\chi$ sur Q , i.e. vers $\nabla_{t,x}\bar{h}$ (rappelons que $\chi = 1$ sur Q et donc $\nabla_{t,x}\chi = 0$).

Ainsi, par convergence uniforme des suites $(h_n(0, \cdot))_n$ et $(h_n(1, \cdot))_n$ vers $h(0, \cdot)$ et $h(1, \cdot)$ on a

$$\int_{\mathbb{R}^d} h_n(0, \cdot) d\mu_0 \xrightarrow{n \rightarrow +\infty} \int_{\mathbb{R}^d} h(0, \cdot) d\mu_0, \quad \text{et} \quad \int_{\mathbb{R}^d} h_n(1, \cdot) d\mu_1 \xrightarrow{n \rightarrow +\infty} \int_{\mathbb{R}^d} h(1, \cdot) d\mu_1,$$

et par la convergence simple presque partout uniformément bornée de $(\nabla_{t,x}h_n)_n$ vers $\nabla_{t,x}\bar{h}$ sur Q , on a par convergence dominée,

$$\begin{aligned} \int_0^1 \int_{\mathbb{R}^d} (\rho \partial_t h_n + m \cdot \nabla_x h_n) dx dt &\xrightarrow{n \rightarrow +\infty} \int_0^1 \int_{\mathbb{R}^d} (\rho \partial_t \bar{h} + m \cdot \nabla_x \bar{h}) dx dt \\ &= \int_0^1 \int_{\mathbb{R}^d} (\rho \partial_t h + m \cdot \nabla_x h) dx dt. \end{aligned} \quad (4.3-59)$$

On peut donc conclure en injectant les h_n dans l'équation (4.3-58) et en passant à la limite. \square

Pour finir, il nous reste avant de passer à la preuve finale du Théorème 4.3-9 (p. 159), que les solutions ρ^* du système (4.3-42) sont en fait portées par le même champ de vitesse que celui défini en (3.3-11) p. 89 qui porte la solution continue du transport optimal de ρ_0 sur μ_1 , tel que cela a été fait dans la sous-section IV.2.2. Il s'agit en fait ici de généraliser la preuve du Lemme 4.2-3 (p. 142).

Lemme 4.3-14.

Soit Ω un ouvert convexe de \mathbb{R}^d . Soit v^* un champ de vitesse sur $[0, 1] \times \Omega$ vérifiant l'ensemble de propriétés (II) p. 75.

Soient ρ_0 une densité de probabilité Lebesgue-mesurable sur Ω , ν_1 une mesure de probabilité définie sur la tribu de Lebesgue de Ω et une fonction ϕ convexe de \mathbb{R}^d vérifiant la propriété (Γ_1) , telles que $\nu_1 = \nabla\phi\#(\rho_0\mathcal{L}^d)$. Soit également $\rho^* \in L^2((0, 1) \times \Omega)$ tel que $\rho^* \geq 0$ et que $\text{supp}(\rho^*)$ soit borné dans $[0, 1] \times \Omega$, tel que (ρ^*, v^*) vérifie au sens des distributions :

$$\begin{cases} \partial_t \rho^* + \text{div}_x(\rho^* v^*) = 0, \\ \rho^*(0, \cdot) = \rho_0, \quad \rho^*(1, \cdot) = \nu_1. \end{cases} \quad (4.3-60)$$

c'est à dire, reformulé au sens faible,

$$\forall h \in C_c^\infty([0, 1] \times \Omega), \quad \int_0^1 \int_{\Omega} (\partial_t h + v^* \cdot \nabla_x h) \rho^* dx dt + \int_{\Omega} h(0, \cdot) \rho_0 dx - \int_{\Omega} h(1, \cdot) d\nu_1 = 0. \quad (4.3-61)$$

Alors, en définissant v_ϕ comme en 3.3-11 (p. 89) pour tout $(t, x) \in \Omega$ par

$$v_\phi(t, x) = \frac{1}{t} \left(\text{id} - \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{\cdot}{1-t} \right) \right) = \frac{\text{id} - \nabla_x(\phi_t)^*}{t}, \quad (4.3-62)$$

(où $(\phi_t)^*$ désigne la transformée de Legendre de $\phi_t = (1-t)|\cdot|^2 + t\phi$),

on a $\rho^* v^* = \rho^* v_\phi$.

Preuve : D'après la Proposition 4.1-8 (p. 135), en définissant la densité ρ_ϕ comme dans la Proposition F.0-6 (rappelé en début de chapitre p. 128) et le Lemme 4.1-5 (p. 132), c'est à dire, rappelons-le, $\rho_\phi : (t, x) \mapsto \rho_t(t, x)$ (mesurable), où pour tout $t \in [0, 1]$, la densité ρ_t vérifie :

$$\rho_t \mathcal{L}^d = [(1-t)\text{id} + t\nabla\phi] \# (\rho_0 \mathcal{L}^d) = \nabla\phi_t \# (\rho_0 \mathcal{L}^d),$$

et en définissant le champ de vitesse v_ϕ par rapport à ϕ par (4.3-62), on a

$$\forall h \in C_c^\infty([0, 1] \times \Omega), \quad \int_0^1 \int_\Omega (\partial_t h + v_\phi \cdot \nabla_x h) \rho_\phi \, dx \, dt + \int_\Omega h(0, \cdot) \rho_0 \, dx - \int_\Omega h(1, \cdot) \, d\nu_1 = 0. \quad (4.3-63)$$

D'après le Lemme 4.3-13 (p. 165), les relations (4.3-61) et (4.3-63) sont encore bien définies et valables pour $h \in W_c^{1,\infty}([0, 1] \times \Omega)$.

Par construction v_ϕ vérifie les hypothèses (II) p. 75 (voir le Théorème 3.1-2 p. 76), de même que v^* par hypothèse. Donc, d'après le Lemme 4.1-10 (p. 137), il existe alors deux potentiels $\psi_\phi, \psi^* \in W_{loc}^{1,\infty}([0, 1] \times \Omega)$ tels que

$$q_\phi = \begin{pmatrix} -\frac{1}{2}|v_\phi|^2 \\ v_\phi \end{pmatrix} = \nabla_{t,x} \psi_\phi \quad \text{et} \quad q^* = \begin{pmatrix} -\frac{1}{2}|v^*|^2 \\ v^* \end{pmatrix} = \nabla_{t,x} \psi^*.$$

Remarquons que $q_\phi, q^* \in L^\infty([0, 1] \times \Omega)^{d+1}$. On pose également $\mu \in L^1([0, 1] \times \Omega)^{d+1}$ et $\mu^* \in L^2([0, 1] \times \Omega)^{d+1}$, définies par $\mu_\phi = (\rho_\phi, \rho_\phi v_\phi)$ et $\mu^* = (\rho^*, \rho^* v^*)$. On introduit la notation $\langle \cdot, \cdot \rangle$ comme suit : pour tous $f \in L^\infty([0, 1] \times \Omega)^d$ et $g \in L^1([0, 1] \times \Omega)^d$, on note :

$$\forall f \in L^\infty([0, 1] \times \Omega)^d, \quad \forall g \in L^1([0, 1] \times \Omega)^d, \quad \langle g, f \rangle = \int_0^1 \int_\Omega g \cdot f \, dx \, dt$$

et, comme dans le chapitre II, on définit l'ensemble parabolöide L^2

$$\mathcal{P} = \left\{ q = (a, b) \in L^2((0, 1) \times \Omega) \times L^2((0, 1) \times \Omega)^d, a + \frac{|b|^2}{2} \leq 0 \right\}.$$

Bien sûr $q_\phi, q^* \in \mathcal{P}$. Par construction, le vecteur μ^* est point par point presque partout orthogonal au parabolöide, définie comme l'ensemble des $(a, b) \in \mathbb{R}^{d+1}$ tels que $a + (1/2)|b|^2 \leq 0$, au point q^* . Ce qui se traduit par la propriété

$$\forall q' \in \mathcal{P}, \quad \langle \mu^*, q' - q^* \rangle \leq 0 \quad (4.3-64)$$

Il faudrait pour conclure notre preuve prouver que μ^* est également orthogonal au parabolöide au point q_ϕ , i.e. il faudrait établir la propriété

$$\forall q' \in \mathcal{P}, \quad \langle \mu^*, q' - q_\phi \rangle \leq 0 \quad (4.3-65)$$

Il serait alors possible d'invoquer la Proposition 2.2-3 (p. 58). Pour prouver (4.3-65) à partir de (4.3-64), il suffirait donc de prouver que $\langle \mu^*, q' \rangle \leq \langle \mu^*, q_\phi \rangle$. Le problème est que cette inégalité n'est pas évidente dans ce sens mais plutôt dans l'autre (il y a égalité entre $\langle \mu^*, q^* \rangle$ et $\langle \mu^*, q_\phi \rangle$, et on peut prouver facilement que $\langle \mu^*, q_\phi \rangle \leq \langle \mu^*, q^* \rangle$). Pour obtenir facilement cette inégalité (c'est à dire dans le sens qui nous intéresse), il suffit d'établir les égalités suivantes : $\langle \mu_\phi, q^* \rangle = \langle \mu^*, q^* \rangle$ et $\langle \mu_\phi, q_\phi \rangle = \langle \mu^*, q_\phi \rangle$.

En prenant $h = \psi^*$ comme fonction test dans (4.3-61) et (4.3-63), on obtient aisément l'égalité $\langle \mu_\phi, q^* \rangle = \langle \mu^*, q^* \rangle$. En effet :

$$\begin{aligned} \langle \mu_\phi, q^* \rangle &= \langle \mu_\phi, \nabla_{t,x} \psi^* \rangle \\ &= \int_0^1 \int_\Omega (\partial_t \psi^* + v_\phi \cdot \nabla_x \psi^*) \rho_\phi dx dt = \int_\Omega \psi^*(1, \cdot) d\nu_1 - \int_\Omega \psi^*(0, \cdot) \rho_0 dx \\ &= \int_0^1 \int_\Omega (\partial_t \psi^* + v^* \cdot \nabla_x \psi^*) \rho^* dx dt = \langle \mu_\phi, \nabla_{t,x} \psi^* \rangle = \langle \mu^*, q^* \rangle \end{aligned}$$

En prenant maintenant $h = \psi_\phi$ comme fonction test dans (4.3-61) et (4.3-63), on obtient l'égalité $\langle \mu_\phi, q_\phi \rangle = \langle \mu^*, q_\phi \rangle$. En effet :

$$\begin{aligned} \langle \mu_\phi, q_\phi \rangle &= \langle \mu_\phi, \nabla_{t,x} \psi_\phi \rangle \\ &= \int_0^1 \int_\Omega (\partial_t \psi_\phi + v_\phi \cdot \nabla_x \psi_\phi) \rho_\phi dx dt = \int_\Omega \psi_\phi(1, \cdot) d\nu_1 - \int_\Omega \psi_\phi(0, \cdot) \rho_0 dx \\ &= \int_0^1 \int_\Omega (\partial_t \psi_\phi + v^* \cdot \nabla_x \psi_\phi) \rho^* dx dt = \langle \mu^*, \nabla_{t,x} \psi_\phi \rangle = \langle \mu^*, q_\phi \rangle \end{aligned}$$

Il est maintenant facile d'obtenir l'inégalité $\langle \mu_\phi, q^* \rangle \leq \langle \mu_\phi, q_\phi \rangle$ comme il suit :

$$\begin{aligned} \langle \mu_\phi, q^* \rangle &\leq \int_0^1 \int_\Omega \left(-\frac{1}{2} |v^*|^2 + v_\phi \cdot v^* \right) \rho_\phi dx dt \\ &\leq \int_0^1 \int_\Omega \left(-\frac{1}{2} |v^*|^2 + \left[\frac{1}{2} |v_\phi|^2 + \frac{1}{2} |v^*|^2 \right] \right) \rho_\phi dx dt \\ &= \int_0^1 \int_\Omega \frac{1}{2} |v_\phi|^2 \rho_\phi dx dt = \int_0^1 \int_\Omega \left(-\frac{1}{2} |v_\phi|^2 + v_\phi \cdot v_\phi \right) \rho_\phi dx dt = \langle \mu_\phi, q_\phi \rangle \end{aligned}$$

En résumé, on a donc bien $\langle \mu^*, q^* \rangle = \langle \mu_\phi, q^* \rangle \leq \langle \mu_\phi, q_\phi \rangle = \langle \mu^*, q_\phi \rangle$, d'où

$$\forall q' \in \mathcal{P}, \quad \langle \mu^*, q' - q_\phi \rangle = \langle \mu^*, q' \rangle - \langle \mu^*, q_\phi \rangle \leq \langle \mu^*, q' \rangle - \langle \mu^*, q^* \rangle = \langle \mu^*, q' - q^* \rangle \leq 0. \quad (4.3-66)$$

D'après la Proposition 2.2-3 (p. 58), on a donc bien $\rho^* v^* = \rho^* v_\phi$. \square

Retour à la preuve du Théorème 4.3-9 (p. 159) : Supposons pour simplifier que ρ_0 est une densité de probabilité (quitte à devoir renormaliser).

1. Commençons par prouver l'existence de μ et de ϕ vérifiant les propriétés de l'énoncé. v^* vérifie les propriétés (II) p. 75, en particulier $v^* \in L_{loc}^\infty([0, 1] \times \Omega)$. ρ^* étant à support borné, le terme $(\rho^*, \rho^* v^*)$ est donc également à support borné, et $(\rho^*, \rho^* v^*) \in L^1([0, 1], \Omega)^{d+1}$. Soit ω un ouvert borné tel que $\text{supp}(\rho^*, \rho^* v^*) \subset \omega$ et $\text{supp}(\rho_0) \subset \omega$: d'après le Lemme 4.3-12 (p. 162), il existe une mesure de probabilité à support borné μ_1 définie sur la tribu de Lebesgue de Ω , vérifiant les propriétés (IV) p. 203, telle que $\text{supp}(\mu_1) \subset \omega$ et

$$\begin{cases} \partial_t \rho^* + \text{div}_x(\rho^* v^*) = 0, \\ \rho^*(0, \cdot) = \rho_0, \quad \rho^*(1, \cdot) = \mu_1. \end{cases}$$

(au sens des distributions) i.e.

$$\forall h \in C_c^\infty([0, 1] \times \Omega), \quad \int_0^1 \int_\Omega (\partial_t h + v^* \cdot \nabla_x h) \rho^* dx dt + \int_\Omega h(0, \cdot) \rho_0 dx - \int_\Omega h(1, \cdot) d\mu_1 = 0.$$

D'après le Théorème 1.1-7 (p. 20) et la Proposition 3.2-3 (p. 80), il existe une fonction ϕ convexe de \mathbb{R}^d vérifiant la propriété (Γ_1) , telle que $\mu_1 = \nabla\phi\#(\rho_0 \mathcal{L}^d)$. Ainsi, en posant pour tout $(t, x) \in [0, 1] \times \Omega$,

$$v_\phi(t, x) = \frac{1}{t} \left(x - \text{Prox}_{\frac{t}{1-t}\phi} \left(\frac{x}{1-t} \right) \right),$$

on obtient, par le Lemme 4.3-14 (p. 166), que

$$\rho^* v^* = \rho^* v_\phi$$

(la forme de v_ϕ définie en (4.3-43) sur $\text{supp}(\rho^*)$ est donc établie). Ainsi, d'après la Proposition 4.2-5 (p. 146), on a alors bien

$$t \mapsto \rho^*(t, \cdot) = \rho_\phi(t, \cdot) = (t\nabla\phi + (1-t)\text{id})\#\rho_0,$$

et $t \mapsto \rho^*(t, \cdot) \in C^0([0, 1[, L^2(\Omega))$.

Certaines des Propositions invoquées s'appliquent dans \mathbb{R}^d tout entier. Cependant, du fait que les densités ρ satisfaisant (4.3-42) sont supposées à support borné (en tant qu'éléments de ${}^bL_+^2((0, 1) \times \Omega)$), on peut alors les prolonger par 0 à \mathbb{R}^d tout entier sans changer les conditions du problème.

2. Montrons maintenant que ρ^* est l'unique solution à support borné du système (4.3-42) dans l'espace ${}^bL_+^2((0, 1) \times \Omega)$.

Supposons qu'il existe deux solutions ρ^1, ρ^2 du système (4.3-42) dans ${}^bL_+^2((0, 1) \times \Omega)$, et montrons qu'alors $\rho^1 = \rho^2$ dans $L^2((0, 1) \times \Omega)$. Introduisons $\bar{\rho} = (\rho^1 + \rho^2)/2$. Un tel $\bar{\rho}$ est également une solution du système (4.3-42) dans ${}^bL_+^2((0, 1) \times \Omega)$. D'après ce qui précède, il existerait donc une fonction $\bar{\phi}$ convexe de \mathbb{R}^d vérifiant la propriété (Γ_1) , telle que, en posant pour tout $(t, x) \in [0, 1] \times \Omega$,

$$\bar{v}(t, x) = \bar{v}_{\bar{\phi}}(t, x) = \frac{1}{t} \left(x - \text{Prox}_{\frac{t}{1-t}\bar{\phi}} \left(\frac{x}{1-t} \right) \right),$$

on ait $\bar{\rho} v^* = \bar{\rho} \bar{v}$, i.e. $(\rho^1 + \rho^2) v^* = (\rho^1 + \rho^2) \bar{v}$.

v^* est donc presque partout égale à \bar{v} sur $\text{supp}(\rho^1) \cup \text{supp}(\rho^2)$ (rappelons que $\rho^1, \rho^2 \geq 0$), d'où $\rho^1 v^* = \rho^1 \bar{v}$ et $\rho^2 v^* = \rho^2 \bar{v}$.

Si on veut être plus rigoureux, multiplions l'égalité $(\rho^1 + \rho^2) v^* = (\rho^1 + \rho^2) \bar{v}$ par les fonctions f_1 et f_2 définies par :

$$f_i(t, x) = \begin{cases} \frac{\rho^i(t, x)}{\rho^1(t, x) + \rho^2(t, x)} & \text{si } \rho^1(t, x) + \rho^2(t, x) > 0, \\ 0, & \text{sinon} \end{cases}$$

pour $i \in \{1, 2\}$, et pour tout $(t, x) \in [0, 1] \times \Omega$ (en choisissant des représentants de ρ^1 et ρ^2 ne prenant pas de valeurs infinies).

Ainsi ρ^1 et ρ^2 vérifient tous deux le système

$$\begin{cases} \partial_t \rho + \text{div}_x(\rho \bar{v}) = \partial_t \rho + \text{div}_x(\rho v^*) = 0, \\ \rho(0, \cdot) = \rho_0, \end{cases}$$

et donc, par la Proposition 4.2-5, on a bien $\rho^1 = \rho^2 : t \mapsto \rho_{\bar{\phi}} = (t\nabla\bar{\phi} + (1-t)\text{id})\#\rho_0$ dans $L^2((0, 1) \times \Omega)$.

□

IV.4 Conclusion du chapitre

Dans ce chapitre, nous avons dans un premier temps pu établir, dans la section IV.1, l'existence d'un point selle pour le lagrangien \mathbf{L} de [6] (Théorème 3.1-2 rappelé et démontré p. 138) de la forme $(\psi_\phi, q_\phi, \mu_\phi)$, avec :

- $\mu_\phi = (\rho_\phi, \rho_\phi v_\phi) \in L^2((0, 1) \times \Omega)^{d+1}$, où $\rho_\phi \in C^0([0, 1], L^2(\Omega))$ est une densité positive, vérifiant pour tout $[0, 1] \times \mathbb{R}^d$:

$$\rho_\phi(t, \cdot) \mathcal{L}^d = X_\phi(t, \cdot) \# (\rho_0 \mathcal{L}^d), \text{ avec } X_\phi(t, \cdot) = \nabla_x \phi_t = (1-t) \text{id} + t \nabla_x \phi,$$

et v_ϕ un champ de vitesse défini, pour tout $(t, x) \in [0, 1] \times \mathbb{R}^d$, par

$$v_\phi(t, x) = \frac{1}{t} \left(x - \text{Prox}_{\frac{t}{1-t} \phi} \left(\frac{x}{1-t} \right) \right) = \frac{x - \nabla_x (\phi_t)^*(x)}{t},$$

où $(\phi_t)^*$ désigne la transformée de Legendre de $\phi_t = (1-t)|\cdot|^2 + t\phi$, et où ϕ est un potentiel vérifiant la propriété (Γ_1) , tel que $\nabla \phi$ soit le plan de transport optimal de $\rho_0 \mathcal{L}^d$ sur $\rho_1 \mathcal{L}^d$ au sens du Théorème 1.1-7 (p. 20),

- $q_\phi = (-(1/2)|v_\phi|^2, v_\phi) \in L^\infty((0, 1) \times \Omega)^{d+1}$,
- $\psi_\phi \in W^{1,\infty}((0, 1] \times \Omega)$ et $q_\phi = \nabla_{t,x} \psi_\phi$.

Nous avons ensuite montré dans la section IV.2, que la composante $\mu_\phi = (\rho_\phi, \rho_\phi v_\phi)$, avec ρ_ϕ et v_ϕ tels que définis ci-dessus, était commune à tous les points selles du lagrangien \mathbf{L} , et ceci indépendamment du choix de ϕ , potentiel de transport optimal de $\rho_0 \mathcal{L}^d$ sur $\rho_1 \mathcal{L}^d$, vérifiant la propriété (Γ_1) .

En ce qui concerne l'algorithme de Benamou-Brenier (sous-section II.1.3), cela signifie que, quand bien même celui-ci, selon son initialisation, peut converger vers différents points selles (nous avons vu dans la sous-section IV.2.1 qu'il n'y avait pas unicité des points selles de \mathbf{L}), la densité ρ et la quantité de mouvement m qui seront alors approchées par l'algorithme seront toujours les mêmes, c'est à dire des solutions du problème de transport optimal dynamique.

Remarquons que nous n'avons pas prouvé l'unicité d'un couple $(\rho, m) \in \Lambda(\rho_0, \rho_1)$ minimisant l'énergie \mathcal{K} sur $\Lambda(\rho_0, \rho_1)$, où le terme \mathcal{K} , défini en 2.2-2 p. 57 désigne, rappelons-le, l'énergie définie sur $\Lambda(\rho_0, \rho_1)$ par

$$\mathcal{K}(\rho, m) = \frac{1}{2} \int_{(0,1) \times \Omega} |v|^2 \rho, \text{ avec } m = \rho v$$

et où $\Lambda(\rho_0, \rho_1)$ (également définie p. 57) désigne l'espace de contrainte dans $L^2((0, 1) \times \Omega)^{d+1}$ imposant :

- $\rho \geq 0$ et $\exists v \in L^2((0, 1) \times \Omega, \rho \mathcal{L}^d)$, $m = \rho v$,
- la conservation de la masse : $\partial_t \rho + \text{div}_x(\rho v) = 0$,
- les conditions initiales et finales $\rho(0, \cdot) = \rho_0$ et $\rho(\cdot, \cdot) = \rho_1$,
- les conditions de Neumann homogènes aux bords.

En effet, nous avons uniquement montré l'unicité d'un tel couple $(\rho, m) \in \Lambda(\rho_0, \rho_1)$ minimisant \mathcal{K} et pouvant s'intégrer dans un point selle de \mathbf{L} , c'est à dire pour lequel il existe un couple $(\psi, q) \in H^1((0, 1) \times \Omega) \times L^2((0, 1) \times \Omega)^{d+1}$ tel que, en posant $\mu = (\rho, m)$, le triplet (ψ, q, μ) soit un point selle de \mathbf{L} , i.e. vérifie les propriétés (I) (voir p. 74 ou au début de ce chapitre).

Pour finir, dans la section IV.3, nous avons entrepris de caractériser un champ de vitesse associé à un transport optimal (Théorème 4.3-9, p. 159) dans $L^2((0, 1) \times \mathbb{R}^d)$: nous avons pu conclure qu'un tel champ de vitesse pouvait être caractérisé par les propriétés (II), à savoir que ce champ de vitesse est localement borné, dérive spatialement d'un potentiel, et vérifie l'équation de Burgers. Autrement dit, plus qualitativement, que les trajectoires déterminées par ce champ de vitesse suivent des lignes droites et ne se croisent jamais.

Ce chapitre met fin à l'étude théorique entreprise au chapitre II. Au cours de celle-ci, nous avons pu établir l'existence d'un point selle pour le lagrangien \mathbf{L} , l'unicité de la composante densité-vitesse, ainsi que la convergence faible de l'algorithme de Benamou-Brenier, et la convergence forte d'une version relaxée de celui-ci. Ce travail aura en particulier été mené autour d'une étude approfondie des propriétés de comportement et de régularité d'un champ de vitesse associé à un transport optimal isotrope.

Le prochain (et dernier) chapitre de ce manuscrit sera quant à lui relativement indépendant des travaux proposés jusqu'à présent. En effet, il s'agira plutôt d'un chapitre d'exploration, se donnant pour objectif de proposer et d'étudier différents modèles, méthodes et algorithmes visant à introduire des contraintes physiques dans les modèles de transport optimal dynamique. Pour les besoins d'une telle exploration, et la liberté souvent nécessaire à une telle démarche, nous nous y autoriserons quelques conjectures et quelques écarts quant à la rigueur mathématique de certaines preuves ou assertions.

Chapitre V

Transport optimal multi-physique et interpolations d'images

Notes pour ce chapitre. *Ce chapitre est une version légèrement revisitée d'un article [46] co-publié avec Emmanuel Maître et Nicolas Papadakis, paru dans le journal ESAIM en 2015 sous le titre "Multi-physics optimal transportation and image interpolation". En conséquence, il est possible que certaines notations diffèrent par rapport aux chapitres précédents.*

L'un des grands intérêts de l'algorithme de Benamou et Brenier, et plus généralement de leur formulation dynamique [6], est qu'au lieu d'implémenter directement l'assignation, le résultat de leur algorithme est le chemin géodésique entre deux densités pour la distance de Wasserstein : c'est à dire l'interpolation de McCann. Une des principales applications de leur algorithme est l'interpolation d'images, et celui-ci a déjà été souvent exploité par la communauté du traitement d'images (voir la sous-section I.3.1 du chapitre I). Cet algorithme détermine le chemin géodésique entre deux images données selon les lois du transport optimal.

Structurellement, le transport optimal ne préserve pas la connectivité des différentes régions (la "topographie") tout au long du chemin d'interpolation (voir la Proposition 1.1-8 p. 23). Ainsi, il n'est pas très difficile d'exhiber des cas tests où de tels algorithmes produisent des chemins d'images (des "films") où des régions de hautes densités sont divisées dès le début avant de re-fusionner (voir [6] par exemple). Un tel comportement n'est pas souhaité pour l'interpolation d'images. Supposons, par exemple, que ces images représentent l'évolution en temps d'une gouttelette d'huile sur de l'eau. La gouttelette peut bien sûr se diviser au cours du temps ; cependant, un tel phénomène possède un coût physique qui n'est pas pris en compte dans une interpolation générée par un authentique transport optimal (une interpolation de McCann). Pour une membrane élastique, nous espérons qu'il n'y ait pas de division du tout !

Ce chapitre se donne ainsi pour objectif d'étudier des façons de prendre en compte de telles contraintes physiques dans la théorie du transport optimal. Nous étudierons le transport optimal anisotrope dans le but de pouvoir modéliser des domaines pouvant contenir des obstacles, des polarisations, des champs de force, etc... Nous considérerons également des contraintes physiques sur le champ de vitesse par lequel la masse est transportée (via l'équation de continuité), dans le but d'empêcher le transfert de masse de ne suivre que des lignes droites ou de se diviser. Dans ce but, nous introduirons plusieurs modèles dérivés du transport optimal qui seront, dans un premier temps, basés sur des variantes du terme d'énergie cinétique classique (voir (2.1-4) p. 52), ou bien, dans un second temps, en ajoutant des contraintes sur l'espace des minimiseurs (c'est à dire en plus de la contrainte de conservation de la masse). Nous nous intéresserons plus particulièrement à des transports convexes sur des variétés riemanniennes et à des transports

non-convexes induis par des contraintes physiques.

La formulation dynamique du transport optimal (i.e. du type "mécanique des fluides") est de plus la seule permettant de prendre en compte des densités qui s'annulent en certaines régions (c'est à dire sans faire d'hypothèse sur les supports des densités), ce qui est aussi d'un grand intérêt en traitement d'images.

Une précédente étude [15] avait déjà étudié la prise en compte de contraintes physiques, mais les outils d'optimisation utilisés ne semblaient pas très satisfaisants en pratique. Par la suite, dans [67], il fut proposé d'utiliser la méthode de splitting proximal. Au regard de ces deux articles, nos contributions seront les suivantes. Nous étudierons, dans une certaine mesure, le caractère bien posé du transport optimal anisotrope et des transports optimaux généralisés incluant des régularisations du champ de vitesse. Nous proposerons des schémas numériques adéquats pour la résolution des problèmes définis. Nous mettrons en application notre modèle pour interpoler des images océanographiques de haute résolution.

Ce chapitre est organisé comme il suit. Premièrement, nous rappellerons dans la section V.1 le schéma numérique introduit dans [67]. Ensuite, nous détaillerons comment un tel cadre peut être étendu pour traiter le transport optimal anisotrope dans la section V.2. Enfin, nous introduirons dans la section V.3 des fonctions non-convexes modélisant des contraintes physiques sur le champ de vitesse transportant la masse. Les validations numériques de la section V.4 démontreront l'intérêt pratique de nos algorithmes sur des exemples d'interpolation d'images.

V.1 Formulation dynamique et méthodes de splitting proximal

V.1.1 Petit rappel récapitulatif sur le transport optimal dynamique

Comme dans les autres chapitres, nous nous restreindrons ici aux plans de transport du type $T : \Omega \subset \mathbb{R}^d \mapsto \Omega$ où $d > 0$ désignera la dimension du problème, et Ω un domaine ouvert, convexe et borné de \mathbb{R}^d . Soient ρ_0, ρ_1 deux fonctions positives L^1 sur Ω de même masse : de même que dans les précédents chapitres, nous pourrions supposer, sans perdre de généralité, que

$$\int_{\Omega} \rho_0(x) dx = \int_{\Omega} \rho_1(x) dx = 1, \quad (5.1-1)$$

Soit T est un plan de transport optimal qui pousse la mesure $\rho_0 \mathcal{L}^d$ sur la mesure $\rho_1 \mathcal{L}^d$.

Rappelons que d'après la Proposition 1.1-8 (p. 23), le chemin géodésique pour la distance de Wasserstein entre deux mesures ayant pour densités ρ_0 et ρ_1 peut être explicité par le chemin $t \mapsto \rho(t, x)$ où $t \in [0, 1]$ est le paramètre du chemin et

$$\rho_0(x) = \rho(t, T_t(x)) |\det(\nabla T_t(x))| \text{ avec } T_t = (1-t) \text{id}_d + tT.$$

Rappelons que formellement, un tel chemin géodésique est solution du problème non-convexe sur les densités $\rho(t, x) \in \mathbb{R}_+$ et les champs de vitesse $v(t, x) \in \mathbb{R}^d$

$$\min_{(\rho, v) \in C_{st}} \frac{1}{2} \int_0^1 \int_{\Omega} \rho(t, x) \|v(t, x)\|^2 dx dt, \quad (5.1-2)$$

sur l'ensemble de contraintes non-linéaires

$$C_{st} = \{(\rho, v), \partial_t \rho + \text{div}_x(\rho v) = 0, \langle v, \vec{n} \rangle = 0, \rho(0, \cdot) = \rho_0, \rho(1, \cdot) = \rho_1\}. \quad (5.1-3)$$

Nous imposons des conditions de Neumann sur le champ de vitesse v , via le vecteur \vec{n} qui est le vecteur unitaire normal au domaine Ω .

D'après [6] et en introduisant le changement de variables $(\rho, v) \mapsto (\rho, m)$, où m désigne la quantité de mouvement $m = \rho v$, nous obtenons un problème d'optimisation convexe sur le couple (ρ, m) :

$$\min_{(\rho, m) \in C} \mathcal{J}(\rho, m) := \int_0^1 \int_{\Omega} J(\rho, m) dx dt$$

avec

$$J(\rho, m) = \begin{cases} \frac{\|m\|^2}{2\rho} & \text{si } \rho > 0 \\ 0 & \text{si } (\rho, m) = (0, 0) \\ +\infty & \text{autrement.} \end{cases} \quad (5.1-4)$$

L'ensemble des contraintes C devient alors linéaire et s'écrit :

$$C(\rho, m) = \{(\rho, m), \partial_t \rho + \operatorname{div}_x m = 0, \langle m, \vec{n} \rangle = 0, \rho(0, \cdot) = \rho_0, \rho(1, \cdot) = \rho_1\} \quad (5.1-5)$$

Notons que $\iota_C(x) = 0$ si $x \in C$ et $+\infty$ sinon ; le modèle convexe de Benamou-Brenier [6] devient alors :

$$(\rho^*, m^*) = \operatorname{argmin}_{\rho, m} \mathcal{J}(\rho, m) + \iota_C(\rho, m). \quad (5.1-6)$$

L'existence d'un minimiseur d'un tel problème dans le cas L^2 a été en grande partie l'objet des chapitres III et IV.

Le problème qui sera traité concrètement, c'est à dire une fois discrétisé, sera en fait légèrement plus complexe. En effet, comme déjà évoqué dans la section 1.2 du chapitre I, la projection sur l'ensemble des contraintes C demande un traitement numérique assez fin (l'inversion d'un laplacien par une méthode de Poisson) et est source d'une forte instabilité numérique. C'est pourquoi l'opérateur divergence sera discrétisé et calculé sur une grille décalée par rapport à celle sur laquelle sera discrétisée la fonctionnelle \mathcal{J} . Cette discrétisation sera celle proposée dans [67], et nous renvoyons donc à cet article pour plus de détails sur cette discrétisation. Le lecteur pourra également se référer à l'article [46] duquel est tiré ce chapitre, et dans lequel cette discrétisation est rappelée.

Notons simplement que ce traitement nécessite un opérateur de recentrage \mathcal{I} , **linéaire et bijectif** depuis la grille décalée sur la grille centrée. Au final, le problème discret 5.1-6 peut être reformulé ainsi :

$$\begin{aligned} (\rho^*, m^*) &= \operatorname{argmin}_{(\rho, m) \in \mathcal{E}_s} \mathcal{J}(\mathcal{I}(\rho, m)) + \iota_C(\rho, m), \\ &= \operatorname{argmin}_{(\rho, m) \in \mathcal{E}_s} \sum_{t=0}^1 \sum_{x \in \Omega} J(\mathcal{I}(\rho(t, x), m(t, x))) + \iota_C(\rho, m). \end{aligned} \quad (5.1-7)$$

Le problème (5.1-7) consiste donc à minimiser la somme de deux fonctions convexes régulières. De même que dans [67], nous pouvons faire usage de la méthode de splitting proximal pour résoudre le problème.

V.1.2 Optimisation avec la méthode de splitting proximal

Dans [67], le problème convexe 5.1-7 est optimisé par l'algorithme primal-dual de [25]. Un tel algorithme est dédié à l'optimisation de problèmes de la forme $G_1(Ku) + G_2(u)$, où G_i est une fonction convexe et K est un opérateur linéaire ($K = \mathcal{I}$ pour la formulation du transport optimal sur grilles décalées et grilles régulières). Dans cette étude, du fait que nous considérerons des

problèmes plus généraux, nous utiliserons plutôt l'algorithme de Forward-Backward généralisé (GFB) proposé dans [71]. Cet algorithme est dédié aux problèmes de la forme

$$\operatorname{argmin}_u \sum_{i=1}^n G_i(u) + F(u), \quad (5.1-8)$$

où le terme additif F est une fonction convexe et différentiable ayant un gradient L -lipschitzien ∇F et où les fonctions G_i sont supposées propres, convexe et simples.

Par simple, nous entendons que l'opérateur proximal des fonctions G_i , défini par

$$\operatorname{Prox}_{\tau G}(u) = \operatorname{argmin}_z \frac{\|z - u\|^2}{2\tau} + G(z), \quad (5.1-9)$$

peut être écrit de façon explicite. L'algorithme GFB s'écrit : Si F est uniformément convexe,

Algorithm 1 Generalized Forward-Backward

Soient $N \in \mathbb{N}^+$, $\epsilon > 0$, $\tau \in]0; 1/L[$, $\lambda \in]0; 1]$ et $\sum_{i=1}^n \omega_i = 1$, $\omega_i \geq 0$ et initialisons $z_i^0 = u^0$, $i = 1 \dots n$

for $k = 0 \dots N$ **do**

while $\|u^k - u^{k-1}\| > \epsilon$ **do**

$$z_i^{k+1} = z_i^k + \lambda \left(\operatorname{Prox}_{\tau/\omega_i G_i} \left(2u^k - z_i^k - \tau \nabla F(u^k) \right) - u^k \right), \quad i = 1 \dots n,$$

$$u^{k+1} = \sum_{i=1}^n \omega_i z_i^{k+1}$$

end while

end for

alors l'algorithme GFB converge fortement vers un minimiseur de (5.1-8); autrement, la convergence est faible. Ceci est une généralisation de l'algorithme Forward-Backward (FB) traitant de problèmes de la forme :

$$\operatorname{argmin}_u G(u) + F(u),$$

où G est une fonction propre, convexe, simple et différentiable ayant un gradient L -lipschitzien. Présentons ici l'algorithme FB (algorithme 2) : Notons que la méthode primale-duale a été

Algorithm 2 Forward-Backward

Soient $N \in \mathbb{N}^+$, $\epsilon > 0$, $\tau \in]0; 1/L[$ et initialisons u^0

for $k = 0 \dots N$ **do**

while $\|u^k - u^{k-1}\| > \epsilon$ **do**

$$u^{k+1} = \operatorname{Prox}_{\tau G}(u^k - \tau \nabla F(u^k))$$

end while

end for

récemment étendue [29, 13, 23] pour traiter le problème (5.1-8).

V.1.3 Application au problème (5.1-7)

Pour pouvoir traiter les différentes grilles numériques, nous devons définir la variable $x = (\bar{\rho}, \bar{m})$ sur la grille régulière et $y = (\rho, m)$ sur la grille décalée. Nous définissons donc l'espace linéaire des contraintes $S(x, y) = \{\mathcal{I}(y) = x\}$ pour coupler ces deux variables. Ainsi, le problème (5.1-7) peut être reformulé :

$$(x^*, y^*) = \operatorname{argmin}_{x, y} \mathcal{J}(x) + \iota_C(y) + \iota_S(x, y).$$

L'algorithme GFB (algorithme 1) peut donc au final être appliqué en prenant $F = 0$, $G_1(x, y) = \mathcal{J}(x) + \iota_C(y)$ et $G_2(x, y) = \iota_S(x, y)$.

Nous allons maintenant détailler la manière d'implémenter les opérateurs proximaux des différents termes concernés par ce problème.

Calcul de $\text{Prox}_{\tau\mathcal{J}}$

Comme la fonctionnelle \mathcal{J} est séparable, l'opérateur proximal peut être implémenté indépendamment de J en chaque point $(\tilde{\rho}, \tilde{m})(t, x) \in \mathbb{R} \times \mathbb{R}^d$ avec $(t, x) \in [0; 1] \times \Omega$. Notons que J est une fonction 1-homogène du fait que $J(k\rho, km) = kJ(\rho, m)$. En conséquence, sa transformée de Legendre est l'indicatrice d'un ensemble convexe. Cela correspond donc à la projection $q^* = P_K(\tilde{q})$ sur l'ensemble convexe

$$K = \left\{ q = (a, b) \in \mathbb{R} \times \mathbb{R}^d \text{ t.q. } a + \frac{|b|^2}{2} \leq 0 \right\}$$

de l'étape B de l'algorithme de [6]. On pourrait montrer [67] que cette projection peut être obtenue de façon équivalente par une application ponctuelle de l'opérateur proximal de J au point $(\tilde{\rho}, \tilde{m})$.

Notons ι_K l'indicatrice de l'ensemble convexe K ($\iota_K(q) = 0$ si $q \in K$ et $+\infty$ sinon). L'opérateur proximal de ι_K représente seulement la projection sur l'ensemble convexe K , d'où

$$q^* = \text{Prox}_{\gamma\iota_K}(\tilde{q}) = P_K(\tilde{q}).$$

Par l'identité de Moreau, nous obtenons que :

$$\text{Prox}_{\gamma F^*}(x) = x - \gamma \text{Prox}_{F/\gamma}(x/\gamma).$$

En notant , il est ensuite possible de montrer que

$$\text{Prox}_{\iota_K^*}(\tilde{\rho}, \tilde{m}) = (\tilde{\rho}, \tilde{m}) - P_K(\tilde{\rho}, \tilde{m}) = \text{Prox}_J(\tilde{\rho}, \tilde{m}),$$

de sorte que

$$P_K(\tilde{\rho}, \tilde{m}) = (\tilde{\rho}, \tilde{m}) - \text{Prox}_J(\tilde{\rho}, \tilde{m}),$$

ce qui montre que la projection sur K peut être calculée par application de l'opérateur proximal.

La Proposition qui suit, présentée dans [67], montre que la fonctionnelle J définie en (5.1-4) est simple, dans le sens où son opérateur proximal peut être calculé de façon explicite.

Proposition 5.1-1.

Pour tout $(\tilde{\rho}, \tilde{m}) \in \mathbb{R} \times \mathbb{R}^d$,

$$\text{Prox}_{\tau J}(\tilde{\rho}, \tilde{m}) = \begin{cases} (\rho^*, \mu(\rho^*)) & \text{si } \rho^* > 0, \\ (0, 0) & \text{sinon.} \end{cases}$$

où

$$\forall \rho \geq 0, \quad \mu(\rho) = \frac{\rho \tilde{m}}{\rho + \tau}$$

et ρ^* désigne la plus grand racine réelle de l'équation polynomiale du troisième ordre en ρ

$$P(\rho) = (\rho - \tilde{\rho})(\rho + \tau)^2 - \frac{\tau}{2} \|\tilde{m}\|^2 = 0.$$

Calcul de $\text{Prox}_{\tau\iota_C}$

L'opérateur proximal de ι_C se révèle être la projection orthogonale Proj_C sur l'ensemble convexe C .

Tout d'abord, nous extrayons les valeurs du bord de la grille décalée en utilisant l'opérateur linéaire b , pour $\tilde{y} = (\tilde{\rho}, \tilde{m})$. Pour $d = 2$ et $m = (m_1, m_2)$, en prenant $\Omega = [0; 1]^2$ et $t \in [0; 1]$, cet opérateur s'écrit :

$$b(\tilde{y}) = (\tilde{\rho}(0, \cdot, \cdot), \tilde{\rho}(1, \cdot, \cdot), \tilde{m}_1(\cdot, 0, \cdot), \tilde{m}_1(\cdot, 1, \cdot), \tilde{m}_2(\cdot, \cdot, 0), \tilde{m}_2(\cdot, \cdot, 1))$$

Prenons les conditions de bord suivantes :

$$b_0 = (\rho_0, \rho_1, 0, 0, 0, 0).$$

L'espace affine C , qui peut être réécrit

$$C = \{y = (\rho, m), Ay = z\}$$

où

$$Ay = (b(y), \text{div}_{t,x}(y)), \text{ et } z = (b_0, 0).$$

Cette projection peut être implémentée en résolvant un système linéaire du type

$$\text{Proj}_C(\tilde{y}) = (\text{id} - A^* \Delta^{-1} A) \tilde{y} + A^* \Delta^{-1} z$$

où l'application $\Delta^{-1} = (AA^*)^{-1}$ requiert la résolution d'une équation de Poisson sur la grille centrée avec les conditions de bord définies ci-dessus. Ceci peut être obtenu par l'application d'une transformée de Fourier rapide (FFT) en $O(NP \log(NP))$ opérations, où N et P représentent le nombre de points en espace et en temps, voir [74].

Calcul de $\text{Prox}_{\tau\iota_S}$

L'opérateur proximal de ι_S se révèle quant à lui être l'opérateur de projection orthogonale Proj_S sur l'ensemble convexe $S(x, y) = \{\mathcal{I}(y) = x\}$. Il est défini par $(x, y) = \text{Proj}_S(\tilde{x}, \tilde{y}) = (\mathcal{I}(z), z)$ où

$$z = (\text{id} + \mathcal{I}^* \mathcal{I})^{-1}(\tilde{y} + \mathcal{I}^*(\tilde{x})), \quad (5.1-10)$$

et où \mathcal{I}^* représente l'adjoint de l'opérateur d'interpolation linéaire. Notons qu'ici, l'implémentation de Proj_S nécessite la résolution d'un système linéaire. Or, ce système est séparable le long de chaque dimension de la grille de discrétisation, c'est pourquoi il n'est en pratique nécessaire que de résoudre une série de petits systèmes linéaires. De plus, du fait que les matrices inverses correspondantes sont les mêmes le long de chaque dimension, nous pouvons pré-implémenter explicitement l'inverse de ces $d + 1$ matrices.

V.1.4 Limitations et motivations

Lors de la résolution d'un problème de transport optimal entre deux densités, nous savons que les trajectoires du transfert de masses seront des lignes droites. Les structures contenues dans les données originales peuvent donc disparaître le long du chemin optimal (on parle bien ici du chemin entre les densités). Un tel comportement est illustré dans la Figure V.1 qui montre le chemin optimal entre deux images contenant chacune deux gaussiennes. Ce résultat a été obtenu par l'algorithme GFB, en prenant $\tau = 0.1$ et $\omega_1 = \omega_2 = \frac{1}{2}$.

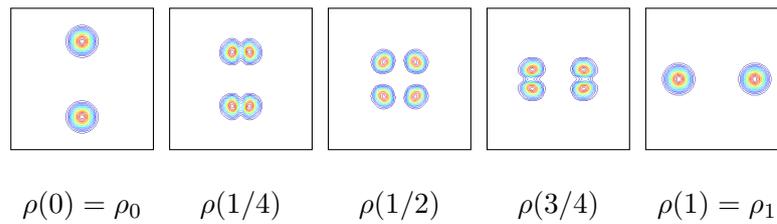


FIGURE V.1 – L'expérience des deux gaussiennes : les gaussiennes se divisent le long du chemin optimal entre ρ_0 et ρ_1 .

Dans cette étude, nous souhaitons incorporer certaines contraintes physiques dans le but de préserver les structures contenues dans les données initiales et finales le long du chemin optimal. Cela pourrait s'avérer être d'un certain intérêt, par exemple en ce qui concerne l'imagerie en géosciences : L'océan et l'atmosphère, observés par satellite, sont soumis à un ensemble complexe de lois physiques. La possibilité de réaliser des interpolations temporelles entre ces images satellites (qui ne peuvent en pratique que capturer des images souvent assez espacées dans le temps) est un problème important pour cette communauté, et ces interpolations doivent être compatibles avec les lois dynamiques sous-jacentes. Du fait que les interpolations réalisées avec un transport optimal classique (en milieu isotrope) correspondent à l'évolution de masses globalement constantes le long de lignes droites, celui-ci s'avère ne pas être une solution satisfaisante.

Nous proposons maintenant quelques généralisation ayant pour but de mieux aborder ce type de problèmes.

V.2 Transport Optimal Anisotrope

Nous proposons tout d'abord d'utiliser cette formulation, afin de définir simplement un problème de transport optimal en milieu anisotrope (obstacles, polarisation, champs de force, etc...) en pénalisant cette énergie de déplacement : on cherche un transport continu de masse (ρ, v) qui minimise une énergie du type

$$\mathcal{K}_A(\rho, v) = \int_0^1 \int_{\Omega} \rho(t, x) v(t, x)^T A(t, x) v(t, x) dx dt, \quad (5.2-11)$$

avec $A(t, x)$ une matrice symétrique et définie positive, représentant la pénalisation anisotrope de l'énergie de déplacement. Un tel modèle peut, comme déjà évoqué plus haut, avoir un intérêt dans beaucoup d'applications, telles que la météorologie ou l'océanographie. Par exemple, si nous souhaitons interpoler ou mesurer la distance (au sens de Wasserstein) entre deux images météorologiques, nous devons prendre en compte les contraintes atmosphériques : déplacement de masses d'air, mouvements rotatifs (cyclones et anticyclones), etc... Nous allons montrer dans cette section qu'une formulation telle que définie en (5.2-11) peut naturellement traiter ce type de problèmes anisotropes.

Nous commencerons par traiter un cas simple et explicite : celui où la matrice A est constante (i.e. $A(t, x) = A$). Puis, dans un second temps et dans un cadre plus formel nous étudierons le cas où la matrice A varie, mais en espace seulement (i.e. $A(t, x) = A(x)$). Pour finir nous implémenterons et illustrerons ce problème.

V.2.1 Cas d'une anisotropie constante

Nous évoquons ici le cas élémentaire où la matrice de pénalisation A est constante. Pour simplifier, plaçons dans un domaine en deux dimensions (le cas des dimensions supérieures sont identiques).

Soient $\rho_0, \rho_1 \in L^2(\Omega)$, et soit A une matrice symétrique et définie positive (donc inversible), soit R une racine carrée de A ($A = R^2$) et définissons $\Omega_R = R\Omega$, la "dilatation du domaine" Ω par la matrice R . On définit le terme d'énergie pénalisée

$$\begin{aligned} \mathcal{E}_A(\rho_0, \rho_1) &= \inf_{(\rho, v) \in C} \mathcal{K}_A(\rho, v) = \inf_{(\rho, v) \in C} \int_0^1 \int_{\Omega} \rho v^T A v \, dx \, dt = \inf_{(\rho, v) \in C} \int_0^1 \int_{\Omega} \rho v^T R^2 v \, dx \, dt \\ &= \inf_{(\rho, v) \in C} \int_0^1 \int_{\Omega} |Rv|^2 \rho \, dx \, dt, \end{aligned} \quad (5.2-12)$$

sur l'espace de contrainte

$$C = \{(\rho, v), \partial_t \rho + \operatorname{div}_x(\rho v) = 0, \langle \rho v, \vec{n} \rangle = 0, \rho(0, \cdot) = \rho_0, \rho(1, \cdot) = \rho_1\}. \quad (5.2-13)$$

Définissons également la distance D_R par :

$$D_R(\rho_0, \rho_1)^2 = \mathcal{W}_2^2 \left((\rho_0 \circ R^{-1}) \mathcal{L}^d, (\rho_1 \circ R^{-1}) \mathcal{L}^d \right)^2 = \inf_{(\tilde{\rho}, \tilde{v}) \in C_R} \int_0^1 \int_{\Omega_R} |\tilde{v}|^2 \tilde{\rho} \, dx \, dt, \quad (5.2-14)$$

avec le nouvel espace de contraintes

$$C_R = \{(\tilde{\rho}, \tilde{v}), \partial_t \tilde{\rho} + \operatorname{div}_x(\tilde{\rho} \tilde{v}) = 0, \langle \tilde{\rho} \tilde{v}, \vec{n}_R \rangle = 0, \tilde{\rho}(0, \cdot) = \rho_0 \circ R^{-1}, \tilde{\rho}(1, \cdot) = \rho_1 \circ R^{-1}\}. \quad (5.2-15)$$

Remarquons qu'ici le vecteur unitaire $\vec{n}_R(z)$ orthogonal à Ω en un point $z \in \partial\Omega_R$ peut s'écrire sous la forme $\vec{n}_R(z) = R^{-1} \vec{n}(R^{-1}z)$, où $\vec{n}(R^{-1}z)$ est le vecteur unitaire orthogonal à Ω_R au point $R^{-1}z \in \partial\Omega$.

Montrons déjà qu'il y a une bijection entre les espaces de contrainte C et C_R , via une dilatation par R de ses éléments. Pour tout $(\rho, v) \in L^2(Q) \times L^2(Q, \rho \mathcal{L}^d)$, on définit les densités "dilatées"

$$\rho_R(t, x) = \rho(t, R^{-1}x) \text{ et } v_R(t, x) = Rv(t, R^{-1}x).$$

Proposition 5.2-2.

⌊ Sous les définitions ci-dessus, si $(\rho, v) \in C$, alors $(\rho_R, v_R) \in C_R$.

Preuve : Soit $(\rho, v) \in C$. Les conditions initiales et finales imposées par C_R sont alors vérifiées par ρ_R par définition même. De même, pour tout $z \in \partial\Omega_R$, on a

$$\begin{aligned} \langle \rho_R(z) v_R(z), \vec{n}_R(z) \rangle &= \langle \rho(R^{-1}z) Rv(R^{-1}z), R^{-1} \vec{n}(R^{-1}z) \rangle \\ &= \langle \rho(R^{-1}z) (R^{-1})^T Rv(R^{-1}z), \vec{n}(R^{-1}z) \rangle \\ &= \langle \rho(R^{-1}z) v(R^{-1}z), \vec{n}(R^{-1}z) \rangle = 0 \end{aligned}$$

(rappelons que R est symétrique, de même donc pour R^{-1} , i.e. $(R^{-1})^T = R^{-1}$).

Pour conclure, il suffit donc de prouver que $\partial_t \rho_R + \operatorname{div}_x(\rho_R v_R) = 0$. Posons

$$R = \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix} \quad \text{et} \quad R^{-1} = \begin{pmatrix} r'_{11} & r'_{12} \\ r'_{21} & r'_{22} \end{pmatrix},$$

pour tout champ de vecteurs $w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ sur $[0, 1] \times \Omega_R$, et en notant "Tr" l'opérateur trace, on a

- $\operatorname{div}_x(Rw) = (r_{11}\partial_{x_1}w_1 + r_{12}\partial_{x_1}w_2) + (r_{21}\partial_{x_2}w_1 + r_{22}\partial_{x_2}w_2) = \operatorname{Tr}(R\nabla_x w)$,
- pour tout $[0, 1]$,

$$\begin{aligned} \nabla_x(w(t, R^{-1}\cdot)) &= \begin{pmatrix} r'_{11}\partial_{x_1}w_1 + r'_{21}\partial_{x_2}w_1 & r'_{12}\partial_{x_1}w_1 + r'_{22}\partial_{x_2}w_1 \\ r'_{11}\partial_{x_1}w_2 + r'_{21}\partial_{x_2}w_2 & r'_{12}\partial_{x_1}w_2 + r'_{22}\partial_{x_2}w_2 \end{pmatrix} (t, R^{-1}\cdot) \\ &= ((\nabla_x w)R^{-1}) (t, R^{-1}\cdot). \end{aligned}$$

De ce fait, on a

$$\begin{aligned} \operatorname{div}_x(\rho_R v_R) &= \operatorname{div}_x(R(\rho v)(t, R^{-1}\cdot)) = \operatorname{Tr}(R\nabla_x(\rho v)R^{-1}) (t, R^{-1}\cdot) \\ &= \operatorname{Tr}(\nabla_x(\rho v)) (t, R^{-1}\cdot) = \operatorname{div}_x(\rho v)(t, R^{-1}\cdot), \end{aligned}$$

$$d'où \partial_t \rho_R + \operatorname{div}_x(\rho_R v_R) = (\partial_t \rho + \operatorname{div}_x(\rho v))(t, R^{-1}\cdot) = 0.$$

□

Proposition 5.2-3.

On a l'égalité

$$\mathcal{E}_A(\rho_0, \rho_1) = \frac{1}{|\det(R)|} D_R(\rho_0, \rho_1)^2$$

et, pour tout couple $(\rho^*, v^*) \in C$, on a

$$\mathcal{K}_A(\rho^*, v^*) = \int_0^1 \int_{\Omega} |Rv^*|^2 \rho^* dx dt = \mathcal{E}_A(\rho_0, \rho_1) \implies \int_0^1 \int_{\Omega_R} |v_R^*|^2 \rho_R^* dx dt = D_R(\rho_0, \rho_1)^2.$$

Preuve :

1. Soit $(\rho, v) \in C$. Par changement de variable $x \leftarrow R^{-1}x$, on a l'égalité

$$\begin{aligned} \mathcal{K}_A(\rho, v) &= \int_0^1 \int_{\Omega} |Rv|^2 \rho dx dt = \frac{1}{|\det(R)|} \int_0^1 \int_{\Omega_R} |v_R|^2 \rho_R dx dt \\ &\geq \frac{1}{|\det(R)|} D_R(\rho_0, \rho_1)^2, \end{aligned}$$

avec $\rho_R(t, \cdot) = \rho(t, R^{-1}\cdot)$ et $v_R(t, x) = Rv(t, R^{-1}x)$ (et donc $(\rho_R, v_R) \in C_R$ d'après la Proposition 5.2-2).

$$\text{On a donc } \mathcal{E}_A(\rho_0, \rho_1) = \inf_{(\rho, v) \in C} \mathcal{K}_A(\rho, v) \geq \frac{1}{|\det(R)|} D_R(\rho_0, \rho_1)^2.$$

2. Soit maintenant $(\tilde{\rho}, \tilde{v}) \in C_R$. Pour tout $t \in [0, 1]$, on pose

$$\rho(t, \cdot) = \tilde{\rho}_{R^{-1}}(t, \cdot) = \tilde{\rho}(t, R\cdot)$$

et

$$v(t, \cdot) = \tilde{v}_{R^{-1}}(t, \cdot) = R^{-1}\tilde{v}_R(t, Rx).$$

On a alors $(\rho, v) \in (C_R)_{R^{-1}} = C$, et en procédant cette fois à un changement de variable $y \leftarrow Ry$, on a

$$\begin{aligned} \int_0^1 \int_{\Omega_R} |\tilde{v}|^2 \tilde{\rho} \, dx \, dt &= \int_0^1 \int_{\Omega_R} (R^{-1}\tilde{v})^T R^2 (R^{-1}\tilde{v}) \tilde{\rho} \, dx \, dt = |\det(R)| \int_0^1 \int_{\Omega} \rho v^T R^2 v \, dx \, dt \\ &\geq |\det(R)| \mathcal{E}_A(\rho_0, \rho_1). \end{aligned}$$

On a donc $D_R(\rho_0, \rho_1)^2 = \inf_{(\tilde{\rho}, \tilde{v}) \in C_R} \int_0^1 \int_{\Omega_R} |\tilde{v}|^2 \tilde{\rho} \, dx \, dt \geq |\det(R)| \mathcal{E}_A(\rho_0, \rho_1)$.

Au final, on a bien $\frac{1}{|\det(R)|} D_R(\rho_0, \rho_1)^2 = \mathcal{E}_A(\rho_0, \rho_1)$.

Pour finir, soit $(\rho^*, v^*) \in C$ tel que

$$\int_0^1 \int_{\Omega} |Rv^*|^2 \rho^* \, dx \, dt = \inf_{(\rho, v) \in C} \int_0^1 \int_{\Omega} |Rv|^2 \rho \, dx \, dt = \mathcal{E}_A(\rho_0, \rho_1).$$

En posant $\rho_R^*(t, x) = \rho^*(t, R^{-1}x)$ et $v_R^*(t, x) = Rv^*(t, R^{-1}x)$, on a

$$\int_0^1 \int_{\Omega} |Rv^*|^2 \rho^* \, dx \, dt = \frac{1}{|\det(R)|} \int_0^1 \int_{\Omega_R} \rho_R^* |v_R^*|^2 \, dx \, dt = \frac{1}{|\det(R)|} D_R(\rho_0, \rho_1)^2,$$

d'où $(\rho_R^*, v_R^*) \in C_R$ (par la Proposition 5.2-2), et $D_R(\rho_0, \rho_1)^2 = \int_0^1 \int_{\Omega_R} \rho_R^* |v_R^*|^2 \, dx \, dt$.

□

Ainsi, réaliser un transport optimal anisotrope entre deux densités ρ_0 et ρ_1 sur un domaine pénalisé par une matrice constante A , symétrique définie positive, revient à réaliser un transport optimal isotrope sur le domaine "déformé" $\Omega_R = R\Omega$, entre les densités $\rho_0 \circ R^{-1}$ et $\rho_1 \circ R^{-1}$, où R est une racine carrée de la matrice A .

V.2.2 Existence d'une solution dans un cas anisotrope en espace seulement

Lorsque l'anisotropie du problème varie seulement spatialement (i.e. $A(t, x) = A(x)$), nous pouvons tenter de nous ramener à un problème classique de Monge-Kantorovitch sur une variété riemannienne. Dans le cas où une telle anisotropie dépend également du temps, quand bien même les algorithmes numériques continuent de nous fournir en pratique des solutions qui semblent correspondre à l'intuition que l'on a pour ce type de problème, il n'y a pas à notre connaissance d'équivalence avec aucun problème de transport optimal, et l'existence et l'unicité d'une solution reste donc un problème ouvert. C'est pourquoi, nous nous intéresserons ici seulement au cas $A(t, x) = A(x)$. Énonçons tout d'abord le Théorème de McCann [55].

Théorème 5.2-4. McCANN

Soit M une variété riemannienne et soit d sa distance géodésique associée. Soit μ une mesure de probabilité telle que μ soit :

1. absolument continue par rapport à la mesure du volume sur M ,
2. également absolument continue par rapport à toutes les autres mesures de probabilité ν sur M qui sont liées à μ par un plan de transport à coût fini.

Alors, pour toute mesure ν vérifiant la deuxième hypothèse énoncé ci-dessus par μ , il existe un unique plan de transport optimal $T(x) = \exp_x(-\nabla\psi(x))$ (défini et uniquement déterminé pour μ -presque tout x , avec $\psi : M \mapsto \mathbb{R}$ une fonction d^2 -concave) résolvant le problème de Monge :

$$\int d(x, T(x))^2 d\mu(x) = \inf_{S, S\#\mu=\nu} \int d(x, S(x))^2 d\mu(x).$$

De plus, pour μ presque tout $x \in M$, le chemin $\gamma(t) = \exp_x(-t\nabla\psi(x))$, $t \in [0, 1]$ est l'unique géodésique reliant x à $T(x)$.

Dans tout ce qui suivra, nous prendrons $M = \Omega \subset \mathbb{R}^d$. Nous munissons alors Ω de la métrique riemannienne g_A qui associe à chaque $x \in \Omega$ le produit scalaire sur $\mathbb{R}^d : (Y, Z) \mapsto Y^\top A(x)Z$. Nous pouvons ensuite définir la distance géodésique associé sur Ω :

$\forall (x, y) \in \Omega$,

$$d_A(x, y) = \inf_{\gamma} \left\{ \int_0^1 \sqrt{g_A(\dot{\gamma}(t), \dot{\gamma}(t))} dt = \int_0^1 (\dot{\gamma}(t)^\top A(\gamma(t)) \dot{\gamma}(t))^{\frac{1}{2}} dt; \gamma(0) = x, \gamma(1) = y \right\}.$$

Nous appliquons le Théorème ci-dessus pour $\mu = \rho_0(x) \sqrt{\det A(x)} dx$ et $\nu = \rho_1(x) \sqrt{\det A(x)} dx$, de façon à ce que μ et ν soient deux mesures de probabilité sur M sous l'hypothèse (5.1-1). Le problème de Monge sur la variété riemannienne (Ω, g_A) pour le coût $c = d_A^2$ consiste à la recherche d'un plan de transport optimal T satisfaisant :

$$\int_{\Omega} d_A(x, T(x))^2 \rho_0(x) \sqrt{\det A(x)} dx = \inf_{S\#_A \rho_0 = \rho_1} \int_{\Omega} d_A(x, S(x))^2 \rho_0(x) \sqrt{\det A(x)} dx,$$

où nous notons $S\#_A \mu = \nu$ comme abréviation de $S\#(\sqrt{\det A} \mu) = \sqrt{\det A} \nu$.

Remarquons que nous avons $S\#_A \rho_0 = S\# \rho_0 = \rho_1$ si $A(x) = \mathbb{I}_d$ pour tout $x \in \text{supp } \rho_0 \cup \text{supp } \rho_1$.

La Proposition qui suit montre que, si A vérifie les hypothèses ci-dessus par rapport au support de ρ_0 et ρ_1 , ce problème de Monge peut être résolu en déterminant $(\rho^*, v^*) \in C_{st}$ (C_{st} étant l'ensemble des contraintes de conservation de la masse défini par la relation (5.1-3)) minimisant le coût de transport (5.2-11) avec $A(t, x) = A(x)$. Le Théorème de McCann nous donne l'existence d'un plan de transport optimal T pour la distance d_A , correspondant à un déplacement à vitesse constante le long des géodésiques de (Ω, d_A) . Nous pouvons montrer **formellement**, comme dans le cas isotrope de [6] (voir l'énoncé 1.1-9 p. 28), qu'il existe un lien entre le champ de vitesse v^* et le chemin géodésique (dans (Ω, d_A)) reliant x à $T(x)$.

Proposition 5.2-5.

Si A vérifie $A(x) = \mathbb{I}_d$ pour tout $x \in \text{supp } \rho_0 \cup \text{supp } \rho_1$, l'estimation du coût du transport optimal sur la variété riemannienne (Ω, g_A) (i.e. la distance de Wasserstein associée à d_A) revient à minimiser l'énergie de déplacement géodésique, c'est à dire : trouver (ρ^*, v^*) tel que,

$$\int_0^1 \int_{\Omega} v^{*T} A v^* \rho^* = \inf_{(\rho, v) \in C_{st}} \int_0^1 \int_{\Omega} v^T A v \rho = \inf_{S \# \rho_0 = \rho_1} \int_{\Omega} d_A(x, S(x))^2 \rho_0(x) dx,$$

Nous énonçons ici une preuve formelle, dans l'esprit de [6], qui pourrait être mieux justifiée sous certaines hypothèses de régularités. Soient $(\rho, v) \in C^0$, et X la solution du problème

$$\begin{cases} \partial_t y(t, x) = v(t, y(t, x)), & \forall (t, x) \in (0, 1) \times \Omega, \\ y(0, x) = x, & \forall x \in \Omega. \end{cases}$$

De part les conditions de bord sur v dans C_{st} , voir (5.1-3), nous avons $X(t, \Omega) = \Omega$ pour tout $t \in [0, 1]$. Nous disposons de la relation classique de Monge-Ampère $\rho(t, X(t, x)) \det(\nabla X(t, x)) = \rho_0(x)$, i.e. $X(t, \cdot) \# \rho_0 = \rho(t, \cdot)$. Par l'inégalité de Jensen, il en suit que

$$\begin{aligned} & \int_0^1 \int_{\Omega} \rho(t, x) v(t, x)^T A(x) v(t, x) dx dt \\ &= \int_0^1 \int_{\Omega} v(t, X(t, x))^T A(X(t, x)) v(t, X(t, x)) \rho(t, X(t, x)) \det(\nabla X(t, x)) dx dt, \\ &= \int_0^1 \int_{\Omega} \partial_t X(t, x)^T A(X(t, x)) \partial_t X(t, x) dt \rho_0(x) dx, \\ &\geq \int_{\Omega} \left(\int_0^1 (\partial_t X(t, x)^T A(X(t, x)) \partial_t X(t, x))^{\frac{1}{2}} dt \right)^2 \rho_0(x) dx, \\ &\geq \int_{\Omega} d_A(x, X(1, x))^2 \rho_0(x) dx, \\ &\geq \int_{\Omega} d_A(x, T(x))^2 \rho_0(x) dx, \end{aligned}$$

où les dernières inégalités proviennent du fait que $X(1, \cdot)$ est tel que $X(1, \cdot) \# \rho_0 = X(1, \cdot) \#_A \rho_0 = \rho_1$.

Inversement, supposons que T soit un plan de transport optimal pour la métrique g_A . Définissons, pour presque tout $x \in \Omega$, l'application $X(\cdot, x) \in C^1([0, 1], \Omega)$ comme étant l'unique géodésique reliant x à $T(x)$ ($X(\cdot, x)$ est régulier en temps pour presque tout $x \in \Omega$). Nous supposons toujours que $X(t, \Omega) = \Omega$ pour tout $t \in [0, 1]$. Définissons également v comme

$$\begin{cases} v(t, X(t, x)) = \partial_t X(t, x) & \forall (t, x) \in (0, 1) \times \Omega, \\ v(t, y) = 0 & \text{ailleurs,} \end{cases}$$

et $\rho(t, \cdot) = \sqrt{\det A(\cdot)} (X(t, \cdot) \#_A \rho_0)$. Ainsi, pour tout $\phi \in C_c^\infty((0, 1) \times \Omega)$, nous obtenons :

$$\begin{aligned} \int_0^1 \int_\Omega (\partial_t \phi + v \cdot \nabla \phi) \rho &= \int_0^1 \int_\Omega (\partial_t \phi + v \cdot \nabla \phi)(t, x) \sqrt{\det A(x)} (X(t, \cdot) \#_A \rho_0)(x) dx dt \\ &= \int_0^1 \int_\Omega (\partial_t \phi(t, X(t, x)) + v(t, X(t, x)) \cdot \nabla \phi(t, X(t, x))) \rho_0(x) \underbrace{\sqrt{\det A(x)}}_{=1} dx dt \\ &= \int_\Omega \int_0^1 (\partial_t \phi(t, X(t, x)) + \partial_t X(t, x) \cdot \nabla \phi(t, X(t, x))) \rho_0(x) dt dx \\ &= \int_\Omega \int_0^1 \frac{d}{dt} (\phi(t, X(t, x))) dt \rho_0(x) dx = 0 \end{aligned}$$

i.e. $\partial_t \rho + \operatorname{div}_x(\rho v) = 0$ dans $\mathcal{D}'((0, 1) \times \Omega)$, d'où $(\rho, v) \in C_{st}$. Par définition nous avons :

$$\int_\Omega d_A(x, T(x))^2 \rho_0(x) dx = \int_\Omega \left(\int_0^1 (\partial_t X(t, x)^T A(X(t, x)) \partial_t X(t, x))^{\frac{1}{2}} dt \right)^2 \rho_0(x) dx.$$

Du fait que $t \mapsto X(t, x)$ soit un chemin géodésique et que A soit indépendante du temps, on a

$$\frac{d}{dt} (\partial_t X(t, x)^T A(X(t, x)) \partial_t X(t, x)) = 0,$$

d'où l'on déduit

$$\left(\int_0^1 (\partial_t X(t, x)^T A(X(t, x)) \partial_t X(t, x))^{\frac{1}{2}} dt \right)^2 = \int_0^1 \partial_t X(t, x)^T A(X(t, x)) \partial_t X(t, x) dt.$$

Nous obtenons finalement, du fait que $(\rho, v) \in C_{st}$,

$$\begin{aligned} \int_\Omega d_A(x, T(x))^2 \rho_0(x) dx &= \int_0^1 \int_\Omega \partial_t X(t, x)^T A(X(t, x)) \partial_t X(t, x) \rho_0(x) dt dx, \\ &= \int_0^1 \int_\Omega v(t, X(t, x))^T A(X(t, x)) v(t, X(t, x)) \rho_0(x) \underbrace{\sqrt{\det A(x)}}_{=1} dx dt, \\ &= \int_0^1 \int_\Omega v(t, x)^T A(x) v(t, x) \sqrt{\det A(x)} (X(t, \cdot) \#_A \rho_0)(x) dx dt, \\ &= \int_0^1 \int_\Omega v(t, x)^T A(x) v(t, x) \rho(t, x) dx dt, \\ &\geq \inf_{(\tilde{\rho}, \tilde{v}) \in C_{st}} \int_0^1 \int_\Omega \tilde{v}(t, x)^T A(x) \tilde{v}(t, x) \tilde{\rho}(t, x) dx dt. \end{aligned}$$

□

V.2.3 Discrétisation, convexification et optimisation

Nous considérons maintenant un domaine discret paramétrisé par $x \in \Omega \subset \mathbb{R}^d$ ($d = 2$ dans ce qui suivra) et $t \in [0, 1]$. Le problème de transport optimal anisotrope s'écrit :

$$\min_{(\rho, v) \in C_{st}} \frac{1}{2} \sum_{t=0}^1 \sum_{x \in \Omega} \rho(t, x) v(t, x)^T A(x) v(t, x),$$

où $\rho(t, x) \in \mathbb{R}$, $v(t, x) = (v_1(t, x), v_2(t, x)) \in \mathbb{R}^2$, et où la 2×2 matrice symétrique définie positive $A(x)$ représente l'anisotropie du domaine Ω .

Par le même changement de variable que précédemment, ($m = \rho v$), nous pouvons rendre convexe le coût du transport anisotrope comme il suit :

$$\mathcal{J}_A(\rho, m) = \sum_{t=0}^1 \sum_{x \in \Omega} J_A(\rho(t, x), m(t, x)),$$

avec

$$J_A(\rho(t, x), m(t, x)) = \begin{cases} \frac{m(t, x)^T A(x) m(t, x)}{\rho(t, x)} & \text{if } \rho(t, x) > 0 \\ 0 & \text{if } (\rho(t, x), m(t, x)) = (0, 0) \\ +\infty & \text{otherwise} \end{cases}$$

où $m(t, x) = (m_1(t, x), m_2(t, x)) \in \mathbb{R}^2$. En prenant $A(x) = w(x) \text{id}_2$, avec $w(x) \in \mathbb{R}$, nous pouvons nous ramener (sous la condition que $w(x) = 1$ sur les supports de ρ_0 et ρ_1) au modèle de transport sur une variété riemannienne caractérisée par w , telle que proposée dans [67]. Le transport optimal anisotrope peut alors être estimé en minimisant la fonctionnelle convexe $J_A(\rho, m) + \iota_C(\rho, m)$.

Pour pouvoir appliquer les algorithmes de splitting proximal à ce nouveau problème, nous devons calculer explicitement l'opérateur proximal pour ce nouveau coût. Ceci peut être aisément réalisé en calculant le point critique définissant $\text{Prox}_{\tau J_A}(\tilde{\rho}, \tilde{m})$ dans (5.1-9). La Proposition suivante montre que la fonctionnelle J_A peut toujours être écrite sous une forme explicite.

Proposition 5.2-6.

Pour tout $(\tilde{\rho}, \tilde{m}) \in \mathbb{R} \times \mathbb{R}^d$,

$$\text{Prox}_{\tau J_A}(\tilde{\rho}, \tilde{m}) = \begin{cases} (\rho^*, \mu_A(\rho^*)) & \text{si } \rho^* > 0, \\ (0, 0) & \text{sinon.} \end{cases}$$

où

$$\forall \rho \geq 0, \quad \mu_A(\rho) = (\rho \text{id}_2 + \gamma(A + A^T))^{-1} \rho \tilde{m}$$

et où ρ^* est la plus grande racine réelle de l'équation polynomiale du cinquième degré en ρ

$$P(\rho) = (\rho - \tilde{\rho}) - \gamma \tilde{m} (\rho \text{id}_2 + \gamma(A + A^T))^{-1} A (\rho \text{id}_2 + \gamma(A + A^T))^{-1} \tilde{m} = 0.$$

La preuve est une simple extension de celle de [67]. En pratique, la densité ρ^* est calculée par une méthode de Newton.

V.2.4 Illustration

Dans cette section, nous considérons toujours un espace-temps discrétisé de taille 32^3 et un algorithme GFB. Dans ce premier exemple, nous présentons un transport d'une gaussienne en deux dimensions dans un domaine anisotrope pour lequel la matrice A est définie pour $x = (x, y) \in [0, 1]^2$ par :

$$A_\lambda(x) = \begin{cases} \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix} & \text{si } x \leq 0.5, \\ \text{id}_2 & \text{sinon.} \end{cases} \quad (5.2-16)$$

La partie gauche du domaine ($x \leq 0.5$) induit alors un transport anisotrope lorsque $\lambda \neq 1$. En fait, en prenant $\lambda > 1$, cette pénalisation se fait dans la direction verticale (voir la sous-section

V.2.1) pour le transfert de masse, et de ce fait le transport déplacera la masse sur la partie droite du domaine de façon à minimiser le coût de transport. Les résultats, pour des valeurs croissantes de $\lambda \in \{1, 3, 20\}$ sont présentés dans la Figure V.2.

Lorsque $\lambda = 1$, cela correspond à un transport optimal classique et nous retrouvons bien la translation de la gaussienne. Lorsque la valeur de λ augmente, nous pouvons observer que le chemin optimal transfère de plus en plus de masse sur le coté droit du domaine. Cependant, notons qu'il n'y a pas à priori de division de masse, comme nous pourrions le penser en observant la Figure V.2 (seulement une dilatation) : les effets de division sont uniquement dus au fait que seul un nombre fini de niveaux de densité est représenté.

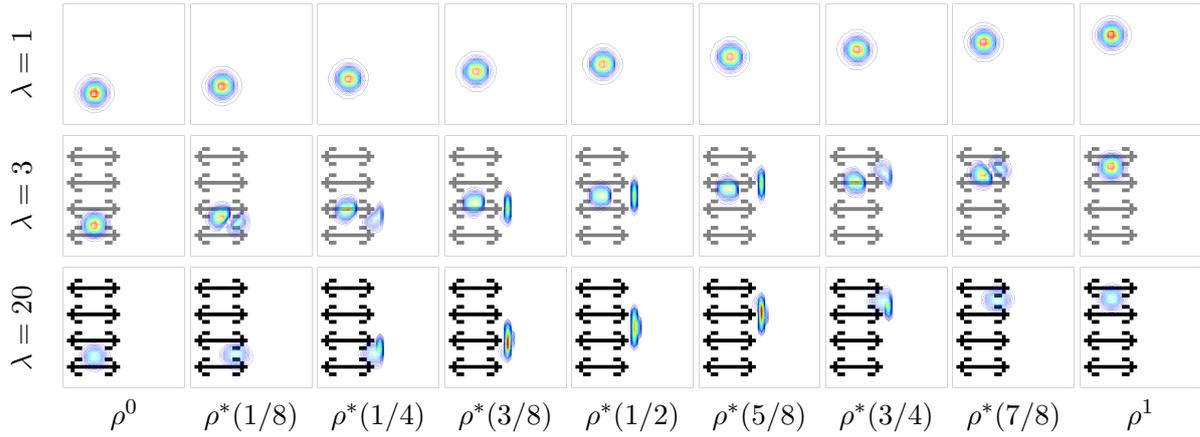


FIGURE V.2 – Illustration de la masse $\rho^*(t)$ estimée entre ρ^0 et ρ^1 , obtenue par un coût de transport généralisé défini par les matrices anisotropes A_λ introduites dans (5.2-16). Les flèches représentent les directions favorisées par l'anisotropie. L'intensité de leur couleur représentent la "force" de cette anisotropie.

Nous reprenons le même exemple en définissant cette fois A de la façon suivante :

$$A_{\lambda,\mu}(x) = \begin{cases} \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix} & \text{si } x \leq 0.5, \text{ et } \begin{bmatrix} 1 & 0 \\ 0 & \mu \end{bmatrix} & \text{sinon.} \end{cases} \quad (5.2-17)$$

Ici, la direction verticale est privilégiée dans la partie droite du domaine. Les résultats, pour des valeurs $\lambda = 20$ et des valeurs décroissantes de $\mu \in \{0.01, 0\}$ sont représentées dans la Figure V.3. Pour $\mu = 0$, nous pouvons observer qu'il se produit alors une "téléportation" verticale de la masse située dans la partie droite du domaine, du fait que le coût de transport y est nul.

Il est donc alors possible de définir une polarisation de l'espace à l'aide de matrices anisotropes pour modéliser les contraintes opérant sur un domaine. Ceci est également illustré par la Figure V.4, qui représente le chemin optimal entre une ligne horizontale et une ligne verticale dans une image carrée. Pour cette expérience, nous avons réalisé un ensemble complexe d'anisotropies, qui sont représentées sur la première image de chaque exemple.

Si les détails des contraintes sont disponibles dans les données du problème, nous pouvons alors construire une variété riemannienne spécifique de telle façon à ce que le modèle anisotrope soit en mesure de simuler des transports rigides ou bien à divergence nulle . Cette approche s'avère néanmoins insuffisante pour aborder des problèmes plus généraux encore, et nous proposons maintenant d'inclure directement des contraintes physiques dans le modèle de transport et non plus sur le domaine.

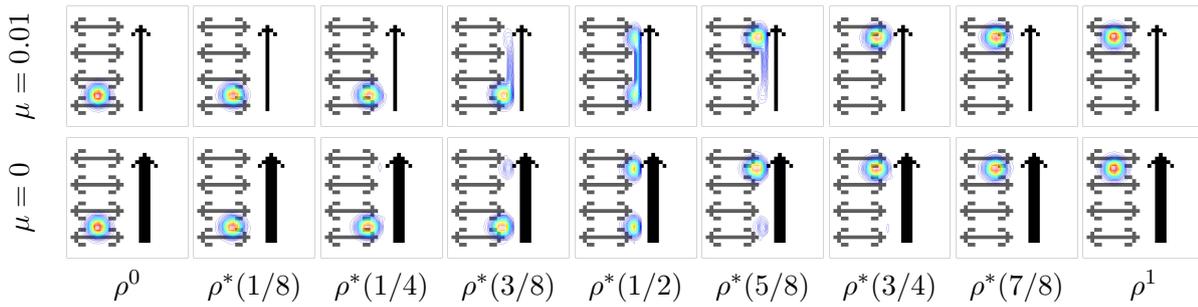


FIGURE V.3 – Illustration de la masse $\rho^*(t)$ estimée entre ρ^0 et ρ^1 , obtenue par un coût de transport généralisé défini par les matrices anisotropes $A_{\lambda,\mu}$ introduites dans (5.2-17), en prenant $\lambda = 20$. Les flèches représentent les directions favorisées par l'anisotropie. L'intensité de leur couleur représentent la force de cette anisotropie.

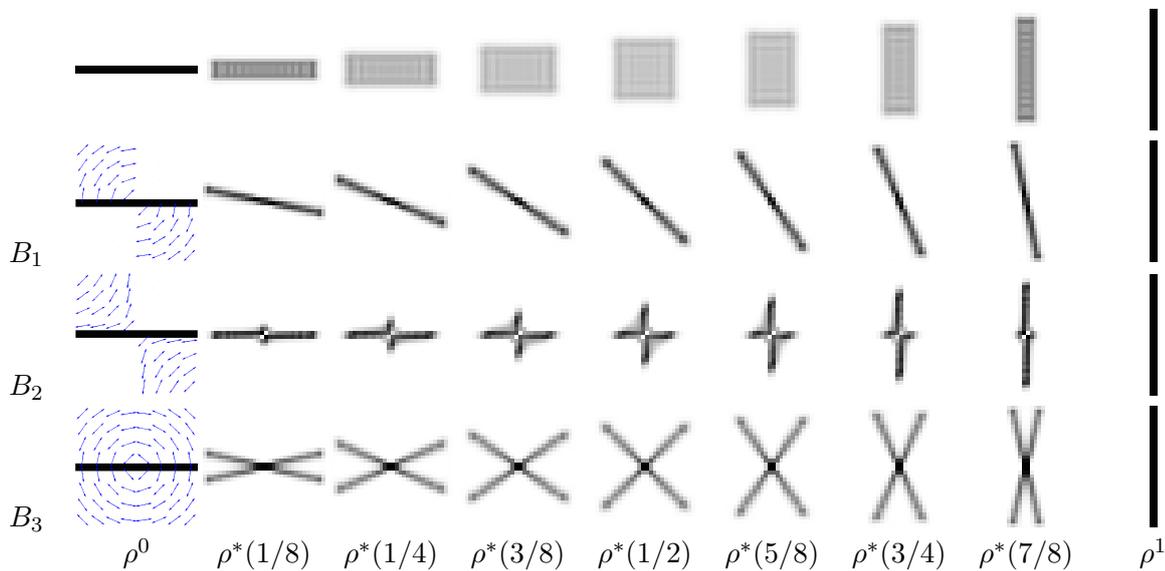


FIGURE V.4 – Illustration de la masse $\rho^*(t)$ estimée entre ρ^0 et ρ^1 , obtenue par des coûts de transport définis par trois différents domaines anisotropes, illustrés par les flèches bleues sur la première colonne. Le premier exemple est un transport isotrope classique.

V.3 Transport optimal avec contraintes physiques

Dans cette étude, nous nous sommes intéressés à l'incorporation de contraintes physiques au sein des modèles de transport optimal. Plus précisément, pour ce qui est des applications au traitement d'images, nous souhaiterions pouvoir considérer des transports incompressibles voire même rigides. De tels transports pourraient par exemple empêcher un objet contenu dans les données de départ d'être divisé le long des chemins de déplacements calculés. Ces transports pourraient être caractérisés par leur champ de vitesse, qui n'est plus une variable du problème (5.1-6) (remplacé par la quantité de mouvement), et qu'il serait alors nécessaire de réintroduire.

V.3.1 Couplage non convexe

Pour obtenir un couplage entre les variables (ρ, m) et v , l'idée qui vient naturellement est de considérer un terme $\iota_D(\rho, m, v) = 0$ de projection orthogonale sur l'ensemble $D = \{m = \rho v\}$. Mais, comme l'ensemble D est alors non convexe et l'opérateur ι_D non régulier, nous choisirons plutôt d'utiliser un couplage différentiable :

$$K(\rho, m, v) = \frac{1}{2} \int_{\Omega} \int_0^1 \|m - \rho v\|^2 dx dt.$$

Dans cette étude, nous nous intéresserons au cas spécifique de transports soumis à des contraintes de vitesse, lesquelles dépendront de l'application souhaitée. De ce fait, nous considérerons donc un nouveau terme fonctionnel, lequel pourra être par exemple défini comme il suit :

- Contrainte de divergence nulle : $R(v) = \iota_{C^v}(v)$, où $C^v = \{\operatorname{div}_x v(t, \cdot) = 0, \langle v(t, \cdot), \vec{n} \rangle = 0, \forall t \in [0; 1]\}$.
- Pénalisation incompressible : $R(v) = \frac{1}{2} \int_0^1 \|\operatorname{div}_x(v(t, \cdot))\|_{L^2(\Omega)}^2 dt$.
- Pénalisation rigide : $R(v) = \frac{1}{2} \int_0^1 \|(\nabla_x v(t, \cdot) + (\nabla_x v(t, \cdot))^T)/2\|_{L^2(\Omega)}^2 dt$.
- Pénalisation en translation : $R(v) = \frac{1}{2} \int_0^1 \|\nabla_x v(t, \cdot)\|_{L^2(\Omega)}^2 dt$.

Par la dénomination "pénalisation en translation", nous entendons une pénalisation sur les champs de vitesse déviant des translations. Ces termes sont convexes, et les trois derniers sont de plus différentiables. Notons que les auteurs de [53] ont proposé en parallèle de nos travaux, de rajouter directement une régularisation du champ de vitesses au problème initial de mécanique des fluides sur le couple (ρ, v) présenté dans (5.1-2), sans considérer la convexification du problème venant de l'introduction de m . Ceci rend néanmoins difficile d'imposer numériquement la contrainte de préservation de la masse (5.1-3), puisqu'il faut pouvoir projeter le couple (ρ, v) sur un ensemble de contraintes non-linéaires.

V.3.2 Existence d'un minimiseur

Dans cette section, nous prouvons l'existence d'au moins un minimiseur pour un problème de transport optimal avec pénalisation du gradient de vitesse (voir pénalisation en translation) et le terme de couplage correspondant. Nous supposons que ρ_0 et ρ_1 appartiennent à $L^\infty(\Omega)$, et posons alors $M = \max(\|\rho_0\|_\infty, \|\rho_1\|_\infty)$. Posons $Q = (0, 1) \times \Omega$, et notons $BL(\Omega)$ l'espace de Beppo-Levi sur $L^2(\Omega)$ [32], défini comme

$$BL(\Omega) = \{u \in \mathcal{D}'(\Omega), \nabla u \in L^2(\Omega)\}.$$

Notre ensemble de contraintes sera, quant à lui, défini par :

$$C_M = \left\{ (\rho, m, v) \in L^\infty(Q) \times L^2(Q)^d \times L^2(0, 1; BL(\Omega))^d, \quad \|\rho\|_{L^\infty(Q)} \leq M, \right. \\ \left. \forall \psi \in C^\infty(Q), \int_0^1 \int_{\Omega} \rho \partial_t \psi + m \cdot \nabla \psi dx dt + \int_{\Omega} \psi(0, x) \rho_0(x) - \psi(1, x) \rho_1(x) dx = 0 \right\}.$$

Notons que cet ensemble est non vide du fait que $((1-t)\rho_0 + t\rho_1, m, 0) \in C_M$ pour une quantité de mouvement m constante en temps telle que $\operatorname{div} m = \rho_1 - \rho_0$ et $\langle m \cdot \vec{n} \rangle = 0$ sur $\partial\Omega \times (0, 1)$ (l'existence d'un tel m est assurée par le fait que $\int_{\Omega} \rho_0 dx = \int_{\Omega} \rho_1 dx$, voir [38], page 24).

Remarque 5.3-1. Remarquons que si $\rho \in C_M$, alors $\int_{\Omega} \rho(t, x) dx = 1$ pour presque tout $t \in (0, 1)$. Ceci peut être facilement obtenu en prenant $\psi(t, x) = t$ dans la définition de C_M .

Proposition 5.3-7.

Le problème

$$\inf_{(\rho, m, v) \in C_M} \int_{\Omega} \int_0^1 J(\rho, m) + \frac{1}{2} \|m - \rho v\|^2 + \frac{1}{2} \|\nabla_x v\|^2 dx dt$$

possède au moins un minimiseur.

Preuve : Soit $(\rho_n, m_n, v_n) \in C_M$ une suite minimisante. Comme (ρ_n) est uniformément bornée par M , nous avons $J(m_n, \rho_n) = \frac{m_n^2}{2\rho_n} \epsilon \frac{1}{2M} m_n^2$. Nous avons alors l'estimation a priori suivante : il existe $C > 0$ telle que pour tout $n \in \mathbb{N}$,

$$\|\rho_n\|_{L^\infty(Q)} \leq M, \quad \|m_n\|_{L^2(Q)} \leq C, \quad \|\nabla_x v_n\|_{L^2(Q)} \leq C, \quad \|m_n - \rho_n v_n\|_{L^2(Q)} \leq C. \quad (5.3-18)$$

La première et la troisième estimation impliquent :

$$\exists C > 0, \quad \|\rho_n v_n\|_{L^2(Q)} \leq C, \quad (5.3-19)$$

et du fait que (ρ_n, m_n) vérifie $\partial_t \rho_n + \operatorname{div}_x m_n = 0$ dans $\mathcal{D}'(Q)$, nous obtenons également

$$\exists C > 0, \quad \|\partial_t \rho_n\|_{L^2(0,1; H^{-1}(\Omega))} \leq C. \quad (5.3-20)$$

Ainsi, quitte à extraire une sous-suite, nous pouvons supposer qu'il existe $(\rho, m, v) \in L^\infty(Q) \times L^2(Q)^d \times L^2(0,1; BL(\Omega))^d$ tel que

$$\rho_n \overset{*}{\rightharpoonup} \rho \text{ dans } L^\infty(Q), \quad m_n \rightharpoonup m \text{ dans } L^2(Q)^d, \quad \nabla_x v_n \rightharpoonup \nabla_x v \text{ dans } L^2(Q)^{d \times d}. \quad (5.3-21)$$

Notons tout d'abord que C_M est fermé pour la topologie liée à ces convergences, et donc la limite faible (ρ, m, v) appartient à C_M .

La fonctionnelle intégrale est convexe et semi-continue inférieurement pour la topologie faible de $L^2(Q) \times L^2(Q)$ (voir [4], p. 138). Le terme de régularisation en v est la semi-norme associée à $L^2(0,1; BL(\Omega))$, et est également semi-continue inférieurement pour ce qui est de la convergence de $(\nabla_x v_n)$ dans $L^2(Q)^{d \times d}$. Le terme de couplage étant également une norme $L^2(Q)$, il suffit pour conclure de prouver la convergence faible de $(\rho_n v_n)$ vers ρv dans $L^2(Q)$. Nous ferons appel à un Lemme de compacité par compensation, en utilisant le fait que la dérivée temporelle de ρ_n et que les dérivées spatiales de v_n soient bornées, comme dans [33]. L'énoncé qui suit est un Corollaire du Lemme de [61] :

Lemme 5.3-8.

Supposons (a_n) convergeant faiblement vers a dans $L^2(0,1; H^1(\Omega))$, et (b_n) convergeant faiblement b dans $L^2(Q)$. Si $(\partial_t b_n)$ est bornée dans $L^2(0,1; H^{-1}(\Omega))$, alors quitte à en extraire une sous-suite, $(a_n b_n)$ converge faiblement vers ab dans $\mathcal{M}(Q)$ (c'est à dire le dual de $\mathcal{C}_c^0(Q)$).

Le problème est qu'ici nous n'avons pas de contrôle sur v_n dans le terme de régularisation, mais seulement sur $\nabla_x v_n$: la suite (v_n) évolue par conséquent dans l'espace $BL(\Omega)$ au lieu de l'espace $H^1(\Omega)$ requis.

Nous espérons pouvoir obtenir ce contrôle grâce au terme de couplage, ce qui n'est pas trivial lorsque ρ_n est seulement positif, et non strictement positif.

Par [32], p. 312, du fait que Q est étoilé, la convergence faible de (v_n) vers v dans $L^2(0,1; BL(\Omega))$

implique la convergence faible de $(v_n + c_n)$ vers v dans $L^2(Q)$ pour des suites constantes en espace $(c_n(t))$.

Comme ρ_n est bornée dans $L^\infty(Q)$, la suite $\rho_n(v_n + c_n)$ est en fait bornée dans $L^2(Q)$, et grâce à (5.3-19), nous déduisons que $\rho_n c_n$ est bornée dans $L^2(Q)$. Par définition,

$$\|\rho_n c_n\|_{L^2(Q)} = \sup_{\psi \in L^2(Q), \|\psi\|_{L^2(Q)}=1} \int_0^1 \int_\Omega \rho_n c_n \psi dx dt,$$

ce qui, en restreignant le supremum aux fonctions constantes en espace $\psi \in L^2(0,1)$ donne

$$\|\rho_n c_n\|_{L^2(Q)} \sup_{\psi \in L^2(0,1), \|\psi\|_{L^2(0,1)}^2 = \frac{1}{|\Omega|}} \int_0^1 \int_\Omega \rho_n c_n \psi dx dt.$$

De plus, du fait que c_n est également constante en espace, et comme $\int_\Omega \rho_n dx = 1$ (voir la Remarque 5.3-1), nous obtenons

$$\|\rho_n c_n\|_{L^2(Q)} \sup_{\psi \in L^2(0,1), \|\psi\|_{L^2(0,1)}^2 = \frac{1}{|\Omega|}} \int_0^1 c_n \psi dt = \frac{1}{\sqrt{|\Omega|}} \|c_n\|_{L^2(0,1)}$$

Cela signifie que la suite (c_n) est bornée dans $L^2(0,1)$ et également dans $L^2(Q)$ en tant que suite (constante en x) de cet espace. Cela implique également que la suite (v_n) est bornée dans $L^2(Q)$, et donc, quitte à devoir en extraire une sous-suite, converge faiblement vers une fonction w . Nous déduisons alors nécessairement de (5.3-21) que $w = v$ du fait que l'opérateur ∇ est faiblement continu. Ainsi, par les estimations (5.3-18) et (5.3-20), nous sommes en mesure d'appliquer le Lemme 5.3-8 énoncé ci-dessus à v_n et ρ_n , et ainsi conclure :

$$\rho_n v_n \rightarrow \rho v \text{ dans } \mathcal{M}(Q).$$

Or, $\rho_n v_n$ est bornée dans $L^2(Q)$, et admet en conséquence une sous-suite convergeant dans cet espace. Ainsi, quitte à devoir extraire de nouveau une sous-suite, nous obtenons

$$\rho_n v_n \rightharpoonup \rho v \text{ dans } L^2(Q)$$

(convergence faible dans $L^2(Q)$). □

Remarque 5.3-2. En ce qui concerne les autres pénalisations, nous verrons que si les algorithmes numériques se comportent, comme prévu du point de vue théorique, la situation semble un peu plus complexe. Le cas de la pénalisation rigide, pourtant conceptuellement très proche de celle analysée ci-dessus du fait des inégalités de Korn, nécessite cependant une étude plus approfondie du fait que la simplification obtenue ici grâce aux constantes c_n ne peut se généraliser directement au cas rigide. Une alternative simple pourrait bien sûr être d'ajouter une norme $L^2(Q)$ sur v multiplié par un petit coefficient strictement positif.

Le cas de la pénalisation incompressible nécessiterait quant à lui une adaptation du Lemme de compacité.

Les opérations nécessaires à la résolution de ce problème étant similaire à celles du problèmes précédent (pénalisation du domaine de transport), en particulier en ce qui concerne le calcul d'opérateurs de divergence, nous avons utilisé une discrétisation similaire pour traiter le problème présent : à savoir que l'opérateur de divergence sera de nouveau calculé sur une grille décalée. De même que précédemment, nous renvoyons à l'article [46] pour les détails de cette discrétisation. Notons tout de même que nous faisons de nouveau appel à un opérateur de recentrage, noté \mathcal{I} , linéaire et bijectif depuis la grille décalée sur la grille centrée.

V.3.3 Optimisation

Le problème de transport optimal généralisé auquel nous nous intéressons ici est le problème suivant :

$$(\rho^*, m^*, v^*) = \underset{\rho, m, v}{\operatorname{argmin}} F(\rho, m, v) \quad (5.3-22)$$

où

$$F(\rho, m, v) = \mathcal{J}(\mathcal{I}(\rho, m)) + \iota_C(\rho, m) + \lambda K(\mathcal{I}(\rho, m), \mathcal{I}^v(v)) + \alpha R(v),$$

pour $\lambda, \alpha \geq 0$ pondérant respectivement le couplage des variables et le terme de régularisation de la vitesse. Le problème (5.3-22) n'est pas convexe en (ρ, m, v) , mais est cependant séparément convexe en (ρ, m) et en v . Notons que le terme de couplage est différentiable et que les termes non réguliers sont séparables. Du fait de ces propriétés et d'après [75], nous pouvons donc opérer une méthode de descente ds coordonnées par blocs et minimiser alternativement chaque problème convexe, et ainsi approcher un point critique du problème (5.3-22). L'algorithme d'optimisation s'écrit :

$$\begin{cases} (\rho^{k+1}, m^{k+1}) = \underset{(\rho, m)}{\operatorname{argmin}} F(\rho, m, v^k) \\ v^{k+1} = \underset{v}{\operatorname{argmin}} F(\rho^{k+1}, m^{k+1}, v) \end{cases} \quad (5.3-23)$$

Pour pouvoir être en mesure d'opérer sur les différentes grilles numériques, nous définissons de nouveau une variable $x = (\bar{\rho}, \bar{m})$ sur la grille régulière et une variable y sur la grille décalée, pour le couple (ρ, m) . Nous faisons de même pour la vitesse et notons w la variable sur la grille centrée, et v la vitesse sur la grille décalée.

Au final, nous souhaitons donc résoudre le problème :

$$(x^*, y^*, v^*, w^*) = \underset{x, y, v, w}{\operatorname{argmin}} \tilde{F}(x, y, v, w) \quad (5.3-24)$$

où

$$\tilde{F}(x, y, v, w) = \mathcal{J}(x) + \iota_C(y) + \iota_S(x, y) + \lambda K(x, w) + R(v) + \iota_{S^v}(v, w),$$

et l'ensemble convexe S^v est défini en fonction des deux variables de vitesse par $S^v(v, w) = \{\mathcal{I}^v(v) = w\}$. Il est donc possible d'estimer un point critique de (5.3-24) par l'algorithme suivant :

$$\begin{cases} (x^{k+1}, y^{k+1}) = \underset{(x, y)}{\operatorname{argmin}} \tilde{F}(x, y, w^k), \\ (v^{k+1}, w^{k+1}) = \underset{(v, w)}{\operatorname{argmin}} \tilde{F}(x^{k+1}, v, w), \end{cases} \quad (5.3-25)$$

lequel effectue donc une alternance de deux problèmes d'optimisation. Nous détaillons maintenant la résolution de ces deux problèmes.

Optimisation de (ρ, m) .

La première étape de (5.3-25) s'écrit :

$$(x^*, y^*) = \underset{x, y}{\operatorname{argmin}} \mathcal{J}(x) + \iota_C(y) + \iota_S(x, y) + \lambda K(x, w^k). \quad (5.3-26)$$

Nous pouvons de nouveau faire appel à l'algorithme GFB 1 pour résoudre ce problème, en prenant $F(x, y) = K(x, w^k)$, $G_1(x, y) = \mathcal{J}(x) + \iota_C(y)$ et $G_2(x, y) = \iota_S(x, y)$. La constante de Lipschitz de ∇F est donnée par $L = 2\lambda \max(1, \|w^k\|_\infty^2)$. Par conséquent, lorsque le couplage entre les variables est renforcé par une valeur élevée de λ , cela diminue la valeur du pas de temps de l'algorithme d'optimisation τ , lequel est inversement proportionnel à L .

Optimisation de v .

Nous allons maintenant détailler la manière de traiter les différents modèles de régularisation de la vitesse que nous avons évoqué.

Contrainte de divergence nulle

Lorsque la vitesse v est contrainte à l'incompressibilité, la seconde étape de (5.3-25) s'écrit :

$$(v^*, w^*) = \underset{v, w}{\operatorname{argmin}} \lambda K(x^{k+1}, w) + \iota_{C^v}(v) + \iota_{S^v}(w, v). \quad (5.3-27)$$

L'algorithme GFB 1 peut alors être appliqué pour résoudre ce problème en prenant les fonctionnelles $F(v, w) = \lambda K(x^{k+1}, w)$, $G_1(v, w) = \iota_{C^v}(v)$ et $G_2(v, w) = \iota_{S^v}(w, v)$. La constante de Lipschitz de ∇F est alors donnée par $L = \lambda \|\bar{\rho}^{k+1}\|_\infty^2$.

Le calcul de $\operatorname{Prox}_{\iota_{C^v}}(\tilde{v})$ correspond à la projection orthogonale de $\tilde{v}(t, \cdot, \cdot)$, sur un champ de vitesse à divergence nulle $2D$, pour tout $t \in [0, 1]$. Pour chaque instant t , nous extrayons tout d'abord les valeurs de bords de la vitesse sur la grille décalée grâce à l'opérateur linéaire b^v , définie pour $d = 2$, $\Omega = [0; 1]^2$ et $\tilde{v} = (\tilde{v}^1, \tilde{v}^2)$ par :

$$b^v(\tilde{v}(t, \cdot, \cdot)) = (\tilde{v}^1(t, 0, \cdot), \tilde{v}^1(t, 1, \cdot), \tilde{v}^2(t, \cdot, 0), \tilde{v}^2(t, \cdot, 1))$$

En imposant des conditions de Neumann homogènes aux bords par $b_0^v = (0, 0, 0, 0)$, l'ensemble affine C^v peut être réécrit

$$C^v = \{v, B_t v(t, \cdot, \cdot) = z, \forall t \in [0; 1]\}$$

où

$$B_t v(t, \cdot, \cdot) = (\operatorname{div}_x(v(t, \cdot, \cdot)), b^v(v(t, \cdot, \cdot))), \text{ et } z = (0, b_0^v).$$

Le projeté v de \tilde{v} peut également être calculé en résolvant un système linéaire pour chaque $t \in [0; 1]$:

$$v(t, \cdot, \cdot) = \operatorname{Proj}_{C^v}(\tilde{v}(t, \cdot, \cdot)) = (\operatorname{id} - B_t^* \Delta^{-1} B_t) \tilde{v}(t, \cdot, \cdot) + B_t^* \Delta_t^{-1} z$$

avec l'opérateur $\Delta_t^{-1} = (B_t B_t^*)^{-1}$. Pour ce qui est de $\operatorname{Prox}_{\tau_{\iota_{S^v}}}$, nous obtenons une expression similaire à celle de $\operatorname{Prox}_{\tau_{\iota_S}}$, définie en (5.1-10), en remplaçant \mathcal{I} par \mathcal{I}^v .

Pénalisation

Si nous souhaitons pénaliser une certaine norme de la vitesse par un terme $\alpha R(v)$, pour R propre, convexe et différentiable avec gradient lipschitzien, nous obtenons le problème suivant :

$$(v^*, w^*) = \underset{v, w}{\operatorname{argmin}} \lambda K(x^{k+1}, w) + \alpha R(v) + \iota_{S^v}(w, v), \quad (5.3-28)$$

lequel peut être résolu grâce à l'algorithme FB 2, en prenant $G(v, w) = \iota_{S^v}(w, v)$ et $F(v, w) = \lambda K(x^{k+1}, w) + \alpha R(v)$. La constante de Lipschitz de ∇F est donnée par $L = \lambda \|\bar{\rho}^{k+1}\|_\infty^2 + \alpha L_R$, où L_R désigne la constante de Lipschitz de ∇R .

Pour une pénalisation incompressible $R(v) = \frac{1}{2} \sum_{t=0}^1 \|\operatorname{div}_x(v(t, \cdot, \cdot))\|_{L^2(\Omega)}^2$, de même que pour la pénalisation rigide ou pour la pénalisation en translation $R(v) = \frac{1}{2} \sum_{t=0}^1 \|(\nabla_x v(t, \cdot, \cdot) + (\nabla_x v(t, \cdot, \cdot))^T)/2\|_{L^2(\Omega)}^2$, nous avons $L_R = 4\sqrt{2}$ pour notre gradient discrétisé et nos opérateurs de divergence calculés sur la grille régulière et sur la grille décalée, pour des coefficients dans $[1, -1]$.

V.4 Expériences et tests

Nous allons maintenant illustrer l'influence de la régularisation et de la pénalisation du champ de vitesse par différents exemples. Dans toutes nos simulations, nous choisirons $\lambda = 10^3$ comme terme de couplage et $\alpha = 2.10^3$ pour la régularisation des vitesses. Le pas de temps de l'algorithme GFB sera donc choisi à $\tau = 10^{-4}$. Dans la pratique, selon la dimension du problème, nous réaliserons jusqu'à 1000 itérations externes de l'algorithme alterné (5.3-25) et 10 itérations internes pour chacun des problèmes (5.3-26) et (5.3-28) lorsque sera impliquée une pénalisation de la vitesse. Notons qu'une seule étape est requise pour le modèle (5.3-27), c'est à dire pour la projection de v sur l'ensemble de divergence nulle C^v . De tels paramètres sont suffisant pour obtenir une convergence du coût de transport et une cohérence des chemins calculés.

V.4.1 Tests synthétiques

Avant toute chose, nous comparons dans la figure V.5 les résultats obtenus avec les pénalisations incompressibles, rigides et de translation sur les exemples de gaussiennes. De telles contraintes physiques empêchent la division de la masse le long des chemins calculés. Avec les contraintes de divergence nulle, les gaussiennes sont déformées, mais on peut cependant voir que la longueur des lignes de niveau des densités sont préservées le long du chemin. Il est important de souligner également qu'à la fois la pénalisation rigide et la pénalisation en translation conservent la forme exacte des deux gaussiennes au cours du temps. De plus, nous pouvons voir sur la figure V.6 représentant les deux chemins calculés, que la pénalisation rigide opère véritablement une rotation et non pas une translation, de telle façon que le transport optimal ne se compose plus de trajectoires en lignes droites.

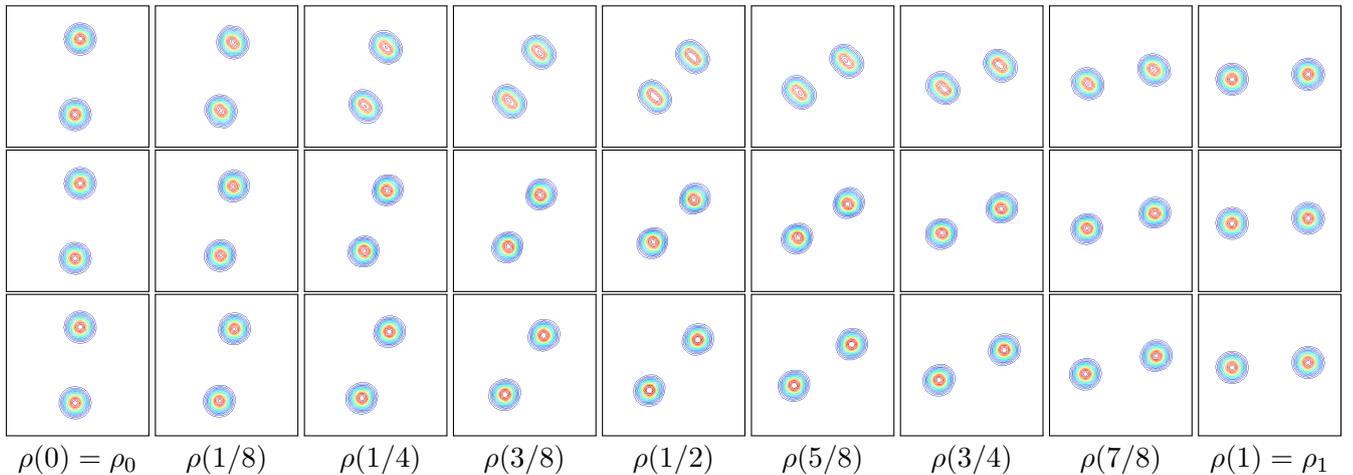


FIGURE V.5 – Simulation avec deux gaussiennes avec pénalisation et initialisation nulle. Représentation des isoclines de la densité $\rho(t)$ le long du chemin optimal calculé avec les différents modèles de pénalisation. La ligne du haut a été réalisée via une pénalisation incompressible, la seconde via une pénalisation en translation (par une minimisation de la norme L^2 de chaque composante du champ de vitesse), et la dernière via une pénalisation rigide.

Dans l'exemple de la Figure V.7 représentant une barre rotative, la pénalisation rigide (dernière ligne) opère une quasi-rotation, et préserve donc mieux les contraintes physiques qu'un transport optimal pur (première ligne). Comme prévu, nous pouvons également observer sur la Figure V.8 que les longueurs des lignes de niveau de la densité calculée sont préservées pour les approches

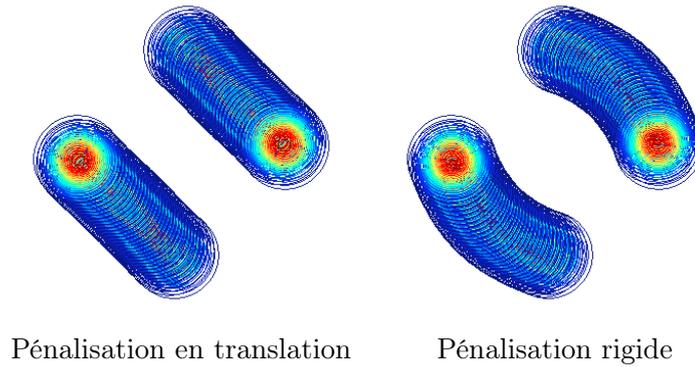


FIGURE V.6 – Simulation avec deux gaussiennes. Représentation de la trajectoire calculée par les modèles de pénalisation en translation (sur la gauche) et de pénalisation rigide (sur la droite). La pénalisation rigide opère ici une véritable rotation.

par pénalisation incompressible ou rigide. Nous renvoyons le lecteur à [15] pour d'autres exemples impliquant de telles pénalisations.

Notons que sur les exemples synthétiques présentés, nous considérons une grille de discrétisation spatio-temporelle 31^3 points. Pour de telles dimensions, seulement 100 itérations externes de l'algorithme de descente de coordonnées par blocs (5.3-25) ont été nécessaires.

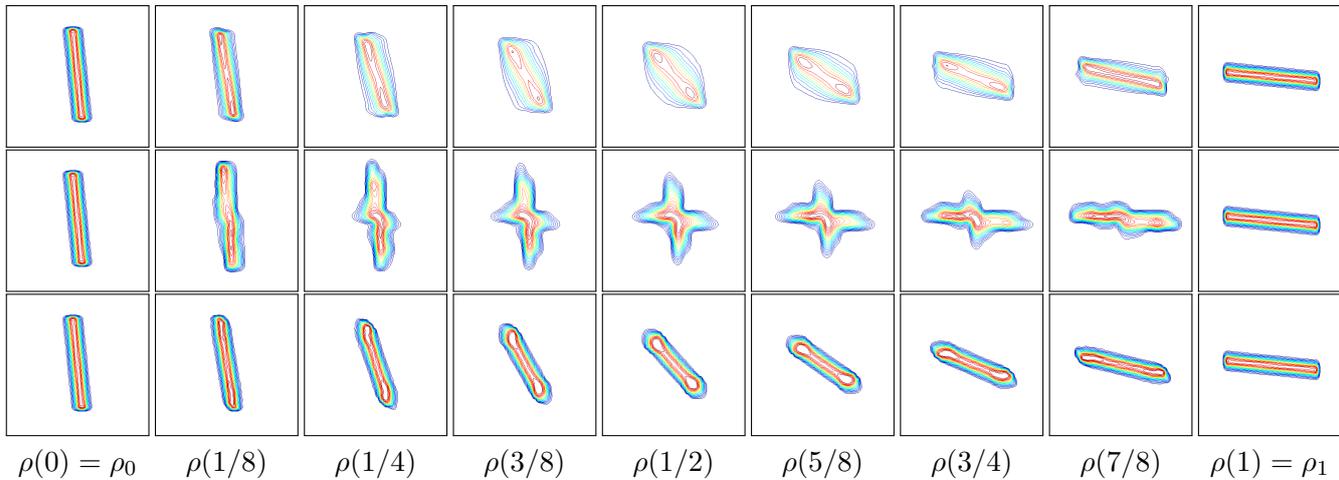


FIGURE V.7 – Barre rotative. Représentation des isoclines de la densité $\rho(t)$ le long du chemin optimal calculé par les différentes approches. La première ligne est le transport optimal de [6], la seconde est l'approche proposée à l'aide d'une pénalisation incompressible, et la dernière à l'aide d'une pénalisation rigide laquelle conserve la nature de l'objet à transporter.

V.4.2 Interpolation d'images en océanographie

Pour pouvoir étudier l'état des océans, des images instantanées sont produites par des codes numériques tels que le modèle de circulation océanique NEMO. Les processus de synthèse d'image à partir des variables numériques physiques dont cependant chronophages, et les océanographes souhaiteraient parfois n'avoir à produire que peu d'images et pouvoir ensuite réaliser des interpolations temporelles entre ces images. Ceci pourrait être d'un intérêt particulier pour visualiser

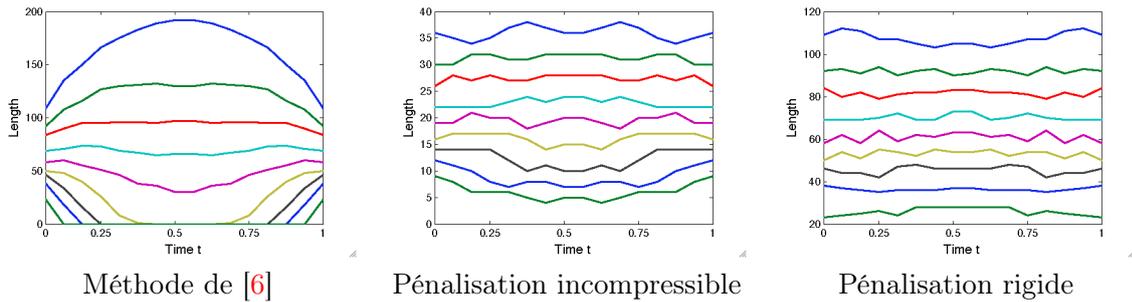


FIGURE V.8 – Barre rotative. Évolution de la longueur des lignes de niveau les plus élevées de la densité calculée au cours du temps $t : |\rho(t, x) > i/10|$, pour $i = 1 \dots 9$. Le graphique de gauche est un transport optimal classique de [6], celui du milieu est l’approche proposée à l’aide d’une pénalisation incompressible et celui de droite correspond à l’approche proposée à l’aide d’une pénalisation rigide. Pénaliser la norme de la vitesse tend donc à préserver les lignes de niveau le long du chemin calculé.

et diffuser des animations des structures dynamiques de l’océan. Le problème principal vient souvent des côtes qui apparaissent dans beaucoup des zones intéressantes, comme illustré à la Figure V.9. Celles-ci rendent alors inapplicables les techniques classiques d’interpolation d’images utilisant le flot optiques ou bien les estimations de difféomorphismes, qui ne peuvent en général pas traiter des domaines si complexes.

Par notre formulation anisotrope, nous pouvons représenter le domaine de l’océan dans l’image par une variété riemannienne, en prenant $A(t, x) = w(x) \text{id}_2$. La variable $w(x)$ décrit alors la variété : celle-ci prend la valeur 1 au niveau de l’océan et au contraire une valeur très élevée au niveau de la terre dans le but de restreindre le transport à l’océan seul (on procède à une pénalisation de la terre).

Les chemins optimaux successifs entre 10 paires d’images de Hatuer de Surface de l’Eau (connu sous l’acronyme de SSH pour Sea Surface Height) de taille 843×516 produites par le modèle NEMO ont ainsi pu être calculés pour une discrétisation en temps de $P = 9$ pas (les données complètes peuvent être consultées [ici](#) et [là](#)). Les couleurs claires correspondent aux zones de plus hautes hauteur d’eau. Cette suite d’images représente le courant des Aiguilles et la création de tourbillons à la Pointe du Cap de Bonne-Espérance.

Comme cela est illustré sur la Figure V.9, laquelle montre le chemin calculé entre plusieurs images consécutives, en ajoutant notre Proposition de pénalisation "divergence nulle", nous pouvons induire des contraintes de rotation dans l’estimation du transport, et ainsi parfaitement simuler la création de tourbillons. Dans cette simulation, nous considérons 1000 itérations externes du processus alterné (5.3-25).

V.5 Conclusion du chapitre

Dans ce chapitre, nous avons étudié, d’une façon plus formelle que dans les chapitres précédents, des modèles de transport optimal généralisés et adaptés aux modèles d’imagerie multi-physiques. Cela est d’un intérêt particulier pour les problèmes d’interpolation d’images, domaine

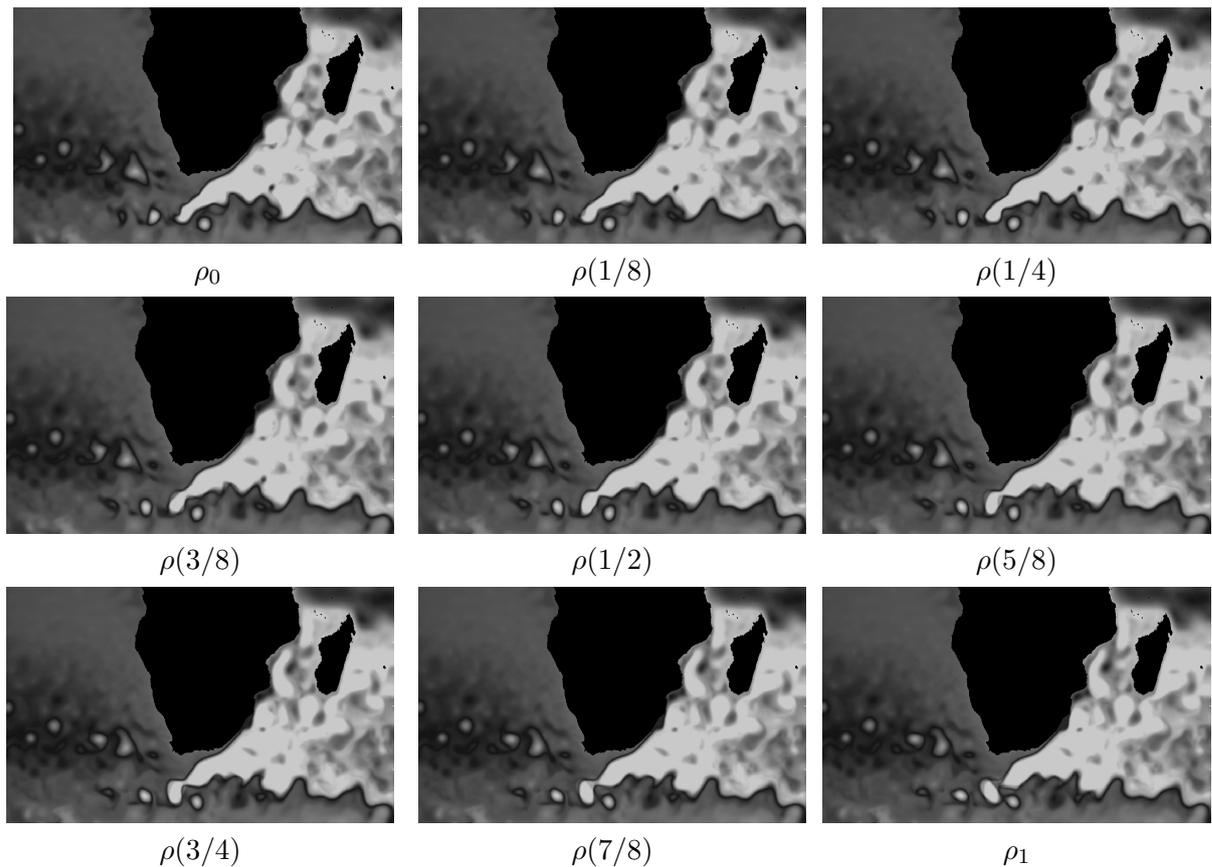


FIGURE V.9 – Interpolation entre images de hauteur d'eau en océanographie

dans lequel les résultats obtenus par simple minimisation d'une énergie cinétique sous certaines contraintes ne préserve pas les caractéristiques des images le long des chemins d'interpolation, ce qui ne correspond en général pas à la réalité physique. Nous avons ensuite proposé des méthodes de splitting proximal adéquates pour résoudre ces nouveaux problèmes d'optimisation. Ces généralisations ne sont pas limitées aux expressions physiques que nous avons considérées ici. En effet, d'autres termes physiques auraient pu être considérés en prenant compte un environnement plus complexe.

Des résultats prometteurs ont été obtenus sur des images océanographiques de haute résolution. Nous souhaiterions en perspective de ces résultats, pouvoir collaborer dans de futurs travaux avec des océanographes, et ainsi tenter de prendre en compte des contraintes physiques plus précises.

D'un point de vue optimisation, nous nous sommes également intéressé à la minimisation de fonctionnelles non convexes et non régulières. Contrairement à la méthode de descente de coordonnées par blocs que nous avons considérée (5.3-23), pouvoir définir un algorithme sans boucle interne (sur (ρ, m) et v) serait d'un intérêt certain pour accélérer le calcul des minima de nos problèmes non convexes. Des avancées récentes sur les méthodes de splitting proximal ont déjà été réalisées pour des problèmes proches [2], mais ils sont cependant encore limités à des fonctionnelles plus simples et ne peuvent en l'état traiter notre formulation de transport optimal généralisé.

Pour finir, en ce qui concerne le transport optimal anisotrope, nous souhaiterions pouvoir étudier l'existence d'un minimiseur dans le cas de variations anisotropes de façon plus rigoureuse que

nous l'avons fait dans ce chapitre. Nous souhaiterions également pouvoir prendre en compte des variations anisotropes dépendant du temps : le défi pour un tel problème est qu'il n'existe à notre connaissance aucune correspondance avec un problème de Monge.

Chapitre VI

Conclusion et perspectives

Le point de départ de nos travaux s'est trouvé être une étude de consistance de l'algorithme de lagrangien augmenté de Benamou-Brenier. Nous avons montré, dans le chapitre II, la convergence faible- L^2 de l'algorithme de Benamou-Brenier vers un point selle du lagrangien \mathbf{L} (voir page 53), lequel modélise la formulation dynamique du problème de transport optimal formulée dans [6]. Cette preuve de convergence étant basée sur une reformulation de l'algorithme en tant que suite itérative d'un opérateur non-expansif \mathbf{B} (1-lipschitzien), dont les points fixes sont équivalents aux points selles de \mathbf{L} . Cette formulation nous aura permis de plus de pouvoir exploiter la littérature associée à la théorie des opérateurs non-expansifs, via laquelle nous avons alors pu proposer une version relaxée de l'algorithme de Benamou-Brenier, à convergence forte, vers un point selle de \mathbf{L} . Les expérimentations numériques n'auront cependant pas révélé une réelle amélioration liée aux versions relaxées, en ce qui concerne la vitesse de convergence et la précision des calculs, par rapport à l'algorithme d'origine.

Cette convergence de l'algorithme et de ses versions relaxées restaient tout de même conditionnée à l'existence d'un point selle pour le lagrangien \mathbf{L} (et donc d'un point fixe de l'opérateur \mathbf{B}). Nous avons alors entrepris, dans les chapitres III et IV, de prouver l'existence d'un tel point selle, tout en montrant également l'unicité de l'évolution de la densité et de la quantité de mouvement résultant d'un tel transport. Nous avons voulu prouver tout cela dans les conditions les plus générales possibles, en particulier dans les cas où les densités de départ et d'arrivée ρ_0 et ρ_1 s'annulent en certains endroits du domaine de transport (l'existence d'un point selle dans le cas de densités strictement positives ayant déjà été traitée dans [40]).

De telles conditions impliquent en particulier de pouvoir prendre en compte le cas où le nombre de composantes connexes des supports des densités ρ_0 et ρ_1 ne sont pas les mêmes. De tels cas révèlent en général des plans de transport optimaux non réguliers, c'est pourquoi une grande partie de nos travaux auront eu pour objet une étude approfondie de la régularité et du comportement d'un champ de vitesse associé à de tels plans de transport. Ces travaux d'étude du champ de vitesse associé à un plan de transport optimal auront été le principal objet du chapitre III (voir en particulier la Proposition-résumé 3.4-38 p. 123 à la fin de ce chapitre). Le chapitre IV aura quant à lui été consacré à exploiter les résultats du chapitre III pour finalement prouver l'existence et l'unicité attendues. Ce chapitre aura également été l'occasion, en particulier dans sa dernière section (section IV.3) de tenter de caractériser de façon générale les propriétés d'un champ de vitesse associé à un transport optimal dans l'espace L^2 (voir le Théorème 4.3-9 p. 159).

Pour finir nous avons pu explorer, dans le chapitre V, différentes approches relatives à l'introduction de contraintes physiques dans la formulation dynamique de [6], telles que l'introduction d'une anisotropie du domaine de transport ou bien de différentes méthodes de pénalisation du

champ de vitesse. Nous avons dans ce même chapitre pu expérimenter avec succès ces explorations à travers différents algorithmes basés sur des méthodes de splitting proximal, et avons pu les appliquer sur divers exemples d'interpolations d'images.

Nos perspectives pour compléter ce travail sont assez larges. En ce qui concerne l'algorithme de Benamou-Brenier, et plus généralement les algorithmes traitant du transport optimal dynamique, nous souhaiterions pouvoir mettre en lumière des outils et des méthodes d'analyse précise des propriétés de convergence de ces algorithmes : critères d'arrêt ou de distance des étapes calculées par rapport à la solution théorique, informations théoriques sur la vitesse convergence, etc...

En ce qui concerne les études menées dans le chapitre III et IV, concernant l'existence et l'unicité de solutions au problème de transport optimal dynamique dans L^2 , nous souhaiterions pouvoir étendre celles-ci aux cas que nous avons explorés dans le chapitre V, ce qui implique, en particulier pour le cas du transport optimal anisotrope, de pouvoir étendre ces études dans le cadre d'un transport sur des variétés riemanniennes.

Annexe A

Liste de Théorèmes généraux

A.1 Convolutions, distributions, espaces L^p et $W^{1,p}$

Lemme A.1-1. PROPRIÉTÉS DU PRODUIT DE CONVOLUTION [17]

- Soient $f \in L^1(\mathbb{R}^d)$ et $g \in L^p(\mathbb{R}^d)$ avec $1 \leq p \leq +\infty$. Alors pour presque tout $x \in \mathbb{R}^d$, la fonction $y \mapsto f(x-y)g(y)$ est intégrable sur \mathbb{R}^d . On pose alors

$$(f * g)(x) = \int_{\mathbb{R}^d} f(x-y)g(y) dy.$$

On a alors $(f * g) \in L^p(\mathbb{R}^d)$, et $\|f * g\|_{L^p} \leq \|f\|_{L^1} \|g\|_{L^p}$.

- Soit $(r_n)_n$ une suite régularisante sur \mathbb{R}^{d+1} définie par

$$\forall n \in \mathbb{N}, \quad r_n \in C_c^\infty, \quad \text{supp}(r_n) \subset B\left(0, \frac{1}{n}\right), \quad \int_{\mathbb{R}^{d+1}} r_n = 1, \quad r_n \geq 0.$$

Alors,

1. Pour tout $f \in C^0(\mathbb{R}^d)$, la suite $(r_n * f)_n$ converge uniformément vers f sur tout compact de \mathbb{R}^d .
2. Pour tout $f \in L^p(\mathbb{R}^d)$, avec $1 \leq p < +\infty$, la suite $(r_n * f)_n$ converge uniformément vers f dans $L^p(\mathbb{R}^d)$.

Théorème A.1-2. THÉORÈME D'URYSOHN (RESTREINT À \mathbb{R}^d) ([72] et [80])

1. Soient V un ouvert de \mathbb{R}^d , et K un compact de \mathbb{R}^d tel que $K \subset V$. Alors il existe $f \in C_c^\infty(\mathbb{R}^d)$ tel que :

$$\mathbf{1}_K \leq f \leq \mathbf{1}_V.$$

2. Soient V un ouvert de \mathbb{R}^d , et K un compact de \mathbb{R}^d tel que $K \subset V$. Alors il existe un ouvert U à fermeture compact, tel que

$$K \subset U \subset \bar{U} \subset V.$$

Théorème A.1-3. PARTITION DE L'UNITÉ [39]

Soit $K \subset \mathbb{R}^d$ un compact, et soient U_1, \dots, U_n une série d'ouverts de \mathbb{R}^d tels que

$$K \subset U_1 \cup \dots \cup U_n.$$

Alors il existe des fonctions $\xi_1, \dots, \xi_n \in C_c^\infty(\mathbb{R}^d)$ telles que pour tout $i \in \llbracket 1, n \rrbracket$, on ait

$$0 \leq \xi_i \leq 1, \quad \text{supp}(\xi_i) \subset U_i,$$

et telles qu'il existe un ouvert U avec $K \subset U \subset U_1 \cup \dots \cup U_n$ et

$$\sum_{i=1}^n \xi_i = 1 \quad \text{sur } U.$$

Théorème A.1-4. DUALITÉ DES ESPACES L^p [80]

Soient ω un ouvert de \mathbb{R}^d , $1 \leq p < +\infty$ et soit $\varphi \in (L^p(\omega))'$. Alors en prenant $q \in]1, +\infty]$ tel que :

— $1/p + 1/q = 1$, si $1 < p < +\infty$,

— $q = +\infty$, si $p = 1$,

il existe $u \in L^q(\omega)$ telle que

$$\forall f \in L^p(\omega), \quad \varphi(f) = \int_{\omega} u f,$$

et $\|u\|_{L^q(\omega)} = \|\varphi\|_{(L^p(\omega))'}$.

Théorème A.1-5. UNIFORME CONVEXITÉ DES ESPACES L^p [80]

Soient $1 < p < +\infty$ et Ω un ouvert de \mathbb{R}^d . Soit $f \in L^p(\Omega)$, et soit $(f_n)_{n \in \mathbb{N}}$ une suite dans $L^p(\Omega)$ convergeant faiblement (dans $L^p(\Omega)$) vers f telle que

$$\limsup_n \|f_n\|_{L^p(\Omega)} \leq \|f\|_{L^p(\Omega)}.$$

Alors $(f_n)_{n \in \mathbb{N}}$ converge fortement vers f dans $L^p(\Omega)$.

Théorème A.1-6. THÉORÈME DE PROLONGEMENT D'UN ESPACE DE SOBOLEV [17]

Soit Ω un espace ouvert régulier (les ouverts auxquels nous appliquerons ce Théorème le seront assez) et borné. Alors, pour tout $p \in [1, +\infty]$, il existe une constante C ne dépendant que de Ω et un opérateur linéaire de prolongement

$$P : W^{1,p}(\Omega) \rightarrow W^{1,p}(\mathbb{R}^d)$$

tel que pour tout $f \in W^{1,p}(\Omega)$,

1. $P(f)|_{\Omega} = f$,
2. $\|P(f)\|_{L^p(\mathbb{R}^d)} \leq C \|f\|_{L^p(\Omega)}$,
3. $\|P(f)\|_{W^{1,p}(\mathbb{R}^d)} \leq C \|f\|_{W^{1,p}(\Omega)}$.

Proposition A.1-7. (Corollaire 9.14 [17])

Soit Ω un ouvert borné de \mathbb{R}^d suffisamment régulier (de classe C^1), et soit $1 \leq p \leq +\infty$.

1. Si $1 \leq p < d$, alors $W^{1,p}(\Omega) \subset L^{p^*}(\Omega)$ où $\frac{1}{p^*} = \frac{1}{p} - \frac{1}{d}$, avec injection continue.

2. Si $p = d$, alors $W^{1,p}(\Omega) \subset L^q(\Omega) \forall q \in [p, +\infty[$, avec injection continue.

3. Si $p > d$, alors $W^{1,p}(\Omega) \subset L^\infty(\Omega)$. De plus, pour tout $f \in W^{1,p}(\Omega)$, et pour presque tous $x, y \in \Omega$,

$$|u(x) - u(y)| \leq \frac{2^{2-\frac{d}{p}}}{1 - \frac{d}{p}} \|\nabla u\|_{L^p(\Omega)} |x - y|^{1-\frac{d}{p}}.$$

Théorème A.1-8. (Théorème 4.2 p.149 [80])

Soit Ω un ouvert borné de \mathbb{R}^d de classe C^∞ . Pour tout $k \in \mathbb{N} \cup \{-1\}$, l'opérateur $-\Delta$ est un isomorphisme de $H^{k+2}(\Omega) \cap H_0^1(\Omega)$ sur $H^k(\Omega)$.

Proposition A.1-9. (Proposition 9.18 [17])

Soit Ω un ouvert de \mathbb{R}^d assez régulier (de classe C^1), et soit $f \in L^p(\Omega)$ avec $1 < p < +\infty$. On définit, pour tout $g \in L^p(\Omega)$,

$$\bar{g} : \mathbb{R}^d \rightarrow \mathbb{R} \\ x \mapsto \begin{cases} g(x) & \text{si } x \in \Omega, \\ 0 & \text{sinon.} \end{cases}$$

On a l'équivalence entre

1. $f \in W_0^{1,p}(\Omega)$,
2. $\bar{f} \in W^{1,p}(\mathbb{R}^d)$ et $\nabla \bar{f} = \overline{\nabla f}$.

A.2 Éléments de théorie de la mesure

On écrira qu'une mesure μ (positive ou non) vérifie les propriétés (III) (p. 76) si elle vérifie les propriétés suivantes :

Propriétés (IV).

Soit Ω un ouvert de \mathbb{R}^d , et μ une mesure sur la tribu de Lebesgue de Ω .

1. $\mu(K) < +\infty$ pour tout compact $K \subset \Omega$.
2. Pour tout ensemble Lebesgue-mesurable E de Ω , on a

$$\mu(E) = \inf\{\mu(V), E \subset V, V \text{ ouvert}\}.$$

3. Pour tout E ouvert et pour tout ensemble Lebesgue-mesurable E de Ω tel que $\mu(E) < +\infty$,

$$\mu(E) = \sup\{\mu(K), K \subset E, K \text{ compact}\}.$$

Théorème A.2-1. THÉORÈME DE REPRÉSENTATION DE RIESZ [72]

Soit Ω un ouvert \mathbb{R}^d , et soit Λ une forme linéaire positive sur $C_c(\Omega)$. Alors on peut construire mesure positive μ_1 sur la tribu de Lebesgue de Ω représentant Λ , i.e.

$$\forall f \in C_c(\Omega), \quad \Lambda(f) = \int_{\Omega} f d\mu. \quad (\text{A.2-1})$$

Cette mesure est définie comme il suit :

1. Pour tout V ouvert de Ω , $\mu(V) = \sup\{\Lambda(f), f \in C_c(\Omega), f \leq \mathbb{1}_V\}$.
2. Pour tout ensemble Lebesgue-mesurable E de Ω , $\mu(E) = \inf\{\mu(V), E \subset V, V \text{ ouvert}\}$.

Cette mesure μ vérifie les propriétés (IV).

Théorème A.2-2. [72]

Soient (X, \mathcal{A}) et (Y, \mathcal{B}) deux espaces mesurables (i.e. \mathcal{A} et \mathcal{B} sont respectivement des σ -algèbres sur X et Y). On munit $X \times Y$ de la σ -algèbre $\mathcal{A} \times \mathcal{B}$ (i.e. la plus petite σ -algèbre contenant tous les ensembles de la forme $A \times B$, avec $A \in \mathcal{A}$ et $B \in \mathcal{B}$).

Pour tout $E \in \mathcal{A} \times \mathcal{B}$, on définit pour tout $x \in X$ et $y \in Y$ les ensembles

$$E_x = \{y \in Y, (x, y) \in E\} \quad \text{et} \quad E^y = \{x \in X, (x, y) \in E\}.$$

Alors pour tout $x \in X$ et $y \in Y$, $E_x \in \mathcal{B}$ et $E^y \in \mathcal{A}$.

Théorème A.2-3. [72]

Soient (X, \mathcal{A}, μ) et (Y, \mathcal{B}, ν) deux espaces mesurés σ -finis (réunions dénombrables de sous-ensembles de mesure finie). Soit $Q \in \mathcal{A} \times \mathcal{B}$ (définie dans le Théorème A.2-2 p. 204). On pose pour tout $x \in X$ et tout $y \in Y$,

$$\Phi(x) = \nu(Q_x), \quad \text{et} \quad \Psi(y) = \mu(Q^y).$$

Alors la fonction Φ est mesurable relativement à \mathcal{A} et la fonction Ψ est mesurable relativement à \mathcal{B} . De plus

$$\int_X \Phi d\mu = \int_Y \Psi d\nu. \quad (\text{A.2-2})$$

Ce dernier Théorème est un cas particulier du Théorème de Fubini qui suit (mais entrant en fait dans sa démonstration). Nous utilisons très souvent le Théorème de Fubini de façon implicite, et lorsqu'il s'agit d'une inversion d'intégrales classique et sans ambiguïté. Nous faisons en revanche explicitement appel à ce Théorème que dans les cas plus complexe, où les notions classiques de théorie de la mesure peuvent poser problème, en particulier dans la section IV.3 du chapitre IV, ainsi que dans l'annexe F.

Théorème A.2-4. THÉORÈME DE FUBINI [72]

Soient (X, \mathcal{A}, μ) et (Y, \mathcal{B}, ν) deux espaces mesurés σ -finis. Soit f une fonction à valeurs réels définie sur $X \times Y$, et mesurable relativement à $\mathcal{A} \times \mathcal{B}$.

- Pour tout $x \in X$, on définit sur Y la fonction $y \mapsto f_x(y) = f(x, y)$.
- Pour tout $y \in Y$, on définit sur X la fonction $x \mapsto f^y(x) = f(x, y)$.

Les assertions suivantes sont vérifiées :

1. Pour tout $x \in X$, f_x est mesurable relativement à \mathcal{B} et, pour tout $x \in X$, f^y est mesurable relativement à \mathcal{A} .
2. Si $0 \leq f \leq +\infty$, et en posant pour tous $x \in X$ et $y \in Y$,

$$\Phi(x) = \int_Y f_x d\nu, \quad \text{et} \quad \Psi(y) = \int_X f^y d\mu, \quad (\text{A.2-3})$$

alors la fonction Φ est mesurable relativement à \mathcal{A} et la fonction Ψ est mesurable relativement à \mathcal{B} . De plus

$$\int_X \Phi d\mu = \int_{X \times Y} f (d\mu \otimes d\nu) = \int_Y \Psi d\nu. \quad (\text{A.2-4})$$

3. Si f est une fonction à valeurs complexes et si l'on a

$$\int_X \Phi^* d\mu < +\infty \quad \text{avec} \quad \Phi^*(x) = \int_Y |f_x| d\nu \quad \text{pour tout } x \in X,$$

ou bien

$$\int_Y \Psi^* d\nu < +\infty \quad \text{avec} \quad \Psi^*(y) = \int_X |f^y| d\mu \quad \text{pour tout } y \in Y,$$

alors $f \in L^1(\mu \otimes \nu)$.

4. Si $f \in L^1(\mu \otimes \nu)$, alors $f_x \in L^1(\nu)$ pour presque tout $x \in X$ et $f^y \in L^1(\mu)$ pour presque tout $y \in Y$. De plus, si l'on définit Φ et Ψ comme dans les relations (A.2-3) (définies donc ici seulement pour, respectivement presque tout $x \in X$ et presque tout $y \in Y$), alors Φ et Ψ appartiennent respectivement à $L^1(\mu)$ et $L^1(\nu)$. De plus la relation (A.2-4) est toujours vérifiée par Φ et Ψ .

Théorème A.2-5. THÉORÈME DE RADON-NIKODYM [72]

Soient Ω un ouvert \mathbb{R}^d et \mathcal{A} une σ -algèbre sur Ω . Soient μ et λ deux mesures sur \mathcal{A} , telles μ soit une mesure positive bornée et λ une mesure quelconque (positive ou complexe). Si λ est absolument continue par rapport à μ (noté $\lambda \ll \mu$), alors il existe un unique $h \in L^1(\mu)$ tel que $\lambda = h \cdot \mu$ sur \mathcal{A} , i.e.

$$\forall E \in \mathcal{A}, \quad \lambda(E) = \int_E h d\mu.$$

Théorème A.2-6. THÉORÈME DE RADON-NIKODYM (VARIANTE SPÉCIFIQUE) ([72] et [34])

Soit λ une mesure positive sur la tribu de Lebesgue, finie sur tout compact de \mathbb{R}^d , et absolument continue par rapport à la mesure de Lebesgue \mathcal{L}^d . Alors en posant $\overline{D}\lambda$ la fonction de \mathbb{R}^d dans \mathbb{R} , appelée dérivée supérieure de λ par rapport à \mathcal{L}^d , et définie par

$$\forall x \in \mathbb{R}^d, (\overline{D}\lambda)(x) = \limsup_{r \rightarrow 0} f_r(x), \quad \text{avec} \quad f_r(x) = \frac{1}{\mathcal{V}_d(B_1)r^d} \lambda(B(x, r)), \quad r > 0$$

(où $B(x, r)$ désigne la boule ouverte de centre x et de rayon r , et $\mathcal{V}_d(B_1)$ le volume de la boule unité en dimensions d), on a les propriétés suivantes :

1. Pour \mathcal{L}^d -presque tout $x \in \mathbb{R}^d$, la famille $(f_r(x))_{r>0}$ converge lorsque r tend vers 0 (et on a donc $\lim_{r \rightarrow 0, r>0} f_r(x) = (\overline{D}\lambda)(x)$). On dit alors que λ est différentiable en x par rapport \mathcal{L}^d .
2. $\overline{D}\lambda$ est Lebesgue-mesurable sur \mathbb{R}^d et $\overline{D}\lambda \in L^1(\mathbb{R}^d)$.
3. Pour tout ensemble E Lebesgue-mesurable,

$$\lambda(E) = \int_E (\overline{D}\lambda)(x) d\mathcal{L}^d(x).$$

Proposition A.2-7. [72]

Soit Ω un ouvert \mathbb{R}^d , et soit μ une mesure sur une σ -algèbre \mathcal{A} de Ω . Alors il existe une fonction mesurable f telle que $|f(x)| = 1$ pour tout $x \in \Omega$, et telle que

$$d\mu = h d|\mu|, \tag{A.2-5}$$

où $|\mu|$ désigne la variation totale de la mesure μ .

Théorème A.2-8. THÉORÈME DE REPRÉSENTATION DE RIESZ GÉNÉRALISÉ [72]

Soit Ω un ouvert \mathbb{R}^d , et soit T une forme linéaire continue sur $C(\Omega)$. Alors il existe une unique mesure définie sur la tribu de Lebesgue (non nécessairement positive), vérifiant les propriétés (IV) p. 203, représentant T , i.e.

$$\forall f \in C(\Omega), \quad T(f) = \int_{\Omega} f d\mu \tag{A.2-6}$$

Avec $\|T\| = |\mu|(\Omega)$.

Théorème A.2-9. THÉORÈME DE LUSIN [72]

Soit Ω un ouvert \mathbb{R}^d , et soit μ une mesure positive vérifiant les propriétés (IV) p. 203. $A \subset \Omega$ mesurable tel que $\mu(A) < +\infty$, et soit f une fonction mesurable sur Ω (pour la tribu de Lebesgue) nulle en dehors de A . Alors, pour tout $\epsilon > 0$, il existe $g \in C_c(\Omega)$ telle que

$$\mu(\{x, f(x) \neq g(x)\}) < \epsilon, \quad \text{et} \quad \sup_{x \in \Omega} |g(x)| \leq \sup_{x \in \Omega} |f(x)|. \quad (\text{A.2-7})$$

Corollaire A.2-10. [72]

Sous les hypothèses du Théorème A.2-9, si de plus il existe une constante $M > 0$ telle que $|f| \leq M$, alors il existe une suite $(g_n)_n \in C_c(\Omega)^{\mathbb{N}}$, avec $|g_n| \leq M$ pour tout $n \in \mathbb{N}$, telle que

$$f(x) = \lim_{n \rightarrow +\infty} g_n(x), \quad \mu - p.p.$$

Corollaire A.2-11.

Sous les hypothèses du Théorème A.2-9, si de plus $f \in L^1_{loc}(\Omega, \mu)$, alors

$$\forall g \in C_c(\Omega), \quad \int_{\Omega} f g d\mu = 0 \implies f = 0, \quad \mu - p.p..$$

Preuve : Soit K un compact et soit g définie pour tout $x \in \Omega$,

$$g(x) = \begin{cases} 0 & \text{si } x \notin K, \\ -1 & \text{si } x \in K \text{ et } f(x) < 0, \\ 1 & \text{si } x \in K \text{ et } f(x) \geq 0. \end{cases}$$

D'après le Corollaire A.2-10, il existe une suite $(g_n)_n \in C_c(\Omega)^{\mathbb{N}}$, avec $|g_n| \leq 1$ pour tout $n \in \mathbb{N}$, telle que

$$g(x) = \lim_{n \rightarrow +\infty} g_n(x), \quad \mu - p.p.$$

Du fait que $|f \cdot g_n| \leq |f| \cdot \mathbf{1}_K$, et que $\int_K |f| d\mu < +\infty$, par convergence dominée on conclut que

$$0 = \lim_{n \rightarrow +\infty} \int_{\Omega} f \cdot g_n d\mu = \int_K |f| d\mu.$$

Ainsi $f = 0$ μ -presque partout sur tout compact K de Ω , donc sur Ω . □

Annexe B

Convergence des opérateurs non-expansifs

B.1 Convergence faible

Dans cette section, nous démontrons la Proposition 2.3-6, énoncée dans le chapitre II (p. 65) et rappelée ci dessous :

Proposition 2.3-6.

Soit \mathbf{M} un opérateur quasi-fermement non-expansif tel que que l'ensemble $\text{Fix}(\mathbf{M})$ des points fixes de \mathbf{M} soit non vide, et soit $(x_n)_n$ une suite d'éléments de H vérifiant pour tout $n \in \mathbb{N}$, l'estimation : $\|\mathbf{M}(x_n) - x_{n+1}\| \leq \epsilon_n$, où $(\epsilon_n)_n$ est une suite réelle positive vérifiant $\sum_n \epsilon_n < +\infty$.

Si $(x_{\varphi_1(n)})_n$ et $(x_{\varphi_2(n)})_n$ sont deux sous-suites de $(x_n)_n$ convergeant faiblement respectivement vers x_1 et x_2 (points fixes de \mathbf{M} par le Lemme d'Opial), alors $x_1 = x_2$.

Nous commençons par établir les deux résultats suivants.

Lemme B.1-1.

Supposons $(u_n)_n$ et $(\epsilon_n)_n$ deux suites réelles positives vérifiant $\sum_n \epsilon_n < +\infty$ et la relation de récurrence :

$$u_{n+1} \leq u_n + \epsilon_n,$$

alors $(u_n)_n$ converge dans \mathbb{R} .

Preuve du Lemme B.1-1 : Pour tout $p, q \in \mathbb{N}$, $p > q$, on a l'inégalité :

$$u_p - u_q \leq \sum_{k=q+1}^p \epsilon_k$$

(u_n) est donc bornée. De toute sous-suite de (u_n) on peut donc extraire une sous-suite convergente. Il suffit pour conclure de montrer que toutes les sous-suites convergentes de (u_n) convergent vers la même limite.

Soient $(u_{\varphi_1(n)})_n$ et $(u_{\varphi_2(n)})_n$ deux sous-suites convergentes de $(u_n)_n$, convergeant respectivement vers l_1 et l_2 dans \mathbb{R} .

Soit $\delta > 0$, comme $\lim_m \sum_{k=m}^{\infty} \epsilon_k = 0$, il existe $N \in \mathbb{N}$ tel que, pour tout $p > q > N$,

$$u_p - u_q \leq \delta.$$

et ainsi pour n, m assez grands, $u_{\varphi_2(n)} - u_{\varphi_1(m)} \leq \delta$.

Par passage à la limite, on obtient alors $l_2 - l_1 \leq \delta$. On peut de même montrer que $l_1 - l_2 \leq \delta$ et ce quel que soit δ , d'où $l_1 = l_2$. \square

Corollaire B.1-2.

Si x^* est un point fixe de \mathbf{M} , alors la suite de terme général $u_n = \|x_n - x^*\|$ est convergente (et $(x_n)_n$ bornée).

Preuve :

$$\|x_{n+1} - x^*\| \leq \|x_{n+1} - \mathbf{M}x_n\| + \|\mathbf{M}x_n - x^*\| \leq \epsilon_n + \|\mathbf{M}x_n - \mathbf{M}x^*\| \leq \epsilon_n + \|x_n - x^*\|.$$

On peut donc conclure par le Lemme [B.1-1](#) précédent. \square

Retour à la preuve de la Proposition B.1 : De part le Corollaire [B.1-2](#), on sait que les suites $(\|x_n - x_1\|^2)_n$ et $(\|x_n - x_2\|^2)_n$ sont convergentes. Soient α_1 et α_2 leurs limites respectives, on a alors

$$\|x_n - x_1\|^2 - \|x_n - x_2\|^2 = \langle x_2 - x_1, 2x_n - x_1 - x_2 \rangle.$$

En faisant converger avec $n \leftarrow \varphi^1(n)$ puis $n \leftarrow \varphi^2(n)$, on obtient :

$$\begin{cases} \alpha_1 - \alpha_2 = -\|x_2 - x_1\|^2 \\ \alpha_1 - \alpha_2 = \|x_2 - x_1\|^2 \end{cases}$$

d'où $x_1 = x_2$. \square

B.2 Convergence forte

Dans cette section, nous démontrons le Théorème [2.3-8](#), énoncé dans le chapitre [II](#) (p. [66](#)). Nous reprendrons ici la démonstration de H. Bauschke dans [\[3\]](#), en ajoutant la prise en compte d'erreurs numériques contrôlées (les ϵ_n dans ce qui suit).

Théorème 2.3-8.

Soit T un opérateur non-expansif et supposons que l'ensemble $\text{Fix}(T)$ des points fixes de T soit non vide, et soit $(\lambda_n)_{n \geq 0}$ une suite de paramètres de $[0, 1[$ vérifiant :

1. $\lambda_n \rightarrow 0$
2. $\prod_n (1 - \lambda_n) = 0 \iff \sum_n \lambda_n = +\infty$
3. $\sum_n |\lambda_{n+1} - \lambda_n| < +\infty$.

Soit H un espace de Hilbert, et T un opérateur non-expansif sur H .

Etant donnés a, x_0 dans H , on définit la suite (x_n) débutant en x_0 par :

$$\forall n \geq 0, \quad x_{n+1} = \lambda_n a + (1 - \lambda_n) T x_n$$

où pour tout $n \in \mathbb{N}$ est vérifiée l'estimation $\|T x_n - T x_{n-1}\| \leq \epsilon_n$, avec $\sum_n \epsilon_n < +\infty$.

Alors la suite $(x_n)_n$ converge fortement vers P_{F_A} (où $F = \text{Fix}(T)$) est l'ensemble des points fixes de T : c'est un ensemble convexe fermé d'après le Lemme 2.3-7).

En préambule de la preuve, nous commencerons par démontrer le Lemme suivant :

Lemme B.2-3.

Pour tout $n \in \mathbb{N}$ et $0 \leq m \leq n$:

$$\lambda_{n+1} + \sum_{k=m}^n \lambda_k \prod_{i=k}^n (1 - \lambda_{i+1}) \leq 1.$$

Preuve : Soit $n \in \mathbb{N}$. Supposons que pour tout $m \leq n - 1$:

$$\lambda_n + \sum_{k=m}^{n-1} \lambda_k \prod_{i=k}^{n-1} (1 - \lambda_{i+1}) \leq 1.$$

Soit $m \leq n$

1. $m \leq n - 1$:

$$\begin{aligned} \lambda_{n+1} + \sum_{k=m}^n \lambda_k \prod_{i=k}^n (1 - \lambda_{i+1}) &\leq \lambda_{n+1} + \lambda_n \prod_{i=n}^n (1 - \lambda_{i+1}) + \sum_{k=m}^{n-1} \lambda_k \prod_{i=k}^n (1 - \lambda_{i+1}) \\ &\leq \lambda_{n+1} + \lambda_n (1 - \lambda_{n+1}) + (1 - \lambda_{n+1}) \sum_{k=m}^{n-1} \lambda_k \prod_{i=k}^{n-1} (1 - \lambda_{i+1}) \\ &\leq \lambda_{n+1} + (1 - \lambda_{n+1}) \underbrace{\left(\lambda_n + \sum_{k=m}^{n-1} \lambda_k \prod_{i=k}^{n-1} (1 - \lambda_{i+1}) \right)}_{\leq 1} \leq 1. \end{aligned}$$

2. $m = n$:

$$\lambda_{n+1} + \lambda_n (1 - \lambda_{n+1}) \leq 1.$$

On a donc la preuve par récurrence ($\lambda_1 + \lambda_0(1 - \lambda_1) \leq 1$). \square

Retour à la preuve du Théorème 2.3-8 :

Commençons par traiter le cas $x_0 = a$.

1. Montrons que (x_n) et $(T_n x_n)$ sont bornées :

Soit x^* un point fixe de T . On montre par récurrence : $\forall n \in \mathbb{N}^*$, $\|x_n - x^*\| \leq \|a - x^*\| + \sum_{k=0}^{n-1} \epsilon_k$. Cette inégalité est clairement réalisée pour $n = 1$ en ayant pris $x_0 = a$ (En effet $\|x_1 - x^*\| \leq \|x_1 - Tx_0\| + \|Tx_0 - x^*\| \leq \epsilon_0 + \|a - x^*\|$).

Prenons maintenant $n \in \mathbb{N}$ et supposons que $\|x_n - x^*\| \leq \|a - x^*\| + \sum_{k=0}^{n-1} \epsilon_k$. On a alors

$$\begin{aligned} \|x_{n+1} - x^*\| &\leq \lambda_{n+1} \|a - x^*\| + (1 - \lambda_{n+1}) \|T_n x_n - x^*\| \\ &\leq \lambda_{n+1} \|a - x^*\| + (1 - \lambda_{n+1}) (\|T_n x_n - Tx_n\| + \|Tx_n - x^*\|) \\ &\leq \lambda_{n+1} \|a - x^*\| + (1 - \lambda_{n+1}) \|x_n - x^*\| + \epsilon_n \\ &\leq \lambda_{n+1} \|a - x^*\| + (1 - \lambda_{n+1}) (\|a - x^*\| + \sum_{k=0}^{n-1} \epsilon_k) + \epsilon_n \\ &\leq \|a - x^*\| + \sum_{k=0}^n \epsilon_k \leq \sum_n \epsilon_n < +\infty. \end{aligned}$$

On en conclut que (x_n) est bornée, de même pour $(T_n x_n)$.

2. On montre ensuite que $x_{n+1} - T_n x_n \xrightarrow[n]{} 0$, du fait que $\|x_{n+1} - T_n x_n\| \leq \lambda_{n+1} \|a - T_n x_n\| \leq C \lambda_{n+1}$, avec $\lambda_{n+1} \xrightarrow[n]{} 0$.
3. Montrons que $x_{n+1} - x_n \xrightarrow[n]{} 0$ et $x_n - Tx_n \xrightarrow[n]{} 0$:

Soit $L > 0$ une constante telle que $\|x_{n+1} - x_n\| + \|a - T_n x_n\| \leq L$. On a

$$\begin{aligned} \|x_{n+1} - x_n\| &= \|(\lambda_{n+1} - \lambda_n)(a - T_{n-1} x_{n-1}) + (1 - \lambda_{n+1})(T_n x_n - T_{n-1} x_{n-1})\| \\ &\leq L |\lambda_{n+1} - \lambda_n| + (1 - \lambda_{n+1}) (\|Tx_n - Tx_{n-1}\| \\ &\quad + \|Tx_n - T_n x_n\| + \|Tx_{n-1} - T_{n-1} x_{n-1}\|) \\ &\leq L |\lambda_{n+1} - \lambda_n| + \epsilon_n + \epsilon_{n-1} + (1 - \lambda_{n+1}) \|x_n - x_{n-1}\|. \end{aligned}$$

Par induction de ce qui précède, on obtient pour tout $n \geq m \geq 0$:

$$\begin{aligned} \|x_{n+1} - x_n\| &\leq L \sum_{k=m+1}^n |\lambda_{k+1} - \lambda_k| + \sum_{k=m+1}^n \epsilon_k + \sum_{k=m+1}^n \epsilon_{k-1} + \|x_{m+1} - x_m\| \prod_{k=m+1}^n (1 - \lambda_{k+1}) \\ &\leq L \sum_{k=m+1}^{\infty} |\lambda_{k+1} - \lambda_k| + \sum_{k=m+1}^{\infty} \epsilon_k + \sum_{k=m+1}^{\infty} \epsilon_{k-1} + L \prod_{k=m+1}^n (1 - \lambda_{k+1}) \end{aligned}$$

Soit $\delta > 0$. Sachant que

$$L \sum_{k=m+1}^{\infty} |\lambda_{k+1} - \lambda_k| + \sum_{k=m+1}^{\infty} \epsilon_k + \sum_{k=m+1}^{\infty} \epsilon_{k-1} \xrightarrow[m]{} 0,$$

il existe donc $m \in \mathbb{N}$ tel que pour tout $n > m$,

$$\|x_{n+1} - x_n\| \leq \delta + L \prod_{k=m+1}^n (1 - \lambda_{k+1}).$$

Comme de plus $\prod_{k=m+1}^n (1 - \lambda_{k+1}) \xrightarrow[n]{} 0$, on a alors

$$\overline{\lim}_n \|x_{n+1} - x_n\| \leq \delta,$$

et ainsi $x_{n+1} - x_n \xrightarrow[n]{} 0$. On en déduit

$$\begin{aligned} \|x_n - Tx_n\| &\leq \|x_n - x_{n+1}\| + \|x_{n+1} - T_n x_n\| + \|T_n x_n - Tx_n\| \\ &\leq \|x_n - x_{n+1}\| + \|x_{n+1} - T_n x_n\| + \epsilon_n, \end{aligned}$$

et finalement $x_n - Tx_n \xrightarrow[n]{} 0$.

4. $\overline{\lim}_n \langle T_n x_n - P_F a, a - P_F a \rangle \leq 0$:

Soit $(x_{\varphi(n)})$ une sous-suite de (x_n) telle que :

$$\lim_n \langle T_{\varphi(n)} x_{\varphi(n)} - P_F a, a - P_F a \rangle = \overline{\lim}_n \langle T_n x_n - P_F a, a - P_F a \rangle$$

Quitte à devoir ré-extraire une sous-suite de $(x_{\varphi(n)})$, on suppose que celle-ci converge faiblement vers un élément x^* de H . Les conditions du Lemme d'Opial étant réunies (i.e. $x_n - Tx_n \xrightarrow[n]{} 0$ fortement), x^* est un point fixe de T . Comme

$$\begin{aligned} \langle T_{\varphi(n)} x_{\varphi(n)} - P_F a, a - P_F a \rangle &= \langle T_{\varphi(n)} x_{\varphi(n)} - x_{\varphi(n)+1}, a - P_F a \rangle + \langle x_{\varphi(n)+1} - P_F a, a - P_F a \rangle \\ &\xrightarrow[n]{} \langle x^* - P_F a, a - P_F a \rangle \leq 0, \end{aligned}$$

on a alors $\overline{\lim}_n \langle T_n x_n - P_F a, a - P_F a \rangle \leq 0$.

5. Conclusion : Soit $\delta > 0$ et $m \in \mathbb{N}^*$ tel que pour tout $n > m$,

$$\langle \tilde{T}_n x_n - P_F a, a - P_F a \rangle \leq \delta, \quad \lambda_n \|a - P_F a\|^2 \leq \delta, \quad \text{et} \quad \sum_{k=m}^{\infty} \epsilon_k \leq \delta.$$

On observe

$$\begin{aligned} \|x_{n+1} - P_F a\|^2 &= \|\lambda_{n+1}(a - P_F a) + (1 - \lambda_{n+1})(T_n x_n - P_F a)\|^2 \\ &= \lambda_{n+1}^2 \|a - P_F a\|^2 + 2\lambda_{n+1}(1 - \lambda_{n+1}) \langle T_n x_n - P_F a, a - P_F a \rangle \\ &\quad + (1 - \lambda_{n+1})^2 (\|T_n x_n - Tx_n\|^2 + \|Tx_n - P_F a\|^2 \\ &\quad + 2\langle T_n x_n - Tx_n, Tx_n - P_F a \rangle) \\ &\leq \lambda_{n+1} \delta + (1 - \lambda_{n+1}) \|x_n - P_F a\|^2 + 2\lambda_{n+1} \delta + C\epsilon_n, \end{aligned}$$

et par induction, on obtient pour tout $n > m$,

$$\begin{aligned} \|x_{n+1} - P_F a\|^2 &\leq 3\delta \left(\lambda_{n+1} + \sum_{k=m}^n \lambda_k \prod_{i=k}^n (1 - \lambda_{i+1}) \right) \\ &\quad + \|x_m - P_F a\|^2 \prod_{k=m}^n (1 - \lambda_{k+1}) + C \sum_{k=m}^n \epsilon_k \\ &\leq 3\delta + \|x_m - P_F a\|^2 \prod_{k=m}^n (1 - \lambda_{k+1}) + C\delta. \quad (\text{Lemme B.2-3}) \end{aligned}$$

On fait tendre n vers l'infini pour obtenir $\overline{\lim}_n \|x_{n+1} - P_F a\|^2 \leq (3 + C)\delta$, pour tout $\delta > 0$, d'où $x_n \xrightarrow[n]{} P_F a$.

6. Généralisation :

Considérons maintenant le cas général (x_0 n'est plus forcément égal à a). Soit $(y_n)_n$ la suite de H définie par $y_0 = a$ et la relation de récurrence : $\forall n \geq 0, x_{n+1} = \lambda_n a + (1 - \lambda_n)Tx_n$. Par ce qui précède on a $y_n \xrightarrow[n]{} P_F a$. Or pour tout $n \in \mathbb{N}$,

$$\begin{aligned} \|x_{n+1} - y_{n+1}\| &\leq (1 - \lambda_{n+1})\|T_n x_n - T y_n\| \\ &\leq (1 - \lambda_{n+1})(\|T_n x_n - T x_n\| + \|T x_n - T y_n\|) \\ &\leq (1 - \lambda_{n+1})\|x_n - y_n\| + \epsilon_n. \end{aligned}$$

Par induction, pour tout $n > m > 0$,

$$\|x_{n+1} - y_{n+1}\| \leq \|x_m - y_m\| \prod_{k=m}^n (1 - \lambda_{k+1}) + \sum_{k=m}^n \epsilon_k.$$

Soit $m \in \mathbb{N}$ tel que $\sum_{k=m}^{\infty} \epsilon_k \leq \delta$. On a $\|x_{n+1} - y_{n+1}\| \leq \|x_m - y_m\| \prod_{k=m}^n (1 - \lambda_{k+1}) + \delta$, on fait tendre n vers l'infini, on obtient $\overline{\lim}_n \|x_n - y_n\| \leq \delta$, et ce pour tout $\delta > 0$, d'où $x_n - y_n \xrightarrow[n]{} 0$ et ainsi $x_n \xrightarrow[n]{} P_F a$.

□

Annexe C

Preuves annexes de la section III.2

L'objectif de cette annexe est de prouver les Lemmes 3.2-5, 3.2-6 et 3.2-7, énoncés dans la section III.2, laquelle a pour but de prouver qu'il est toujours possible de supposer que le potentiel ϕ du plan de transport optimal vérifie l'hypothèse (Γ_1) .

Lemme 3.2-5.

Soit f une fonction convexe sur \mathbb{R}^n différentiable en un point $x \in \mathbb{R}^n$. Alors il existe $\epsilon > 0$ tel que f soit lipschitzienne sur $B(x, \epsilon)$, avec pour constante de lipschitz $L = 4(|\nabla f(x)| + 1)$.

Preuve : f est différentiable en x , donc pour z au voisinage de x ,

$$f(z) = f(x) + \langle \nabla f(x), z - x \rangle + \sigma_x(z - x), \text{ avec } \sigma_x(h) = o(h).$$

Il existe donc $\epsilon > 0$, tel que pour tout h vérifiant $|h| \leq 2\epsilon$, $\sigma_x(h) \leq |h|$. Ainsi, pour tout $y \in B(x, 2\epsilon)$,

$$f(y) \leq f(x) + |\nabla f(x)||y - x| + |y - x| \leq f(x) + 2(|\nabla f(x)| + 1)\epsilon \quad (\text{C.0-1})$$

Pour tout $y \in B(x, 2\epsilon)$, soit $y' \in B(x, 2\epsilon)$ le symétrique de y par rapport à x .

Comme $x = \frac{1}{2}(y + y')$, l'inégalité de convexité nous donne, en majorant $f(y')$ comme dans (C.0-1),

$$f(x) \leq \frac{1}{2}f(y) + \frac{1}{2}f(y') \leq \frac{1}{2}f(y) + \frac{1}{2}f(x) + (|\nabla f(x)| + 1)\epsilon,$$

d'où

$$f(x) - 2(|\nabla f(x)| + 1)\epsilon \leq f(y) \quad (\text{C.0-2})$$

Soient $y_1, y_2 \in B(x, \epsilon)$ supposés distincts. On pose $y'_1 = y_1 - \epsilon \frac{y_2 - y_1}{|y_2 - y_1|}$ et $y'_2 = y_2 + \epsilon \frac{y_2 - y_1}{|y_2 - y_1|}$. Il est clair que $y'_1, y'_2 \in B(x, 2\epsilon)$ et

$$y_1 = \lambda y'_1 + (1 - \lambda)y_2, \text{ et } y_2 = \lambda y'_2 + (1 - \lambda)y_1$$

avec $\lambda = \frac{|y_2 - y_1|}{\epsilon + |y_2 - y_1|} \in [0, 1]$. Par l'inégalité de convexité, on obtient

$$f(y_1) \leq \lambda f(y'_1) + (1 - \lambda)f(y_2)$$

i.e.

$$f(y_1) - f(y_2) \leq \lambda(f(y'_1) - f(y_2))$$

Or, par (C.0-1) : $f(y'_1) \leq f(x) + 2(|\nabla f(x)| + 1)\epsilon$; et par (C.0-2) : $-f(y_2) \leq -f(x) + 2(|\nabla f(x)| + 1)\epsilon$. En observant de plus que $\lambda \leq \frac{|y_2 - y_1|}{\epsilon}$, on obtient

$$f(y_1) - f(y_2) \leq 4(|\nabla f(x)| + 1)|y_2 - y_1|$$

En procédant de même en considérant cette fois y_2 comme interpolation de y'_2 et y_1 , on obtient symétriquement :

$$f(y_2) - f(y_1) \leq 4(|\nabla f(x)| + 1)|y_2 - y_1|$$

On peut donc conclure

$$|f(y_2) - f(y_1)| \leq 4(|\nabla f(x)| + 1)|y_2 - y_1|$$

Ceci pour tout $y_1, y_2 \in B(x, \epsilon)$ (on peut trivialement y inclure le cas $y_1 = y_2$). \square

Lemme 3.2-6.

Soit f une fonction convexe sur \mathbb{R}^n différentiable en un point $x \in \mathbb{R}^n$. Alors il existe $\delta > 0$ tel que pour tout $y \in B(x, \delta)$, et pour tout $\alpha \in \partial f(y)$, $|\alpha| \leq 4(|\nabla f(x)| + 1)$.

Preuve (fortement inspiré de celle de la Proposition 16.14 (iii) de [4]) :

D'après le Lemme 3.2-5, il existe $\delta > 0$ tel que f soit lipschitzienne sur $B(x, 2\delta)$, avec pour constante de Lipschitz $L = 4(|\nabla f(x)| + 1)$. Soit $y \in B(x, \delta)$, et soit $\epsilon > 0$ tel que $\overline{B}(y, \epsilon) \subset B(x, 2\delta)$. Soit $\alpha \in \partial f(y)$. Pour tout $h \in \mathbb{R}^n$,

$$f(y + h) \geq f(y) + \langle \alpha, h \rangle,$$

et donc pour $|h| \leq \epsilon$,

$$\langle \alpha, h \rangle \leq f(y + h) - f(y) \leq 4(|\nabla f(x)| + 1)|h|$$

en posant $h = \frac{\epsilon}{|\alpha|}\alpha$ (si $\alpha \neq 0$), on a bien $|h| = \epsilon$ et ainsi :

$$|\alpha| \leq 4(|\nabla f(x)| + 1)$$

(et si $\alpha = 0$ l'inégalité précédente est triviale). \square

Lemme 3.2-7.

Si f est convexe et continue et superlinéaire sur \mathbb{R}^d tel qu'il existe $a, b, c \geq 0$ avec $a < 1$ tels que

$$\forall x \in \mathbb{R}^d, f(x) \leq a|x|^2 + b|x| + c,$$

alors f^* est convexe continue et superlinéaire sur \mathbb{R}^d (et admet un minimum sur \mathbb{R}^d par coercivité).

Preuve :

1. f étant superlinéaire, il est clair que pour tout $x \in \mathbb{R}^d$, il existe $r_x > 0$ tel que si $|y| \geq r_x$, alors $f(y)/|y| \geq |x|$. Ainsi pour tout $y \in \mathbb{R}^d$ tel que $|y| \geq r_x$, on a

$$\langle x, y \rangle - f(y) \leq |y|(|x| - f(y)/|y|) \leq 0.$$

De plus, du fait que f est continue sur \mathbb{R}^d , $M = \sup_{B(0, r_x)} (\langle x, y \rangle - f(y))$ est finie, donc au final

$$\sup_{y \in \mathbb{R}^d} (\langle x, y \rangle - f(y)) \leq \max\{0, M\} < +\infty.$$

f^* est ainsi finie et donc continue (car localement lipschitzienne) sur \mathbb{R}^d .

2. Enfin, pour tout $x \in \mathbb{R}^d$,

$$f^*(x) \geq |x|^2 - f(x) \geq (1 - a)|x|^2 - b|x| - c$$

Or $1 - a > 0$ et ainsi $(1 - a)|\cdot|^2 - b|\cdot| - c$ est superlinéaire, et donc il en est de même pour f^* qui est du fait coercitive et, étant convexe et continue, admet alors un minimum sur \mathbb{R}^d .

□

Annexe D

Calcul d'un opérateur proximal

Dans cette annexe, nous proposons de donner le détail du calcul de l'opérateur \mathbf{p}_ϕ à partir duquel avait été calculé le champ de vitesse v_ϕ explicité dans l'exemple 3.3-22, dont l'objet était la construction d'un champ de vitesse de transport optimal ne vérifiant pas une régularité H^1 .

Rappelons que le champ de vitesse v_ϕ est calculé à partir de l'opérateur \mathbf{p}_ϕ via la formule suivante :

$$v_\phi(t, \cdot) = \frac{\text{id} - \mathbf{p}_\phi(t, \cdot)}{t}, \text{ pour tout } t \in]0, 1[$$

(voir la définition 3.3-9 p. 89).

Dans cet exemple, nous considérons le potentiel de transport ϕ défini pour tout $(x, y) \in \mathbb{R}^2$ par :

$$\phi(x, y) = \frac{1}{2}(|x|^2 + |y|^2) + |y|,$$

et nous déclarions que l'opérateur \mathbf{p}_ϕ associé à ce potentiel était alors défini, pour tout $t \in]0, 1[$ et pour tout $(x, y) \in \mathbb{R}^2$, de la manière suivante :

$$\mathbf{p}_\phi(t, x, y) = \begin{cases} (x, y) + (0, -t), & \text{si } y > t, \\ (x, 0), & \text{si } -t \leq y \leq t, \\ (x, y) + (0, t), & \text{si } y < -t. \end{cases}$$

(voir la Remarque 3.3-23 p. 122).

Nous allons donc à présent détailler le calcul à aboutissant à l'expression de \mathbf{p}_ϕ énoncée ci-dessus.

On rappelle que d'après la définition 3.3-9 (p. 89) de l'opérateur \mathbf{p}_ϕ , pour tous $t \in]0, 1[$ et $(x, y), (x', y') \in \mathbb{R}^2$, on a

$$(x', y') = \mathbf{p}_\phi(t, x, y) \Leftrightarrow (x, y) \in (1 - t)(x', y') + t\partial\phi(x', y').$$

Le calcul de la sous-différentielle $\partial\phi$ de ϕ donne le résultat suivant : pour tout $(x, y) \in \mathbb{R}^2$, on a

- $\partial\phi(x, y) = \{(x, y) + (0, 1)\}$ (singleton) si $y > 0$,
- $\partial\phi(x, y) = \{(x, y) + (0, -1)\}$ (singleton) si $y < 0$,
- $\partial\phi(x, y) = \{(x, 0)\} + \{0\} \times [-1, 1] = \{x\} \times [-1, 1]$ si $y = 0$.

Le potentiel ϕ est donc différentiable en tous points des ensemble $\mathbb{R} \times]-\infty, 0[$ et $\mathbb{R} \times]0, +\infty[$, mais non-différentiable en chacun des points de l'ensemble $\mathbb{R} \times \{0\}$. Passons maintenant au calcul

de \mathbf{p}_ϕ : soit $t \in]0, 1[$, et soient $(x, y), (x', y') \in \mathbb{R}^2$ tels que $(x', y') = \mathbf{p}_\phi(t, x, y)$. On distingue trois cas : $y' > 0$, $y' < 0$ et $y' = 0$. Développons ces trois cas :

1. On a tout d'abord :

$$\begin{aligned} y' > 0 &\Leftrightarrow \begin{cases} y' > 0, \\ (x', y') = \mathbf{p}_\phi(t, x, y) \end{cases} \Leftrightarrow \begin{cases} y' > 0, \\ (x, y) \in (1-t)(x', y') + t\partial\phi(x', y') \\ \quad = \{(1-t)(x', y') + t((x', y') + (0, 1))\} \\ \quad = \{(x', y') + (0, t)\} \end{cases} \\ &\Leftrightarrow \begin{cases} y' > 0, \\ (x, y) = (x', y') + (0, t) \end{cases} \Leftrightarrow \begin{cases} y > t, \\ (x', y') = (x, y) + (0, -t) = \mathbf{p}_\phi(t, x, y). \end{cases} \end{aligned}$$

2. De même, on a

$$\begin{aligned} y' < 0 &\Leftrightarrow \begin{cases} y' < 0, \\ (x', y') = \mathbf{p}_\phi(t, x, y) \end{cases} \Leftrightarrow \begin{cases} y' < 0, \\ (x, y) \in (1-t)(x', y') + t\partial\phi(x', y') \\ \quad = \{(1-t)(x', y') + t((x', y') + (0, -1))\} \\ \quad = \{(x', y') + (0, -t)\} \end{cases} \\ &\Leftrightarrow \begin{cases} y' < 0, \\ (x, y) = (x', y') + (0, -t) \end{cases} \Leftrightarrow \begin{cases} y < -t, \\ (x', y') = (x, y) + (0, t) = \mathbf{p}_\phi(t, x, y). \end{cases} \end{aligned}$$

3. Enfin, pour finir on a

$$\begin{aligned} y' = 0 &\Leftrightarrow \begin{cases} y' = 0, \\ (x', y') = \mathbf{p}_\phi(t, x, y) \end{cases} \\ &\Leftrightarrow \begin{cases} y' = 0, \\ (x, y) \in (1-t)(x', y') + t\partial\phi(x', y') \\ \quad = \{(1-t)(x', 0) + t((x', 0) + (0, \lambda)), \lambda \in [-1, 1]\} \\ \quad = \{x'\} \times [-t, t] \end{cases} \\ &\Leftrightarrow \begin{cases} y \in [-t, t], \\ (x', y') = (x, 0) = \mathbf{p}_\phi(t, x, y). \end{cases} \end{aligned}$$

Annexe E

Preuves annexes de la section IV.1

Notes pour ce chapitre.

- Pour tout hyper-pavé ouvert $Q =]a_1, b_1[\times \dots \times]a_d, b_d[\subset \mathbb{R}^d$, on définit le diamètre $\text{diam}(Q)$ de Q par :

$$\text{diam}(Q) = \max\{|a_1 - b_1|, \dots, |a_d - b_d|\}.$$

- On appelle fonction à variables séparées sur Q , une fonction du type

$$\chi : x = (x_1, \dots, x_d) \in Q \mapsto \chi_1(x_1) \dots \chi_d(x_d),$$

avec χ_k définie sur $]a_k, b_k[$ pour tout $k \in \llbracket 1, d \rrbracket$.

- On note $|\cdot|_\infty$ la norme sur \mathbb{R}^d définie par

$$|x|_\infty = \max\{|x_1|, \dots, |x_d|\}, \text{ pour } x = (x_1, x_2) \in \mathbb{R}^d.$$

- Pour toute fonction $f = (f_1, \dots, f_d) \in L^1(\omega)^d$ (avec ω un ouvert de \mathbb{R}^d), on définit $\|f\|_{L^1(\omega)}$ par

$$\|f\|_{L^1(\omega)} = \max\{\|f_1\|_{L^1(\omega)}, \dots, \|f_d\|_{L^1(\omega)}\}.$$

Le but de cette section est de démontrer le Lemme 4.1-11 que nous rappelons ici :

Lemme 4.1-11.

Soit Ω un ouvert borné de \mathbb{R}^d , et soit une distribution $T \in \mathcal{D}'([0, 1] \times \Omega)$, telle que l'on ait $\nabla_{t,x} T \in L_{loc}^\infty([0, 1] \times \Omega)^d$ (ou du moins un de ses représentants de classe \mathcal{L}^d -mesurable que l'on notera g).

Alors $T \in W_{loc}^{1,\infty}([0, 1] \times \Omega)$, ou plus précisément il existe $f \in W_{loc}^{1,\infty}([0, 1] \times \Omega)$ tel que :

- $\forall h \in C_c^\infty([0, 1] \times \Omega), \quad \langle T, h \rangle = \int_0^1 \int_\Omega f h \, dx \, dt,$
- $\forall h \in C_c^\infty([0, 1] \times \Omega)^d, \quad \langle T, \text{div}(h) \rangle = \int_0^1 \int_\Omega f \cdot \text{div}(h) \, dx \, dt = - \int_0^1 \int_\Omega g \cdot h \, dx \, dt.$

De plus f est lipschitzienne (ou du moins un de ses représentants) sur Ω .

Dans ce but, nous démontrerons dans un premier temps le Lemme suivant :

Lemme E.0-1.

Soient $a_1, \dots, a_d, b_1, \dots, b_d \in \mathbb{R}$ tels que pour tout $k \in \llbracket 1, d \rrbracket$, on ait $a_k < b_k$, et soit l'ensemble $Q =]a_1, b_1[\times \dots \times]a_d, b_d[$, un hyper-pavé ouvert de \mathbb{R}^d .

Pour tout $k \in \llbracket 1, d \rrbracket$, soit $\chi_k \in C_c^\infty(]a_k, b_k[)$, tel que pour tout $k \in \llbracket 1, d \rrbracket$, on ait $\int_{a_k}^{b_k} \chi_k = 1$. On définit la fonction à variables séparées $\chi \in C_c^\infty(Q)$ par :

$$\chi : x = (x_1, \dots, x_d) \in Q \longmapsto \chi_1(x_1) \dots \chi_d(x_d).$$

Alors pour tout $h \in C_c^\infty(Q)$, il existe $\varphi \in C_c^\infty(Q)^d$ telle que :

$$h = \left(\int_Q h(t) dt \right) \chi + \operatorname{div}(\varphi), \quad (\text{E.0-1})$$

et

$$\|\varphi\|_{L^1(Q)} \leq 2 \operatorname{diam}(Q) \|h\|_{L^1(Q)}. \quad (\text{E.0-2})$$

La preuve de ce dernier Lemme se révèle relativement simple, bien qu'assez longue (et assez lourde). C'est pourquoi, dans le but de donner au lecteur une idée claire de la preuve, nous choisissons dans un premier temps d'énoncer et de démontrer une version simplifiée de ce Lemme (en dimension 1). La preuve complète du Lemme E.0-1 sera exposée à la suite de celui-ci.

Lemme E.0-2.

Soit $\chi \in C_c^\infty(\mathbb{R})$ telle que $\int_{\mathbb{R}} \chi(t) dt = 1$. Alors pour tout $f \in C_c^\infty(\mathbb{R})$, il existe $\psi \in C_c^\infty(\mathbb{R})$ telle que

$$f = \left(\int_{\mathbb{R}} f(t) dt \right) \chi + \psi'. \quad (\text{E.0-3})$$

et

$$\sup_{x \in \mathbb{R}} |\psi(x)| \leq 2 \int_{\mathbb{R}} |f(t)| dt. \quad (\text{E.0-4})$$

Preuve : Soit $f \in C_c^\infty(\mathbb{R})$, on définit $\psi \in C_c^\infty(\mathbb{R})$ par

$$\forall x \in \mathbb{R}, \psi(x) = \int_{-\infty}^x \left(f(t) - \chi(t) \int_{\mathbb{R}} f(s) ds \right) dt. \quad (\text{E.0-5})$$

ψ est de classe C^∞ et vérifie avec f les relations (E.0-3) et (E.0-4). En effet, pour tout $x \in \mathbb{R}$, on vérifie immédiatement

$$\psi'(x) = f(x) - \chi(x) \int_{\mathbb{R}} f(s) ds,$$

et

$$\begin{aligned} |\psi(x)| &\leq \int_{\mathbb{R}} \left(f(t) - \chi(t) \int_0^1 f(s) ds \right) dt \leq \int_{\mathbb{R}} |f(t)| dt + \left(\int_{\mathbb{R}} |f(t)| dt \right) \int_0^1 f(s) ds \\ &\leq 2 \int_{\mathbb{R}} |f(t)| dt. \end{aligned}$$

De plus ψ est à support compact dans \mathbb{R} . En effet :

- la fonction $t \mapsto f(t) - \chi(t) \int_0^1 f(s) ds$ est à support compact, donc uniformément nulle sur $] - \infty, x[$ si x est assez petit, d'où l'on déduit que ψ est uniformément nulle sur $] - \infty, x_0[$ pour un x_0 assez petit,
- la fonction $t \mapsto f(t) - \chi(t) \int_0^1 f(s) ds$ est également uniformément nulle sur $]x, +\infty[$ si x est assez grand. Donc pour x assez grand on a :

$$\begin{aligned} \psi(x) &= \int_{-\infty}^x \left(f(t) - \chi(t) \int_0^1 f(s) ds \right) dt = \int_{\mathbb{R}} \left(f(t) - \chi(t) \int_0^1 f(s) ds \right) dt \\ &= \int_{\mathbb{R}} f(t) dt - \underbrace{\left(\int_{\mathbb{R}} \chi(t) dt \right)}_{=1} \cdot \left(\int_{\mathbb{R}} f(s) ds \right) = 0. \end{aligned}$$

On a donc $\psi \in C_c^\infty(0, 1)$. □

Passons donc maintenant à la preuve du Lemme E.0-1 (p. 222) qui sera, rappelons-le, une généralisation en dimension d de la preuve du Lemme E.0-2.

Retour à la preuve du Lemme E.0-1 : Pour i allant de m à n , on pose

$$(t_i)_i = (t_m, \dots, t_n) \tag{E.0-6}$$

$$\bigotimes_{i=m}^n \int_{a_i}^{b_i} h((t_i)_i) (dt_i)_i = \int_{a_m}^{b_m} \dots \int_{a_n}^{b_n} h(t_m, \dots, t_n) dt_n \dots dt_m \tag{E.0-7}$$

Soit $h \in C^\infty(Q)$. Alors, pour tout $k \in \llbracket 2, d \rrbracket$, la fonction

$$(x_k, \dots, x_d) \mapsto \bigotimes_{i=1}^{k-1} \int_{a_i}^{b_i} h((t_i)_i, x_k, \dots, x_d) (dt_i)_i \tag{E.0-8}$$

est de classe C^∞ sur $]a_k, b_k[\times \dots \times]a_d, b_d[$, et ceci en invoquant la règle de dérivation sous le signe intégrale.

On introduit alors les fonctions de classe C^∞ sur Q :

$$\varphi_1 : x = (x_1, \dots, x_d) \in Q \mapsto \int_{a_1}^{x_1} \left(h(s_1, x_2, \dots, x_d) - \chi_1(s_1) \int_{a_1}^{b_1} h(t_1, x_2, \dots, x_d) dt_1 \right) ds_1,$$

pour tout $k \in \llbracket 2, d-1 \rrbracket$,

$$\begin{aligned} \varphi_k : x \in Q &\mapsto \chi_1(x_1) \dots \chi_{k-1}(x_{k-1}) \int_{a_k}^{x_k} \left(\bigotimes_{i=1}^{k-1} \int_{a_i}^{b_i} h((t_i)_i, s_k, x_{k+1}, \dots, x_d) (dt_i)_i \right. \\ &\quad \left. - \chi_k(s_k) \bigotimes_{j=1}^{k-1} \int_{a_j}^{b_j} \int_{a_k}^{b_k} h((\tau_j)_j, u_k, x_{k+1}, \dots, x_d) du_k (d\tau_j)_j \right) ds_k \\ &= \chi_1(x_1) \dots \chi_{k-1}(x_{k-1}) \bigotimes_{i=1}^{k-1} \int_{a_i}^{b_i} \left[\int_{a_k}^{x_k} (h((t_i)_i, s_k, x_{k+1}, \dots, x_d) \right. \\ &\quad \left. - \chi_k(s_k) \int_{a_k}^{b_k} h((t_i)_i, u_k, x_{k+1}, \dots, x_d) du_k \right) ds_k \Big] (dt_i)_i, \end{aligned}$$

et enfin,

$$\begin{aligned} \varphi_d : x \in Q &\mapsto \chi_1(x_1) \cdots \chi_{d-1}(x_{d-1}) \int_{a_d}^{x_d} \left(\bigotimes_{i=1}^{d-1} \int_{a_i}^{b_i} h((t_i)_i, s_d) (dt_i)_i - \chi_d(s_d) \bigotimes_{j=1}^d \int_{a_j}^{b_j} h((\tau_j)_j) (d\tau_j)_j \right) ds_d \\ &= \chi_1(x_1) \cdots \chi_{d-1}(x_{d-1}) \bigotimes_{i=1}^{d-1} \int_{a_i}^{b_i} \left[\int_{a_d}^{x_d} \left(h((t_i)_i, s_d) - \chi_d(s_d) \int_{a_d}^{b_d} h((t_i)_i, u_d) du_d \right) ds_d \right] (dt_i)_i \end{aligned}$$

Il est clair que :

$$\partial_1 \varphi_1(x) = \partial_1 \varphi_1(x_1, \dots, x_d) = h(x_1, \dots, x_d) - \chi_1(x_1) \int_{a_1}^{b_1} h(t_1, x_2, \dots, x_d) dt_1,$$

pour tout $k \in \llbracket 2, d-1 \rrbracket$,

$$\begin{aligned} \partial_k \varphi_k(x) &= \chi_1(x_1) \cdots \chi_{k-1}(x_{k-1}) \bigotimes_{i=1}^{k-1} \int_{a_i}^{b_i} h((t_i)_i, x_k, \dots, x_d) (dt_i)_i \\ &\quad - \chi_1(x_1) \cdots \chi_k(x_k) \bigotimes_{i=1}^k \int_{a_i}^{b_i} h((t_i)_i, x_{k+1}, \dots, x_d) (dt_i)_i \end{aligned}$$

et,

$$\partial_d \varphi_d(x) = \chi_1(x_1) \cdots \chi_1(x_{d-1}) \bigotimes_{i=1}^{d-1} \int_{a_i}^{b_i} h((t_i)_i, x_d) (dt_i)_i - \chi(x) \bigotimes_{i=1}^d \int_{a_i}^{b_i} h((t_i)_i) (dt_i)_i,$$

Ainsi, en posant $\varphi = (\varphi_1, \dots, \varphi_d) \in C^\infty(Q)^d$, on a :

$$\begin{aligned} h(x) &= h(x_1, \dots, x_d) \\ &= \chi_1(x_1) \int_{a_1}^{b_1} h(t_1, x_2, \dots, x_d) dt_1 + \left(h(x_1, \dots, x_d) - \chi_1(x_1) \int_{a_1}^{b_1} h(t_1, x_2, \dots, x_d) dt_1 \right), \end{aligned}$$

c'est à dire

$$h(x) = \chi_1(x_1) \int_{a_1}^{b_1} h(t_1, x_2, \dots, x_d) dt_1 + \partial_1 \varphi_1(x).$$

En continuant (si $d \geq 3$) :

$$\begin{aligned} h(x) &= \chi_1(x_1) \chi_2(x_2) \int_{a_2}^{b_2} \int_{a_1}^{b_1} h(t_1, t_2, x_3, \dots, x_d) dt_1 dt_2 + \partial_1 \varphi_1(x) \\ &\quad + \left(\chi_1(x_1) \int_{a_1}^{b_1} h(t_1, x_2, \dots, x_d) dt_1 - \chi_1(x_1) \chi_2(x_2) \int_{a_2}^{b_2} \int_{a_1}^{b_1} h(t_1, t_2, x_3, \dots, x_d) dt_1 dt_2 \right) \\ &= \chi_1(x_1) \chi_2(x_2) \int_{a_2}^{b_2} \int_{a_1}^{b_1} h(t_1, t_2, x_3, \dots, x_d) dt_1 dt_2 + \partial_1 \varphi_1(x) + \partial_2 \varphi_2(x), \end{aligned}$$

etc... En procédant ainsi, par télescopage successif, on obtient finalement :

$$\begin{aligned} h(x) &= \chi_1(x_1) \cdots \chi_d(x_d) \bigotimes_{i=1}^d \int_{a_i}^{b_i} h((t_i)_i) (dt_i)_i + \sum_{k=1}^d \partial_k \varphi_k(x) \\ &= \chi(x) \int_Q h(t) dt + \sum_{k=1}^d \partial_k \varphi_k(x) \\ &= \chi(x) \int_Q h(t) dt + \operatorname{div}_x(\varphi)(x). \end{aligned} \tag{E.0-9}$$

On a donc prouvé (E.0-1) Montrons maintenant que pour tout $k \in \llbracket 1, d \rrbracket$, on a la majoration :

$$\int_Q |\varphi_k(x)| dx \leq 2|b_k - a_k| \int_Q |h(x)| dx. \quad (\text{E.0-10})$$

En effet, pour tout $k \in \llbracket 2, d - 1 \rrbracket$,

$$\begin{aligned} \int_Q |\varphi_k(x)| dx &= \prod_{l=1}^{k-1} \left(\int_{a_l}^{b_l} \chi_l(x_l) dx_l \right) \bigotimes_{i=k+1}^d \int_{a_i}^{b_i} \left[\bigotimes_{j=1}^{k-1} \int_{a_j}^{b_j} \left[\int_{a_k}^{b_k} \int_{a_k}^{x_k} (h((t_j)_j, s_k, (x_i)_i) \right. \right. \\ &\quad \left. \left. - \chi_k(s_k) \int_{a_k}^{b_k} h((t_j)_j, u_k, (x_i)_i) du_k \right) ds_k \right] (dt_j)_j \Big| dx_k (dx_i)_i \\ &\leq \bigotimes_{i=k+1}^d \int_{a_i}^{b_i} \left[\bigotimes_{j=1}^{k-1} \int_{a_j}^{b_j} \left[\int_{a_k}^{b_k} \int_{a_k}^{x_k} |h((t_j)_j, s_k, (x_i)_i) \right. \right. \\ &\quad \left. \left. - \chi_k(s_k) \int_{a_k}^{b_k} h((t_j)_j, u_k, (x_i)_i) du_k \right) ds_k dx_k \right] (dt_j)_j \Big| (dx_i)_i \\ &\leq |b_k - a_k| \bigotimes_{i=k+1}^d \int_{a_i}^{b_i} \left[\bigotimes_{j=1}^{k-1} \int_{a_j}^{b_j} \left[\int_{a_k}^{b_k} |h((t_j)_j, s_k, (x_i)_i)| ds_k \right. \right. \\ &\quad \left. \left. + \int_{a_k}^{b_k} \chi_k(s_k) \left| \int_{a_k}^{b_k} h((t_j)_j, u_k, (x_i)_i) du_k \right| ds_k \right] (dt_j)_j \Big| (dx_i)_i \\ &= |b_k - a_k| \bigotimes_{i=k+1}^d \int_{a_i}^{b_i} \left[\bigotimes_{j=1}^{k-1} \int_{a_j}^{b_j} \left[\int_{a_k}^{b_k} |h((t_j)_j, s_k, (x_i)_i)| ds_k \right. \right. \\ &\quad \left. \left. + \left(\int_{a_k}^{b_k} \chi_k(s_k) ds_k \right) \left| \int_{a_k}^{b_k} h((t_j)_j, u_k, (x_i)_i) du_k \right| \right] (dt_j)_j \Big| (dx_i)_i \\ &\leq 2|b_k - a_k| \bigotimes_{j=1}^{k-1} \int_{a_j}^{b_j} \left[\int_{a_k}^{b_k} \left[\bigotimes_{i=k+1}^d \int_{a_i}^{b_i} |h((t_j)_j, s_k, (x_i)_i)| (dx_i)_i \right] ds_k \right] (dt_j)_j \\ &= 2|b_k - a_k| \int_Q |h(x)| dx. \end{aligned}$$

On procède de même pour φ_1 et φ_d . On a donc bien montré (E.0-2), à savoir que :

$$\|\varphi\|_{L^1(Q)} \leq 2 \text{diam}(Q) \|h\|_{L^1(Q)}.$$

Montrons pour finir que, pour tout $k \in \llbracket 1, d \rrbracket$, on a $\varphi_k \in C_c^\infty(Q)$. Soient $\alpha_1, \dots, \alpha_d, \beta_1, \dots, \beta_d \in \mathbb{R}$ tels que pour tout $p \in \llbracket 1, d \rrbracket$, $\text{supp}(\chi_p) \subset [\alpha_p, \beta_p]$, et $\text{supp}(h) \subset [\alpha_1, \beta_1] \times \dots \times [\alpha_d, \beta_d]$.

Soit $x = (x_1, \dots, x_d) \in]a_p, b_p[\times \dots \times]a_d, b_d[\setminus ([\alpha_1, \beta_1] \times \dots \times [\alpha_d, \beta_d])$. Il existe donc $k \in \llbracket 1, d \rrbracket$ tel que $x_p \in]a_p, b_p[\setminus [\alpha_p, \beta_p]$.

Ainsi, pour tout $l \in \llbracket 1, p - 1 \rrbracket$ (s'ils existent), on a $h(t_1, \dots, t_l, x_{l+1}, \dots, x_d) = 0$ pour tous $t_1, \dots, t_l \in]a_1, b_1[\times \dots \times]a_l, b_l[$.

1. On a ainsi :

$$\bigotimes_{i=1}^l \int_{a_i}^{b_i} h((t_i)_i, x_{l+1}, \dots, x_d) (dt_i)_i = \bigotimes_{i=1}^{l-1} \int_{a_i}^{b_i} \int_{a_l}^{b_l} h((t_i)_i, u_l, x_{l+1}, \dots, x_d) du_l (dt_i)_i = 0,$$

et

$$\bigotimes_{j=1}^{l-1} \int_{a_j}^{b_j} h((\tau_j)_j, s_l, x_{l+1}, \dots, x_d) (d\tau_j)_j = 0,$$

pour tout $s_l \in]a_l, b_l[$. D'où $\varphi_l(x) = 0$.

2. On a $x_p \in]a_p, b_p[\setminus [\alpha_p, \beta_p]$, donc $\chi_p(x_p) = 0$ et ainsi, pour tout $l \in \llbracket p+1, d \rrbracket$ (s'ils existent), on a $\varphi_l(x) = 0$.

3. Pour le cas de φ_p , distinguons deux cas :

— Pour tout $s_p \in]a_p, \alpha_p[\cup]\beta_p, b_p[$, on a $\chi_p(x_p) = 0$ et $h(t_1, \dots, t_{p-1}, s_p, x_{p+1}, \dots, x_d) = 0$ pour tous $t_1, \dots, t_{p-1} \in]a_1, b_1[\times \dots \times]a_{p-1}, b_{p-1}[$. On a donc immédiatement : si $x_p \in]a_p, \alpha_p[$ alors

$$\int_{a_p}^{x_p} \left(h((t_i)_i, s_p, x_{k+1}, \dots, x_d) - \chi_p(s_p) \int_{a_k}^{b_k} h((t_i)_i, u_p, x_{k+1}, \dots, x_d) du_p \right) ds_p = 0$$

(avec $h(x_{k+1}, \dots, x_d) = \emptyset$ si $p = d$), pour tous $t_1, \dots, t_{p-1} \in]a_1, b_1[\times \dots \times]a_{p-1}, b_{p-1}[$. D'où $\varphi_p(x) = 0$.

— Si au contraire $x_p \in]\beta_p, b_p[$, on a

$$\begin{aligned} & \int_{a_p}^{x_p} \left(h((t_i)_i, s_p, x_{k+1}, \dots, x_d) - \chi_p(s_p) \int_{a_p}^{b_p} h((t_i)_i, u_p, x_{k+1}, \dots, x_d) du_p \right) ds_p \\ &= \int_{a_p}^{b_p} \left(h((t_i)_i, s_p, x_{k+1}, \dots, x_d) - \chi_p(s_p) \int_{a_p}^{b_p} h((t_i)_i, u_p, x_{k+1}, \dots, x_d) du_p \right) ds_p \\ &= \int_{a_p}^{b_p} h((t_i)_i, s_p, x_{k+1}, \dots, x_d) ds_p \\ & \quad - \left(\int_{a_p}^{b_p} \chi_p(s_p) ds_p \right) \int_{a_p}^{b_p} h((t_i)_i, u_p, x_{k+1}, \dots, x_d) du_p \\ &= \int_{a_p}^{b_p} \left(h((t_i)_i, s_p, x_{k+1}, \dots, x_d) ds_p - \int_{a_p}^{b_p} h((t_i)_i, u_p, x_{k+1}, \dots, x_d) du_p \right) \\ &= 0, \end{aligned}$$

(E.0-11)

pour tous $t_1, \dots, t_{p-1} \in]a_1, b_1[\times \dots \times]a_{p-1}, b_{p-1}[$. D'où $\varphi_p(x) = 0$.

En conclusion, on a $\text{supp}(\varphi_k) \subset [\alpha_1, \beta_1] \times \dots \times [\alpha_d, \beta_d]$ et donc $\varphi_k \in C_c^\infty(Q)$ pour tout $k \in \llbracket 1, d \rrbracket$, c'est à dire $\varphi \in C_c^\infty(Q)^d$. \square

Nous sommes maintenant en mesure de démontrer le Lemme 4.1-11 : nous procéderons par recouvrement fini des compacts de Ω par des hyper-pavés, et nous pourrons appliquer le Lemme E.0-1 sur chacun d'eux par partition de l'unité.

Rappel de notation : On note $|\cdot|_\infty$ la norme sur \mathbb{R}^d définie par

$$|x|_\infty = \max\{|x_1|, \dots, |x_d|\}, \text{ pour } x = (x_1, x_2) \in \mathbb{R}^d.$$

Retour à la preuve du Lemme 4.1-11 : Soit $K \subset \Omega$ un compact. On pose $\text{Dist}(K, {}^c\Omega) =$

$\inf_{x \in K, y \in {}^c\Omega} |x - y|_\infty$. L'ensemble Ω étant ouvert, on a $\text{Dist}(K, {}^c\Omega) > 0$. On définit \mathcal{U}_K l'ensemble des hyper-pavés ouverts de diamètre strictement inférieur à $\text{Dist}(K, {}^c\Omega)$, autrement dit :

$$\mathcal{U}_K = \left\{ Q =]a_1, b_1[\times \dots \times]a_d, b_d[\subset \mathbb{R}^d, \forall k \in \llbracket 1, d \rrbracket \ a_k < b_k, \text{diam}(Q) < \text{Dist}(K, {}^c\Omega) \right\}.$$

\mathcal{U}_K est un recouvrement ouvert de K . Par compacité de K , il existe donc $n \in \mathbb{N}$ et $Q_1, \dots, Q_n \in \mathcal{U}_K$ tels que

$$K \subset \bigcup_{i=1}^n Q_i.$$

Pour tout $i \in \llbracket 1, n \rrbracket$, on définit comme dans l'énoncé du Lemme E.0-1 (p. 222) une fonction $\chi^i \in C_c^\infty(]0, 1[\times \Omega)$ à variables séparées sur Q_i , telle que $\text{supp}(\chi^i) \subset]0, 1[\times Q_i$ (avec $\int_0^1 \int_{Q_i} \chi^i dx dt = 1$). Soit la fonction $\xi \in C_c^\infty(\Omega)$ telle que $\text{supp}(\xi) \subset \cdot$. D'après le Théorème A.1-3 (p. 202), il existe des fonctions $\xi_1, \dots, \xi_n \in C_c^\infty(\Omega)$ telles que pour tout $i \in \llbracket 1, n \rrbracket$, on ait

$$0 \leq \xi_i \leq 1, \quad \text{supp}(\xi_i) \subset Q_i,$$

et telles qu'il existe un ouvert U avec $K \subset U \subset Q_1 \cup \dots \cup Q_n$ (\bar{U} est donc compact) et $\sum_{i=1}^n \xi_i = 1$.

Soit $h \in C_c^\infty(]0, 1[\times \Omega)$ telle que $\text{supp}(h) \subset]0, 1[\times U$. D'après le Lemme E.0-1, il existe des fonctions $\varphi^1, \dots, \varphi^n \in C_c^\infty(]0, 1[\times \Omega)^d$ telles que, pour tout $i \in \llbracket 1, n \rrbracket$:

1. $\text{supp}(\varphi^i) \subset]0, 1[\times Q_i$
2. $h\xi_i = \left(\int_0^1 \int_\Omega h\xi_i dx dt \right) \chi^i + \text{div}_{t,x}(\varphi^i),$
3. $\int_{(0,1) \times \Omega} |\varphi^i| \leq 2 \text{diam}(]0, 1[\times Q_i) \int_{(0,1) \times \Omega} |h| \xi_i$

(on a ici pris la liberté, par rapport au Lemme E.0-1, de prolonger les diverses fonctions en jeu par 0 sur tout Ω ou $]0, 1[\times \Omega$).

On a donc

$$\begin{aligned} \langle T, h \rangle &= \sum_{i=1}^n \langle T, h\xi_i \rangle \\ &= \sum_{i=1}^n \left(\int_0^1 \int_\Omega h\xi_i dx dt \right) \langle T, \chi^i \rangle + \langle T, \text{div}_{t,x}(\varphi^i) \rangle \\ &= \sum_{i=1}^n \left(\int_0^1 \int_\Omega h\xi_i dx dt \right) \langle T, \chi^i \rangle - \langle \nabla_{t,x} T, \varphi^i \rangle \\ &= \sum_{i=1}^n \left(\int_0^1 \int_\Omega h\xi_i dx dt \right) \langle T, \chi^i \rangle - \int_0^1 \int_\Omega g \cdot \varphi^i dx dt. \end{aligned} \tag{E.0-12}$$

Ainsi :

$$\begin{aligned}
|\langle T, h \rangle| &\leq \sum_{i=1}^n \left(\int_{(0,1) \times \Omega} |h| \xi_i \right) |\langle T, \chi^i \rangle| + \int_{(0,1) \times \Omega} |g| \cdot |\varphi^i| \\
&\leq \sum_{i=1}^n \left(\int_{(0,1) \times \Omega} |h| \right) |\langle T, \chi^i \rangle| + \|g\|_{L^\infty(]0,1[\times U)} \int_{(0,1) \times \Omega} |\varphi^i| \\
&\leq \sum_{i=1}^n \left(\int_{(0,1) \times \Omega} |h| \right) |\langle T, \chi^i \rangle| + 2\|g\|_{L^\infty(]0,1[\times U)} \cdot \text{diam}(]0,1[\times Q_i) \int_{(0,1) \times \Omega} |h| \xi_i \\
&\leq \left(\sum_{i=1}^n \left(\int_{(0,1) \times \Omega} |h| \right) |\langle T, \chi^i \rangle| \right) \\
&\quad + 2\|g\|_{L^\infty(]0,1[\times U)} \cdot \max\{1, \text{Dist}(K, {}^c\Omega)\} \int_{(0,1) \times \Omega} \left(\sum_{i=1}^n |h| \xi_i \right) \\
&\leq \left(2\|g\|_{L^\infty(]0,1[\times U)} \cdot \max\{1, \text{Dist}(K, {}^c\Omega)\} + \sum_{i=1}^n |\langle T, \chi^i \rangle| \right) \int_{(0,1) \times \Omega} |h|.
\end{aligned} \tag{E.0-13}$$

On peut donc conclure que, pour tout $h \in C_c^\infty(]0,1[\times \Omega)$ telle que $\text{supp}(h) \subset]0,1[\times U$, on a

$$|\langle T, h \rangle| \leq C_U \|h\|_{L^1((0,1) \times \Omega)}, \tag{E.0-14}$$

avec $C_U = 2\|g\|_{L^\infty(]0,1[\times U)} \cdot \max\{1, \text{Dist}(K, {}^c\Omega)\} + \sum_{i=1}^n |\langle T, \chi^i \rangle|$.

Soit $T_U : h \in C_c^\infty(]0,1[\times U) \mapsto \langle T, p_0(h) \rangle$, où p_0 désigne l'opérateur de prolongement de $]0,1[\times U$ à $]0,1[\times \Omega$ par 0, c'est à dire :

$$p_0 : h \in C_c^\infty(]0,1[\times U) \mapsto \left(x \in \Omega \mapsto \begin{cases} h(x), & \text{si } x \in]0,1[\times U \\ 0 & \text{sinon} \end{cases} \right) \in C_c^\infty(]0,1[\times \Omega).$$

On a donc l'estimation

$$|\langle T, h \rangle| \leq C_U \|h\|_{L^1((0,1) \times U)}, \tag{E.0-15}$$

pour tout $h \in C_c^\infty(]0,1[\times U)$. T_U est donc une forme linéaire sur $C_c^\infty(]0,1[\times U)$, continue pour la norme $\|\cdot\|_{L^1((0,1) \times U)}$. Ainsi, par densité de $C_c^\infty(]0,1[\times U)$ dans $L^1(]0,1[\times U)$ (toujours pour la norme $\|\cdot\|_{L^1((0,1) \times U)}$), T_U se prolonge en une forme linéaire continue \tilde{T}_U sur $L^1(]0,1[\times U)$, vérifiant toujours par prolongement l'estimation (E.0-15) pour tout $h \in L^1(]0,1[\times U)$.

D'après le Lemme A.1-4 (p. 202), il existe donc une fonction $f_U \in L^\infty((0,1) \times U)$, telle que pour tout $h \in L^1(]0,1[\times U)$, on ait

$$\tilde{T}_U(h) = \int_0^1 \int_U f_U h \, dx \, dt, \tag{E.0-16}$$

et donc telle que pour tout $h \in C_c^\infty(]0,1[\times \Omega)$ avec $\text{supp}(h) \subset]0,1[\times U$, on ait

$$\langle T, h \rangle = \tilde{T}_U(h|_U) = \int_0^1 \int_U f_U h|_U \, dx \, dt, \tag{E.0-17}$$

En écrivant Ω comme une réunion croissante de compacts (voir le Lemme F.0-1 p. 231), et en appliquant le raisonnement précédent sur chacun de ces compacts, on peut donc conclure à

l'existence d'une fonction $f \in L_{loc}^\infty([0, 1] \times \Omega)$ représentant la distribution T ; c'est à dire telle que pour tout $h \in C_c^\infty(]0, 1[\times \Omega)$, on ait :

$$\langle T, h \rangle = \int_0^1 \int_\Omega f h \, dx \, dt. \quad (\text{E.0-18})$$

Par définition de $\nabla_{t,x}T$ et de son représentant de classe \mathcal{L}^d -mesurable $g \in L_{loc}^\infty([0, 1] \times \Omega)$, on a ainsi

$$\int_0^1 \int_\Omega f \cdot \operatorname{div}(h) \, dx \, dt = \langle T, \operatorname{div}(h) \rangle = -\langle \nabla_{t,x}T, h \rangle = -\int_0^1 \int_\Omega g \cdot h \, dx \, dt,$$

pour tout $h \in C_c^\infty(]0, 1[\times \Omega)$. On a donc bien $f \in W_{loc}^{1,\infty}([0, 1] \times \Omega)$.

Pour finir, d'après le dernier point de la Proposition A.1-7, il existe un représentant continue de f (que nous notons également f par abus de notation) tel que, pour tout $x, y \in \Omega$,

$$|f(x) - f(y)| \leq 4\|\nabla f\|_{L^\infty(\Omega)}|x - y|.$$

□

Remarque E.0-1. Il serait possible de généraliser le Lemme 4.1-11 à tout $p \in]1, +\infty]$: c'est à dire que de montrer que $\nabla_{t,x}T \in L_{loc}^p([0, 1] \times \Omega)^d$ implique que $T \in W_{loc}^{1,p}([0, 1] \times \Omega)$, et ceci avec une démonstration similaire. Pour cela il faudrait montrer dans le Lemme E.0-1 une estimation du type $\|\varphi\|_{L^q} \leq C\|h\|_{L^q}$ (pour q tel que $1/p + 1/q = 1$) plutôt que $\|\varphi\|_{L^1} \leq C\|h\|_{L^1}$.

Annexe F

Étude de la notion de poussée de mesure dans le cadre du transport optimal

Dans cette annexe, nous étudierons en détail la notion de poussée de mesure, plus précisément les propriétés utiles à notre étude qui en découle dans le cadre du transport optimal et de l'interpolation de McCann, à savoir principalement les Propositions [F.0-3](#) et [F.0-6](#). Comme nous ferons souvent appel aux propriétés [\(IV\)](#) p. [203](#), nous nous en rappelons l'énoncé ci-dessous.

Propriétés (IV).

Soit Ω un ouvert de \mathbb{R}^d , et μ une mesure sur la tribu de Lebesgue de Ω .

1. $\mu(K) < +\infty$ pour tout compact $K \subset \Omega$.
2. Pour tout ensemble Lebesgue-mesurable E de Ω , on a

$$\mu(E) = \inf\{\mu(V), E \subset V, V \text{ ouvert}\}.$$

3. Pour tout E ouvert et pour tout ensemble Lebesgue-mesurable E de Ω tel que $\mu(E) < +\infty$,

$$\mu(E) = \sup\{\mu(K), K \subset E, K \text{ compact}\}.$$

Commençons par définir plus rigoureusement la notion de poussée de mesure (Proposition [F.0-3](#) plus bas). Pour cela nous aurons besoin des deux Lemmes suivant : les Lemmes [F.0-1](#) et [F.0-2](#).

Lemme F.0-1.

Soit V un ouvert de \mathbb{R}^d . Il existe une suite $(K_n)_{n \in \mathbb{N}}$ de compacts de \mathbb{R}^d telle que pour tout $n \in \mathbb{N}$, $K_n \subset K_{n+1}$, et

$$V = \bigcup_{n \in \mathbb{N}} K_n.$$

Preuve : Supposons dans un premier temps que V est borné. Son adhérence \bar{V} est donc un compact de \mathbb{R}^d . Donc, d'après la propriété de Borel-Lebesgue, pour tout $n \in \mathbb{N}$, il existe $k_n \in \mathbb{N}$

et $x_0^n, \dots, x_{k_n}^n \in \bar{V}$ tel que

$$\bar{V} \subset \bigcup_{i=0..k_n} \bar{B}\left(x_i^n, \frac{1}{n+1}\right) \quad (\text{F.0-1})$$

(où $\bar{B}(x, r)$ désigne la boule fermée de centre x et de rayon r). Pour tout $n \in \mathbb{N}$, on pose

$$A_n = \left\{ i \in \llbracket 0, k_n \rrbracket, \bar{B}\left(x_i^n, \frac{1}{n+1}\right) \subset V \right\} \quad \text{et} \quad K_n = \bigcup_{i \in A_n} \bar{B}\left(x_i^n, \frac{1}{n+1}\right) \subset V$$

(K_n étant compact du fait que A_n est fini). Montrons que V est la réunion des K_n . Tout d'abord, il est clair que

$$\bigcup_{n \in \mathbb{N}} K_n \subset V.$$

Soient $x \in V$ et $r > 0$ tels que $B(x, r) \subset V$ ($B(x, r)$ étant la boule ouverte de centre x et de rayon r), et soit $n \in \mathbb{N}$ tel que $1/(n+1) < r/2$. D'après (F.0-1), il existe donc x_i^n tel que $x \in \bar{B}(x_i^n, 1/(n+1))$. Ainsi pour tout $y \in \bar{B}(x_i^n, 1/(n+1))$, on a

$$|x - y| \leq |x - x_i^n| + |x_i^n - y| \leq \frac{1}{n+1} + \frac{1}{n+1} < r,$$

donc

$$\bar{B}\left(x_i^n, \frac{1}{n+1}\right) \subset B(x, r) \subset V,$$

d'où $i \in A_n$ et $\bar{B}(x_i^n, 1/(n+1)) \subset K_n$. Il s'en suit que $x \in K_n$, et on peut ainsi conclure que

$$V \subset \bigcup_{n \in \mathbb{N}} K_n.$$

Pour finir, si l'on ne suppose plus que V est borné, on peut toujours se ramener au fait que V peut s'écrire comme une réunion dénombrable d'ouverts bornés (par exemple la réunion des $B(0, n) \cap V$). Donc en tant que réunion dénombrable de réunions dénombrables de compacts, V est la réunion d'une famille dénombrable $(K_n)_{n \in \mathbb{N}}$ de compacts de \mathbb{R}^d .

Quitte à poser pour tout $n \in \mathbb{N}$,

$$K'_n = \bigcup_{j=0}^n K_j,$$

on peut considérer que cette famille est croissante au sens de l'inclusion. □

Lemme F.0-2.

Soit K un compact de \mathbb{R}^d . Il existe une suite $(V_n)_{n \in \mathbb{N}}$ d'ouverts de \mathbb{R}^d telle que pour tout $n \in \mathbb{N}$, $V_{n+1} \subset V_n$, et

$$K = \bigcap_{n \in \mathbb{N}} V_n.$$

Il existe de plus un ouvert borné V' de \mathbb{R}^d tel que pour tout $n \in \mathbb{N}$, $V_n \subset V'$.

Preuve : De part la propriété de Borel-Lebesgue, pour tout $n \in \mathbb{N}$, il existe $k_n \in \mathbb{N}$ et $x_0^n, \dots, x_{k_n}^n \in K$ tel que

$$K \subset \bigcup_{i=0..k_n} B\left(x_i^n, \frac{1}{n+1}\right) \quad (\text{F.0-2})$$

(où $B(x, r)$ désigne la boule ouverte de centre x et de rayon r). On pose ainsi, pour tout $n \in \mathbb{N}$

$$V_n = \bigcup_{i=0..k_n} B\left(x_i^n, \frac{1}{n+1}\right).$$

Il est clair que

$$K \subset \bigcap_{n \in \mathbb{N}} V_n.$$

Soit $x \in \bigcap_{n \in \mathbb{N}} V_n$. Pour tout $n \in \mathbb{N}$, il existe $i_n \in \llbracket 0, k_n \rrbracket$ tel que $x \in B\left(x_{i_n}^n, \frac{1}{n+1}\right)$.

La suite $(x_{i_n}^n)_{n \in \mathbb{N}}$ (d'éléments de K) converge donc vers x , et donc $x \in K$. D'où

$$\bigcap_{n \in \mathbb{N}} V_n \subset K.$$

Quitte à poser pour tout $n \in \mathbb{N}$

$$V'_n = \bigcap_{j=0}^n V_j,$$

on peut considérer que la famille $(V_n)_{n \in \mathbb{N}}$ est décroissante au sens de l'inclusion.

On Remarquera en outre que pour tout $n \in \mathbb{N}$, $V_n \subset V'$, V' étant l'ouvert borné défini par

$$V' = \bigcup_{x \in K} B(x, 1).$$

□

Proposition F.0-3.

On munit \mathbb{R}^d de la tribu de Lebesgue. Soient μ une mesure positive σ -finie sur la de Lebesgue, et $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$ mesurable (toujours sur la tribu de Lebesgue). On suppose que l'une des deux hypothèses suivantes est vérifiée :

1. La mesure μ est finie,
2. L'application T est coercive par rapport à μ , i.e. $T^{-1}(A)$ est borné μ -presque partout pour tout ensemble Lebesgue-mesurable borné A dans \mathbb{R}^d .

Alors il existe une mesure positive sur la tribu de Lebesgue ν sur \mathbb{R}^d , vérifiant les propriétés (IV) (voir le début de l'annexe), telle que

$$\forall f \in C_c^0(\mathbb{R}^d), \quad \int_{\mathbb{R}^d} f d\nu = \int_{\mathbb{R}^d} f(Tx) d\mu \quad (\text{F.0-3})$$

De plus, pour tout ensemble Lebesgue-mesurable $A \subset \mathbb{R}^d$, on a $\nu(A) = \mu(T^{-1}(A))$.

On dit alors que ν est la **poussée de la mesure** μ par l'opérateur T , noté

$$\nu = T\#\mu.$$

Preuve : L'existence d'une mesure positive sur la tribu de Lebesgue ν sur Ω , vérifiant les propriétés (IV), et la relation (F.0-3) provient de l'application du Théorème A.2-1 (p. 204), à la forme linéaire positive sur $C_c^0(\mathbb{R}^d)$ définie par

$$f \mapsto \int_{\mathbb{R}^d} f \, d\nu = \int_{\mathbb{R}^d} f(Tx) \, d\mu.$$

Cette forme linéaire est bien définie si l'une des deux hypothèses de l'énoncé est vérifiée (μ est finie ou T est coercive).

La partie moins triviale de cette preuve concernera donc le dernier point de l'énoncé. Nous montrerons tout d'abord la propriété pour les ouverts bornés et les compacts de \mathbb{R}^d , puis nous généraliserons à tous les ensembles Lebesgue-mesurables.

1. Soit V un ouvert borné de \mathbb{R}^d . D'après le Lemme F.0-1 (p. 231), il existe une suite $(K_n)_{n \in \mathbb{N}}$ de compacts, croissante au sens de l'inclusion, tel que V soit la réunion des K_n . Pour tout $n \in \mathbb{N}$, le Théorème A.1-2 (p. 201) nous donne l'existence d'un $f_n \in C_c^0(\Omega)$ tel que $\mathbb{1}_{K_n} \leq f_n \leq \mathbb{1}_V$. La suite d'indicatrices $(\mathbb{1}_{K_n})_n$ convergent simplement vers $\mathbb{1}_V$, il en est donc de même pour la suite $(f_n)_n$, et ainsi pour la suite $(f_n \circ T)_n$ vers $\mathbb{1}_V \circ T$. D'après le premier point des propriétés (IV), on a $\nu(V) \leq \nu(\overline{V}) < +\infty$. Soit V' un ouvert borné de \mathbb{R}^d tel que $\overline{V} \subset V'$. D'après le Théorème A.1-2, il existe une fonction $g \in C_c^0(\Omega)$ tel que $\mathbb{1}_{\overline{V}} \leq g \leq \mathbb{1}_{V'}$. D'où

$$f_n \circ T \leq \mathbb{1}_V \circ T \leq \mathbb{1}_{\overline{V}} \circ T \leq g \circ T$$

pour tout $n \in \mathbb{N}$, et

$$\int_{\mathbb{R}^d} g(Tx) \, d\mu = \int_{\mathbb{R}^d} g \, d\nu \leq \nu(\overline{V'}) < +\infty.$$

Ainsi par convergence dominé, on a

$$\begin{aligned} \nu(V) &= \int_{\mathbb{R}^d} \mathbb{1}_V \, d\nu = \lim_{n \rightarrow +\infty} \int_{\mathbb{R}^d} f_n \, d\nu = \lim_{n \rightarrow +\infty} \int_{\mathbb{R}^d} f_n(Tx) \, d\mu \\ &= \int_{\mathbb{R}^d} \mathbb{1}_V(Tx) \, d\mu = \mu(T^{-1}(V)). \end{aligned} \quad (\text{F.0-4})$$

2. Soit K un compact de \mathbb{R}^d . On raisonne comme dans le point précédent : le Lemme F.0-2 nous donne l'existence d'une suite $(V_n)_{n \in \mathbb{N}}$ d'ouverts, décroissante au sens de l'inclusion, tel que K soit l'intersection des V_n , ainsi que d'un ouvert borné V' tel que pour tout $n \in \mathbb{N}$, $V_n \subset V'$. De plus pour tout $n \in \mathbb{N}$, le Théorème A.1-2 nous donne l'existence d'un $f_n \in C_c^0(\Omega)$ tel que $\mathbb{1}_K \leq f_n \leq \mathbb{1}_{V_n} \leq \mathbb{1}_{V'}$. La suite d'indicatrices $(\mathbb{1}_{V_n})_n$ convergent simplement vers $\mathbb{1}_K$, et ainsi donc de même pour la suite $(f_n)_n$, et pour la suite $(f_n \circ T)_n$ vers $\mathbb{1}_K \circ T$, avec la domination $f_n \circ T \leq \mathbb{1}_{V'} \circ T$, avec d'après le point précédent

$$\int_{\mathbb{R}^d} \mathbb{1}_{V'}(Tx) \, d\mu = \int_{\mathbb{R}^d} \mathbb{1}_{V'} \, d\nu \leq \nu(V') \leq \nu(\overline{V'}) < +\infty.$$

On conclut alors par convergence dominé que $\nu(K) = \mu(T^{-1}(K))$.

3. Soit A un ensemble Lebesgue-mesurable borné de \mathbb{R}^d . ν vérifie les propriétés (IV), donc en particulier par les points 2 et 3 :

$$\nu(A) = \inf\{\nu(V), A \subset V, V \text{ ouvert borne}\}$$

et

$$\nu(A) = \sup\{\nu(K), K \subset A, K \text{ compact}\}. \quad (\text{F.0-5})$$

(On peut considérer que l'on minimise sur les ouverts bornés du fait que A est supposé borné). Pour tous compacts K et ouverts V tel que $K \subset A \subset V$, on a

$$\nu(K) = \mu(T^{-1}(K)) \leq \mu(T^{-1}(A)) \leq \mu(T^{-1}(V)) = \nu(V),$$

d'où

$$\begin{aligned} \nu(A) &= \sup\{\nu(K), K \subset A, K \text{ compact}\} \\ &= \sup\{\mu(T^{-1}(K)), K \subset A, K \text{ compact}\} \leq \mu(T^{-1}(A)) \\ &\leq \inf\{\mu(T^{-1}(V)), A \subset V, V \text{ ouvert borne}\} \\ &= \inf\{\nu(V), A \subset V, V \text{ ouvert borne}\} = \nu(A). \end{aligned} \quad (\text{F.0-6})$$

d'où $\nu(A) = \mu(T^{-1}(A))$. Pour finir, si l'on considère maintenant que A est un ensemble Lebesgue-mesurable quelconque, A peut s'écrire comme une réunion dénombrable $(A_n)_n$ d'ensembles Lebesgue-mesurables bornés (par exemple la réunion des $A_n = B(0, n) \cap A$). D'où

$$\begin{aligned} \nu(A) &= \lim_{n \rightarrow +\infty} \nu\left(\bigcup_{k=0}^n A_k\right) = \lim_{n \rightarrow +\infty} \mu\left(T^{-1}\left(\bigcup_{k=0}^n A_k\right)\right) = \lim_{n \rightarrow +\infty} \mu\left(\bigcup_{k=0}^n T^{-1}(A_k)\right) \\ &= \mu\left(\bigcup_{k \in \mathbb{N}} T^{-1}(A_k)\right) = \mu\left(T^{-1}\left(\bigcup_{k \in \mathbb{N}} A_k\right)\right) = \mu(T^{-1}(A)). \end{aligned} \quad (\text{F.0-7})$$

□

Passons maintenant à l'interpolation de McCann : la Proposition F.0-6 plus bas. Comme nous souhaitons toujours travailler sous l'hypothèse (Γ_1) , en particulier lorsque nous traitons des interpolés du potentiel de transport ϕ , énonçons avant tout le petit Lemme suivant.

Lemme F.0-4.

Soit $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ vérifiant la propriété (Γ_1) . Alors, pour tout $t \in [0, 1]$, $t\phi + (1-t)|\cdot|^2/2$ vérifie également la propriété (Γ_1) .

Preuve : Évidente (rappelons au passage que $(t\phi + (1-t)|\cdot|^2/2)^* = 1-t(t\phi)$ pour tout $t \in [0, 1]$).

□

Avant d'énoncer la Proposition F.0-6, énonçons le Lemme suivant, qui sera nécessaire d'invoquer dans la preuve de la première. De façon abusive, pour f et g deux fonctions mesurables, nous noterons souvent $f = T\#g$ au lieu de $f\mathcal{L}^d = T\#(g\mathcal{L}^d)$.

Lemme F.0-5.

Soit $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ vérifiant la propriété (Γ_1) , et \mathbf{p} défini par rapport à ϕ à la définition 3.3-9 (p. 89), c'est à dire pour rappel :

$$\mathbf{p}(t, \cdot) = \mathbf{p}_\phi(t, \cdot) = \text{Prox}_{\frac{t}{1-t}\phi}\left(\frac{x}{1-t}\right) = \nabla_x(\phi_t)^*$$

où $(\phi_t)^*$ désigne la transformée de Legendre de $\phi_t = (1-t)|\cdot|^2/2 + t\phi$.

Pour tout $r > 0$, la fonction

$$\mathcal{G} : (t, x, y) \mapsto \mathbf{1}_{\mathbf{p}(t, \overline{B(x, r)})}(y)$$

est Lebesgue-mesurable sur $[0, 1[\times \mathbb{R}^d \times \mathbb{R}^d$.

Preuve : Soient $r > 0$. Remarquons tout d'abord que $\mathcal{G} = \mathbf{1}_F$ avec $F = \{(t, x, y) \in [0, 1[\times \mathbb{R}^d \times \mathbb{R}^d, y \in \mathbf{p}(t, \overline{B(x, r)})\}$. Ainsi, pour prouver que \mathcal{G} est Lebesgue-mesurable, il suffit de prouver que F est un élément de la tribu de Lebesgue. Nous allons prouver plus précisément que F est fermé dans $[0, 1[\times \mathbb{R}^d \times \mathbb{R}^d$. Soit $(t_n, x_n, y_n)_n$ une suite d'éléments de F convergeant vers $(t^*, x^*, y^*) \in [0, 1[\times \mathbb{R}^d \times \mathbb{R}^d$. Montrons que $(t^*, x^*, y^*) \in F$. Pour tout $n \in \mathbb{N}$,

$$\begin{aligned} (t_n, x_n, y_n) \in F &\Leftrightarrow y_n \in \mathbf{p}\left(t_n, \overline{B(x_n, r)}\right) \\ &\Rightarrow \exists z_n \in \overline{B(x_n, r)}, y_n = \mathbf{p}(t_n, z_n) \\ &\Leftrightarrow \exists z_n \in \overline{B(x_n, r)}, z_n \in t_n \partial\phi(y_n) + (1-t_n)y_n \\ &\Leftrightarrow \exists z_n \in \overline{B(x_n, r)}, \exists \xi_n \in \partial\phi(y_n), z_n = t_n \xi_n + (1-t_n)y_n. \end{aligned} \tag{F.0-8}$$

On considère donc des suites $(z_n)_n$ et $(\xi_n)_n$ telles que définies dans la relation (F.0-8). La suite $(y_n)_n$ est convergente, donc borné. Soit ω un ouvert borné contenant tous les y_n pour $n \in \mathbb{N}$. ϕ vérifie la propriété (Γ_1) , donc en particulier ϕ est localement lipschitzienne sur \mathbb{R}^d , donc lipschitzienne sur ω . D'après le Lemme 3.3-1 (p. 86), $\partial\phi(\omega)$ est borné dans \mathbb{R}^d : la suite $(\xi_n)_n$ est donc bornée dans \mathbb{R}^d , il est donc de même pour la suite $(z_n)_n$. Quitte à devoir extraire une sous-suite on peut donc considérer que $(\xi_n)_n$ et $(z_n)_n$ convergent dans \mathbb{R}^d . On notes ξ^* et z^* leurs limites respectives. D'après le Lemme 3.3-3 (p. 87), on a $\xi^* \in \partial\phi(y^*)$, d'où

$$z^* = t^* \xi^* + (1-t^*)y^* \in t^* \partial\phi(y^*) + (1-t^*)y^* \Rightarrow y^* = \mathbf{p}(t^*, z^*).$$

De plus, pour tout $n \in \mathbb{N}$, on a $\|x_n - z_n\| \leq r$. Par passage à la limite, on a donc $\|x^* - z^*\| \leq r$, i.e. $z^* \in \overline{B(x^*, r)}$. Ainsi, $y^* \in \mathbf{p}(t^*, \overline{B(x^*, r)})$, i.e. $(t^*, x^*, y^*) \in F$. \square

Proposition F.0-6.

Soient $\rho_0 \in L^1(\mathbb{R}^d, d\mathcal{L}^d)$ à support compact, tel que $\rho_0 \geq 0$, et $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ vérifiant la propriété (Γ_1) .

Alors, pour tout $t \in [0, 1[$, il existe une mesure positive ν_t sur la tribu de Lebesgue de \mathbb{R}^d , à support dans $(t\nabla\phi + (1-t)\text{id})(\text{supp}(\rho_0))$ (donc borné), vérifiant les propriétés (IV) (voir le début de l'annexe), telle que $\nu_t = (t\nabla\phi + (1-t)\text{id})\#(\rho_0\mathcal{L}^d)$, i.e.

$$\forall f \in C_c^0(\mathbb{R}^d), \quad \int_{\mathbb{R}^d} f d\nu_t = \int_{\mathbb{R}^d} f(t\nabla\phi(x) + (1-t)x) \rho_0(x) d\mathcal{L}^d(x) \tag{F.0-9}$$

De plus, ν_t est absolument continue par rapport à la mesure de Lebesgue \mathcal{L}^d sur \mathbb{R}^d (i.e. $\nu_t \ll \mathcal{L}^d$), autrement dit il existe $\rho_t \in L^1(\mathbb{R}^d, \mathcal{L}^d)$ tel que $\nu_t = \rho_t \mathcal{L}^d$. Il est en outre possible, pour tout $t \in [0, 1[$, de choisir un représentant de ρ_t de tel façon à ce que $(t, x) \mapsto \rho_t(x)$ soit mesurable sur $[0, 1[\times \mathbb{R}^d$.

Pour finir, les propriétés suivantes sont vérifiées :

1. Pour tout $h \in L_{loc}^\infty(\mathbb{R}^d, \mathcal{L}^d)$, on a
 - $h \circ (t\nabla\phi + (1-t)\text{id}) \in L_{loc}^\infty(\mathbb{R}^d, \mathcal{L}^d)$,
 - $\int_{\mathbb{R}^d} h d\nu_t = \int_{\mathbb{R}^d} h(x)\rho_t(x) d\mathcal{L}^d(x) = \int_{\mathbb{R}^d} h(t\nabla\phi(x) + (1-t)x)\rho_0(x) d\mathcal{L}^d(x)$.
2. Pour tout $h \in L_{loc}^\infty([0, 1[\times\mathbb{R}^d, \mathcal{L}^{d+1})$, on a
 - $(t, x) \mapsto h(t, t\nabla\phi(x) + (1-t)x) \in L_{loc}^\infty([0, 1] \times \mathbb{R}^d)$,
 - $\int_0^1 \int_{\mathbb{R}^d} h(t, x)\rho_t(x) d\mathcal{L}^d(x) dt = \int_0^1 \int_{\mathbb{R}^d} h(t, t\nabla\phi(x) + (1-t)x)\rho_0(x) dx dt$.
3. Pour tout $h \in C_c^0([0, 1[\times\mathbb{R}^d)$, $t \mapsto \int_{\mathbb{R}^d} h(t, \cdot) d\nu_t = \int_{\mathbb{R}^d} h \rho_t d\mathcal{L}^d$ est continue sur $[0, 1[$, autrement dit $t \mapsto \nu_t$ est continue sur $[0, 1[$ dans $\mathcal{D}'(\mathbb{R}^d)$.

Preuve : Pour tout $t \in [0, 1[$, on applique la Proposition F.0-3 (p. 233) pour $\mu = \rho_0\mathcal{L}^d$ et $T = t\nabla\phi + (1-t)\text{id}$. Ici, du fait que $\rho_0 \in L^1(\mathbb{R}^d, d\mathcal{L}^d)$, $\rho_0\mathcal{L}^d$ est une mesure finie. La Proposition F.0-3 nous donne donc l'existence pour tout $t \in [0, 1[$ d'une mesure positive ν_t sur la tribu de Lebesgue de \mathbb{R}^d (définie comme dans le Théorème A.2-1 p. 204), vérifiant les propriétés (IV), et la relation (F.0-9) :

$$\forall f \in C_c^0(\mathbb{R}^d), \quad \int_{\mathbb{R}^d} f d\nu_t = \int_{\mathbb{R}^d} f(t\nabla\phi(x) + (1-t)x)\rho_0(x) d\mathcal{L}^d(x)$$

Ainsi, pour tout $t \in [0, 1[$ et pour tout $f \in C_c^0(\mathbb{R}^d)$ tel que :

$$\text{supp}(f) \subset \mathbb{R}^d \setminus (t\nabla\phi + (1-t)\text{id})(\text{supp}(\rho_0)),$$

on a

$$\int_{\mathbb{R}^d} f d\nu_t = \int_{\mathbb{R}^d} f(t\nabla\phi(x) + (1-t)x)\rho_0(x) d\mathcal{L}^d(x) = 0.$$

Donc $\text{supp}(\nu_t) \subset (t\nabla\phi + (1-t)\text{id})(\text{supp}(\rho_0))$ et, $\text{supp}(\rho_0)$ étant borné, il en est de même pour $(t\nabla\phi + (1-t)\text{id})(\text{supp}(\rho_0))$ en vertu du Lemme 3.3-2 (p. 86).

De plus, pour tout ensemble Lebesgue-mesurable A de \mathbb{R}^d , on a

$$\nu_t(A) = (\rho_0\mathcal{L}^d) \left((t\nabla\phi + (1-t)\text{id})^{-1}(A) \right) = (\rho_0\mathcal{L}^d) (\mathbf{p}(t, A)) = \int_{\mathbf{p}(t, A)} \rho_0 d\mathcal{L}^d(x) \quad (\text{F.0-10})$$

En effet, rappelons que pour tout $x \in \mathbb{R}^d$ où ϕ est différentiable, on a d'après (3.3-10),

$$x = \mathbf{p}(t, y) \Leftrightarrow y = (1-t)x + t\nabla\phi(x).$$

Or pour tout $t \in [0, 1[$, $\mathbf{p}(t, \cdot)$ est lipschitzien (d'après la définition 3.3-9 p. 89), donc d'après le Lemme 3.3-21 (p. 103), si $\mathcal{L}^d(A) = 0$, alors $\mathcal{L}^d(\mathbf{p}(t, A)) = 0$, d'où $\nu_t(A) = \int_{\mathbf{p}(t, A)} \rho_0 d\mathcal{L}^d(x) = 0$. La mesure ν_t est donc bien absolument continue par rapport à la mesure de Lebesgue \mathcal{L}^d .

Pour tout $t \in [0, 1[$ et tout $x \in \mathbb{R}^d$, on pose

$$\rho_t(x) = (\overline{D\nu_t})(x) = \limsup_{r \rightarrow 0} \frac{1}{\mathcal{V}_d(B_1)r^d} \nu_t(B(x, r)), \quad r > 0 \quad (\text{F.0-11})$$

(telle décrite dans le Théorème A.2-6 p. 206).

Ici, d'après (F.0-10), pour tous $r > 0$ et $(t, x) \in [0, 1[\times \mathbb{R}^d$,

$$\nu_t(B(x, r)) = \int_{\mathbf{p}(t, B(x, r))} \rho_0 d\mathcal{L}^d = \int_{\mathbb{R}^d} \mathbb{1}_{\mathbf{p}(t, B(x, r))}(y) \rho_0(y) d\mathcal{L}^d(y).$$

Or, remarquons que pour tout $(t, x) \in [0, 1[\times \mathbb{R}^d$,

$$\mathcal{L}^d(\mathbf{p}(t, B(x, r))) \leq \mathcal{L}^d\left(\mathbf{p}\left(t, \overline{B(x, r)}\right)\right) \leq \mathcal{L}^d(\mathbf{p}(t, B(x, r))) + \mathcal{L}^d(\mathbf{p}(t, S(x, r))).$$

avec $S(x, r) = \overline{B(x, r)} \setminus B(x, r)$. Or $\mathcal{L}^d(S(x, r)) = 0$, et $\mathbf{p}(t, \cdot)$ est lipschitzien sur \mathbb{R}^d , donc d'après le Lemme 3.3-21 (p. 103), $\mathcal{L}^d(\mathbf{p}(t, S(x, r))) = 0$. Donc

$$\mathcal{L}^d(\mathbf{p}(t, B(x, r))) = \mathcal{L}^d\left(\mathbf{p}\left(t, \overline{B(x, r)}\right)\right), \quad \text{et alors} \quad \mathbb{1}_{\mathbf{p}(t, B(x, r))}(y) = \mathbb{1}_{\mathbf{p}(t, \overline{B(x, r)})}(y)$$

pour \mathcal{L}^d -presque tout $y \in \mathbb{R}^d$, d'où

$$\nu_t(B(x, r)) = \int_{\mathbb{R}^d} \mathbb{1}_{\mathbf{p}(t, B(x, r))}(y) \rho_0(y) d\mathcal{L}^d(y) = \int_{\mathbb{R}^d} \mathbb{1}_{\mathbf{p}(t, \overline{B(x, r)})}(y) \rho_0(y) d\mathcal{L}^d(y).$$

D'après le Lemme F.0-5 (p. 235), la fonction $(t, x, y) \mapsto \mathbb{1}_{\mathbf{p}(t, \overline{B(x, r)})}(y)$ est Lebesgue-mesurable sur $[0, 1[\times \mathbb{R}^d \times \mathbb{R}^d$. D'après le Théorème A.2-4 (p. 205), la fonction

$$(t, x) \mapsto \nu_t(B(x, r)) = \int_{\mathbb{R}^d} \mathbb{1}_{\mathbf{p}(t, \overline{B(x, r)})}(y) \rho_0(y) d\mathcal{L}^d(y)$$

est donc mesurable sur $[0, 1[\times \mathbb{R}^d$. Ainsi, par passage à la limite supérieure simple d'une suite dénombrable (qui pourrait l'être) de fonctions mesurables, la fonction $(t, x) \mapsto \rho_t(x)$ est **mesurable** sur $[0, 1[\times \mathbb{R}^d$. Pour finir, d'après le Théorème A.2-4, pour tout $t \in [0, 1[$ la fonction $x \mapsto \rho_t(x)$ est mesurable sur \mathbb{R}^d et, d'après le Théorème A.2-6 (p. 206), $\nu_t = \rho_t \mathcal{L}^d$, i.e.

$$\forall t \in [0, 1[, \nu_t(E) = \int_E \rho_t(x) d\mathcal{L}^d(x),$$

pour tout E sous-ensemble de \mathbb{R}^d Lebesgue-mesurable. Pour finir,

$$\int_{[0, 1[\times \mathbb{R}^d} \rho_t(x) (dx \otimes dt) = \int_0^1 \int_{\mathbb{R}^d} \rho_t(x) dx dt = \int_0^1 \nu_t(\mathbb{R}^d) dt = \int_{\mathbb{R}^d} \rho_0 d\mathcal{L}^d < +\infty.$$

En effet, d'après la définition 3.3-9 (p. 89), pour tout $t \in [0, 1[$, $\mathbf{p}(t, \cdot)$ est surjective de \mathbb{R}^d dans \mathbb{R}^d , d'où d'après (F.0-10) :

$$\nu_t(\mathbb{R}^d) = \int_{\mathbf{p}(t, \mathbb{R}^d)} \rho_0 d\mathcal{L}^d = \int_{\mathbb{R}^d} \rho_0 d\mathcal{L}^d.$$

Ainsi $(t, x) \mapsto \rho_t(x) \in L^1([0, 1[\times \mathbb{R}^d)$. L'unicité d'un tel $(t, x) \mapsto \rho_t(x)$ découle du Corollaire A.2-11 (p. 207).

Vérifions maintenant les propriétés définies dans l'énoncé :

1. Soit $h \in L_{loc}^\infty(\mathbb{R}^d, \mathcal{L}^d)$,

- Soient V un ouvert borné de \mathbb{R}^d et $t \in [0, 1[$. L'ensemble image $(t\nabla\phi + (1-t)\text{id})(V)$ est borné toujours en vertu du Lemme 3.3-2 (p. 86) et fait que ϕ vérifie la propriété (Γ_1) . Il existe donc une constante $M > 0$ telle que, en posant

$$A = \{y \in (t\nabla\phi + (1-t)\text{id})(V), |h(y)| > M\},$$

on ait $\mathcal{L}^d(A) = 0$. Or pour tout $x \in V$ tel que ϕ soit différentiable en x (rappelons que l'ensemble des points pour lesquels ce n'est pas le cas est de mesure nulle),

$$|h(t\nabla\phi(x) + (1-t)x)| > M \Leftrightarrow t\nabla\phi(x) + (1-t)x \in A \Leftrightarrow x \in \mathbf{p}(t, A)$$

Ici $\mathcal{L}^d(\mathbf{p}(t, A)) = 0$ (toujours en vertu du Lemme 3.3-21 p. 103, du fait que $\mathbf{p}(t, \cdot)$ est lipschitzienne sur \mathbb{R}^d). Donc $|h(t\nabla\phi(x) + (1-t)x)| \leq M$ pour \mathcal{L}^d -presque tout $(t, x) \in V$, d'où $h \circ (t\nabla\phi + (1-t)\text{id}) \in L_{loc}^\infty(\mathbb{R}^d, \mathcal{L}^d)$.

- Soient toujours V un ouvert borné de \mathbb{R}^d et $t \in [0, 1[$. ρ_0 étant supposé à support compact, si l'on suppose de plus que $\text{supp}(\rho_0) \subset V$, d'après le Corollaire A.2-10 p. 207, il existe une suite $(h_n)_n \in C_c(\mathbb{R}^d)^\mathbb{N}$, avec $|h_n| \leq M$ pour tout $n \in \mathbb{N}$, convergent \mathcal{L}^d -presque partout vers $h \cdot \mathbf{1}_{(t\nabla\phi + (1-t)\text{id})(V)}$.

L'ensemble B des points de $(t\nabla\phi + (1-t)\text{id})(V)$ tels que $(h_n(x))_n$ ne converge pas vers $h(x)$ ou ne soit pas borné en norme euclidienne par M est donc de mesure nulle pour \mathcal{L}^d .

En raisonnant comme dans le point précédent, on peut montrer que l'ensemble des points x de V où ϕ est différentiable, et tels que $(h_n(t\nabla\phi(x) + (1-t)x))_n$ ne converge pas vers $h(t\nabla\phi(x) + (1-t)x)$ et ne soit pas borné en norme euclidienne par M est inclus dans $\mathbf{p}(t, B)$, ensemble de mesure pour \mathcal{L}^d , ceci en vertu du Lemme 3.3-21 (p. 103) et du fait que $\mathbf{p}(t, \cdot)$ est lipschitzienne sur \mathbb{R}^d . D'après la relation (F.0-9), on a pour tout $n \in \mathbb{N}$,

$$\int_{\mathbb{R}^d} h_n d\nu_t = \int_{\mathbb{R}^d} h_n(x)\rho_t(x) d\mathcal{L}^d(x) = \int_{\mathbb{R}^d} h_n(t\nabla\phi(x) + (1-t)x)\rho_0(x) d\mathcal{L}^d(x).$$

Par convergence dominée, on obtient

$$\int_{\mathbb{R}^d} h d\nu_t = \int_{\mathbb{R}^d} h(x)\rho_t(x) d\mathcal{L}^d(x) = \int_{\mathbb{R}^d} h(t\nabla\phi(x) + (1-t)x)\rho_0(x) d\mathcal{L}^d(x),$$

et ceci pour tout $h \in L_{loc}^\infty(\mathbb{R}^d, \mathcal{L}^d)$ et $t \in]0, 1[$.

2. Soit $h \in L_{loc}^\infty([0, 1[\times \mathbb{R}^d, \mathcal{L}^{d+1})$,

- De même que dans le point précédent, on suppose V un ouvert borné cette fois de $[0, 1[\times \mathbb{R}^d$.

De même que précédemment, on peut déduire que $(t \mapsto (t, t\nabla\phi + (1-t)\text{id}))(V)$ est borné. Il existe donc une constante $M > 0$ telle que, en posant

$$A = \{(t, y) \in (t \mapsto (t, t\nabla\phi + (1-t)\text{id}))(V), |h(t, y)| > M\},$$

on ait $\mathcal{L}^{d+1}(A) = 0$. Or pour tout $(t, x) \in V$ tel que ϕ soit différentiable en x (rappelons que l'ensemble des points pour lesquels ce n'est pas le cas est de mesure nulle),

$$\begin{aligned} |h(t, t\nabla\phi(x) + (1-t)x)| > M &\Leftrightarrow (t, t\nabla\phi(x) + (1-t)x) \in A \\ &\Leftrightarrow x \in \mathbf{p}(t, A_t) \\ &\Leftrightarrow (t, x) \in B = \{(t', x') \in V, x' \in \mathbf{p}(t', A_{t'})\}, \end{aligned}$$

où pour tout $t \in [0, 1]$, A_t est définie par $A_t = \{x \in \mathbb{R}^d, (t, x) \in A\}$ et est un ensemble Lebesgue-mesurable de \mathbb{R}^d en vertu du Théorème A.2-2 p. 204 (et remarquons que $\mathbf{1}_A(t, x) = \mathbf{1}_{A_t}(x)$ pour tout $(t, x) \in [0, 1] \times \mathbb{R}^d$). Avec la même définition, on peut également définir $B_t = \{x \in \mathbb{R}^d, (t, x) \in B\} = \mathbf{p}(t, A_t)$ pour tout $t \in [0, 1[$, de même Lebesgue-mesurable. Pour finir, d'après le Théorème A.2-4 (p. 205), $t \mapsto \mathcal{L}^d(A_t)$ et $t \mapsto \mathcal{L}^d(B_t)$ sont mesurables.

Si $\mathcal{L}^{d+1}(A) = 0$, alors

$$\begin{aligned} 0 = \mathcal{L}^{d+1}(A) &= \int_0^1 \int_{\mathbb{R}^d} \mathbf{1}_A(t, x) d\mathcal{L}^d(x) dt = \int_0^1 \int_{\mathbb{R}^d} \mathbf{1}_{A_t}(x) d\mathcal{L}^d(x) dt \\ &= \int_0^1 \mathcal{L}^d(A_t) dt. \end{aligned} \quad (\text{F.0-12})$$

Donc $\mathcal{L}^d(A_t) = 0$ pour presque tout $t \in [0, 1]$. Or pour tout $t \in [0, 1[$, $\mathbf{p}(t, \cdot)$ est lipschitzien, donc d'après le Lemme 3.3-21 (p. 103), si $\mathcal{L}^d(A_t) = 0$, alors $\mathcal{L}^d(\mathbf{p}(t, A_t)) = 0$. Ainsi, $\mathcal{L}^d(B_t) = \mathcal{L}^d(\mathbf{p}(t, A_t)) = 0$ pour presque tout $t \in [0, 1]$. Donc d'après le Théorème A.2-3 (p. 204),

$$\mathcal{L}^{d+1}(B) = \int_0^1 \mathcal{L}^d(B_t) dt = 0, \quad (\text{F.0-13})$$

d'où $|h(t, t\nabla\phi(x) + (1-t)x)| \leq M$ pour \mathcal{L}^{d+1} -presque tout $(t, x) \in V$. On peut donc en conclure que $(t, x) \mapsto h(t, t\nabla\phi(x) + (1-t)x) \in L_{loc}^\infty([0, 1] \times \mathbb{R}^d, \mathcal{L}^{d+1})$.

- Pour le deuxième point de la propriété 2, on raisonne de même que dans le deuxième point de la propriété 1, via le Théorème de Lusin (la propriété étant valable pour $h \in C_c^0([0, 1] \times \mathbb{R}^d)$).
3. De par le Lemme de continuité sous le signe intégrale, il est clair que pour tout $h \in C_c^0([0, 1] \times \mathbb{R}^d)$, la fonction

$$t \mapsto \int_{\mathbb{R}^d} h(t, \cdot) d\nu_t = \int_{\mathbb{R}^d} h(t, t\nabla\phi(x) + (1-t)x) \rho_0(x) d\mathcal{L}^d(x)$$

est continue sur $[0, 1]$.

□

Dans la suite, pour simplifier la notation, nous noterons souvent dx à la place de $d\mathcal{L}^d$, et $dx dt$ (par abus) à la place $dx \otimes dt$.

Annexe G

Démonstration du Lemme 4.2-4

L'objectif de cette annexe est de démontrer le Lemme suivant, énoncé au début de la sous-section IV.2.3 à la page 145 :

Lemme 4.2-4.

Soit $\mathcal{H} = \{f, f \in W_{loc}^{1,p}([0, 1] \times \mathbb{R}^d) \cap L^\infty((0, 1) \times \mathbb{R}^d), \forall p \in [1, 2]\}$. Soient de plus $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ convexe vérifiant la propriété (Γ_1) , et un champ de vitesse v définie par rapport à ϕ comme en (3.3-11) p. 89.

Pour tout $\varphi \in C_c^\infty((0, 1) \times \mathbb{R}^d)$, il existe $\tilde{h} \in \mathcal{H}$, définie pour tout $(t, x) \in]0, 1[\times \mathbb{R}^d$ par

$$\tilde{h}(t, x) = - \int_t^1 \varphi(s, (s-t)v(t, x) + x) ds, \quad (\text{G.0-1})$$

et vérifiant

$$\partial_t \tilde{h} + v \cdot \nabla_x \tilde{h} = \varphi \quad (\text{G.0-2})$$

et telle qu'il existe $0 < t_m < 1$ ne dépendant que de φ (en particulier ne dépendant pas de ϕ et donc de v), tel que $\tilde{h}(t, \cdot) = 0$ pour tout $t_m \leq t \leq 1$.

De plus, on a :

$$\nabla_{t,x} \tilde{h} = {}^t(\nabla_{t,x} v) \alpha + \beta \quad (\text{G.0-3})$$

($\nabla_{t,x} v$ étant de taille $d \times (d+1)$), avec $\alpha \in L^\infty([0, 1] \times \mathbb{R}^d)^d$, et $\beta \in L_{loc}^\infty([0, 1] \times \mathbb{R}^d)^{d+1}$, tels que :

- $\|\alpha\|_{L^\infty([0,1] \times \mathbb{R}^d)} \leq \|\nabla_x \varphi\|_{L^\infty([0,1] \times \mathbb{R}^d)}$,
- pour tout ouvert borné $\omega \subset \mathbb{R}^d$,

$$\|\beta\|_{L^\infty([0,1] \times \omega)} \leq (\|v\|_{L^\infty([0,1] \times \omega)} + 1) \|\nabla_x \varphi\|_{L^\infty([0,1] \times \omega)} + \|\varphi\|_{L^\infty([0,1] \times \omega)},$$

- $\alpha(t, \cdot) = 0$ et $\beta(t, \cdot) = 0$ pour tout $t_m \leq t \leq 1$,
- $\left| \nabla_x \tilde{h}(t, x) \right| \leq (|\nabla_x v(t, x)| + 1) \|\nabla_x \varphi\|_{L^\infty([0,1] \times \mathbb{R}^d)}$, pour presque tout $(t, x) \in]0, 1[\times \mathbb{R}^d$.

Pour cela, nous commencerons par démontrer le résultat suivant :

Lemme G.0-1.

Soient $n, m \in \mathbb{N}$, et $\omega \subset \mathbb{R}^m$. Soient $f \in C_c^\infty((0, 1) \times \mathbb{R}^n)$ et $g :]0, 1[\times \mathbb{R}^m \rightarrow \mathbb{R}^n$, et soit $y \in \omega$ tel que :

1. Pour tout $s \in]0, 1[$, $g(s, \cdot)$ est Fréchet-différentiable au point y .
2. Cette différentiabilité est uniforme par rapport à s : autrement dit, il existe une fonction $R_{g,y} :]0, 1[\times V_0 \rightarrow \mathbb{R}^n$ (où $V_0 \subset \mathbb{R}^m$ est un voisinage de 0 adapté) tel que pour tout $s \in]0, 1[$ et $\lambda \in V_0$,

$$|R_{g,y}(s, \lambda)| \leq \sigma_{g,y}(\lambda)|\lambda|$$

où $\sigma_{g,y} : V_0 \rightarrow \overline{\mathbb{R}}$ est indépendante de s , et $\sigma_{g,y}(\lambda) \rightarrow 0$ ($\lambda \rightarrow 0$), et

$$g(s, y + \lambda) = g(s, y) + \nabla_y g(s, y)\lambda + R_{g,y}(s, \lambda)$$

Alors il existe une fonction $R_y :]0, 1[\times V_0 \rightarrow \mathbb{R}$ tel que pour tout $s \in]0, 1[$ et $\lambda \in V_0$,

$$|R_y(s, \lambda)| \leq \sigma_y(\lambda)|\lambda|$$

où $\sigma_y : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ est indépendante de s , et $\sigma_y(\lambda) \rightarrow 0$ ($\lambda \rightarrow 0$), et

$$f(s, g(s, y + \lambda)) = f(s, g(s, y)) + {}^t\lambda {}^t(\nabla_y g(s, y)) \nabla_y f(s, g(s, y)) + R_y(s, \lambda)$$

Preuve : $f \in C_c^\infty((0, 1) \times \mathbb{R}^n)$, donc en effectuant un développement de Taylor par rapport à la seconde variable,

$$\forall z, \lambda' \in \mathbb{R}^n, \forall s \in]0, 1[, f(s, z + \lambda') = f(s, z) + {}^t\lambda' \nabla_y f(s, z) + R_{f,y}(s, \lambda')$$

avec $R_y :]0, 1[\times \mathbb{R}^n \rightarrow \mathbb{R}$ tel que pour tout $s \in]0, 1[$, et $\lambda' \in \mathbb{R}^n$, $R_y(\lambda') \leq \frac{1}{2} \|\text{Hess}(f)\|_\infty \cdot |\lambda'|^2$. Donc pour tout $s \in]0, 1[$, et pour tout $\lambda \in \mathbb{R}^m$ au voisinage de 0, en prenant $z = g(s, y)$ et $\lambda' = \nabla_y g(s, y)\lambda + R_{g,y}(s, \lambda)$.

$$\begin{aligned} f(s, g(s, y + \lambda)) &= f(s, g(s, y) + \nabla_y g(s, y)\lambda + R_{g,y}(s, \lambda)) \\ &= f(s, g(s, y)) + {}^t\lambda {}^t(\nabla_y g(s, y)) \nabla_y f(s, g(s, y)) + {}^tR_{g,y}(s, \lambda) \nabla_y f(s, g(s, y)) \\ &\quad + R_{f,y}(s, \nabla_y g(s, y)\lambda + R_{g,y}(s, \lambda)) \end{aligned}$$

On pose $R_y(s, \lambda) = {}^tR_{g,y}(s, \lambda) \nabla_y f(s, g(s, y)) + R_{f,y}(s, \nabla_y g(s, y)\lambda + R_{g,y}(s, \lambda))$. Pour tout $s \in]0, 1[$, et pour tout $\lambda \in \mathbb{R}^m$ au voisin de 0 assez petit pour que $\sigma_{g,y}(\lambda) \leq 1$ (et donc $|R_{g,y}(s, \lambda)| \leq |\lambda|$),

$$\begin{aligned} |R_y(s, \lambda)| &\leq \|\nabla_y f\|_\infty \sigma_{g,y}(\lambda) |\lambda| + \frac{1}{2} \|\text{Hess}(f)\|_\infty (|\nabla_y g(s, y)| \cdot |\lambda| + |R_{g,y}(s, \lambda)|)^d \\ &\leq \|\nabla_y f\|_\infty \sigma_{g,y}(\lambda) |\lambda| + \frac{1}{2} \|\text{Hess}(f)\|_\infty (A \cdot |\lambda| + |\lambda|)^d \\ &\leq \sigma_y(\lambda) |\lambda| \end{aligned}$$

en posant $\sigma_y(\lambda) = \|\nabla_y f\|_\infty \sigma_{g,y}(\lambda) + \frac{1}{2} \|\text{Hess}(f)\|_\infty (A + 1)^d \cdot |\lambda|$. σ_y est bien indépendante de s et $\sigma_y(\lambda) \rightarrow 0$ lorsque $\lambda \rightarrow 0$. Au final on a bien :

$$f(s, g(s, y + \lambda)) = f(s, g(s, y)) + {}^t\lambda {}^t(\nabla_y g(s, y)) \nabla_y f(s, g(s, y)) + R_y(s, \lambda)$$

□

Nous sommes maintenant en mesure de démontrer le Lemme 4.2-4 :

Retour à la preuve du Lemme 4.2-4 : On fait le choix de poser

$$\tilde{h}(t, x) = - \int_t^1 \varphi(s, (s-t)v(t, x) + x) ds,$$

pour tout $(t, x) \in]0, 1[\times \mathbb{R}^d$, par une application formelle de la méthode des caractéristiques, au regard de la Remarque 3.3-16 (p. 95), appuyée de la Proposition 3.3-14 (p. 96).

φ étant bornée, \tilde{h} est bien définie, et il est clair que $\tilde{h} \in L^\infty((0, 1) \times \mathbb{R}^d)$.

Commençons par démontrer que \tilde{h} est localement lipschitzienne sur $]0, 1[\times \mathbb{R}^d$. Soit $\omega \subset\subset \mathbb{R}^d$ (i.e. $\bar{\omega}$ est compact) et soient $0 < a < b < 1$. D'après la Proposition 3.3-15 (p. 96), il existe une constante L telle que v soit L -lipschitzienne sur $]a, b[\times \mathbb{R}^d$, et d'après la Proposition 3.3-12 (p. 93) : $v \in L^\infty((0, 1) \times \omega)^d$. On pose $M = \|v\|_{L^\infty((0, 1) \times \omega)}$. Pour tout $(t_1, x_1), (t_2, x_2) \in]a, b[\times \mathbb{R}^d$,

$$\begin{aligned} |\tilde{h}(t_1, x_1) - \tilde{h}(t_2, x_2)| &\leq \left| \int_{t_1}^1 (\varphi(s, (s - t_1)v(t_1, x_1) + x_1) - \varphi(s, (s - t_2)v(t_2, x_2) + x_2)) ds \right| \\ &\quad + \left| \int_{t_1}^{t_2} \varphi(s, (s - t_2)v(t_2, x_2) + x_2) ds \right| \\ &\leq \|\nabla\varphi\|_\infty \int_{t_1}^1 |((s - t_1)v(t_1, x_1) + x_1) - ((s - t_2)v(t_2, x_2) + x_2)| ds \\ &\quad + \|\varphi\|_\infty |t_2 - t_1| \\ &\leq \|\nabla\varphi\|_\infty \int_0^1 (|s - t_1| |v(t_1, x_1) - v(t_2, x_2)| + |t_2 - t_1| |v(t_2, x_2)|) ds \\ &\quad + \|\nabla\varphi\|_\infty |x_2 - x_1| + \|\varphi\|_\infty |t_2 - t_1| \\ &\leq \|\nabla\varphi\|_\infty |v(t_1, x_1) - v(t_2, x_2)| + (\|\nabla\varphi\|_\infty M + \|\varphi\|_\infty) |t_2 - t_1| \\ &\quad + \|\nabla\varphi\|_\infty |x_2 - x_1| \\ &\leq L \|\nabla\varphi\|_\infty (|t_2 - t_1| + |x_2 - x_1|) + (\|\nabla\varphi\|_\infty M + \|\varphi\|_\infty) |t_2 - t_1| \\ &\quad + \|\nabla\varphi\|_\infty |x_2 - x_1| \end{aligned}$$

Au final on a bien :

$$|\tilde{h}(t_1, x_1) - \tilde{h}(t_2, x_2)| \leq L_h(a, b, \omega) (|t_2 - t_1| + |x_2 - x_1|)$$

avec $L_h(a, b, \omega) = \max(\|\nabla\varphi\|_\infty(L + M) + \|\varphi\|_\infty, \|\nabla\varphi\|_\infty(1 + L))$. \tilde{h} est donc localement lipschitzienne sur $]0, 1[\times \mathbb{R}^d$.

Donc, toujours d'après le Théorème de Rademacher, \tilde{h} est Fréchet-différentiable presque partout sur $]0, 1[\times \mathbb{R}^d$, et sa différentielle au sens de Fréchet coïncide avec sa différentielle au sens des distributions. Nous allons maintenant calculer la différentielle de \tilde{h} et montrer que l'égalité $\partial_t \tilde{h} + v \cdot \nabla_x \tilde{h} = \varphi$ est bien vérifiée au sens des distributions.

On a vu dans l'exemple 3.3-18 (p. 99), que v était presque partout différentiable et que $\partial_t v(t, x) + v(t, x) \cdot \nabla_x v(t, x) = 0$ pour presque tout $(t, x) \in]0, 1[\times \mathbb{R}^d$.

On se réfère au Lemme G.0-1 : on prendra $f = \varphi$ et $g(s, (t, x)) = (s - t)v(t, x) + x$ pour tout $s, t \in]0, 1[$ et $x \in \mathbb{R}^d$. Soit $(t, x) \in]0, 1[\times \mathbb{R}^d$ un point en lequel v est différentiable. Pour tout $s \in]0, 1[$, $g(s, \cdot)$ est différentiable en (t, x) et

$$\begin{cases} \partial_t g(s, (t, x)) = (s - t)\partial_t v(t, x) - v(t, x) \\ \nabla_x g(s, (t, x)) = (s - t)\nabla_x v(t, x) + \mathbf{I}_{\mathbb{R}^d} \end{cases} \quad (\text{G.0-4})$$

Plus précisément, pour $\delta = (\delta_t, \delta_x) \in \mathbb{R} \times \mathbb{R}^d$ suffisamment proche de 0, on a

$$v(t + \delta_t, x + \delta_x) = v(t, x) + \nabla_{t,x} v(t, x) \delta + R_v(\delta)$$

avec $R_v(\delta) = o(\delta)$. Donc pour tout $s \in]0, 1[$:

$$\begin{aligned} g(s, (t + \delta_t, x + \delta_x)) &= g(s, (t, x)) + ((s - t) \nabla_{t,x} v(t, x) \delta - \delta_t v(t, x) + \delta) + (s - t) R_v(\delta) \\ &= g(s, (t, x)) + \nabla_{t,x} g(s, (t, x)) \delta + R_{g,(t,x)}(\delta) \end{aligned}$$

avec $|R_{g,(t,x)}(\delta)| = |(s - t) R_v(\delta)| \leq |R_v(\delta)| = o(\delta)$ en remarquant bien sûr que R_v est indépendant de s .

D'après le Lemme G.0-1 (p. 241), il existe une fonction $R_{(t,x)} :]0, 1[\times V_0 \rightarrow \mathbb{R}$ (où $V_0 \subset \mathbb{R}^m$ est un voisinage de 0 adapté, i.e. suffisamment restreint) tel que pour tout $s \in]0, 1[$ et $\delta \in V_0$,

$$|R_{(t,x)}(s, \delta)| \leq \sigma_{(t,x)}(\delta) |\delta|$$

où $\sigma_{(t,x)} : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ est indépendante de s , et $\sigma_{(t,x)}(\delta) \rightarrow 0$ ($\delta \rightarrow 0$), et

$$\begin{aligned} \varphi(s, g(s, (t, x) + \delta)) &= \varphi(s, g(s, (t, x)) + \nabla_{t,x} g(s, (t, x)) \delta + R_{g,(t,x)}(\delta)) \\ &= \varphi(s, g(s, (t, x))) + {}^t \delta^t (\nabla_{(t,x)} g(s, (t, x))) \nabla_{(t,x)} \varphi(s, g(s, (t, x))) + R_{(t,x)}(s, \delta) \end{aligned}$$

Ainsi en intégrant selon s entre $t + \delta_t$ et 1, on obtient :

$$\begin{aligned} \tilde{h}(t + \delta_t, x + \delta_x) &= - \int_{t + \delta_t}^1 \varphi(s, g(s, (t, x) + \delta)) ds \\ &= - \int_t^1 \varphi(s, g(s, (t, x) + \delta)) ds + \int_t^{t + \delta_t} \varphi(s, g(s, (t, x) + \delta)) ds \end{aligned} \tag{G.0-5}$$

Si on analyse les deux termes de l'égalité ci-dessus :

$$\begin{aligned} \int_t^1 \varphi(s, g(s, (t, x) + \delta)) ds &= \int_t^1 \varphi(s, g(s, (t, x))) ds + \int_t^1 R_{(t,x)}(s, \delta) ds \\ &\quad + {}^t \delta \int_t^1 {}^t (\nabla_{(t,x)} g(s, (t, x))) \nabla_{(t,x)} \varphi(s, g(s, (t, x))) ds \\ &= -\tilde{h}(t, x) + \int_t^1 R_{(t,x)}(s, \delta) ds \\ &\quad + {}^t \delta \int_t^1 {}^t (\nabla_{(t,x)} g(s, (t, x))) \nabla_{(t,x)} \varphi(s, g(s, (t, x))) ds \end{aligned}$$

et

$$\left| \int_t^1 R_{(t,x)}(s, \delta) ds \right| \leq \int_0^1 |R_{(t,x)}(s, \delta)| ds \leq \sigma_{(t,x)}(\delta) |\delta|, \text{ et donc } \int_t^1 R_{(t,x)}(s, \delta) ds = o(\delta)$$

De plus, pour tout $s \in]0, 1[$,

$$|\varphi(s, g(s, (t, x) + \delta)) - \varphi(s, g(s, (t, x)))| \leq |{}^t \delta^t (\nabla_{(t,x)} g(s, (t, x))) \nabla_{(t,x)} \varphi(s, g(s, (t, x))) + R_{(t,x)}(s, \delta)|$$

En remarquant d'après l'équation (G.0-4) p. 243 que :

$$|\nabla_{t,x} g(s, (t, x))| \leq |\nabla_x v(t, x)| + |v(t, x)| + d,$$

on a l'estimation :

$$|\varphi(s, g(s, (t, x) + \delta)) - \varphi(s, g(s, (t, x)))| \leq (|\nabla_x v(t, x)| + |v(t, x)| + d) \|\nabla \varphi\|_\infty + \sigma_{(t, x)}(\delta) |\delta|$$

Donc pour en revenir au deuxième terme de l'équation (G.0-5),

$$\left| \int_t^{t+\delta_t} \varphi(s, g(s, (t, x) + \delta)) ds - \int_t^{t+\delta_t} \varphi(s, g(s, (t, x))) ds \right| \leq (|\nabla_x v(t, x)| + |v(t, x)| + d) \|\nabla \varphi\|_\infty + \sigma_{(t, x)}(\delta) |\delta| \cdot |\delta_t|$$

$$d'où \int_t^{t+\delta_t} \varphi(s, g(s, (t, x) + \delta)) ds - \int_t^{t+\delta_t} \varphi(s, g(s, (t, x))) ds = o(\delta).$$

Or la fonction $s \mapsto \varphi(s, g(s, (t, x))) = \varphi(s, (s-t)v(t, x) + x)$ est continue car $\varphi \in C_c^\infty((0, 1) \times \mathbb{R}^d)$, donc

$$\int_t^{t+\delta_t} \varphi(s, g(s, (t, x))) ds = \varphi(t, g(t, (t, x))) \delta_t + o(\delta)$$

c'est à dire, du fait que $\varphi(t, g(t, (t, x))) = \varphi(t, (t-t)v(t, x) + x) = \varphi(t, x)$,

$$\int_t^{t+\delta_t} \varphi(s, g(s, (t, x))) ds = \varphi(t, x) \delta_t + o(\delta)$$

Au final

$$\tilde{h}(t + \delta_t, x + \delta_x) = \tilde{h}(t, x) + o(\delta) - \delta \int_t^1 {}^t \nabla_{t,x} g(s, (t, x)) \nabla_x \varphi(s, g(s, (t, x))) ds + \varphi(t, x) \delta_t$$

\tilde{h} est donc différentiable en (t, x) , et en se référant à l'équation (G.0-4) :

$$\begin{cases} \partial_t \tilde{h}(t, x) = - \int_t^1 ((s-t) \partial_t v(t, x) - v(t, x)) \cdot \nabla_x \varphi(s, g(s, (t, x))) ds + \varphi(t, x), \\ \nabla_x \tilde{h}(t, x) = - \int_t^1 (s-t) \nabla_x v(t, x) \cdot \nabla_x \varphi(s, g(s, (t, x))) ds - \int_t^1 \nabla_x \varphi(s, g(s, (t, x))) ds \end{cases} \quad (\text{G.0-6})$$

et donc, pour presque tout $(t, x) \in]0, 1[\times \mathbb{R}^d$,

$$\nabla_{t,x} h(t, x) = {}^t (\nabla_{t,x} v(t, x)) \alpha(t, x) + \beta(t, x),$$

avec

$$\alpha(t, x) = - \int_t^1 (s-t) \nabla_x \varphi(s, g(s, (t, x))) ds \in \mathbb{R}^d \quad (\text{G.0-7})$$

et

$$\beta(t, x) = \begin{pmatrix} \int_t^1 v(t, x) \cdot \nabla_x \varphi(s, g(s, (t, x))) ds + \varphi(t, x) \\ - \int_t^1 \nabla_x \varphi(s, g(s, (t, x))) ds \end{pmatrix} \in \mathbb{R}^{d+1} \quad (\text{G.0-8})$$

(en effet, on a ici $\nabla_{t,x} v(t, x) \in \mathcal{M}_{d \times (d+1)}(\mathbb{R})$).

On déduit également directement de (G.0-6) la majoration :

$$\left| \nabla_x \tilde{h}(t, x) \right| \leq (|\nabla_x v(t, x)| + 1) \|\nabla_x \varphi\|_{L^\infty([0,1] \times \mathbb{R}^d)}, \quad (\text{G.0-9})$$

pour presque tout $(t, x) \in]0, 1[\times \mathbb{R}^d$. De plus, du fait que $\partial_t v(t, x) + v(t, x) \cdot \nabla_{t,x} v(t, x) = 0$ pour presque tout $(t, x) \in]0, 1[\times \mathbb{R}^d$,

$$\begin{aligned} & \partial_t \tilde{h}(t, x) + v(t, x) \cdot \nabla_x \tilde{h}(t, x) \\ &= - \int_t^1 (s - t) (\partial_t v(t, x) + v(t, x) \cdot \nabla_x v(t, x)) \cdot \nabla_{(t,x)} \varphi(s, g(s, (t, x))) ds + \varphi(t, x) \\ &= \varphi(t, x) \end{aligned}$$

et ceci donc pour presque tout $(t, x) \in]0, 1[\times \mathbb{R}^d$. Comme nous l'avons vu plus haut, \tilde{h} est Fréchet-différentiable presque partout sur $]0, 1[\times \mathbb{R}^d$, et sa différentielle au sens de Fréchet coïncide avec sa différentielle au sens des distributions. Ainsi

$$\partial_t \tilde{h} + v \cdot \nabla_x \tilde{h} = \varphi \quad \text{et} \quad \nabla_{t,x} h = {}^t(\nabla_{t,x} v) \alpha + \beta$$

pour $\nabla_{t,x} h$ définie au sens de \mathcal{H} (i.e. au sens des distributions).

Comme $\varphi \in C_c^\infty((0, 1) \times \mathbb{R}^d)$, il existe $0 < t_m < 1$, tel que $\varphi(s, \cdot) = \nabla_x \varphi(s, \cdot) = 0$ pour tout $t_m < s \leq 1$, d'où

$$\alpha(t, \cdot) = 0 \quad \text{et} \quad \beta(t, \cdot) = 0, \quad \text{et ainsi} \quad \tilde{h}(t, \cdot) = 0,$$

pour tout $t_m \leq t \leq 1$.

Les propriétés de majoration de α et β découlent quant à elles directement de (G.0-7) et (G.0-8).

□

Petit index des notations

S	page 56
$\Lambda(\rho_0, \rho_1)$	page 57
$\tilde{\mathcal{K}}$	page 57
$\tilde{\mathcal{P}}$	page 57
L	page 53
$\mathbf{L}^{ps}(\rho_0, \rho_1, \Omega)$	page 74
$\Upsilon(\rho_0)$	page 81
$\Theta(\rho_0)$	page 82
$\Phi(\rho_0 \mathcal{L}^d, \mu_1)$	page 80
$\mathcal{P}_p(\mathbb{R}^d)$	page 19
\mathcal{W}_p	page 19
γf	page 88
Prox_f	page 35, 88
$\mathbf{P}, \mathbf{P}_\phi, v_\phi$	page 89
v_γ	page 109
${}^bL_+^2((0, 1) \times \Omega)$	page 158
Propriété (Γ_1)	page 76
Propriété (Γ_2)	page 98
Propriété (I)	page 74
Propriété (I'_α)	page 77
Propriété (I'_β)	page 79
Propriété (II)	page 75
Propriété (III)	page 76
Propriété (IV)	page 203

Bibliographie

- [1] L. Ambrosio, N. Gigli, and G. Savare. *Gradient Flows : In Metric Spaces and in the Space of Probability Measures*. Birkhäuser Verlag, 2005.
- [2] H. Attouch, J. Bolte, and B. Svaiter. Convergence of descent methods for semi-algebraic and tame problems : proximal algorithms, forward-backward splitting, and regularized gauss-seidel methods. *Mathematical Programming*, 137(1-2) :91–129, 2013.
- [3] H. Bauschke. The approximation of fixed points of compositions of nonexpansive mappings in hilbert space. *Journal of Mathematical Analysis and Applications*, 202(1) :150–159, 1996.
- [4] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, 2011.
- [5] J. Benamou, B. Froese, and A. Oberman. Numerical solution of the second boundary value problem for the elliptic Monge-Ampère equation. Technical report, Rapport de recherche INRIA, 2012.
- [6] J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution of the Monge-Kantorovich mass transfer problem. *Numerische Mathematik*, 84(3) :375–393, 2000.
- [7] J.-D. Benamou, G. Carlier, and M. Laborde. An augmented Lagrangian approach to Wasserstein gradient flows and applications. 2015.
- [8] J.-D. Benamou, G. Carlier, Q. Mérigot, and E. Oudet. Discretization of functionals involving the Monge-Ampère operator. *Numerische Mathematik*, pages 1–26, 2015.
- [9] J.-D. Benamou, B. D. Froese, and A. M. Oberman. Numerical solution of the optimal transportation problem via viscosity solutions for the Monge-Ampere equation. *CoRR*, abs/1208.4873, 2, 2012.
- [10] J.-D. Benamou, B. D. Froese, and A. M. Oberman. Numerical solution of the optimal transportation problem using the Monge-Ampère equation. *Journal of Computational Physics*, 260 :107–126, 2014.
- [11] N. Bonnotte. From Knothe’s rearrangement to Brenier’s optimal transport map. *SIAM Journal on Mathematical Analysis*, 45(1) :64–87, 2013.
- [12] R. I. Bot and E. R. Csetnek. Inducing strong convergence into the asymptotic behaviour of proximal splitting algorithms in hilbert spaces. *arXiv preprint arXiv :1609.01627*, 2016.
- [13] R. I. Bot, E. R. Csetnek, and A. Heinrich. On the convergence rate improvement of a primal-dual splitting algorithm for solving monotone inclusion problems. *CoRR*, abs/1303.2875, 2013.
- [14] A. Bouharguane, A. Iollo, and L. Weynans. Numerical solution of the Monge-Kantorovich problem by Picard iterations. Technical Report RR-8477, INRIA, 2014.
- [15] A. Bouharguane, E. Maitre, E. Oudet, and N. Papadakis. Multiphysics optimal transportation and image analysis. Technical Report hal-00740671, 2012.

- [16] Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Communications on Pure and Applied Mathematics*, 44(4) :375–417, 1991.
- [17] H. Brezis, P. G. Ciarlet, and J. L. Lions. *Analyse fonctionnelle : théorie et applications*, volume 91. Dunod Paris, 1999.
- [18] C. Brune. *4D imaging in tomography and optical nanoscopy*. PhD thesis, University of Münster, Germany, 2010.
- [19] L. A. Caffarelli. Boundary regularity of maps with convex potentials. *Communications on pure and applied mathematics*, 45(9) :1141–1151, 1992.
- [20] L. A. Caffarelli. Boundary regularity of maps with convex potentials–ii. *Annals of mathematics*, 144(3) :453–496, 1996.
- [21] G. Carlier, V. Duval, G. Peyré, and B. Schmitzer. Convergence of entropic schemes for optimal transport and gradient flows. *arXiv preprint arXiv :1512.02783*, 2015.
- [22] G. Carlier, A. Galichon, and F. Santambrogio. From Knothe’s transport to Brenier’s map and a continuation method for optimal transport. *SIAM Journal On Mathematical Analysis*, 41(6) :2554–2576, 2010.
- [23] A. Chambolle. Rates of convergence for accelerated primal-dual algorithms. Presented in SIAM Conference on Imaging Sciences, May 2014.
- [24] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1) :120–145, 2011.
- [25] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1) :120–145, 2011.
- [26] L. Chizat, G. Peyré, B. Schmitzer, and F.-X. Vialard. Unbalanced optimal transport : geometry and Kantorovich formulation. *arXiv preprint arXiv :1508.05216*, 2015.
- [27] P. Clarysse, B. Delhay, M. Picq, and J. Pousin. Optimal extended optical flow subject to a statistical constraint. *Journal of Computational and Applied Mathematics*, 234(4) :1291–1302, 2010.
- [28] P. L. Combettes and J.-C. Pesquet. A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery. *IEEE Journal of Selected Topics in Signal Processing*, 1(4) :564–574, 2007.
- [29] L. Condat. A primal-dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms. *Journal of Optimization Theory and Applications*, 158(2) :460–479, 2013.
- [30] D. Cordero-Erausquin. Sur le transport de mesures périodiques. *Comptes Rendus de l’Académie des Sciences-Series I-Mathematics*, 329(3) :199–202, 1999.
- [31] E. J. Dean and R. Glowinski. Numerical solution of the two-dimensional elliptic Monge–Ampère equation with dirichlet boundary conditions : an augmented Lagrangian approach. *Comptes rendus Mathématique*, 336(9) :779–784, 2003.
- [32] J. Deny and J.-L. Lions. Les espaces du type de beppo levi. *Annales de l’Institut Fourier*, 5 :305–370, 1954.
- [33] R. J. DiPerna and P.-L. Lions. Ordinary differential equations, transport theory and sobolev spaces. *Inventiones mathematicae*, 98(3) :511–547, 1989.
- [34] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions*. CRC PRESS, 1992.
- [35] N. Feyeux. *Transport optimal pour l’assimilation de données images*. PhD thesis, Université de Grenoble, 2016.

- [36] N. Feyeux, A. Vidard, and M. Nodet. Optimal transport for data assimilation. 2016.
- [37] M. Fortin and R. Glowinski. *Augmented Lagrangian Methods : Applications to the Numerical Solution of Boundary-Value Problems*. Studies in Mathematics and its Applications. Elsevier Science, 2000.
- [38] V. Girault and P.-A. Raviart. *Finite Element Methods for Navier-Stokes Equations : Theory and Algorithms*. Springer-Verlag, 2011.
- [39] F. Golse. Distributions, analyse de fourier, équations aux dérivées partielles. *Cours de l'École Polytechnique*. Accessible à l'adresse <http://www.math.polytechnique.fr/~golse/mat431.html>, 2010.
- [40] K. Guittet. On the time-continuous mass transport problem and its approximation by augmented Lagrangian techniques. *SIAM Journal on Numerical Analysis*, 41(1) :382–399, 2004.
- [41] S. Haker, L. Zhu, A. Tannenbaum, and S. Angenent. Optimal mass transport for registration and warping. *International Journal of Computer Vision*, 60(3) :225–240, Dec. 2004.
- [42] B. He and X. Yuan. Convergence analysis of primal-dual algorithms for a saddle-point problem : From contraction perspective. *SIAM Journal on Imaging Sciences*, 5(1) :119–149, 2012.
- [43] M. Henry. *Transport Optimal et Ondelettes : Nouveaux Algorithmes et Applications à l'Image*. PhD thesis, Université de Grenoble, 2016.
- [44] M. Henry, E. Maitre, and V. Perrier. Optimal transport using Helmholtz-Hodge decomposition and first-order primal-dual algorithms. In *Image Processing (ICIP), 2015 IEEE International Conference on*, pages 4748–4752. IEEE, 2015.
- [45] M. Henry, E. Maitre, and V. Perrier. Primal-dual formulation of the dynamical optimal transport using Helmotz-Hodge decomposition. 2016.
- [46] R. Hug, E. Maitre, and N. Papadakis. Multi-physics optimal transportation and image interpolation. *ESAIM : Mathematical Modelling and Numerical Analysis*, 49(6) :1671–1692, 2015.
- [47] A. Iollo and D. Lombardi. A Lagrangian scheme for the solution of the optimal mass transfer problem. *Journal of Computational Physics*, 230(9) :3430 – 3442, 2011.
- [48] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the fokker-planck equation. *SIAM Journal On Mathematical Analysis*, 29(1) :1–17, 1998.
- [49] L. Kantorovitch. On the translocation of masses. *Management Science*, 5(1) :1–4, 1958.
- [50] P. L. Lions and B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *Siam J. Numer. Anal.*, 16 :964–979, 1979.
- [51] G. Loeper and F. Rapetti. Numerical solution of the Monge–Ampère equation by a newton’s algorithm. *Comptes Rendus Mathématique*, 340(4) :319–324, 2005.
- [52] D. Lombardi and E. Maitre. Eulerian models and algorithms for unbalanced optimal transport. *ESAIM : Mathematical Modelling and Numerical Analysis*, 49(6) :1717–1744, 2015.
- [53] J. Maas, M. Rumpf, C. Schönlieb, and S. Simon. A generalized model for optimal transport of images including dissipation and density modulation. Preprint arXiv : 1504.01988, 2015.
- [54] B. Maury, A. Roudneff-Chupin, F. Santambrogio, and J. Venel. Handling congestion in crowd motion modeling. *arXiv preprint arXiv :1101.4102*, 2011.
- [55] R. McCann. Polar factorization of maps on riemannian manifolds. *Geometric & Functional Analysis GAFA*, 11(3) :589–608, 2001.

- [56] R. J. McCann. A convexity principle for interacting gases. *Advances in mathematics*, 128(1) :153–179, 1997.
- [57] G. Monge. Mémoire sur la théorie des déblais et de remblais. *Académie Royale des Sciences de Paris*, page 666–704, 1781.
- [58] G. Monge. *Géométrie descriptive : leçons données aux écoles normales, l’an 3 de la république*. Baudouin, imprimeur du corps législatif et de l’institut national, 1798.
- [59] A. Monteil. Uniform estimates for a modica-mortola type approximation of branched transportation. *arXiv preprint arXiv :1503.03735*, 2015.
- [60] A. Monteil. *Approximations elliptiques d’énergies singulières sous contrainte de divergence*. PhD thesis, Université Paris-Sud 11, 2016.
- [61] A. Moussa. Some variants of the classical aubin-lions lemma. *arXiv preprint arXiv :1401.7231*, 2014.
- [62] A. M. Oberman. Wide stencil finite difference schemes for the elliptic Monge-Ampere equation and functions of the eigenvalues of the hessian. *Discrete Contin. Dyn. Syst. Ser. B*, 10(1) :221–238, 2008.
- [63] V. Oliker and L. Prussner. On the numerical solution of the equation and its discretizations, i. *Numerische Mathematik*, 54(3) :271–293, 1989.
- [64] F. Otto. The geometry of dissipative evolution equations : the porous medium equation. 2001.
- [65] E. Oudet and F. Santambrogio. A modica-mortola approximation for branched transport and applications. *Archive for rational mechanics and analysis*, 201(1) :115–142, 2011.
- [66] N. Papadakis. *Optimal Transport for Image Processing*. PhD thesis, Université de Bordeaux, 2015.
- [67] N. Papadakis, G. Peyré, and E. Oudet. Optimal transport with proximal splitting. *SIAM Journal on Imaging Sciences*, 7(1) :212–238, 2014.
- [68] G. Peyré. Entropic approximation of Wasserstein gradient flows. *SIAM Journal on Imaging Sciences*, 8(4) :2323–2351, 2015.
- [69] J. Rabin and N. Papadakis. Convex color image segmentation with optimal transport distances. In *International Conference on Scale Space and Variational Methods in Computer Vision*, pages 256–269. Springer, 2015.
- [70] H. Raguét, J. Fadili, and G. Peyré. Generalized forward-backward splitting. *SIAM Journal on Imaging Sciences*, 6(3) :1199–1226, 2013.
- [71] H. Raguét, J. Fadili, and G. Peyré. A generalized forward-backward splitting. *SIAM Journal on Imaging Sciences*, 6(3) :1199–1226, 2013.
- [72] W. Rudin, N. Dhombres, and F. Hoffman. *Analyse réelle et complexe*. Masson et Cie, Department of Mathematics, 1975.
- [73] C. Smith and M. Knott. On hoeffding-fréchet bounds and cyclic monotone relations. *Journal of multivariate analysis*, 40(2) :328–334, 1992.
- [74] G. Strang. *Linear Algebra and Its Applications*. Brooks Cole, 1988.
- [75] P. Tseng. Convergence of a block coordinate descent method for nondifferentiable minimization. *Journal of Optimization Theory and Applications*, 109(3) :475–494, 2001.
- [76] H. Uzawa. Iterative methods for concave programming. *Studies in linear and nonlinear programming*, 6, 1958.
- [77] C. Villani. *Topics in Optimal Transportation*. American Mathematical Society, 2003.

- [78] Q. Xia. Optimal paths related to transport problems. *Communications in Contemporary Mathematics*, 5(02) :251–279, 2003.
- [79] Q. Xia. Interior regularity of optimal transport paths. *Calculus of Variations and Partial Differential Equations*, 20(3) :283–299, 2004.
- [80] C. Zuily. *Éléments de distributions et d'équations aux dérivées partielles : cours et problèmes résolus*. Dunod, 2002.