Mémoire de synthèse pour l'obtention de l'

HABILITATION À DIRIGER DES RECHERCHES

en **Mathématiques Appliquées**

présenté par **Filippo Santambrogio**

sous la coordination de **Ivar Ekeland**

Problèmes classiques et moins classiques en transport optimal

Régularité, approximation, EDP et applications

Soutenance le 25 novembre 2009

Après avis des rapporteurs :

·	M. J.-A. Carrillo	Profesor Investigador	ICREA, Universitat Autònoma de Barcelona
·	M. L. C. Evans	Professeur	University of California at Berkeley
·	M. C. Villani	Professeur	École Normale Supérieure de Lyon

Devant le jury composé de :

·	M. Y. Brenier	Directeur de Recherches	CNRS, Université de Nice-Sophia-Antipolis
·	M. G. Buttazzo	Professeur	Università di Pisa
·	M. G. Carlier	Professeur	Université Paris-Dauphine
·	M. J.-A. Carrillo	Profesor Investigador	ICREA, Universitat Autònoma de Barcelona
·	M. I. Ekeland	Professeur	U. Paris-Dauphine et U. of British Columbia
·	M. P.-L. Lions	Professeur	Collège de France
·	M. J.-M. Morel	Professeur	École Normale Supérieure de Cachan
·	M. C. Villani	Professeur	École Normale Supérieure de Lyon

CEREMADE – Université Paris-Dauphine

École Doctorale Décision, Informatique, Mathématiques, Organisation

Al nonno

Mi spiace nonnino, questa volta tocca a te : è la dura legge dei nonni, quando se ne vanno si meritano che gli si dedichi una tesi. Ma lassù hai tempo per leggertela tutta, mi sa… e se hai domande, non esitare !

Ce mémoire est dédicacé à mon grand-père, que j'ai perdu l'an dernier

Remerciements

J'écris ce mémoire d'habilitation dix ans après avoir quitté Milan pour commencer à étudier des maths, et transformer au fur et à mesure ces maths en mon travail, à Pise d'abord et à Paris après. Pendant ces années j'ai beaucoup appris et j'ai beaucoup collaboré. Je n'ai jamais été seul en faisant des maths. Cela impose évidemment que je profite de ce moment pour remercier plein de monde.

Même si je vais me restreindre à la période entre ma thèse et cette HDR, je pense que je ne peux que commencer par mon directeur de thèse, Giuseppe Buttazzo, qui m'a appri le goût pour le calcul des variations et avec qui je continue à travailler avec un plaisir énorme. Je le remercie de tout ça et d'avoir accepté de venir à Paris pour prendre part au Jury.
Si le nom de Giuseppe n'apparaît pas parmi les collaborations qui ont donné lieu aux travaux inclus dans ce mémoire (même si on voit son attention derrière à ce genre de problèmes), la deuxième personne que je dois mentionner y apparait beaucoup, et il s'agit de Guillaume Carlier, que je remercie de tout coeur pour tout ce qu'on a partagé pendant ces dernières années, dans le travail et en dehors. La plus part des collaborations que j'ai nouées au CEREMADE sont passées par son élan et son envie de m'impliquer dans des projets divers. Sa participation au Jury de cette HDR n'est qu'une partie infime du rôle qu'il a eu courant ma formation et ma carrière.
Le coordinateur du travail que je présente aujourd'hui en vu de l'obtention de l'habilitation est Ivar Ekeland, justement le maître de Guillaume, qui m'a fait un grand honneur en acceptant cette tâche. Je me souviens avoir toujours - depuis que Giuseppe m'en avait parlé en 2003, lors de ma "Laurea" - associé le nom d'Ivar à ce genre de problèmes d'optimisation que j'ai essayé de traiter, et le choix de lui comme coordinateur a été pour moi très naturel. Je le remercie de l'avoir trouvé naturel aussi.
Jean-Michel Morel est quelqu'un d'autre à qui je dois beaucoup, qui m'a vraiment pris en charge à mon arrivée à Paris en 2006, et qui m'a aidé à compléter ma migration de Pise en France. Les collaborations avec lui ont été toujours enrichissantes, et j'espère qu'elle continueront en se renouvelant encore. Dans son cas aussi, le remercier simplement de sa participation au Jury serait injustement reductif.

Quand on organise un jury pour une soutenance, il faut bien qu'il y ait quelqu'un qu'on connaît bien, soit-il le directeur ou un collaborateur étroit - et qui va être le destinataire des premiers remerciements - mais il ne faut pas non plus oublier qu'une partie majeure du travail est effectuée par d'autres collègues qui ont exactement la péculiariété d'être un peu moins proches au niveau professionnel. C'est par exemple le cas des rapporteurs. Cedric Villani, grand spécialiste du transport optimal, m'a fait un grand honneur en acceptant d'être rapporteur et de participer au jury. José-Antonio Carrillo m'a fait aussi un honneur, en démontrant en plus une grande disponibilité à se mettre efficacement au travail sur ce mémoire, bien que son domaine n'y soit pas si proche que ça.

On n'en doute pas, l'honneur d'avoir Craig Evans comme rapporteur n'est pas moindre, et le fait qu'il n'ait pas pu se déplacer pour la soutenance ne le dévalorise pas du tout. D'ailleurs, j'ai pas mal harcelé ces trois rapporteurs avec beaucoup de mails et de communications, mais c'était surtout à lui que je pensais quand je leur ai écrit à propos d'une erreur de régularité elliptique - qu'on a quand même su résoudre - dans l'un des papiers qui composent ce mémoire. Je remercie donc tous les rapporteurs de leur disponibilité et de leur attention.

Il est peut-être un peu injuste d'insérer Yann Brenier dans cette partie dédiée aux membres du jury qui sont "un peu moins proches" : depuis qu'il a été rapporteur de ma thèse de doctorat on a eu l'occasion de se côtoyer beaucoup, et j'en suis très fier. Je le remercie d'avoir accepté de rajouter cette nouvelle occasion, la participation à mon Jury d'HDR, à la liste, et j'espère que ça continuera dans cette direction.

Pour conclure avec le Jury, je pense vraiment qu'il est inutile de dire que je suis très honoré que Pierre-Louis Lions ait accepté d'y prendre part, et je le remercie de tout coeur.

Si en ce jour de soutenance je n'ai pas pu résister à la tentation de concentrer mes remerciements sur les membres du Jury, il ne faut pas oublier que la soutenance même n'aurait pas été possible sans tout le travail qui il y a derrière. Et ce travail a été rendu possible seulement grâce aux multiples collaborations que j'ai eues lors de ces années. Je vais me limiter aux collaborations dont les résultats ont abouti à un article qui est inclus dans ce mémoire, et ça fait déjà une liste bien longue : Alfred Galichon, Gigi De Pascale, Chloé Jimenez, Fethallah Benmansour, Gabriel Peyré, Vincenzo Vespri, Alessio Figalli, Marc Bernot, Bertrand Maury, Aude Chupin... et si je rajoute les travaux en cours, les projets, les articles sur d'autres sujets, cette liste ne terminera vraiment pas, avec Alessio Brancolini, Eugene Stepanov, Lorenzo Brasco, François-Xavier Vialard, Fadoua Balabdaoui, Kaspar Rufibach, Lambert Piozin, Pascal Mossay, Edouard Oudet... pour ne citer que les plus récents. Je les remercie tous de ce qu'on a fait ensemble, et je suis aussi très content des relations amicales qu'on a su et qu'on sait toujours entretenir, au labo et en dehors. C'est quelque chose de très important, qui aide dans le travail de tous les jours, et qui me fait souvenir que beaucoup d'autres collègues du CEREMADE méritent un petit mot de remerciement pour la bonne ambiance que, grace à eux, j'ai constamment trouvé au labo. Je ne pense pas seulement aux co-bureaux, aux voisins et aux autres membre permanents du labo, mais par exemple à tous les thésards que j'ai eu la chance de connaître un peu mieux.

Cela pourrait paraître réthorique, mais je voudrais bien, avant de terminer ces deux pages de remerciements, dediées à beaucoup de personnes (à qui une phrase, à qui un paragraphe, à qui une fraction de proposition), dédier quelques lignes à un pays, plutôt qu'à une personne. Disons-le, il faut quand même que je remercie la France. Ce pays m'a accueilli d'une manière exquise et c'est tautologique de dire que, sans la France, je ne serais pas là où je suis. Je n'aurais pas ce genre de poste, je n'aurait pas fait ce genre de travaux ; je ne serais pas propriétaire, je pense, et (là c'est sûr) je ne serais pas pacsé. Tout cela contribue à ma félicité, je la remercie.

J'ai glissé quelques petites mentions à mes affaires personnelles dans ces remerciements plutôt académiques. Ce n'est pas mon intention d'y rentrer plus dans les détails. Tout ceux qui méritent d'être remerciés pour leur rôle dans ma vie privée le savent, et soit ils sont là ce jour, soit ils y sont avec l'esprit.

Contents

Introduction

Je recueille dans ce mémoire que j'écris en vue de l'obtention de l'Habilitation à Diriger des Recherches les principaux travaux autour du transport optimal qui ont suivi ma thèse de doctorat. Il s'agit des recherches que j'ai menées pendant mes derniers mois de doctorat à la Scuola Normale Superiore di Pisa, quand j'avais déjà déposé ma thèse et j'attendais de la soutenir, pendant le post-doctorat que j'ai commencé à l'École Normale Supérieure de Cachan, et que j'ai abandonné lors de mon recrutement à l'Université de Paris-Dauphine, et pendant les deux premières années en tant que Maître de Conférences.

Bien que les collaborations avec mon ancien directeur de thèse continuent régulièrement, on peut voir, en vue des sujets que je présente dans ce mémoire, que j'ai élargi beaucoup mes collaborations, en particulier en France, et j'ai touché des sujets d'optimisation qui s'éloignent du transport et que j'ai choisi de ne pas inclure ici par souci d'unité thématique. Je tâcherai dans cette introduction de souligner tous les aspects importants de mon activité de recherche en général, et je laisserai après les quatre chapitres suivants parler pour eux-mêmes quant aux détails techniques des articles. Cette partie plus technique a été rédigée en anglais pour que tous les rapporteurs ainsi que toute personne intéressée de manière purement scientifique (sans qu'elle soit concernée par une "évaluation" de mon activité, mais seulement pour voir quels sont les résultats, les stratégies, les lemmes...) puisse les lire. Au début de chaque chapitre il y aura aussi un résumé en français.

Ce mémoire se compose de quatre chapitres, qui se concentrent sur quatre aspects différents de ma recherche. Chaque chapitre est composé ensuite par plusieurs sections, à peu près une par article ou groupe des travaux, et il est conclu par une section dédiée aux perspectives (souvent numériques) mais aussi à tous ces lemmes et résultats partiels qui sont présents dans les travaux du chapitre mais qui n'y ont pas un rôle central.

Les trois premiers chapitres se différencient quant aux modèles de transport qui sont utilisés et aux éventuels critères de concentration qui apparaissent. Pour être plus clair, dans le deuxième chapitre je présente des problèmes de congestion, là où le transport est plus cher si trop de monde passe au même endroit, et dans le troisième c'est le tour des problèmes de branchement, là où, au contraire, le transport conjoint est encouragé, ce qui donne lieux à une structure ramifiée du réseau. Le premier chapitre est au contraire neutre par rapport à la concentration, ce qui revient à dire que je parle de la théorie classique de Monge-Kantorovitch. Les mots "classiques et moins classiques" dans le titre se réfèrent justement à ce dépassement de la théorie de Monge-Kantorovitch, pour analyser des modèles aux applications différentes et voir ce que la théorie du transport en elle-même peut y rapporter, et comment il faut la changer pour en dire plus. Or, en fait dans mon parcours de recherche j'ai suivi la démarche inverse, car j'ai commencé ma thèse en m'occupant de critères de concentration pour les mesures (en planification urbaine, avec trois travaux qui ont constitué les trois premiers chapitres de ma thèse et qui ont profité de la collaboration et de la direction - plus ou moins officielle - de G. Buttazzo et G. Carlier), puis je me suis occupé de plusieurs questions et modèles et en branchement et en congestion, et ce n'est que récemment que je suis arrivé à des sujets où le problème est classique:

$$\min\Big\{\int c(x,y)\,d\gamma\ :\ \gamma\in\Pi(\mu,\nu)\Big\},$$

où $\Pi(\mu,\nu)$ est l'ensemble des plans de transport de μ à ν (mesures sur le produit dont les deux marges sont μ et ν, respectivement, comme Kantorovitch nous l'enseigne, [85]). Je tâcherai dans cette introduction de montrer quels sont les points communs de ces trois variantes du problème (normale, congestionnée, branchée), sous le point de vue des problèmes à divergence fixée ou des mesures sur l'espace des chemins.

Le quatrième chapitre est un peu différent: j'ai voulu y mettre tout ce qui concerne de problèmes d'évolution - qui gardent quand même un certain esprit variationnel - et qui ont en commun des contraintes sur les densités réalisées par les particules à chaque instant. Le problème des fluides incompressibles, dont la formalisation passe par les équations d'Euler et qui a été introduit comme un problème variationnel par V. Arnold et puis Y. Brenier, demande à ce que la densité soit uniforme à chaque instant; d'autre côté un modèle de mouvement de foule dont B. Maury m'a parlé la première fois il y a deux ans (sans savoir qu'on pouvait l'aborder par le transport), demande à ce que la densité ne dépasse jamais un certain seuil. Dans les deux cas, les contraintes activent des multiplicateurs de Lagrange qui jouent le rôle d'une pression et qui influencent l'évolution.

Je ne résumerai ici que très brièvement les contenus des chapitres, en renvoyant aux résumés qui les précèdent un par un. Dans le premier, trois travaux sont présentés:

- l'un sur le transport avec coût quadratique et ses relations avec le transport monotone de Knothe (qui est vu comme une limite de transport optimaux quand les coûts associés à chaque coordonnée dégénèrent), en collaboration avec G. Carlier et A. Galichon;

- le deuxième sur la densité de transport, qui est en fait un concept qui n'existe que dans le cas où le coût de transport est homogène de degré 1 (et que je considère dans le cas de la norme euclidienne $c(x,y)=|x-y|$);

- le dernier est tout récent et présente, en collaboration avec G. Carlier et L. DePascale, d'une part une stratégie générale pour approcher des problèmes de transport avec coût non strictement convexe, d'autre part applique cette stratégie à un cas nouveau, et notamment le cas d'une puissance plus grande que 1 d'une norme qualconque sur $\mathbb{R}^2$: $c(x,y)=||x-y||^p$, $p>1$.

Le deuxième chapitre porte sur la congestion et, après une brève introduction au problème discret et aux relations entre les équilibres et l'optimisation, cinq travaux sont présentés:

- dans la première section je présente celui où, avec G. Carlier et C. Jimenez, nous avons introduit la même théorie dans un cadre continu, avec des densités de trafic vues comme des métriques sur un domaine donné: la formalisation passe par des mesures sur les courbes mais on verra ensuite que dans certains cas on peut se réduire à un problème plus simple, de minimisation d'une fonctionnelle convexe sur des champs de vecteurs dont la divergence est prescrite; la réduction à ce problème vectoriel et les ingrédients pour l'équivalence viennent d'un article qui a suivi avec L. Brasco et G. Carlier, qui conclut la section, en précisant que l'équivalence démontrée demande une certaine régularité du champ de vecteur optimal;

- la deuxième section porte sur ces questions de régularité, qui présentent un grand intérêt en soi, car dans les cas qui sont raisonnable en modélisation, le champ de vecteur optimal résout une équation elliptique très dégénérée; dans l'article avec Brasco et Carlier on montre ce dont on a besoin pour les applications à la congestion, mais dans un autre article, en collaboration avec V. Vespri, on donne aussi un résultat de continuité qui a un intérêt en soi (et qui a en outre des applications et à la congestion et à d'autres cadres);

- la troisième section présente ce qu'on a fait, avec F. Bemansour, G. Carlier et G. Peyré, sous le point de vue numérique, ce qui a engendré deux articles, l'un où l'on explique comment calculer le sous-gradient des distance géodésique calculé par Fast Marching, ce qui a plusieurs applications, l'autre où l'on applique tout cela aux problèmes de congestion, en démontrant la convergence.

Le troisième chapitre ne se réfère qu'à deux travaux complets (il y en a d'autres qui sont en cours et qui sont mentionnés dans la partie "perspectives"), qui sont des contributions à la théorie du transport branché un peu différentes de qui avait été fait auparavant (d'où le mot "last" dans le titre, car cela fait un peu une césure, qui conclut l'étude théorique de ces problèmes, pour laisser ouvertes d'autres questions vers le numérique et la théorie des jeux):

- dans le premier, avec J.-M. Morel, nous étudions la régularité des réseaux optimaux en vue d'établir rigoureusement des "lois d'angle" aux points de jonction: nous donnons un résultat de régularité si la mesure cible est équivalente à Lebesgue, nous donnons un contre-exemple avec une mesure atomique;

- le deuxième donne un résultat de $\Gamma-$convergence pour une suite de fonctionnelles elliptiques à la Modica-Mortola qui convergent bien vers la fonctionnelle optimisée dans le cas du transport branché. Ceci ouvre des possibilités numériques très intéressantes et dans la partie de perspectives j'en offre un aperçu.

Du quatrième chapitre j'ai déjà parlé plus en détail, je me limite ici à dire qu'il contient deux travaux

- l'un en collaboration avec M. Bernot et A. Figalli, sur les modèles variationnels à la Brenier (configuration initiale et finale fixée, contrainte d'incompressibilité, minimisation de l'énergie cinétique), où l'on donne une caractérisation partielle des solutions en dimension un et deux, en trouvant des solutions inconnues auparavant;

- l'autre en collaboration avec B. Maury et A. Roudneff-Chupin, sur un problème de flot-gradient avec contraintes de densité maximale (ici seule a configuration initiale est fixée), où l'on étudie la convergence d'un schéma discret.

Je veux maintenant présenter un peu plus en détails certains aspects de ma recherche, ses fils conducteurs, et ce que j'aurais voulu développer plus mais que je n'ai pas encore pu faire de manière satisfaisante: il est certain qu'il s'agira d'un bon point de départ pour mes recherches futures.

Optimisation à divergence fixée. Ceci est l'un des fils conducteurs de mes recherches et de ce mémoire, et il l'était déjà à l'époque de ma thèse de doctorat. Il est connu que le problème de Monge

$$\min\Big\{ \int |x-y|\, d\gamma \; : \; \gamma \in \Pi(\mu,\nu), \Big\},$$

est équivalent à ce qui a été introduit par Beckmann sous le nom de *Continuous transportation problem*:

$$\min\Big\{ \int |v(x)|\, dx \; : \; \nabla \cdot v = \mu - \nu \Big\},$$

c'est-à-dire la minimisation de la norme L^1 à divergence fixée. En fait, pour être précis, il faudrait se mettre dans les mesures vectorielles et minimiser la variation totale. La densité de transport que je traite dans la Section 1.2 coïncide justement avec la variation totale $|v|$ du champ de vecteur optimal, et ses estimations L^1 et L^p permettent, entre autres, de résoudre ces problèmes dans le cadre des fonctions plutôt que des mesures.

Or, Beckmann avait introduit ce problème dans [41] de manière plus générale, en minimisant une fonctionnelle convexe

$$v \mapsto \int H(|v(x)|)dx,$$

sous les mêmes contraintes de divergence ($H : \mathbb{R} \to \mathbb{R}$ étant une fonction convexe et croissante). Le formalisme de Beckmann manquait hélas de rigueur mathématique et, même si on en soupçonnait l'équivalence, la construction qu'on a introduite avec Carlier et Jimenez pour étudier les problèmes de congestion et faire le lien avec les équilibres de Wardrop était fine et compliquée et passait par les mesures sur l'espace des chemins. Ce n'est que dans le papier avec Brasco qu'on a pu démontrer l'équivalence entre le problème vectoriel à divergence fixée et e problème scalaire lié à l'équilibre de Wardrop.

Mais il est intéressant de voir que le transport branché présenté au Chapitre 3 aussi relève de ce cadre: dans sa formulation proposée par Q. Xia ([107]), le problème qui est obtenu par relaxation du cas discret (introduit par Gilbert dans [83]) est

$$\min\big\{ M^\alpha(v) \; : \; \nabla \cdot v = \mu - \nu \big\},$$

où les v admissibles sont forcement des mesures (et en fait ils se concentreront sur des graphes, finis ou rectifiables) et

$$M^\alpha(v) = \begin{cases} \int_M \theta^\alpha d\mathcal{H}^1 & \text{si } v = U(M,\theta,\xi), \\ +\infty & \text{sinon,} \end{cases}$$

où l'écriture $v = \theta\xi \cdot \mathcal{H}^1_{|M}$ signifie que v est une mesure vectorielle concentrée sur l'ensemble $1-$rectifiable M et que sa densité par rapport à la mesure de Hausdorff $\mathcal{H}^1$ sur M est donnée par la multiplicité réelle $\theta : M \to \mathbb{R}^+$ fois une orientation (vecteur unité tangent) $\xi : M \to \mathbb{R}^d$. L'exposant α est entre 0 et 1, de manière à avoir un comportement sous-additif de la fonctionnelle et favoriser le mouvement conjoint.

Ce que j'ai proposé dans la Section 3.2 se base justement sur cette formulation: plutôt que minimiser la fonctionnelle M^α qui est très singulière, l'approcher par des fonctionnelles "elliptiques" M^α_ε du type

$$M^\alpha_\varepsilon(v) = \varepsilon^{\gamma_1} \int |v(x)|^\beta dx + \varepsilon^{\gamma_2} \int |\nabla u(x)|^2 dx,$$

pour un exposant $\beta \in]0,1[$.

De cette manière le problème de congestion et celui de branchement ne sont que les correspondants convexe et concave, respectivement, du problème de Monge. Il est utile de remarquer que toute cette machinerie à divergence fixée n'existe que dans le cas de Monge (coût = distance, et non $|x-y|^p$).

Mesures sur l'espace des chemins. Un autre outil qui apparaît souvent dans ce mémoire est celui des mesures sur les chemins. Je vais tâcher de l'introduire de manière la plus générale possible.

Soit $\mathbb{T} \subset \mathbb{R}$ un ensemble de "temps" et C l'ensemble de "courbes", c'est-à-dire des fonctions continues définies sur $\mathbb{T}$ et à valeur dans un espace Ω donné. Évidemment, si $\mathbb{T}$ est discret (un nombre fini de points, par exemple), toute trajectoire sera continue et donc admise dans C. De plus, on préfère parfois imposer d'autres conditions aux courbes (plus que continues, Lipschitziennes...). On cherche une mesure Q sur C, qui minimise une fonctionnelle $J(Q)$ parmi toutes les mesures qui satisfont des contraintes sur leurs projections $(\pi_t)_\# Q$ et/ou sur des projections couplées $(\pi_s, \pi_t)_\# Q$.

Si on considère $\mathbb{T} = \{0,1\}$ on tombe exactement sur le problème de Kantorovitch: on cherche une mesure sur des courbes à deux temps, donc sur $\Omega \times \Omega$, en fixant les deux marges. Évidemment on ne fixe pas la marge couplée $(\pi_0, \pi_1)_\# Q$ sinon il n'y aurait qu'une seule mesure Q admissible. Cet ensemble de temps est celui qui donne la description la moins détaillée possible d'un mouvement: il ne voit que le point de départ et le point d'arrivé.

Par contre, pour plusieurs applications, des descriptions plus détaillées sont nécessaires car on veut voir le parcours suivi par chaque particule. On prend alors $\mathbb{T} = [0,1]$, et les contraintes peuvent être sur $(\pi_0)_\# Q$ et $(\pi_1)_\# Q$ ou sur $(\pi_0, \pi_1)_\# Q$, comme dans les problèmes de congestion. La fonctionnelle ne sera plus dans cas linéaire mais elle s'exprimera comme $Q \mapsto \int_C Z_Q(\omega) Q(d\omega)$ où $Z_Q(\omega)$ répresente le coût pour la courbe $\omega \in C$ associé à la configuration Q (qui peut donner lieu à une intensité de trafic plus ou moins élevée sur ω). Une fonctionnelle de la même forme se trouve dans les problèmes de transport branché, mais pour une autre expression de Z_Q, qui favorise dans ce cas le groupement des courbes et le fait que ω soit chargé par Q.

Un autre cas qui rentre dans ce cadre est celui des problèmes des fluides incompressibles: dans ce cas la fonctionnelle est linéaire (l'intégrale de l'énergie cinétique de chaque ω) et les contraintes portent sur tous les $(\pi_t)_\# Q$ ainsi que sur $(\pi_0, \pi_1)_\# Q$.

Pour ce qui est des autres ensembles de temps $\mathbb{T}$, il arrive parfois de considérer $[0, +\infty[$ (en transport branché, on impose souvent que les courbes soit $1-$Lipschitz et on ne met pas de bornes a priori sur leur longueur), ou bien $\{t_0, t_1, \dots, t_n\}$, un autre ensemble discret que $\{0,1\}$. Ce dernier choix est utile parfois dans des démonstrations, comme quand on démontre que les distances de

Wasserstein sont en fait des distances (et on utilise alors trois temps, dont un intermédiaire) ou quand on veut construire une mesure sur les courbes en partant d'une courbe des mesures (voir [88], mais on s'occupe de cela dans [17] aussi); là, la partition $\{t_0, t_1, \ldots, t_n\}$ deviendra à la limite de plus en plus fine et tendra vers $[0, T]$.

Il est intéressant de voir comment souvent des problèmes posés naturellement sur les mesures sur les chemins peuvent devenir des problèmes à divergence fixée (ce qui est le cas de la congestion et du transport branché, et l'on peut dire d'une manière un peu souple que les problèmes de mécanique des fluides qui font intervenir les équations d'Euler concernent aussi des conditions de divergence). De plus, il est parfois possible de faire le lien avec une description intermédiaire, qui est celle des courbes de mesures. Par exemple le modèle du transport branché presenté en [15] essayait justement d'expliquer ce phénomène de branchement à travers une métrique riemanienne sur les mesures favorisant celles qui sont plus concentrées (le modèle n'étant pas équivalent aux modèles de transport branché du Chapitre 3, il y a du travail à faire pour comprendre les différences et les équivalences, ce qui est l'objet du travail en cours avec Brasco, [17]).

EDP. Même si les travaux que je présente sont des travaux principalement de calcul des variations, il est bien connu que cette branche des mathématiques est profondément liée aux équations aux dérivées partielles, et celles-ci apparaissent aussi en force dans mes études.

Pour commencer, les équations qui apparaissent le plus sont celles qu'on peut trouver en tant qu'*Équation d'Euler-Lagrange* pour une fonctionnelle à optimiser. Dans le cas du transport, il s'agit de l'équation de *Monge-Ampère*, qui prend une forme assez simple dans le cas du coût quadratique. Ses extensions à d'autres coûts ont été récemment étudiées, en répondant à des questions pointues de régularité, par Ma-Trudinger-Wang, Figalli, Loeper...

Dans ma recherche, j'ai souvent utilisé - sans y prouver des nouveaux résultats, simplement en appliquant les anciens - des résultats sur cette équation elliptique et sur d'autre. Ceci est le cas de l'étude de certains modèles de planification urbaine que j'ai fait au début de ma thèse, ou d'autres travaux, principalement en optimisation de forme, où les fonctionnelles concernée font expliciment intervenir la solution d'une EDP. Ces travaux d'optimisation de forme, n'ont malheureusement trouvé la place ni dans ma thèse de doctorat ni ici, à cause de leur marginalité par rapport à la composante transport de ma recherche.

J'ai donc eu l'occasion de toucher à l'équation de Monge-Ampère ainsi qu'à *l'équation de Laplace* $\Delta u = f$ *ou du* $p-$*Laplacien* $\Delta_p u = f$. Ces équations sont toutes elliptiques dans un certain sens et n'ont pas d'aspects évolutionnaires. Dans le même cadre on peut voir l'équation que j'ai réellement étudiée, en collaboration avec Brasco, Carlier et Vespri, et qui apparaissait comme équation d'Euler d'une fonctionnelle convexe de congestion:

$$\nabla \cdot (\nabla K(\nabla u)) = f,$$

où K est une fonction convexe mais très dégénérée (il faut penser à $K(z) = \big(|z| - 1\big)_+^p$, une fonction qui est complètement platte sur une boule). Le type de problème est le même que le $p-$Laplacien mais la dégénérescence beaucoup plus forte. On a démontré sur cette équation des résultats de régularité H^1, L^∞ et C^0 (ce dernier en dimension deux seulement) sur $\nabla K(\nabla u)$ que je présente dans la Section 2.2.

De plus, il y a dans la théorie du transport un côté *équations d'évolution*, principalement grâce à l'*équation de continuité*

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho v) = 0$$

qui régit l'évolution d'une densité ρ_t. Des équations où le champ de vitesse v dépend de la densité ρ, en se donnant la densité initiale ρ_0 sont souvent étudiées par la théorie du transport en tant que *flots gradients* d'une fonctionnelle dans l'espace de Wasserstein. C'est par exemple le cas de l'équation de la chaleur, de l'équation des milieux poreux ou de Fokker-Planck. Sinon, il se peut que l'équation sur (ρ, v) ci-dessus soit couplée avec une autre faisant intervenir $\partial v/\partial t$, ce qui est typiquement associé à des conditions initiales et finales sur ρ (et qui correspond donc à un problème géodésique).

Ces deux aspects sont présents dans le dernier chapitre: l'équation étudiée dans la Section 4.2 peut être interprétée comme une forme dégénérée de celle des milieux poreux, et l'équation d'Euler pour les *fluides incompressibles* de la Section 4.1 rentre sans problème dans le cadre géodésique. Il est d'ailleurs bien connu qu'elle constitue l'équation des géodésiques dans le groupe des difféomorphismes qui préservent le volume. Par ailleurs, j'ai étudié ce type d'équations à l'époque de ma thèse aussi, mais plutôt en regardant des fluides compressibles, en collaboration avec L. Ambrosio dans [14], où l'on analyse les conditions d'optimalité d'un problème géodésique introduit avec A. Brancolini et G.. Buttazzo dans [15].

Et, pour finir, il y a la classe des EDP d'Hamilton-Jacobi qui est en relation avec le transport. On trouve une *équation d'Hamilton-Jacobi* avec le temps quand on regarde le problème dual avec les potentiels de Kantorovitch (voir par exemple la formulation dynamique des problèmes de transport avec coût $|x-y|^p$, [84], ou l'approche Benamou-Brenier, [44, 45]); par contre, dans le cas du coût $c(x,y) = |x-y|$, l'équation satisfaite par le potentiel u est une *équation Eikonale* $|\nabla u| = 1$. Le potentiel et la densité de transport satisfont ensemble le *Système de Monge-Kantorovitch*

$$\begin{cases} \nabla \cdot (\sigma \nabla u) = \nu - \mu & \text{dans } \Omega \\ |\nabla u| \leq 1 & \text{dans } \Omega, \\ |\nabla u| = 1 & \text{p.p. sur } \sigma > 0. \end{cases}$$

Et l'équation Eikonale apparaît aussi dans l'étude qu'on fait des problèmes de congestion, car dans le problème dual on a besoin d'associer à toute métrique ξ la distance géodésique c_ξ qui lui est associée. Il s'agit donc de trouver la *solution de viscosité* de l'équation $|\nabla u| = \xi$. Dans le Chapitre 2 on verra non seulement l'équivalence entre plusieurs définitions de solution dans le cas de $\xi \in L^p$, mais on utilisera aussi la méthode du Fast Marching pour l'analyser numériquement.

Numérique. Je profite du fait que je viens de nommer la méthode du Fast Marching pour me concentrer un peu sur l'aspect numérique de mes recherches. Ceci est un aspect complètement nouveau, qui n'existait pas du tout à l'époque de ma thèse. Il était absent de ma formation (qui est plutôt de maths pures) et c'est sûrement grâce au fait que je me suis installé en France que j'ai commencé à le côtoyer et à l'apprécier.

Comme on peut voir, il y a du numérique pratiquement dans tous les chapitres de ce mémoire. Le dernier est peut-être le plus susceptible d'en contenir, d'un côté à cause des personnes qui y sont impliquées (je pense à Y. Brenier et B. Maury), de l'autre parce que ces sujets ont bien fait l'objet de simulations numériques, mais c'est le seul où je n'ai pas voulu mettre des images dans le mémoire. C'est parce que je n'y ai malheureusement - pour l'instant - pas participé du tout, et tout le développement du numérique pour Euler Incompressible et pour le mouvement de foules selon les modèles présentés dans le chapitre est dû à Brenier et Maury ainsi qu'à M. Bernot, J. Venel et A. Roudenff-Chupin.

Par contre il y a trois méthodes numériques dans les trois premiers chapitres. Dans le premier il s'agit de quelque chose de très spécifique: on approche le problème du transport avec coût quadratique d'une densité diffuse vers une mesure atomique finie en partant du Transport de Knothe et en passant par les problèmes avec coût $\sum_{i=1}^{d} |x_i - y_i|\varepsilon^{d-i}$ (Kntohe étant obtenu comme limite pour $\varepsilon \to 0$ et le transport usuel avec $\varepsilon = 1$). On écrit une EDO pour l'évolution des potentiels et on l'approche numériquement (après avoir montré le comportement Lipschitz de tous les ingrédients, ce qui montre le caractère bien-posé de l'équation et justifie la convergence).

Dans le deuxième chapitre les perspectives numériques, bien qu'en cours, sont déjà beaucoup plus élargies: il s'agit d'un travil en collaboration avec E. Oudet pour utiliser l'approximation elliptique à la Modica-Mortola du transport branché. Oudet a déjà utilisé des approximations comme celle-ci pour des problèmes qui font intervenir le perimètre (voir [96]) et avec d'excellents résultats. Pareillement, ceux qu'on a obtenu en transport branché sont très prometteurs. L'idée serait aussi de modifier les fonctionnelles pour couvrir d'autres cas, dont le *Problème de Steiner*. On est en train d'y travailler et j'espère montrer un résultat de $\Gamma-$convergence; par contre, j'avoue que, étant donnée la non-convexité des fonctionnelles, la justification de la convergence vers le minimiseur n'est pas achevé.

Au chapitre 2 on s'est retrouvé dans la situation suivante: en vue des applications à la congestion, on avait une fonctionnelle définie sur des métriques ξ discrétisées (définies sur les points d'une grille carrée), qui passait par les valeurs $c_\xi(x_i, y_j)$, c'est-à-dire les distances entre des points x_i et des points y_j, calculée par rapport à cette métrique ξ, ce qui se fait en discret par la méthode du Fast Marching. Comme on devait optimiser par rapport à ξ, on voulait dériver ces c_ξ par rapport à ξ. On a trouvé une manière de calculer le gradient de ces quantités par rapport à ξ (en fait, c'est plutôt le sous-gradient, ou mieux le sur-gradient, comme il s'agit de fonctions concaves non lisses de ξ) e intégrant ce calcul dans la routine même du Fast Marching. On l'a appelé Subgradient Marching Algorithm. Après, on l'a appliqué aux problèmes de congestion (dans [7], avec démonstrations de convergence). Dans [6], par contre, on donne le détail de l'algorithme du calcul de sous-gradient et on l'applique à d'autres problèmes. Quant il s'agit de problèmes convexes la convergence ne pose pas de difficultés - par un algorithme de sous-gradient projecté - et les résultats sont satisfaisants. On a essayé aussi de l'utiliser sur un problème de tomographie (d'une métrique inconnue on connaît les distance d'un nombre élévé de couples de points, par exemple sur le bord: on cherche à trouver la métrique en minimisant un critère de moindre carrés plus une pénalisation régularisante) mais seulement pour montrer que d'autres applications étaient possibles.

Dans tous les cas, quand j'ai dit que j'ai été impliqué dans le calcul numérique, je veux dire que j'ai discuté profondément des détails de l'algorithme avec les numériciens, qui se sont après occupés du codage en lui même. Il est intéressant de voir que l'expérience en algorithmes de sous-gradient pour des problèmes convexes m'a permis aussi, récemment, de collaborer avec des statisticiens pour un problème de régression (un travail conjoint avec F. Balabdaoui et K. Rufibach, voir sur ma page `www.ceremade.dauphine.fr/~filippo`, car il n'est pas présent en bibliographie à cause de son éloignement excessif par rapport aux sujets de ce mémoire. . .).

Théorie des jeux. Le transport optimal permet bien, si on le veut, de se rapprocher des sujets de la théorie des jeux, et mon insertion dans l'équipe d'économie mathématique du CEREMADE m'y a aidé encore plus. Je suis rentré en contact avec une équipe active et intéressé par les sujets dont je pouvais parler. Évidemment, comme ce lien est quelque chose de nouveau pour moi, ce que je peux présenter en théorie des jeux est encore limité.

Il y a dans ce mémoire deux sujets qui rentrent dans le cadre, et les deux sont liés à des équilibres sur des réseaux de transport. L'étude des *équilibres de Wardrop* est classique dans la théorie des jeux: un réseau est donné et l'on connaît, pour chaque arrête, quel sera son temps de parcours d'après le nombre d'agents qui y passent (un temps qui empire s'il y a plus de monde); on connaît les origines et les destinations de chacun et on se demande comment il se répartissent. Évidemment le choix de chacun ira sur le parcours qui lui proposera un temps minimal pour aller de son origine à sa destination, mais son choix influencera aussi tous les temps de parcours de tout le monde. On trouvera un équilibre si la configuration est telle que personne n'a envie de changer de parcours. C'est un *équilibre de Nash* (le nom de Wardrop est associé à l'équilibre de Nash dans ce type de jeu). On peut démontrer (voir le début du Chapitre 2) qu'une configuration est d'équilibre si et seulement si elle optimise une certaine fonctionnelle de congestion totale.

Tout cela est bien connu dans le cadre discret et la nouveauté du Chapitre 2 est le cadre continu, où le réseau est remplacé par un domaine qui admet tout parcours possible.

L'autre problème d'équilibre sur les réseaux que je nomme se trouve au Chapitre 3 et il est de nature complètement différente. Il s'agit cette fois là d'un *jeu coopératif.* Dans le problème du transport branché on veut relier plusieurs origines à plusieurs destinations (une, dans le cas le plus simple), en minimisant un coût qui encourage la collaboration. Mettons nous dans un cas simple: un certain nombre de petites villes doivent bâtir leur réseau routier pour se connecter à la capitale. Elles le feront de manière optimale, pour minimiser le coût total. Et après, comment se répartir ce prix à payer? Il s'agit exactement d'un problème de recherche de *noyau* (*core*) dans un jeu coopératif où chaque coalition a son propre gain (ou coût: ici c'est le coût pour connecter à la capitale un certain sous-ensemble de villes). On peut imaginer plusieurs manières raisonnables de partager le coût, mais il n'est même pas évident qu'il y en ait une qui marche (telle qu'il n'y ait pas une sous-coalition qui puisse abandonner la grosse coalition pour aller seule, et payer moins cher). J'ai étudié ce problème avec un étudiant de Dauphine, en lui donnant des cas précis à analyser pour son mémoire de L3. J'en parle à la fin du Chapitre 3 comme travail à continuer.

Économie Mathématique. L'économie mathématique serait l'étape suivante à la quelle se dédier (en appliquant éventuellement la théorie du transport optimal) après l'optimisation et la théorie des jeux, mais dans ce contexte je ne peux que parler de deux sujets qui font l'objet de travaux en cours. Mon avancement dans ce sujet est plus lent, mais celle-ci constituera sans aucun doute l'une des directions futures de me recherches. Je dirai donc très brièvement quels sont les deux problèmes auxquels je me réfère.

Le premier est le problème de Rochet-Choné, ou *principal-agent problem*, un problème de théorie des contrats (disons qu'il s'agit d'une firme monopoliste qui doit choisir sa production et ses prix, pour optimiser ses revenus) qui peut soit s'exprimer comme un problème d'optimisation dans l'ensemble des fonctions convexes, soit comme un problème posé sur les mesures, à travers la correspondance qui associe à chaque fonction convexe la mesure image de son gradient. Cette correspondance est bijective grâce au théorème de Brenier sur le transport optimal dans le cas quadratique. La solution de ce problème n'est pas connu dans le cas général, ni d'ailleurs les bonnes conditions d'optimalité et plusieurs questions de régularité restent ouvertes. On envisage une étude à ce propos avec G. Carlier.

D'autre côté, j'ai commencé récemment à regarder avec A. Lachapelle, qui est étudiant en thèse à Dauphine, un cas précis, qui est la comparaison entre un Rochet-Choné deterministe (le cas usuel) ou aléatoire (où la firme peut vendre des contrats qui sont en fait des loteries: d'après la repartition d'agents selon les types, et leurs préférences et aversion au risque, existe-t-il des cas où la firme à intérêt à vendre des loteries?).

Et l'autre problème est le problème de Hotelling. Il s'agit d'un problème où un certain nombre des firmes doivent choisir leurs positions et les prix à pratiquer pour maximiser leurs revenus. Les firmes étant en concurrence, c'est un équilibre de Nash qu'on cherche. Dans un travail conjoint avec P.Mossay, toujours en cours (et qui a été relancé par la récente rencontre qu'on a organisée à Paris à l'École des Mines autour du transport et de l'économie), on donne une approche par potentiels de Kantorovitch. De plus, on a une conjecture qui a passé plusieurs tests pour l'instant, suivant laquelle la configuration spatiale optimale des firmes est celle qui optimise la somme des revenus des firmes moins le coût total de transport des agents, en prenant bien entendu en compte que pour chaque configuration spatiale seul les prix qui donnent un équilibre sont retenus.

Calcul variationnel au-delà du transport optimal. Comme je suis en train de faire une présentation générale des aspects divers et variés de ma recherche, je veux dédier quelques lignes à ce qui n'apparaît pas dans ce mémoire, car il ne touche pas, ou pas assez, au transport optimal.

J'ai déjà vaguement mentionné les problèmes d'optimisation de forme: j'ai étudié, en collaboration avec G. Buttazzo, des problèmes d'optimisation de *compliance*. Rapidement: dans un domaine $\Omega \subset \mathbb{R}^d$ soumis à une force $f \in L^2(\Omega)$ la compliance $C(\Sigma)$ d'un sous-ensemble Σ est la valeur

$$\int_\Omega f u_{\Sigma,f,\Omega} dx \quad \text{où} \quad \begin{cases} -\Delta u_{\Sigma,f,\Omega} = f & \text{en } \Omega, \\ u_{\Sigma,f,\Omega} = 0 & \text{sur } \partial\Omega \cup \Sigma. \end{cases}$$

Le problème consistant à optimiser cette quantité sous des contraintes sur Σ est un problème typique de l'optimisation de forme. Nous nous sommes surtout occupés du cas où Σ devait être constitué par un nombre fixé de boules où par des points (si l'on remplace le Laplacien avec un $p-$Laplacien, pour des raisons capacitaires) ou par une courbe connexe de longueur fixée. On a surtout étudié l'asymptotique pour voir comment les boules ou la courbe se distribuent dans Ω si les nombre de boules ou la longueur augmentent.

Cela est en relation avec des problèmes asymptotiques de *positionnement optimal*, pour minimiser typiquement la fonctionnelle

$$\Sigma \mapsto \int d(x,\Sigma) f(x) dx.$$

Dans [16] on a étudié des variantes intéressantes appélées *short term* ou *sequential location*, où le nombre de points tend vers l'infini mais il faut les placer l'un après l'autre, sans déplacer ceux qui ont déjà été placés. Sinon, une autre question très intriguante qui est en cours d'étude est celle où f n'est pas seulement positive: il y a des régions qui veulent se rapprocher de Σ et d'autres qui veulent s'en éloigner. C'est le sujet d'un mémoire de DEA à Pise que G. Buttazzo a confié à une étudiante, I. Lucardesi, mais on envisage de travailler ensembles sur les questions qui resteront ouvertes après.

Un autre sujet dans le calcul variationnel mais plus éloigné du transport est celui que j'ai approché en collaboration avec F.-X. Vialard lors de mon post-doc à Cachan (là aussi, voir ma page web pour les références de l'article). Pour des applications à l'appariement d'images via des groupes de difféomorphismes on s'est intéressé à la quantité $\int f(\phi_t) g dx$ où f et g sont des fonctions peu régulières (disons BV) et ϕ_t est une famille de difféomorphismes obtenue en intégrant un champ de vecteurs X. Donner une formule pour la dérivée (par rapport à t) de cette quantité est évident si tout est régulier mais pas trivial dans le cadre BV. La démontrer a fait sortir des questions intéressantes sur les fonctions BV.

Encadrement et enseignement au niveau recherche J'ai déjà mentionné de manière indirecte ce sujet une ou deux fois au cours de cette introduction, mais je tiens à le résumer pour que l'aspect "diriger des recherches" prenne sa place.

Le seul étudiant e thèse dont j'ai un pourcentage officiel d'encadrement est Lorenzo Brasco, doctorant en cotutelle entre Pise et Dauphine, co-encadré par G. Buttazzo (34%), G. Carlier (33%) et moi même (33%). Lorenzo est basé à Pise et vient passer des périodes à Dauphine, qui se sont avéré plutôt productives. Les sections 2.1 et 2.2 de ce mémoire se basent sur un article co-écrit avec lui et G. Carlier, qui est celui où l'on montre des résultats de régularité elliptique (sujet sur lequel il est très compétent) et on utilise la théorie de DiPerna-Lions pour définir un flot dans le cadre d'un problème congestionné. De plus, on a un autre article en cours, qui penche plutôt dans la direction du transport branché. Je crois en plus que, en ayant étudié à Pise mais étant Maître de Conférences à Dauphine, mon encadrement joue un rôle plutôt important dans sa thèse, en tant que lien et intermédiaire entre les autres directeurs de thèse.

Une autre étudiante de doctorat que je cotoie régulièrement est Aude Roudneff-Chupin, qui est en thèse à Orsay avec B.Maury. Bien que je n'aie aucun rôle officiel dans son encadrement, sa thèse porte spécifiquement sur les modèles macroscopiques en mouvement de foule et c'est bien ce qu'on est en train de faire ensemble ([13]).

Comme j'ai précisé avant, j'ai commencé récemment à travailler avec Aimé Lachapelle, étudiant en thèse avec G. Carlier à Dauphine, qui m'a demandé personnellement si je voulais le suivre sur une petite partie de sa thèse. Ce travail qu'on fait ensemble est encore bien à définir, mais pour l'instant cela se propose plutôt comme une direction d'une partie de ses recherches (je lui conseille des voies à suivre pour la résolution d'un problème, et on en discute), plutôt qu'une simple collaboration sur un article. Je tiens quand même à préciser que, surtout sur d'autres sujets liés à sa thèse, et notamment les *jeux à champ moyen*, c'est sûrement moi qui ai eu l'occasion d'apprendre quelque chose d'Aimé.

Pour terminer, un encadrement au niveau beaucoup moins pointu: j'ai récemment dirigé un stage de Licence d'un briliant étudiant dauphinois, Lambert Piozin. Ce stage, à la différence des autres qui ont été menés par ses collègues à la même occasion, a abouti à des résultats nouveaux et intéressants.

Je lui ai confié des calculs qui visaient à étudier les équilibres en jeux coopératifs dans un réseau branché et il a su rentrer dans le sujet.

Je termine ce paragraphe, et cette introduction, en parlant d'enseignement. Ce n'est pas par hasard, et ce n'est pas une contradiction, c'est une partie de mon travail que j'aime bien et qui complète très bien le côté recherche. Or, ce n'est pas ici que je vais parler de mon enseignement des fonctions d'une variable au jeunes dauphinois du 1er cycle, et je veux seulement nommer ce que j'ai fait au niveau plus élévé. Un cours dispensé devant un public de thésards et des collègues est une autre manière de rentrer dans les directions que prennent les recherches des autres. Ce n'est pas vraiment "diriger" une recherche mais cela joue un rôle que je trouve passionnant et important.

Après quelques premières expériences, commencés en tant que postdoc quand j'ai partagé avec J.-M. Morel un mini-cours sur le transport branché au Wolfgang Pauli Institut de Vienne, j'ai participé avec F. Bolley et B. Nazaret à un cours d'école doctorale sur le transport à l'IHP (2008) et à Dauphine (2009). De plus, ce juin 2009 j'ai eu l'honneur d'être invité à dispenser deux cours différents (voir [24] and [25] pour les *lecture notes*) lors de l'école d'été organisée par l'Institut Fourier de Grenoble. J'ai donné, parallèlement aux cours dispensés par d'autres collègues, une série de leçons sur des problèmes de modélisation en économie, trafic et planification urbaine qui font intervenir le transport optimal et j'ai aussi ouvert les trois semaines de l'école par un cours introductif sur la théorie générale du transport. Le public était assez international et cela a été une expérience enrichissante .

Et je pense que c'est l'heure d'arrêter cette présentation qui n'est pas trop mathématique et de rentrer dans les détails scientifiques, les théorèmes, les stratégies. . .

Chapter 1

Contributions on true Monge-Kantorovich problems

Résumé Je présente dans ce chapitre trois articles qui ont comme point commun qu'ils concernent la vraie théorie de Monge-Kantorovitch, c'est-à-dire le problème de minimisation (1.3.1) ci-dessous, sans variantes de congestion ou de branchement.

Dans le premier on considère la relation entre le transport de Knothe (qui est grosso modo le transport qui réarrange de manière monotone une composante, puis la composante suivante en regardant les désintegrations par rapport à celle d'avant... en pratique il fait ce qu'un trasport optimal ferait en dimension un, en donnant un ordre de priorité aux composantes) et des transports optimaux qui donnent des poids différents aux différentes composantes, à la limite où chaque poids devient négligeable par rapport aux précédents. On en démontre la convergence et on en tire une méthode numérique pour réconstruire un transport optimal en partant de celui de Knothe, qui est beaucoup plus simple à calculer.

Le deuxième article concerne la densité de transport et redémontre des estimations L^p qui étaient déjà connues mais avec une technique complètement différente. Cette technique, analytique et non géométrique, permet d'obtenir des résultats plus simples et parfois plus forts.

Le troisième démontre l'existence d'un transport optimal pour le coût $c(x,y) = ||x-y||^p$, pour une norme quelconque sur $\mathbb{R}^2$ et $p > 1$, en mettant le résultat dans le cadre d'une stratégie générale pour des problèmes de transport avec coût convexe mais non strictement convexe. Cette stratégie fait intervenir des problèmes avec contraintes ($x - y$ qui doit appartenir à un convexe donné), c'est à dire avec des coûts qui peuvent valoir $+\infty$ et, pour qu'elle marche dans le cas général, il y a encore beaucoup d'étapes à franchir

In this chapter I want to present my researches on what we could call the "true" transport theory, i.e. those problems that do not exit out of the framework of the classical theory of Monge-Kantorovich minimization

$$\min\left\{\int_{\Omega\times\Omega} c(x,y)\,\gamma(dx,dy)\ :\ \gamma\in\Pi(\mu,\nu)\right\}, \tag{1.1}$$

where $\Pi(\mu,\nu)$ is the set of transport plans with fixed marginals μ and ν. In particular, no congestioned or branching variant is considered. This is a recent part of my researches, since during my thesis and right after most of my works included concentration criteria giving rise to congestion or branching effects. Yet, even if recent, this is a subject I'm more and more involved in.

I will mention three problems, one that I studied in collaboration with G. Carlier and A. Galichon ([2]), one alone ([3]), one with G. Carlier and L. De Pascale ([4]).

The first deals with quadratic costs with vanishing weights on the different coordinates and the relation between the corresponding optimal transport and the so-called Knothe rearrangement, giving an answer to a conjecture by Y. Brenier. As a byproduct of the limit result we proved, that can be interpreted as a continuity result, we get a continuation method for getting the usual optimal transport starting from the Knothe one. The latter being much easier to compute, a numerical method is obtained.

The second deals with the concept of transport density, which is a scalar measure associated to the problem of Monge, i.e. when $c(x,y)=|x-y|$. This measure gives the quantity of mass passing through each region during the transport. Summability estimates on it find various applications.

The third one is mainly a strategy, which is able to give some partial result on the classical problem of existence of an optimal transport map in the case $c(x,y)=\phi(x-y)$ when ϕ ic convex but not strictly convex. It is successfully applied to the case of a cost which is a convex and increasing function of an arbitrary norm in $\mathbb{R}^2$.

1.1 From Brenier to Knothe and from Knothe to Brenier

1.1.1 The transport of Knothe and the convergence result

The Knothe transport from a measure μ to a measure ν in $\mathbb{R}^d$ is roughly defined in the following way: let us denote by μ^d and ν^d the d-th marginal of μ and ν; et $T_d=T_d(x_d)$ be the monotone nondecreasing map transporting μ^d to ν^d (such a map is well-defined and unique as soon as μ_d has no atoms). Then disintegrate μ and ν according to x_d and y_d, respectively. For every value of x_d and y_d, call $\mu^{d-1}_{x_d}$ and $\nu^{d-1}_{y_d}$ the $(d-1)-$th marginals of these disintegrated measures. Let $T_{d-1}=T_{d-1}(x_{d-1},x_d)$ be such that $T_d(.,x_d)$ is monotone and maps $\mu^{d-1}_{x_d}$ to $\nu^{d-1}_{T_d(x_d)}$. Then go on with successive disintegrations and marginals: for any $k\leq d-1$ the measures $\mu^k_{(x_d,\dots,x_{k+1})}$ are defined as the $k-$th marginals of the disintegration of μ with respect to $(x_d,x_{d-1},\dots,x_{k+1})$ and $\nu^k_{(y_d,\dots,y_{k+1})}$ are defined analogously. What we are doing is defining a map which looks first at the x_d and y_d coordinates, then at the $(d-1)-$th and so on, repeating the construction iteratively until we define $T_1(x_1,x_2,...,x_d)$, which is monotone in x_1 and transports $\mu^1_{(x_d,\dots,x_2)}$ onto $\nu^1_{T_2(x_2,\dots,x_d)}$. Finally, the *Knothe rearrangement* T is defined by $T(x)=(T_1(x_1,x_2,...,x_d),...,T_{d-1}(x_{d-1},x_d),T_d(x_d))$. Obviously, T is a transport map from μ to ν, i.e. $T_{\#}\mu = \nu$. By construction, the Knothe transport T has a triangular Jacobian matrix with nonnegative entries on its diagonal. Note also that the computation of the Knothe transport only involves one-dimensional monotone rearrangements and that it is well defined as soon the measures one transports have no atoms. Before the age of optimal transport, which provided a map pushing μ onto ν which is the gradient of a convex functions, several proofs (for instance in functional or geometric inequalities) made use of this map instead of the Brenier one $T = \nabla\psi$. Not only, in numerical applications, for instance in image processing, the Knothe transport and in general monotone re-arrangement, often performed with respect to several sets of coordinate axis, so as to avoid a strong anisotropy in the output, are used to produce "reasonable" transport maps, instead of optimal ones. Finally, from an economical point of view, it is quite natural to consider a rearrangement like Knothe's one, since its idea is "first arrange the x_d criterion; then, in case of equality, the x_{d-1},..., and so on up to x_1" and it corresponds to a hierarchical structure of preferences, i.e. a lexicographical order.

Yann Brenier confirmed to us what we had been told, namely that he conjectured that the Knothe map should be the limit of the optimal transports T_ε, obtained with respect to the costs

$$c_\varepsilon(x,y) := \sum_{i=1}^{d} \lambda_i(\varepsilon)(x_i - y_i)^2$$

where $\lambda_i(\varepsilon) = \varepsilon^{d-i}$ or, more generally, $\lambda_k(\varepsilon)/\lambda_{k+1}(\varepsilon) \to 0$.

In [2], we proved this convergence under additional assumptions on ν and gave a counter-exemple if these assumptions were not satisfied. I start from recalling the assumptions we need

Assumption *(H-source)***:** the measure μ^d, as well as μ^d-almost all the measures $\mu^{d-1}_{x_d}$, and almost all the measures $\mu^{d-2}_{x_d,x_{d-1}}$ ... up to almost all the measures $\mu^1_{x_d,\dots,x_2}$, which are all measures on the real line, must have no atoms.

Notice that *(H-source)* is satisfied as soon as μ is absolutely continuous with respect to the Lebesgue measure and that it is a natural assumption if one wants to define monotone rearrangments.

Assumption *(H-target)***:** the measure ν^d, as well as ν^d-almost all the measures $\nu^{d-1}_{y_d}$, and almost all the measures $\nu^{d-2}_{y_d,y_{d-1}}$ ... up to almost all the measures $\nu^2_{y_d,\dots,y_3}$, which are all measures on the real line too, must have no atoms.

Notice that *(H-target)* is not natural as *(H-source)* is. Yet, we will show a counter-example to the convergence result when it is not satisfied. *(H-target)* as well is satisfied should ν be absolutely continuous (actually, this assumption is slightly weaker then *(H-source)*, since the last disintegration measures are not concerned).

The main theoretical result of [2] is the following, which comes from sort of a $\Gamma-$convergence development.

Theorem 1.1. *Let μ and ν be two probability measures on $\mathbb{R}^d$ satisfying (H-source) and (H-target), respectively, with finite second moments, and γ_ε be an optimal transport plan for the costs $c_\varepsilon(x,y) = \sum_{i=1}^d \lambda_i(\varepsilon)(x_i - y_i)^2$, for some weights $\lambda_k(\varepsilon) > 0$. Suppose $\lambda_k(\varepsilon)/\lambda_{k+1}(\varepsilon) \to 0$ for $k < d$, as $\varepsilon \to 0$. Let T be the Knothe map between μ and ν and $\gamma_K \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ the associated transport plan (i.e. $\gamma_K := (id \times T)\#\mu$). Then $\gamma_\varepsilon \rightharpoonup \gamma_K$ as $\varepsilon \to 0$.*

Moreover, should the plans γ_ε be induced by transport maps T_ε, then these maps would converge to T in $L^2(\mu)$ as $\varepsilon \to 0$.

A counterexample when the measures have atoms It is possible to show that the hypothesis *(H-target)* in the above Theorem is crucial. We can see a very simple example in $\mathbb{R}^2$ where μ is absolutely continuous with respect to the Lebesgue measure but ν does not satisfy *(H-target)*, and we show that the conclusion of Theorem 1.1 fails to hold. On the square $\Omega := [-1,1] \times [-1,1]$, take $\mu = \frac{1}{2} 1_{\{x_1 x_2 < 0\}} \cdot \mathcal{L}^2$ so that the measure μ is uniformly spread on the upper left and the lower right quadrants, and $\nu = \mathcal{H}^1_{|S}/2$, being S the segment $[-1,1] \times \{0\}$.

The Knothe map is easily computed as $(y_1, y_2) = T(x) := (2x_1 + sgn(x_2), 0)$. The solution of any transportation problem with $\lambda^\varepsilon = (\varepsilon, 1)$ is $(y_1, y_2) = T^0(x) := (x_1, 0)$ (no transport may do better than this one, which projects on the support of ν). Therefore, in this example the optimal transportation maps fail to tend to the Knothe-Rosenblatt map. The reason is the atom in the measure $\nu^2 = \delta_0$.

Anyway, it is quite difficult to characterize the conditions so as the convergence to Knothe to hold. For instance, the assumption *H-target* is obviously not satisfied if ν itself is purely atomic! Yet, this is precisely the case we considered in the algorithm we proposed later on in the paper. But it is not difficult to check that the same proof of the main theorem may be extended to the following case. Keep the same assumptions on μ but suppose that ν is concentrated on a set S with the property

$$y, z \in S, \quad y \neq z \Rightarrow y_d \neq z_d.$$

This is particularly useful when ν is purely atomic, concentrated on a finite set of points with different y_d components.

1.1.2 A continuation method

The other point we investigated later in the paper is the other direction: from Knothe to Brenier. We studied the dependence $\varepsilon \mapsto T_\varepsilon$ by means of the evolution with respect to ε of the dual variables. This enabled us to design a numerical strategy to approximate all the optimal transports T_ε taking as initial condition the Knothe transport T. The idea behind our numerical scheme is to start with the (cheap to compute) Knothe map as initial value and then approximate an ODE w.r.t. ε for the dual variables until the parameter ε reaches the desired value.

We restricted our analysis to the case where μ is the uniform density on a polyhedral convex domain $\Omega \subset \mathbb{R}^2$ with $|\Omega| = 1$ and ν is composed by N atoms of mass $1/N$ with distinct y_2 components. We expressed everything in terms of the dual variables, i.e. instead of solving the family of optimal transportation problems

$$\inf_{\gamma \in \Pi(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} c_\varepsilon(x,y) d\gamma(x,y) \tag{1.2}$$

we look at the dual formulation in terms of prices:

$$\sup_p \quad \Phi(p, \varepsilon) := \frac{1}{N} \sum_{i=1}^N p_i + \int_\Omega p^*_\varepsilon(x) dx, \tag{1.3}$$

where $p^*_\varepsilon(x) = \min_i\{c_\varepsilon(x, y_i) - p_i\}$ and we impose as a normalization $p_1 = 0$. For each ε, there is a unique maximizer $p(\varepsilon)$. For each (p, ε) we define $C(p, \varepsilon)_i = \{x \in \Omega \,:\, \inf_j c_\varepsilon(x, y_j) - p_j = c_\varepsilon(x, y_i) - p_i\}$. It is easy to check that $\Phi_\varepsilon := \Phi(., \varepsilon)$ is concave differentiable and that the gradient of Φ_ε is given by

$$\frac{\partial \Phi_\varepsilon}{\partial p_i}(p) = \frac{1}{N} - |C(p, \varepsilon)_i|.$$

By concavity of Φ_ε, the solution $p(\varepsilon)$ of (1.3) is characterized by the equation $\nabla\Phi_\varepsilon(p(\varepsilon)) = 0$. The optimal transportation between μ and ν for the cost c_ε is then the piecewise constant map taking the value y_i in the cell $C(p(\varepsilon), \varepsilon))_i$. Our aim is to characterize the evolution of $p(\varepsilon)$ as ε varies. Formally, differentiating the optimality condition $\nabla\Phi(p(\varepsilon), \varepsilon) = 0$, we obtain a differential equation for the evolution of $p(\varepsilon)$:

$$\frac{\partial}{\partial \varepsilon}\nabla_p\Phi(p(\varepsilon), \varepsilon) + D^2_{p,p}\Phi(p(\varepsilon), \varepsilon) \cdot \frac{dp}{d\varepsilon}(\varepsilon) = 0. \tag{1.4}$$

We want to show that the equation (1.4) is well-posed starting with the initial condition $p(0)$, corresponding to horizontal cells of area $1/N$.

The price vector $p(\varepsilon)$, along the evolution, will always be such that all the areas $|C(p, \varepsilon)_i|$ are equal. Yet, to prove that the differential equation is well posed we need to set it in an open set,

$$\mathcal{O} = \left\{(p, \varepsilon) \,:\, \frac{1}{2N} < |C(p, \varepsilon)_i| < \frac{2}{N} \text{ for every } i\right\}. \tag{1.5}$$

The initial datum of the equation will be such that $|C(p(0), 0)_i| = 1/N$ and we will look at the solution only up to the first moment where it exits $\mathcal{O}$. Yet, inside the set it will be well-posed and it will imply conservation of the areas. Hence it will never exit $\mathcal{O}$.

All the quantities we are interested in depend on the position of the vertices of the cells $C(p, \varepsilon)_i$, which are all polygons. Let us call $x(p, \varepsilon)^\pm_{i,j}$ the two extremal points of the common boundary between $C(p, \varepsilon)_i$ and $C(p, \varepsilon)_j$.

The key point is proving the following:

Lemma 1.2. *The positions of the vertices $x(p, \varepsilon)^\pm_{i,j}$ have a Lipschitz dependence on p and ε.*

After that, it is possible to prove

Lemma 1.3. *The function Φ admits pure second derivatives with respect to p and mixed second derivatives with respect to p and ε, and these second derivatives are Lipschitz continuous on $\mathcal{O}$. Moreover, there exists a constant $\lambda > 0$ such that we have $D^2_{p,p} \geq \lambda I_N$ on $\mathcal{O}$.*

From the previous results on the form and the regularity of the derivatives of $\nabla_p\Phi$, we deduce from the Cauchy-Lipschitz Theorem that the ODE (1.4) governing the evolution of the dual variables is well posed and actually characterizes the optimal prices:

Theorem 1.4. *Let $p(\varepsilon)$ be the solution of the dual problem (1.3) (recall the normalization $p_1(\varepsilon) = 0$), then it is the only solution of the ODE:*

$$\frac{dp}{d\varepsilon}(\varepsilon) = -D^2_{p,p}\Phi(p(\varepsilon), \varepsilon)^{-1}\left(\frac{\partial}{\partial \varepsilon}\nabla_p\Phi(p(\varepsilon), \varepsilon)\right) \tag{1.6}$$

with initial condition $p(0)$ such that all the horizontal strips $C(p(0),0)_i$ have area $1/N$. Moreover, usual approximation methods for this ODE may work to compute its solution.

See Figure 1.1 to have an idea of the evolution of the cells as ε moves.

1.2 Summability of transport densities

1.2.1 What is a transport density and what is its interest?

In the case $c(x,y) = |x-y|$ it is classical to associate to any optimal transport plan γ a positive measure σ on Ω, called transport density, which represents the amount of transport taking place in each region of Ω (Ω is, say, a bounded and convex subset of $\mathbb{R}^d$). This density σ is defined by

$$\langle \sigma, \phi \rangle := \int_{\Omega\times\Omega} \gamma(dx,dy) \int_0^1 \phi(\omega_{x,y}(t))|\dot{\omega}_{x,y}(t)|dt \quad \text{for all } \phi \in C^0(\Omega) \tag{1.7}$$

where $\omega_{x,y}$ is a curve parametrizing the straight line segment connecting x to y (the same could be generalized to other Riemannian distances than the euclidean one, and this segment should be replaced by a geodesic curve). Alternatively, if we look at the action of σ on sets, we have, for every Borel set A,

$$\sigma(A) := \int_{\Omega\times\Omega} \mathcal{H}^1(A \cap [[x,y]])\, \gamma(dx,dy),$$

where $[[x,y]]$ is the segment joining the two points x and y.

This positive measure σ is the total variation of a vector measure $\bar{\lambda}$ solving the problem

$$\min \left\{ \int_\Omega |\lambda|(dx), \; : \; \lambda \in \mathcal{M}(\Omega;\mathbb{R}^d),\, \nabla\cdot\lambda = \nu - \mu \right\}, \tag{1.8}$$

which is the so-called continuous transportation problem proposed by Beckmann in [41], where the divergence is to be interpreted in a weak sense, with Neumann boundary conditions.

More precisely, for every transport plan γ we can build a vector measure λ, defined through

$$\langle \lambda, \phi \rangle := \int_{\Omega\times\Omega} \gamma(dx,dy) \int_0^1 \phi(\omega_{x,y}(t)) \cdot \dot{\omega}_{x,y}(t)dt, \qquad \text{for all } \phi \in C^0(\Omega;\mathbb{R}^d)$$

and the $\bar{\lambda}$ associated to an optimal γ turns out to be optimal for (1.8). Thanks to our definitions, it is evident that we have $|\bar{\lambda}| \leq \sigma$, while the equality comes from the fact that transport rays cannot cross: if several segments involved by an optimal transport pass through the same point, then they all share the same direction.

A first natural question is whether the transport density σ is absolutely continuous. This would for instance allow to set the problem (1.8) in a L^1 setting instead of using the space $\mathcal{M}(\Omega)$ of finite vector measures on Ω. Notice that, to this aim, it would be sufficient to state that there exists an optimal transport plan γ such that the corresponding σ (or, equivalently, the corresponding λ) is absolutely continuous (it would not be necessary to prove it for every σ).

Actually, the precise relation between $\bar{\lambda}$ and σ is $\bar{\lambda} = \sigma \nabla u$, where u is a Kantorovich potential in the transportation from μ to ν. The condition $\sigma << \mathcal{L}^d$ would also allow to write the system

$$\begin{cases} \nabla \cdot (\sigma \nabla u) = \nu - \mu & \text{in } \Omega \\ |\nabla u| \leq 1 & \text{in } \Omega, \\ |\nabla u| = 1 & \text{a.e. on } \sigma > 0, \end{cases} \tag{1.9}$$

without passing through the theory of $\sigma-$tangential gradient (see for instance [74] or [53]).

There are several papers, mainly by De Pascale and Pratelli, Evans and Feldmann and McCann ([73, 74, 75, 80]), addressing absolute continuity and more general questions. In [73] the authors show estimates on the dimension of σ in terms of the dimension of μ and ν, and they get in particular $\sigma << \mathcal{L}^d$ whenever one of the two source measures μ or ν is absolutely continuous. In the same paper they also give several L^p estimates, which are then strengthened in [74] and in [75], where they finally get the important result

$$\mu, \nu \in L^p \Rightarrow \sigma \in L^p \quad \text{for all } p \in [1, +\infty]. \tag{1.10}$$

Among the other L^p results, [73] proves

$$\mu \in L^p \Rightarrow \sigma \in L^q \quad \text{for all } q < \min\left\{(2d)', 1 + \frac{p-1}{2}\right\}. \tag{1.11}$$

Estimates on σ may have various applications: for instance, lower bounds could be used to retrieve information on the behavior of u as a solution of (1.9) (and, in order to apply standard elliptic theory, L^∞ estimates as well are needed). Not only, lower bounds would also be useful to prove some density results on the transport set: the techniques developed by Champion and De Pascale in [67] would allow in such a case to derive the existence of an optimal transport map in Monge's problem. This is approximately what the same authors did in a subsequent paper, [68], where they solved the Monge problem for general norms.

To mention, on the other hand, applications of upper bounds and L^p estimates on the transport density, we refer to [5] and to Section 2.1, where these results are used to prove well-posedness for congestion problems and continuous Wardrop equilibria.

Coming back to the simple question of absolute continuity, Feldmann and McCann proved in [80] that there exists a unique L^1 transport density for L^1 sources. This is another absolute continuity result and it is coupled with a uniqueness result which is stated as "two L^1 functions which are transport densities, in a certain sense, between the same L^1 sources must coincide". A more complete uniqueness result may be found in the Lecture Notes by Ambrosio [28], where the links between σ and the other formulations of the Monge transport problem are well underlined. The proof of the uniqueness in [28] is based on a decomposition into transport rays and on a one-dimensional result, and the result reads as "two different optimal transport plans always induce the same measure σ, provided at least one of the two source measures is in L^1".

Yet, the proofs in [73, 74, 75] and [80] are quite complicated and long. This is natural since they are actually the first pioneering works on transport density; moreover they present much wider results (dimensional estimates, existence of the limit of the cost on a ball, uniqueness...). What I proposed in [3] was a series of very simple proofs for some results which were partially already known. The starting point is a proof of $\sigma << \mathcal{L}^d$ that arose during the preparation of a course on Optimal Transport at IHP in Paris. One lecture of the course was devoted to divergence-constrained problems and the goal was to show well-posedness in L^1.

The strategy of the proof passes through the absolute continuity of the interpolation measures μ_t. To prove that these measures are absolutely continuous we pass through the discrete case and then get it at the limit. Actually, the absolute continuity of the interpolations is well known, especially in the case of strictly convex cost function (rather than $|x-y|$). In this last case it is explicitly stated in Theorem 8.7 of Part I of the new book on Optimal Transport by Cedric Villani, [105], where a general Lipschitz result on intermediate transport maps (Theorem 8.5) is used. As far as the distance case $|x-y|$ is concerned, there exists a proof of the same absolute continuity estimates which is presented in [52]. Yet, no application to transport densities is presented, even if this is the most natural framework to apply those estimates passing to the limit from the discrete case, since we know that transport density is independent of the particular transport plan which is selected by the approximation.

Later on, it appeared quite easily that the same technique could be used for L^p estimates, to get again the old results as well as some improved ones.

1.2.2 Estimates

To investigate the properties of the transport density, a key tool will be looking at the interpolating measures from μ to ν. Take an optimal transport plan γ from μ to ν (remember that the transport density σ does not depend on the choice of γ) and define μ_t the standard interpolation between the two measures: $\mu_t = (\pi_t)_\#\gamma$, where $\pi_t(x, y) = (1-t)x + ty$ (this is the same, in this framework, as $\omega_{x,y}(t)$, when the segments are parametrized at constant speed).

The transport density σ may be easily written as

$$\sigma = \int_0^1 (\pi_t)_\#(|x-y| \cdot \gamma)dt.$$

Since we supposed that Ω is bounded it is evident that we have

$$\sigma \leq C \int_0^1 \mu_t \, dt. \tag{1.12}$$

To prove that σ is absolutely continuous, or to give L^p estimates on σ, it will be sufficient to look at almost every measure μ_t an get either absolute continuity or L^p estimates.

Theorem 1.5. *There exists an optimal transport plan γ from μ to ν (which is obtained as a limit from a discrete approximation of ν), such that*

- *if $\mu << \mathcal{L}^d$, then $\mu_t << \mathcal{L}^d$ for all $t < 1$,*
- *if $\mu \in L^p$, then $||\mu_t||_p \leq (1-t)^{-d/p'}||\mu||_p$, where $p' = \frac{p}{p-1}$ is the conjugate exponent of p.*

The same property stays true for every optimal transport plan γ which is the limit (in $\mathcal{P}(\Omega \times \Omega)$) of optimal plans γ_n from μ to some atomic measures ν_n approximating ν.

The above result, as I mentioned in the statement, is obtained by approximation. We first suppose that ν is finitely atomic and we get precise expressions of the densities of μ_t since the optimal transport is in this case a collection of homotheties; then, we approximate ν by means of atomic measures and the result stays true by semicontinuity for the transport plan γ that we obtain as a limit of the optimal ones that we had in the discrete case.

Notice that not every optimal γ may be obtained as a limit of optimal plans transporting μ onto discrete measures, as one can see from the following example: take μ and ν the uniform measures on $[-1, 0]$ and $[0, 1]$, respectively. The transport plan γ corresponding to the map $T(x) = -x$ is optimal for the cost $|x - y|$ since this cost coincides with $y - x$ on $\mathrm{supp}(\mu) \times \mathrm{supp}(\nu)$ and for such a cost every transport plan is optimal. Yet, the measure $\mu_{1/2}$ associated to this γ is δ_0 and is not absolutely continuous, while both μ and ν are. This prevents γ from being approximable by discrete optimal plans since otherwise the previous theorem should apply.

The consequences on the transport density are the following, where the conditions on p come from the integrability of $[0, 1[\ni t \mapsto (1 - t)^{-d/p'}$.

Theorem 1.6. *Suppose $\mu = f \cdot \mathcal{L}^d$. Then the unique transport density σ associated to the transport of μ onto ν is absolutely continuous. Moreover, if $f \in L^p(\Omega)$ and $p < d' := d/(d-1)$, σ also belongs to $L^p(\Omega)$, and if $p \geq d'$ it belongs to any space $L^q(\Omega)$ for $q < d'$.*

After that, it is possible to look at what can be obtained under assumptions on both measures μ and ν. It would be natural to try to exploit the following strategy: use μ for giving estimates on μ_t for $t \leq 1/2$ and ν for $t \geq 1/2$. In this way the factors $(1 - t)^{-d/p'}$ or, conversely, $t^{-d/p'}$ would be never taken close to their degeneracy endpoint.

This strategy works quite well, but it has an extra difficulty: actually, we didn't know a priori that μ_t shared the same behavior of piecewise homotheties of μ, but we got it as a limit from discrete approximations. And, when we pass to the limit, we do not know which optimal transport γ will be selected as a limit of the optimal plans γ_n. When we want to glue together estimates on μ_t for $t \leq 1/2$ which have been obtained by approximating ν, and estimates on μ_t for $t \geq 1/2$ which come from the approximation of μ, we need the two approximations to converge to the same transport plan, otherwise we could not put together the two estimates and deduce anything on σ.

Hence, the main technical point has been proving that one particular optimal transport plan, namely the one which is monotone on transport rays, is approximable in both directions. Actually, it is possible to prove it (see Section 1.4) by changing a little bit the cost functions, so that this transport plan will be the limit of plans γ_n which are optimal for a modified cost $|x - y| + \frac{1}{n}|x - y|^2$, but this is not a problem for letting the same estimates of Theorem 1.5 work. And, once this is obtained, it is not difficult to prove the following:

Theorem 1.7. *Suppose that μ and ν are probability measures on Ω, both belonging to $L^p(\Omega)$, and σ the unique transport density associated to the transport of μ onto ν. Then σ belongs to $L^p(\Omega)$ as well.*

Theorem 1.8. *Suppose $\mu \in L^p(\Omega)$ and $\nu \in L^q(\Omega)$. For notational simplicity take $p > q$. Then, if $p < d/(d-1)$, the transport density σ belongs to L^p and, if $p \geq d/(d-1)$, it belongs to $L^r(\Omega)$ for all the exponents r satisfying*

$$r < r(p,q,d) := \frac{dq(p-1)}{d(p-1)-(p-q)}.$$

1.3 Non-strictly convex costs and convex constraints

1.3.1 Framework and general strategy

In this section I want to present the results and the ideas that arose during a joint work with G. Carlier and L. De Pascale to study the transport problem

$$\min\left\{\int_{\Omega\times\Omega} c(x-y)\,\gamma(dx,dy)\ :\ \gamma \in \Pi(\mu,\nu)\right\},$$

where c is convex but not strictly convex.

A very natural strategy (which is not peculiar to our work, other teams in Trieste and Brest are working on closely related approaches) could be the following:

- take an optimal plan γ and look at its optimality conditions by means of duality arguments: from the fact that

  $$u(x)+v(y) = c(x-y) \text{ on } \operatorname{supp}\gamma \quad \text{and} \quad u(x)+v(y) \leq c(x-y)$$

  one gets as usual, if x is a differentiability point for u,

  $$\nabla u(x) \in \partial c(x-y);$$

  which is equivalent to

  $$x - y \in \partial c^*(\nabla u(x)). \tag{1.13}$$

  Let us define

  $$\mathcal{F}_c := \{\partial c^*(p)\ :\ p \in \mathbb{R}^d\},$$

  which is the set of all values of the subdifferential multi-map of c^*. This values are those convex sets where the function c is affine, and they will be called *faces* of c.

  Thanks to (1.13), for every fixed x, all the points y such that (x, y) belongs to the support of an optimal transport plan are such that the difference $x-y$ belong to a same face of c. Classically, when these faces are singleton (i.e. when c^* is differentiable, which is the same as c being strictly convex), this is the way to obtain a transport map, since one only y is admitted for every x.

  Equation (1.13) also enables to classify the points x as follows. For every $K \in \mathcal{F}_c$ we define the set

  $$X_K := \{x \in \Omega\ :\ \exists \nabla u(x)\ ,\ \partial c^*(\nabla u(x)) = K\}.$$

Hence γ may be decomposed into several subplans γ_K according to the criterion $x \in X_K$, which is the same as $x - y \in K$. If the set $\mathcal{F}_c$ is finite or countable, we can define

$$\gamma_K := \gamma_{|X_K \times \Omega},$$

which is the simpler case. Actually, in this case, the marginal $\mu_K := (\pi_x)_\# \gamma_K$ is a submeasure of μ, and in particular it is absolutely continuous if μ is, which is often useful for proving existence of transport maps.

If this is not the case, if $\mathcal{F}_c$ is uncountable, one extra-possibility for using countable decompositions comes if one considers the set $\mathcal{F}_c^{multi} := \{K \in \mathcal{F}_c \ : \ K \text{ is not a singleton } \}$. If $\mathcal{F}_c^{multi}$ is countable, then one can separate those x such that $\partial c^*(\nabla(\phi(x))$ is a singleton (where a transport already exists) and look at a decomposition for $K \in \mathcal{F}_c^{multi}$ only. Otherwise, in general, one should pass trough a disintegration procedure.

In some cases, it could be useful to bundle together different possible Ks so that the decomposition is countable, even if coarser. This is what I will present in Section 1.3.2

- This reduces the transport problem into a superposition of transport problems

$$\min \left\{ \int_{\Omega \times \Omega} c(x-y)\, \gamma(dx, dy) \ : \ \gamma \in \Pi(\mu_K, \nu_K) \ \text{ and is concentrated on } x - y \in K \right\},$$

and the advantage is that the cost c restricted to K is easier to look at. For instance, if K is a face of c, then c is affine on K and in this case the transport cost does not depend any more on the transport plan.

- In this case, one wants to find a transport map from μ_K to ν_K satisfying the constraint $x - T(x) \in K$, knowing a priori that a transport plan satisfying the same constraint exists. Notice that this problem may be assimilated to an L^∞ transport problem, at least in the case where K is a convex compact set with non-empty interior. In this case, if one denotes by $|| \cdot ||_K$ the norm such that $K = \{x \ : \ ||x||_K \leq 1\}$, one has

$$\min \big\{ \max\{||x-y||_K,\ (x,y) \in \operatorname{supp}(\gamma)\}, \gamma \in \Pi(\mu,\nu) \big\} \leq 1 \tag{1.14}$$

and the question is whether the same result would be true if one restricted the admissible set to transport maps only (passing from Kantorovitch to Monge, say). The answer would be yes if a solution of (1.14) was induced by a transport map T (which is true if $\mu_K << \mathcal{L}^d$ and $K = B(0,1)$, see [69], but is not known in general). Moreover, the answer is obviously yes in $\mathbb{R}$, where the monotone transport solves all the L^p problems, and hence the L^∞ as well.

- The answer would be yes also in the case (and it is actually almost equivalent) where we are able to select, for instance by a secondary minimization, one particular transport plan satisfying $\operatorname{supp}(\gamma) \subset \{x - y \in K\}$ which is induced by a map. This leads to the very natural question of solving

$$\min \left\{ \int \left(\frac{1}{2}|x-y|^2 + I_K(x-y) \right) \gamma(dx, dy) \ : \ \gamma \in \Pi(\mu,\nu) \right\},$$

or, more generally, to transport problems where the cost function involves convex constraints on $x-y$.

- As usual, optimality conditions on optimal transport plans and optimal potentials for this problems read as

$$u(x)+v(y)=\frac{1}{2}|x-y|^2 \text{ on } \operatorname{supp}\gamma \quad \text{and} \quad u(x)+v(y)\leq\frac{1}{2}|x-y|^2 \text{ for all } (x,y)\,:\,x-y\in K$$

  and they lead to

$$\nabla u(x)=x-y+n_K(x-y),$$

  where n_K is a normal vector to K. This is equivalent to saying that $x-y$ is the projection on K of $\nabla u(x)$ (which is a nice counterpart of the non-constrained result) and is sufficient to get the existence of a transport map, provided two facts hold:

  - μ is absolutely continuous, so that we can assure that u, under mild regularity assumptions o, is (possibly approximately) differentiable $\mu-$a.e.;
  - an optimal potential u does actually exist, in a class of functions (Lipschitz, BV...) which are differentiable almost everywhere, at least in a weak sense.

- In order to apply the study of convex-constrained problems to the original problem with $c(x-y)$ the first issue (i.e. absolute continuity) does not give any problem if the decomposition is at most countable, while it is not trivial in case of disintegration, and it presents the same kind of difficulties as in Sudakov's proof for Monge's problem. This is the part of the problem where a group in Trieste (Bianchini, Caravenna, Daneri) is actively working, with some partial results (see for instance [65]) in general and more advanced results for the case of norms.

- As far as the second issue is concerned, this is much more delicate, since in general we have no existence result for potentials with non-finite costs. In particular, a counterexample has been provided by Caravenna when $c(x,y)=|x-y|+I_K(x-y)$ where $K=\mathbb{R}^+\times\mathbb{R}^+\subset\mathbb{R}^2$ and it is easily adapted to the case of quadratic costs with convex constraints. On the other hand, it is easy to think that the correct space to set the dual problem in Kantorovitch theory for this kind of costs would be $BV(\Omega)$ since the the constraints on $x-y$ enable to control the increments of the potentials u and v on some directions, thus getting some sort of monotonicity. Yet, this is not sufficient to find a bound in BV if an L^∞ estimate is not available as well and the counterexample that we mentioned - which gives infinite values for both u and v, exactly proves that this kind of estimates are hard to prove.

1.3.2 The case of the cost $h(||x-y||)$ in $\mathbb{R}^2$, h strictly convex and increasing

I want to show here how this method, which is not completely efficient in general may be used to treat the case of a cost given by a strictly convex and increasing function of a generic norm in $\mathbb{R}^2$. Up to now, the existence of an optimal transport map for generic norms was already known in $\mathbb{R}^2$ since

[33] and has been recently generalized to any dimension by [68] and [64] with different techniques. This proof gives the result for the case $||x-y||^p$, $p>1$, and for more general convex cost functions.

The main intermediate result is the following:

Theorem 1.9. *Suppose that μ,ν are probability measures on a compact domain $\Omega\subset\mathbb{R}^2$, with $\mu<<\mathcal{L}^2$, K is a closed and convex subset of $\mathbb{R}^2$ and $h:\mathbb{R}\to\mathbb{R}$ a strictly convex function. Set $c(z)=h(z_1)+I_K(z)$ and $c(x,y)=c(x-y)$, where z_1 denotes the first component of z. Then the transport problem*

$$\min\left\{\int c\,d\gamma,\ \gamma\in\Pi(\mu,\nu)\right\}$$

admits at least a solution which is induced by a transport map T, provided that a transport plan with finite cost exists and that the dual problem admits Lipschitz solutions, i.e. that there exist γ, u and v such that

$$u(x)+v(y)\le c(x-y),\quad u(x)+v(y)=c(x-y)\ \text{ on }\ \mathrm{supp}(\gamma),\ \gamma\in\Pi(\mu,\nu),\ u,v\in Lip(\Omega).$$

The proof of this result mainly uses the fact that knowing ∇u at a point x is enough to determine a restricted sets of possible y's such that $(x,y)\in\mathrm{supp}(\gamma)$, and that, thanks to the strictly convex behaviour of h, the component y_1 is the same or all these points and depends only on $\nabla u(x)$. Then one can prove that, on every vertical line, all the points x that are associated to more than one single point y may be decomposed as a countable union of sets that are sent to a common vertical line (i.e. the gradient of u is the same, and hence y_1 is the same). On all these sets the problem becomes essentially an L^∞ problem in dimension one and is solved by the monotone transport. Notice that the choice of the monotone one prevents from having measurability problems in gluing the transport on different lines.

From the previous result we may go on and get the following:

Theorem 1.10. *Suppose that μ,ν are probability measures in $\Omega\subset\mathbb{R}^2$, with $\mu<<\mathcal{L}^2$, $||\cdot||$ is an arbitrary norm of $\mathbb{R}^2$ and $h:\mathbb{R}^+\to\mathbb{R}^+$ a strictly convex and increasing function. Then the transport problem*

$$\min\left\{\int h(||x-y||)\,d\gamma,\ \gamma\in\Pi(\mu,\nu)\right\}$$

admits at least a solution which is induced by a transport map T.

For proving this theorem the strategy is almost the one which was described in the previous subsection, with the trick of assimilating some faces of the convex function $h(||x-y||)$. Actually, the boundary of the unit ball of the norm $||\cdot||$ has a countable number of flat parts F_i, to which we associate the cone $K_i=\mathbb{R}^+\times F_i$, and all the other points are strict convexity points. Then one may classify the points x according to the fact that $\nabla u(x)$ belongs either to the strict convexity part $\mathbb{R}^2\setminus\bigcup_i K_i$, and in this case it admits one image point y only, or to K_i. On each set K_i the cost function behaves as a one-variable function, say $h(z_1)$, up to a change of variables so that the direction e_1 is the direction orthogonal to F_i. Then one applies the previous theorem to the transport from μ_{K_i} (which is still absolutely continuous since it is a submeasure of μ) and ν_{K_i}. The assumption that a finite cost transport plan exists is satisfied taking the restriction γ_{K_i} of an optimal γ on $\{x-y \in K_i\}$. The existence of optimal potentials comes from the existence for the original problem, which admitted a pair (u, v) such that

$$u(x) + v(y) \leq h(||x-y||) \leq h(||x-y||) + I_{K_i}(x-y),$$
$$u(x) + v(y) = h(||x-y||) = h(||x-y||) + I_{K_i}(x-y) \text{ on } \operatorname{supp}(\gamma_{K_i}).$$

1.4 Partial results, interesting lemmas and perspectives

As I stressed at the beginning of this chapter, this part of my researches is quite recent and hence it is likely to provide interesting developments in the future.

The paper on the link between Knothe transport and optimal transports for anisotropic quadratic costs gives an answer to a conjecture by Brenier but also provides a numerical method for finding optimal maps (say, optimal partitions) when μ is diffuse and ν is finitely atomic. Such a procedure is probably not very rapid but numerics on optimal transport is far from being solved in a satisfactory way and every method is likely to bring interesting improvements.

Not only, it happened that during the revision of the paper, after a useful remark by T. Mikami who found an error in a preliminary version, we needed to manage the convergence of a term of the following type: μ and ν were given probabilities over $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^{n-k}$ (with variables (x^+, x^-) and (y^+, y^-)) and $\mu = \mu^+ \otimes \mu^{x^+}$ and $\nu = \nu^+ \otimes \nu^{y^+}$ were disintegrations of μ and ν, respectively. The dependence of μ^{x^+} on x^+ and of ν^{y^+} on y^+ was only measurable. A sequence η_n in $\Pi(\mu^+, \nu^+)$ was given and we looked at

$$\lim_{n\to\infty} \int W_2^2(\mu^{x^+}, \nu^{x^+}) d\eta_n(x^+, y^+).$$

The following lemma turned out to be useful and seems to be interesting by itself.

Lemma 1.11. *Let η_n and η be measures on $\Omega \times \Omega$, all with marginals μ and ν, respectively and $a : \Omega \to X$ and $b : \Omega \to Y$ be two measurable maps valued in two Polish spaces X and Y. Let $c : X \times Y \to [0, +\infty]$ be a continuous function with $c(a, b) \leq f(a) + g(b)$ with f, g continuous and $\int f(a(x))d\mu$, $\int g(b(y))d\nu < +\infty$. Then we have*

$$\eta_n \rightharpoonup \eta \Rightarrow \int_{\Omega\times\Omega} c(a(x), b(y)) d\eta_n \to \int_{\Omega\times\Omega} c(a(x), b(y)) d\eta.$$

Another useful lemma coming from this chapter of my researches is the following one, which has already been used by T. Champion and L. De Pascale in [68] in order to get information on the interpolation measures μ_t and, consequently, on the transport set:

Lemma 1.12. *For fixed measures $\mu, \nu \in \mathcal{P}(\Omega)$, consider the family of minimization problems (P_ε):*

$$(P_\varepsilon) = \min \left\{ W_1((\pi_1)_\#\gamma, \nu) + \varepsilon C_1(\gamma) + \varepsilon^2 C_2(\gamma) + \varepsilon^{3d+3} \#((\pi_1)_\#\gamma), \, : \, \gamma \in \mathcal{P}(\Omega \times \Omega), \, (\pi_0)_\#\gamma = \mu \right\},$$

where W_1 is the usual Wasserstein distance, $C_p(\gamma) = \int |x-y|^p \gamma(dx, dy)$ for $p = 1, 2$ and the symbol $\#$ denotes the cardinality of the support of a measure.

Call γ_ε any minimizer of (P_ε) and $\nu_\varepsilon := (\pi_1)_\#\gamma_\varepsilon$ its second marginal. It is straightforward that ν_ε is an atomic measure and that γ_ε is the optimal transport from μ to ν_ε for the cost $C_1 + \varepsilon C_2$.

Then, as $\varepsilon \to 0$ we have $\nu_\varepsilon \rightharpoonup \nu$ and $\gamma_\varepsilon \rightharpoonup \bar\gamma$, where $\bar\gamma$ is the optimal Monge transport plan from μ to ν which is monotone on transport rays, and which is characterized by

$$\bar\gamma = \operatorname{argmin}\left\{C_2(\gamma)\ :\ \gamma\ \textit{ is a } C_1\textit{-optimal transport plan from } \mu \textit{ to } \nu\right\}. \tag{1.15}$$

Concerning the last section of this part of my researches, I think that it is the one which could give the widest developments, since up to now we only have applied the general strategy to the only case where everything worked. Yet, I think that it is important to stress the following result.

Theorem 1.13. *Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ be two probability measures, μ being absolutely continuous, K a closed and convex subset of $\mathbb{R}^d$ and $h : \mathbb{R}^d \to \mathbb{R}$ a strictly convex function. Let $c(z) = h(z) + I_K(z)$: then the transport problem*

$$\min\left\{\int c(x-y)\,d\gamma,\ \gamma \in \Pi(\mu,\nu)\right\}$$

admits a unique solution, which is induced by a transport map T, provided that a transport plan with finite cost exists and that the dual problem admits solutions (u, v) with u which is a.e. approximately differentiable (for instance $u \in BV$).

In order to give just a brief sketch of the proof, the procedure is standard, writing that u, v and $c(x - y)$ satisfy an inequality which becomes an equality on $\operatorname{supp}(\gamma)$. This implies that, for all $(\bar x, \bar y) \in \operatorname{supp}(\gamma)$, $\bar x$ minimizes $x \mapsto c(x - \bar y) - u(x)$. If u is at least approximately differentiable at $\bar x$, this implies

$$\nabla u(\bar x) \in \partial c(\bar x - \bar y) = \partial h(\bar x - \bar y) + N_K(\bar x - \bar y),$$

where N_K is the normal cone to K. Yet, when a vector l and a point $\bar z \in K$ satisfy

$$l \in \partial h(\bar z) + N_K(\bar z),$$

this gives that $\bar z$ minimizes $K \ni z \mapsto h(z) - l \cdot z$. Since h is strictly convex this gives the uniqueness of $\bar z$, which will depend on l. As usual, this implies that γ is induced by a transport and that it is unique. Notice that, in the case $h(z) = \frac{1}{2}|z|^2$, the point $\bar z$ will be the projection of l on K.

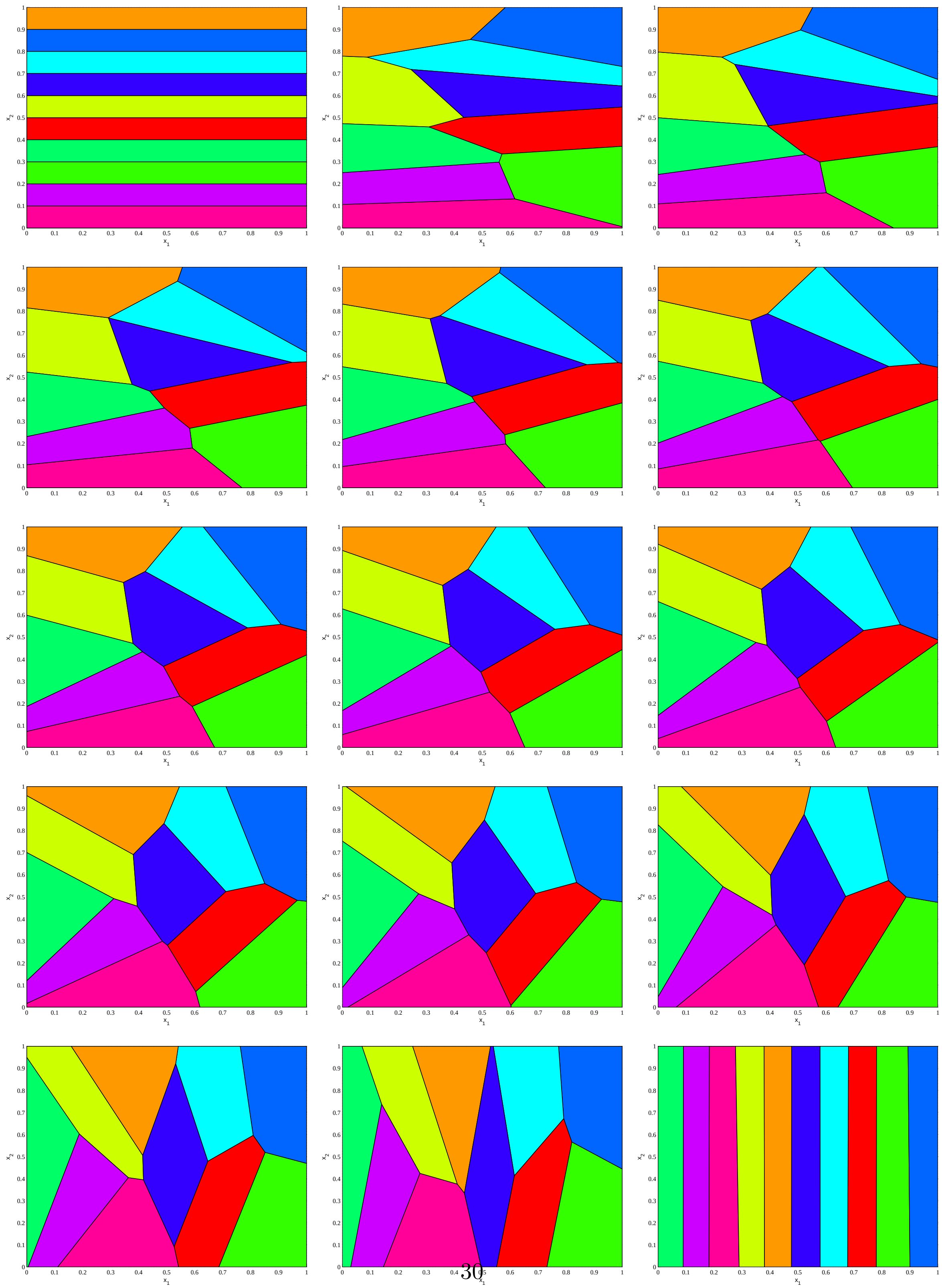


Figure 1.1: Ten sample points: evolution of the tesselation for $\varepsilon = 0$ to $\varepsilon = +\infty$ (from top left to bottom right).

Chapter 2

A continuous theory for congested traffic

Résumé Ce chapitre présente des travaux qui arrivent à "fonder" une théorie continue pour les problèmes de congestion de trafic, qui dans le cas d'un réseaux ont été introduits par Wardrop, et puis étudiés en connexion avec l'optimisation convexe. Non seulement on en donne une formulation continue, basée sur les mesures sur l'espace des chemins et sur la généralisation du concept de densité de trafic à une configuration de chemins quelconque, mais on fait aussi le lien avec les problèmes à divergence fixée proposés par Beckmann. Pour que ce lien ne soit pas seulement formel, il faut pouvoir définir un flot à partir d'un champ de vecteur optimal v dans le problème à divergence fixée. Dans les cas qui sont intéressants au niveau de la modélisation, ce champ de vecteur ne sera pas très régulier, parce qu'il résout une EDP elliptiques très dégénérée. Pourtant, on arrive à démontrer des résultats de régularité Sobolev qui permettent de lui appliquer la théorie de DiPerna-Lions, ce qui nous suffit pour les applications qu'on a en vue. En suite, on a pu pousser les résultats au délà de la régularité H^1 en obtenant, en dimension 2, la continuité.

Cetthe théorie a donc donné l'mpulsion à des nouvelles question de régularité elliptique, mais ce n'est pas la seule chose. Elle a été l'occasion de dévélopper des méthodes numériques intéressantes. Le problème de congestion faisant intervenir des distance géodésiques par rapport à une métrique Riemanienne inconnue, il était important de savoir dériver ces distances par rapport à la métrique elle-même, et de le faire dans sa version discrétisée pour pouvoir faire des calculs numériques. C'est ce qui nous a amenés à modifier la méthode de Fast Marching pour qu'elle sache calculer non seulement les distances, mais aussi les gradients de ces distances. Ce calcul peut être appliqué à plusieurs problèmes variationnels faisant intervenir des métriques inconnues

In this chapter I will present a branch of my researches, which stemmed from transport theory towards the study of its congested variants around the end of my PhD thesis. The idea of this kind of variants is looking at transport problem where movement costs more if several agents (or particles) pass through a common place. During my PhD thesis, this idea appeared already in [15] and [20] (the latter being inspired by the work of Beckmann, [41, 43]). Yet, it is only later, from [5], that we tried, with different collaborators, to give a formal unified theory which, at the same time, generalizes the *road traffic theory* that Wardrop ([106]) presented on discrete networks, and makes the link with the *continuous transportation model* by Beckmann. The result is quite satisfactory: this subject presented unexpected connections with different PDE topics (DiPerna-Lions theory, degenerate elliptic PDEs...), but it also allowed for efficient numerical computations. Such computations were on the other hand the motivation for an interesting extension of the Fast Marching Method for geodesic computations.

To present the whole framework, I will start from a brief outline of the discrete problem.

In the discrete framework, one considers

- A finite graph with edges $e \in E$, representing the network, and a set of sources S and destinations D;
- for every pair $(s, d) \in S \times D$, the set $C(s, d) = \{\sigma \text{ from } s \text{ to } d\}$ of possible paths from s to d: this set it finite, on a finite graph, if only simple paths are taken into account;
- a demand input $\gamma : S \times D \to \mathbb{R}$, where $\gamma(s, d)$ denotes the quantity of commuters from $s \in S$ to $d \in D$; according to the cases γ may be prescribed or a set Γ of possible γ's may be given: this set is usually either a singleton $\Gamma = \{\bar{\gamma}\}$ or $\Gamma = \Pi(\mu, \nu)$;
- an unknown repartition strategy $q = (q_\sigma)_\sigma$ such that $\sum_{\sigma \in C(s,d)} q_\sigma = \gamma(s, d)$, each term q_σ representing the quantity of agents moving on the path σ;
- a consequent traffic intensity (depending on q) $i_q = (i_q(e))_e$, given by $i_q(e) = \sum_{e \in \sigma} q_\sigma$, i.e. the total quantity of agents passing through e;
- an increasing function $g : \mathbb{R}^+ \to \mathbb{R}^+$ such that $g(i_q(e))$ represents the congestioned cost (per unit length) on the edge e (one could also think at costs depending explicitly on e as well, as if different edges had different congestion effects, for instance if one represents a highway and one a country path);
- then, the cost for each path σ is given by $c(\sigma) = \sum_{e \in \sigma} g(i_q(e))\mathcal{H}^1(e)$.

The global strategy q represents the overall distribution on choices of commuters' paths. Imposing a Nash equilibrium condition (no single commuter wants to change his choice, provided all the others keep the same strategy) gives the following condition:

$$\sigma \in C(s, d),\ q_\sigma > 0 \Rightarrow c(\sigma) = \min\{c(\tilde{\sigma}) \,:\, \tilde{\sigma} \in C(s, d)\}.$$

This condition is well-known among geographical economists and game-theorists as *Wardrop equilibrium*. It is just the translation into this framework of the Nash equilibrium. Notice that the framework we described is discretized in the network but not with respect to agents: there is a continuum of agents that can be splitted at will among possible paths. This is what is called a *non-atomic game*, since the weight of each agent is zero and when she compares her own current situation to what would happen if she changed her path, she does not take into account that changing her path would affect q.

Proving the existence of at least an equilibrium is a priori non trivial, but becomes much easier in connection with the following variational principle.

Imagine that, instead of looking for an equilibrium configuration, one wants to minimizing a global quantity

$$\min\Big\{\sum_e H(i_q(e))\mathcal{H}^1(e)\ :\ q\ \text{ admissible}\Big\}$$

($H:\mathbb{R}^+\to\mathbb{R}^+$ being an increasing function: for instance with $H(t)=tg(t)$ we get the total cost for all commuters) among all possible strategies q.

Minimizing $J(q)$ among possible strategies q has obviously a solution and one can look for optimality conditions. Suppose that H and Γ are convex, so that the necessary conditions will also be sufficient: it is easy to see that q minimizes if and only if, for every other admissible $\tilde q$, one has

$$\sum_e H'(i_q(e))(i_{\tilde q}(e)-i_q(e))\mathcal{H}^1(e)\geq 0.$$

Set $\xi(e):=H'(i_q(e)$ and rewrite the right hand side as

$$\sum_e \xi(e)(i_{\tilde q}(e)-i_q(e))\mathcal{H}^1(e)=\sum_e\sum_{\omega\ni e}\xi(e)(\tilde q(\omega)-q(\omega))\mathcal{H}^1(e)=\sum_\omega\left(\sum_{e\in\omega}\xi(e)\mathcal{H}^1(e)\right)(\tilde q(\omega)-q(\omega)).$$

This means that, if one sets $L_\xi(\omega):=\sum_{e\in\omega}\xi(e)\mathcal{H}^1(e)$, the optimal q must minimize as well $\sum_\omega L_\xi(\omega)q(\omega)$, since we got $\sum_\omega L_\xi(\omega)\tilde q(\omega)\geq\sum_\omega L_\xi(\omega)q(\omega)$.

This implies two facts. First, all the curves ω which are charged by q must be optimal for L_ξ among all the curves sharing the same starting and arrival points (since the conditions of admissibility on q only look at those points). If one sets $d_\xi(s,d)=\min_{\omega\in C(s,d)}L_\xi(\omega)$, this means that $q(\omega)>0$ and $\omega\in C(s,d)$ imply $L_\xi(\omega)=d_\xi(s,d)$.

Second, another condition occurs when the demand γ is not fixed. To optimize $\sum_\omega L_\xi(\omega)q(\omega)$ one also needs to choose $\gamma\in\Gamma$ so as to minimize

$$\sum_{s,d} d_\xi(s,d)\gamma(s,d),\quad \gamma\in\Gamma.$$

This second condition is empty if Γ only contains one γ but it is of particular interest when $\Gamma=\Pi(\mu,\nu)$, since it says that γ must solve a Kantorovitch problem for the cost d_ξ.

The first condition, on the other hand, always gives some information on q and exactly says: if q is optimal, then it is a Wardrop equilibrium for $g=H'$.

2.1 Continuous formulation

2.1.1 Optimality and equilibrium of traffic on curves

I try to present now the continuous extension of this model, which is mainly contained in [5]. In a domain $\Omega \subset \mathbb{R}^n$ the demand is represented by probabilities $\gamma \in \mathcal{P}(\Omega \times \Omega)$. We are given a set $\Gamma \subset \mathcal{P}(\Omega \times \Omega)$, the set of admissible demand couplings: usually $\Gamma = \{\bar{\gamma}\}$ or $\Gamma = \Pi(\mu, \nu)$. Let us also set

$$C = \{\text{Lipschitz paths } \omega : [0,1] \to \Omega\} \qquad C(s,d) = \{\omega \in C : \omega(0) = s_{,}\, \omega(1) = d\}.$$

We look for a probability $Q \in \mathcal{P}(C)$ such that $(\pi_{0,1})_\# Q \in \Gamma$, where $\pi_t : C \to \Omega$ denotes the evaluation map at time t: $\pi_t(\omega) := \omega(t)$ and $\pi_{0,1} := (\pi_0, \pi_1)$.

We want to define a traffic intensity $i_Q \in \mathcal{M}^+(\Omega)$ such that, for every set A the quantity $i_Q(A)$ represents "how much " the movement takes place in A. To do so, for $\phi \in C^0(\Omega)$ and $\omega \in C$ set

$$L_\phi(\omega) = \int_0^1 \phi(\omega(t))|\omega'(t)|dt.$$

Then we define i_Q by

$$< i_Q, \phi >= \int_C L_\phi(\omega) Q(d\omega).$$

Notice that this is just a generalization of the definition of transport density, which is obtained if Q is the image of an optimal Monge transport plan γ through the map that associates the segment $\omega_{x,y}$ to every pair (x, y).

In the continuous case it is more delicate to look at the equilibrium condition (because of costs and metrics that are defined a.e.. . .), hence it is better to start from the optimization point of view: we minimize the convex functional

$$F(i_Q) = \begin{cases} \int H(i_Q(x))dx & \text{if } i_Q << \mathcal{L}^n, \\ +\infty & \text{otherwise} \end{cases}$$

among all admissible strategies Q. Here H is a convex, increasing and superlinear function, which is supposed to satisfy an estimate like $at^q \leq H(t) \leq b(1 + t^q)$ (which, for convex functions, also implies $H'(t) \leq C(t^{q-1} + 1)$).

The precise minimization problem reads

$$(W) \qquad \min \{F(i_Q) \,|\, Q \in \mathcal{P}(C), (\pi_{0,1})_\# Q \in \Gamma\}.$$

The minimum may obviously be restricted to the class $\mathcal{Q}^q(\Gamma)$ of measures $Q \in \mathcal{P}(C)$, such that $(\pi_{0,1})_\# Q \in \Gamma$ and $i_Q \in L^q$.

It is possible to prove, similarly to the discrete case (where the integrals are performed w.r.t. to the length measure on the network), the following theorem:

Theorem 2.1. *Suppose that $\mathcal{Q}^q(\Gamma) \neq \emptyset$: then, (W) admits a solution $\overline{Q}$ which is characterized by*

$$\int_\Omega \overline{\xi}\, di_{\overline{Q}} = \inf \left\{ \int_\Omega \overline{\xi}\, di_Q \ : \ Q \in \mathcal{Q}^q(\Gamma) \right\} \ \textit{with } \overline{\xi} := H'(i_{\overline{Q}}) \in L^{q'}. \tag{2.1}$$

Guaranteeing the existence of at least a $Q \in \mathcal{Q}^q(\Gamma)$ may obtained through different considerations, that I will not detail here. In order to provide two examples:

- in the case $\Gamma = \Pi(\mu, \nu)$, the results giving L^q estimates on the transport density (see Section 1.3) may be used to guarantee the existence of at least one Q such that $i_Q \in L^q$, if $\mu, \nu \in L^q$,
- in the case $\Gamma = \{\bar{\gamma}\}$, Yann Brenier pointed out to me that the techniques used in incompressible fluid mechanics to find a Q satisfying $(\pi_{0,1})_\# Q = \bar{\gamma}$ and $(\pi_t)_\# Q = \mathcal{L}^d_{|\Omega}$ and concentrated on uniformly Lipschitz curves may be used to guarantee an L^∞ traffic intensity. Composing with Lipschitz homeomorphisms may give the result also for $\mu, \nu \neq \mathcal{L}^d_{|\Omega}$.

As in the discrete case, one would like to find the equivalence between the optimization and the equilibrium formulation by transforming the expression $\int_\Omega \overline{\xi} i_{\overline{Q}}$ into $\int L_\xi d\overline{Q}$. Yet, this presents a lot of problems due to the fact that ξ is not a continuous function and it is not even defined pointwisely. In [5] we proposed a solution based on the fact that is is still possible to define a distance with a metric $\xi \in L^p$, if $p > d$, due to Sobolev immersion theory.

For a non-negative function $\xi \in C^0(\Omega)$ we define

$$c_\xi(x, y) = \inf\{L_\xi(\omega) \,:\, \omega \in C(x, y)\}.$$

Since the $W^{1,p}$ norms of c_ξ are bounded by the corresponding L^p norms of ξ, one can get Hölder estimates if p is sufficiently large. Here the integrability of $i \in L^q$ implies $\xi = H'(i) \in L^{q'}$ and this is enough to obtain Hölder continuity if we assume that $q < d'$ ($q < 2$ in dimension 2), which will be done from now on.

For a non-negative function $\xi \in L^{q'}(\Omega)$ we then define

$$\overline{c}_\xi(x, y) = \sup\left\{c(x, y) \,:\, c = \lim_n c_{\xi_n} \text{ in } C^0(\overline{\Omega} \times \overline{\Omega}) \,:\, (\xi_n)_n \in C^0(\overline{\Omega}),\, \xi_n \geq 0,\, \xi_n \to \xi \text{ in } L^{q^*}\right\}.$$

It is possible to check that this definition extends the definition of c, in the sense that if $\xi \in C^0$, then $\overline{c}_\xi = c_\xi$. Moreover, for any $\xi \in L^{q'}$ we have $\overline{c}_\xi \in C^0$.

Once we have defined an extension of c_ξ, we need to extend L_ξ as well, so as to be able to state what is an equilibrium.

The first step is the following:

Lemma 2.2. *Let us assume that $q < 2$. Let $Q \in \mathcal{Q}^q(\Gamma)$, ξ be a non-negative element of $L^{q'}$, and $(\xi_n)_n$ be a sequence of non-negative continuous functions that converges to ξ in $L^{q'}$, then we have the following:*

(i) $(L_{\xi_n})_n$ *converges strongly in $L^1(C, Q)$ to some limit which is independent of the approximating sequence $(\xi_n)_n$ and which will again be denoted L_ξ.*

(ii) $\int_\Omega \xi(x) i_Q(x)\, dx = \int_C L_\xi(\omega)\, dQ(\omega)$.

(iii) $L_\xi(\omega) \geq \overline{c}_\xi(\omega(0), \omega(1))$ *for Q-a.e. $\omega \in C$.*

With these extensions of L and c, the characterization of optimal transport strategies then reads as:

Theorem 2.3. *Let us assume that $q < 2$ and that H is strictly convex. A transportation strategy $\overline{Q}$ is optimal if and only if, setting $\overline{\xi} := H'(i_{\overline{Q}})$, one has:*

1. $(\pi_{0,1})_\#\overline{Q}$ *solves*
$$\inf_{\gamma\in\Gamma} \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_{\overline{\xi}}(x,y)d\gamma(x,y), \tag{2.2}$$

2. *for $\overline{Q}$-a.e. $\omega \in C$, one has:*
$$L_{\overline{\xi}}(\omega) = \overline{c}_{\overline{\xi}}(\omega(0),\omega(1)). \tag{2.3}$$

This means exactly that, even in the continuous case, the optimality conditions may be expressed in terms of a formal Wardrop equilibrium where the length of every curve is weighted by the metric $\overline{\xi}(x) = H'(i_Q(x))$.

2.1.2 Vector formulation (Beckmann)

All the results of the previous part are valid for the case $\Gamma = \{\bar{\gamma}\}$ as well as for $\Gamma =$ all the transport plans.

Yet, in this second case, something more may be said, in the direction of the equivalence with Beckmann's problem. This is what has been investigated in [8].

In order to get an idea, instead of defining a scalar traffic intensity i_Q one can define a vector measure v_Q by:

$$\int_{\overline{\Omega}} \varphi(x)\, dv_Q(x) := \int_C \Big(\int_0^1 \varphi(\omega(t))\cdot\omega'(t))dt\Big) dQ(\omega),\ \forall\varphi\in C(\overline{\Omega},\mathbb{R}^d),$$

i.e. sort of a vector version of i_Q. It is immediate to check that $|v_Q| \leq i_Q$, and that

$$\nabla\cdot v_Q = \mu - \nu,\ \ v_Q\cdot\hat{n} = 0 \text{ on } \partial\Omega$$

(in a weak sense: for all C^1 functions ψ we have $-\int\nabla\psi\cdot v_Q = \int\psi\, d(\mu-\nu)$). Since H is increasing, this implies that the infimum of the previous problem with i_Q is larger than that of the minimal flow problem:

$$(B)\quad \min\left\{ \int_\Omega \mathcal{H}(v(x))\ dx\ :\ \nabla\cdot v = \mu-\nu,\ v\cdot\hat{n} = 0 \text{ on } \partial\Omega \right\}, \tag{2.4}$$

where $\mathcal{H}(v) := H(|v|)$. It is a generalization of Beckmann's problem (see Section 1.3).

A natural question, arising for instance from a comparison with the Monge case, where looking for the vector or the scalar transport density was the same, is the possible equivalence of the two problems.

One can see that a minimizer of the scalar problem can be built formally from a minimizer of the vector one in the following way: if $\bar{v}$ is the unique solution of the vector problem (2.4) and μ and ν

are absolutely continuous (so that we will write $\mu = f_0(x)dx$, $\nu = f_1(x)dx$ and $\rho_t = (1-t)f_0 + tf_1$), we consider the non-autonomous Cauchy problem

$$\begin{cases} \omega'(s) &= \mathrm{w}(s, \omega(s)) \\ \omega(0) &= x \end{cases} \tag{2.5}$$

for the non-autonomous vector field

$$\mathrm{w}(t,x) = \frac{\bar{v}(x)}{\rho_t(x)}, \quad (t,x) \in [0,1] \times \Omega. \tag{2.6}$$

The latter will not have any Lipschitz continuity property in general, unless the optimizer $\bar{v}$ itself is regular: anyway, if we assume that one can prove $\bar{v} \in \mathrm{Lip}(\Omega)$, then the flow $X : [0,1] \times \Omega \to \Omega$ of w is well-defined as the solution of (2.5) and we can take μ_t as the image of μ through the map $X(t,\cdot)$. One can see that μ_t must coincide with the linear interpolating curve $(1-t)\mu + t\nu = \rho_t(x)dx$ (because this curve solves the continuity equation thanks to the divergence condition). This yields that $(X(1,\cdot))_\#\mu = \nu$, which ensures that $X(1,\cdot)$ transports μ on ν.

This dynamical system is a typical way, originally due to Moser, to produce a transport map between two given measures, and it has been introduced in optimal transport theory by Evans and Gangbo (see [79]) for the case $c(x,y) = |x-y|$.

If we now consider the probability measure concentrated on the flow, i.e. the image measure Q of μ through the map $x \mapsto X(\cdot, x)$, it is not difficult to see that Q is admissible and $i_Q = |\bar{v}|$. Moreover, this construction provides a transport map (that is $X(1,\cdot)$) from μ to ν, whose transport "rays" evidently do not cross and which is monotone on transport "rays" (as a consequence of Cauchy-Lipschitz Theorem).

This should finally imply that the minima of the two problems coincide and that we can build a minimizer for (W) from a minimizer for (B), up to the fact that $\bar{v}$ is not regular in general and the construction is only formal.

In [8], we tried to prove in a rigorous way the equality of the two minima, and to provide a rigorous construction for a minimizer $\overline{Q}$ from a minimizer $\bar{v}$.

The first issue may be solved thanks to the concept of *superposition solution* introduced in [30]: since the curve of measure $t \mapsto \mu_t := (1-t)\mu + t\nu$ is a solution of the continuity equation

$$\frac{\partial}{\partial t}\mu_t + \nabla \cdot (\mathrm{w}\mu_t) = 0,$$

and the vector field w satisfies some integrability conditions, there exists a measure $Q \in \mathcal{P}(C)$ concentrated on solutions of (2.5), such that $\mu_t = (\pi_t)_\# Q$. This is enough for proving the following:

Theorem 2.4. *Let $\mu, \nu \in \mathcal{P}(\Omega)$ having L^q density w.r.t. to $\mathcal{L}^d$. Then we have $(W) = (B) < +\infty$.*

The second issue, i.e. giving a precise and unique construction for identifying a flow of the vector field w so as to build a measure $\overline{Q}$ concentrated on the integral curves of w itself, solving (W), is solved thanks to the regularity results of next section.

In fact, we said that should w be Lipschitz the construction would give rigorously a unique flow through the standard Cauchy-Lipschitz theorem, but it is not possible in general to hope for Lipschitz regularity of w. Even if one puts strong regularity assumptions on the densities f_0 and f_1, this strongly depends on the form of the function H.

Actually, one may write optimality conditions for v and sees that he has $v = \nabla\mathcal{H}^*(\nabla u)$ where u solves

$$\begin{cases} \nabla \cdot \nabla\mathcal{H}^*(\nabla u) &= \mu - \nu, & \text{in } \Omega, \\ \nabla\mathcal{H}^*(\nabla u) \cdot \hat{n} &= 0, & \text{on } \partial\Omega, \end{cases} \tag{2.7}$$

For $H(t) = t^2$ this is a simple Laplace equation and regularity theory is well-known. For $H(t) = t^p$ this gives a $p'-$Laplace equation and here as well lots of studies have been done. Yer, for modeling reasons, it is important to look at the case $H'(0) > 0$, since $H'(0)$ represents the metric when there is no traffic. In reality, even if nobody is on the road it is not true that we can move at an infinite speed! Hence, a typical case could be

$$\mathcal{H}(\sigma) = \frac{1}{p}|v|^p + |v|, \; v \in \mathbb{R}^N, \tag{2.8}$$

which leads to a function $\mathcal{H}^*$ which vanishes on $\overline{B_1}$. In particular, the corresponding equation for u is very very degenerate and regularity results are less studied.

Next section will present some regularity results for this kind of equations, with the aim of applying them to these congestion problems. The main result that we will be able to prove is the following:

Theorem 2.5. *Suppose that*

(i) $\mu_i = f_i\mathcal{L}^d$, *with* $f_i \in \mathrm{Lip}\ (\Omega)$ *and* $f_i \geq c > 0$, *for* $i = 0, 1$*;*

(ii) Ω *is an open connected bounded subset of* $\mathbb{R}^d$ *having Lipschitz boundary,*

then the vector field w *satisfies the hypotheses of DiPerna-Lions Theory for defining a weak flow of* w*, namely we have* $\mathrm{w} \in W^{1,2}$ *and* $\nabla \cdot \mathrm{w} \in L^\infty$*.*

Notice that the Sobolev regularity of w is equivalent to that of $\bar{v}$, once f_0, and f_1 are Lipschitz, and for the condition on the divergence one may see that we have

$$\nabla \cdot \mathrm{w} = \frac{\nabla \cdot \bar{v}}{\rho_t} - \frac{\bar{v} \cdot \nabla\rho_t}{(\rho_t)^2} = \frac{f_0 - f_1}{\rho_t} - \frac{\bar{v} \cdot \nabla\rho_t}{(\rho_t)^2}.$$

Lipschitz regularity and lower bounds on ρ_t (i.e. on f_0 and f_1) and L^∞ on $\bar{v}$ seem compulsory for getting the assumption on the divergence of w.

2.2 Regularity issues for very degenerate elliptic equations

The elliptic equations that we have been induced to consider for the applications to traffic congestion are of the following form

$$\nabla \cdot (F(\nabla u)) = f, \tag{2.9}$$

where F is the gradient of a convex function $K : \mathbb{R}^d \to \mathbb{R}$. This equation is typically coupled with a Neumann boundary condition, $F(\nabla u) \cdot n = 0$ on $\partial\Omega$. In this way it is exactly the Euler-Lagrange equation of the minimization problem

$$\min \left\{ \int K(\nabla u) + \int f u, \quad u \in \text{ a suitable Sobolev space} \right\}.$$

When $K = \mathcal{H}^*$ and $f = \mu - \nu$, this optimization problem is the dual of the minimization problem (B) and the solutions $\bar{v}$ and u of the two problems are linked by $\bar{v} = F(\nabla u)$.

As we waid, when $K(z) = |z|^2$ or $K(z) = |z|^p$ we get the usual Laplace and $p-$Laplace equations, respectively, which have been investigated intensively in the literature and have provided a lot of regularity results on the solution u (which is a priori supposed to belong to H^1 or $W^{1,p}$ only) according to the regularity of f. Most of the results have been extended to the case of variable coefficients or of different functions K, which share anyway some properties of the square or of the $p-$th power. This latter case is much more difficult than that of the square (which gives a linear equation), mainly because of the degeneracy of D^2K near $z = 0$. Actually, lower bounds on D^2K are often useful to give estimates on the norms of the solution in terms of the norms of f, and if D^2K is allowed to tend to zero for small values for ∇u (which is the case for $p > 2$), extra difficulties arise. Yet, this is - roughly speaking - compensated by the fact that the smallness of the gradient already provides some regularity estimates.

This is what is usually done in elliptic regularity but the equation we wanted to look at was even worse than the $p-$Laplace equation. Our main example is given by the function

$$K_{(p)}(z) = \frac{1}{p}\big(|z| - 1\big)_+^p, \qquad p > 1, \tag{2.10}$$

which vanishes, together with its Hessian, on the whole ball B_1. This function is the Legendre transform of $H(z) = |z| + \frac{|z|^p}{p}$. This means that all the values of ∇u which are smaller than 1 are a source of problems. Obviously, the regularity result that one may expect to prove must concern $F(\nabla u)$ instead of u itself: one can convince himself that the homogeneous equation (the same, with $f = 0$) would be solved by any $1-$Lipschitz function and that nothing more may be said on u. On the contrary, $F(\nabla u)$ is reasonably more regular, since either it vanishes or we are in a zone where $|\nabla u| > 1$ and the equation is more elliptic. But gluing the two zones, which are not open sets with nice boundaries, is not trivial.

In [8] we obtained the following results, which were tuned on the applications we had in mind.

Theorem 2.6. *Take $K = K_{(p)}$ for $p \geq 2$. Let us suppose that Ω has a $C^{3,1}$ boundary and that $f \in W^{1,q}$ with $q = p' = p/(p-1)$. Then, if $u \in W^{1,p}(\Omega)$ is a weak solution of the Neumann boundary problem*

$$\begin{cases} \nabla \cdot (\nabla K(\nabla u)) &= f, \quad \text{in } \Omega, \\ F(\nabla u) \cdot \hat{n} &= 0, \quad \text{on } \partial\Omega, \end{cases} \tag{2.11}$$

and we set $G(z) = (|z| - 1)_+^{\frac{p}{2}} \frac{z}{|z|}$, we have $G(\nabla u) \in W^{1,2}(\Omega)$.

Remark 2.1. Notice that this Sobolev regularity result asks for Sobolev regularity of f itself, which is not at all natural in elliptic regularity. The reason lies in the fact that the proof is obtained by estimating the incremental ratios (say, by translations) and usually the terms of the kind $\int(f_h - f)(u_h - u)$ are estimated putting one more discrete derivative on the u. Yet, such a term is not controlled by the incremental ratio of $G(\nabla u)$ (especially in this case where G vanishes for non-zero values of ∇u).

On the other hand, an L^∞ result on ∇u is easy to obtain (and the idea is that functions of the kind $(\partial u/\partial x_i - (1+\delta))_+$ are subsolution of a uniformly elliptic equation). Hence we get

Theorem 2.7. *Suppose that $\Omega \subset \mathbb{R}^d$ has a $C^{2,1}$ boundary, that $D^2K(z) \geq c_1 I_d$ for $|z| \geq 2$ and that $f \in L^r$ with $r > d$: then any solution of problem* (2.11) *is a Lipschitz function.*

The two above results may be put together to obtain the following, and apply DiPerna-Lions theory as we stated in the previous section.

Corollary 2.8. *Take $K = K_{(p)}$ for $p \geq 2$, suppose that Ω has a $C^{3,1}$ boundary and that $f \in W^{1,q}$ with $q = p' = p/(p-1)$. Then, if u is any solution of problem* (2.11)*, we have $F(\nabla u) \in W^{1,2}(\Omega) \cap L^\infty(\Omega)$.*

Notice that all these results, thanks to the Neumann boundary condition, are global up to $\partial\Omega$, which is what we need to apply DiPerna-Lions theory (a global bound on the divergence is needed).

In a subsequent paper, in collaboration with Vespri, we were able to prove something more in dimension two.

Theorem 2.9. *Let u be a solution of $\nabla \cdot (F(\nabla u)) = f$, in dimension $d = 2$, where $F = \nabla K$ and K is a convex $C^{1,1}_{loc}$ function satisfying*

$$\text{for all } \delta > 0 \text{ there exists } c_\delta > 0 \ : \ D^2K(z) \geq c_\delta I_d \text{ for all } z \text{ such that } |z| \geq 1 + \delta. \tag{2.12}$$

Suppose $f \in L^{2+\varepsilon}$ and $F(\nabla u) \in H^1 \cap L^\infty$. Then $g(\nabla u)$ is locally continuous on Ω for any continuous function $g : \mathbb{R}^2 \to \mathbb{R}$ such that $g = 0$ on B_1. In particular, if $K = K_{(p)}$, then $F(\nabla u)$ is continuous.

There are also estimations of the modulus of continuity, but they are not that explicit (we start from proving a local logarithmic modulus of continuity $\omega(R) = \sqrt{|\log R|}$ for functions like $(\partial u/\partial x_i - (1+\delta))_+$). The techniques are inspired by what developed has been by DiBenedetto and Vespri for more general but less degenerate equations in [78].

This has some applications: for instance the curves which are charged by Q are C^1 classical solutions of the dynamical system, and the metric ξ associated to the optimal Q is actually continuous, which gives a cleaner interpretation in terms of continuous Wardrop equilibria of what presented in [5].

2.3 Numerical approach and results

For applications, being able to compute or approximate equilibrium and optimal traffic repartitions is crucial. This is often performed in the case where $\bar{\gamma}$ is prescribed, since it seems more important for real traffic applications. Even in the discrete (network) case, this is a classical issue, mainly studied in operational research. Up to computational complexity, it is a kind of problems which is not that hard to approach, due to its convexity. For reducing the complexity, it is quite typical to look at the dual problem: we actually have

$$\min_q \sum_e H(i_q(e))\mathcal{H}^1(e) = \max_\xi \big(\sum_{s,d} d_\xi(s,d)\bar{\gamma}(s,d) - (-\sum_e H^*(\xi(e))\mathcal{H}^1(e)\big).$$

Moreover, the optimal $\overline{\xi}$ in the dual problem will be given by $\overline{\xi} = H'(i_q)$(for an optimal q), which will allow to recover the traffic intensity i_q from $\overline{\xi}$. The advantage of the dual formulation is that we have one variable per edge instead of one per path!

The continuous congested problem as well admits, in the same spirit, a dual formulation which is easier to be solved. of last Section admits a dual formulation. By using the same notation as before, the dual problem we look at is

$$\min \left\{ J(\xi) := \int_\Omega H^*(\xi(x))dx - \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_\xi(x,y)d\bar{\gamma}(x,y) \,:\, \xi \in L^{q'}, \xi \geq \xi_0 := H'(0) \right\} \tag{2.13}$$

For numerical purposes, this problem has to be discretized on a grid, and the delicate point is to define discretizations of the geodesic distances $\overline{c}_\xi(x,y)$. This is done thanks to the so-called Fast Marching Method.

Let me give a short description of this method, which is much more accurate than a simple computation of the distances on a network associated to the grid (say, a Djistra-like algorithm). For a fixed source S_0 we will write $\mathcal{U}_\xi(.) = c_\xi(S_0,.)$ and notice that the function $\mathcal{U}_\xi$ is the viscosity solution of the Eikonal equation $|\nabla\mathcal{U}| = \xi$, with $\mathcal{U}(S_0) = 0$.

Adopting standard notation, we will note $\mathcal{U}_{i,j}$ the value of $\mathcal{U}$ at the lattice vertex (i,j) (forgetting for a while the dependence on ξ). A discrete version of the Eikonal equation is solved in order to compute $\mathcal{U} = (\mathcal{U}_{i,j})$. Rouy and Tourin [98] showed, with generalizations to other Hamilton-Jacobi equations, that the viscosity solution on a square gird of step h is selected if one finds a discrete set of values for $\mathcal{U}_{i,j}$ satisfying

$$D\mathcal{U}_{i,j} = \xi_{i,j}, \tag{2.14}$$

where we noted

$$D_x\mathcal{U}_{i,j} := \max\{(\mathcal{U}_{i,j} - \mathcal{U}_{i-1,j}), (\mathcal{U}_{i,j} - \mathcal{U}_{i+1,j}), 0\}/h,$$
$$D_y\mathcal{U}_{i,j} := \max\{(\mathcal{U}_{i,j} - \mathcal{U}_{i,j-1}), (\mathcal{U}_{i,j} - \mathcal{U}_{i,j+1}), 0\}/h$$
$$D\mathcal{U}_{i,j} = \sqrt{D_x\mathcal{U}_{i,j}^2 + D_y\mathcal{U}_{i,j}^2}.$$

The Fast Marching Method (FMM) is a numerical method introduced by Sethian in [99] for efficiently solving the isotropic Eikonal equation on a cartesian grid, i.e. equation (2.14). Values of $\mathcal{U}$ are obtained recrusively one from the others, selecting in a clever way the order to visit the points.

Let us call $c^h_\xi(T,S_0) = \mathcal{U}(T)$ the value of this solution (computed in this exact way) at the target point T when the vanishing boundary datum is fixed at S_0 and the metric is ξ.

The functional to be minimized will hence be the following

$$J^h(\xi) = h^2 \sum_{i,j} H^*(i, j; \xi_{i,j}) - \sum_{r,s} c^h_\xi(S_\alpha, T_\beta)\gamma^h_{\alpha,\beta},$$

where the weights $\gamma^h_{\alpha,\beta}$ represent the coupling on the set of pairs sources S_α - targets T_β and $\sum_{\alpha,\beta} \gamma_{\alpha,\beta} = 1$.

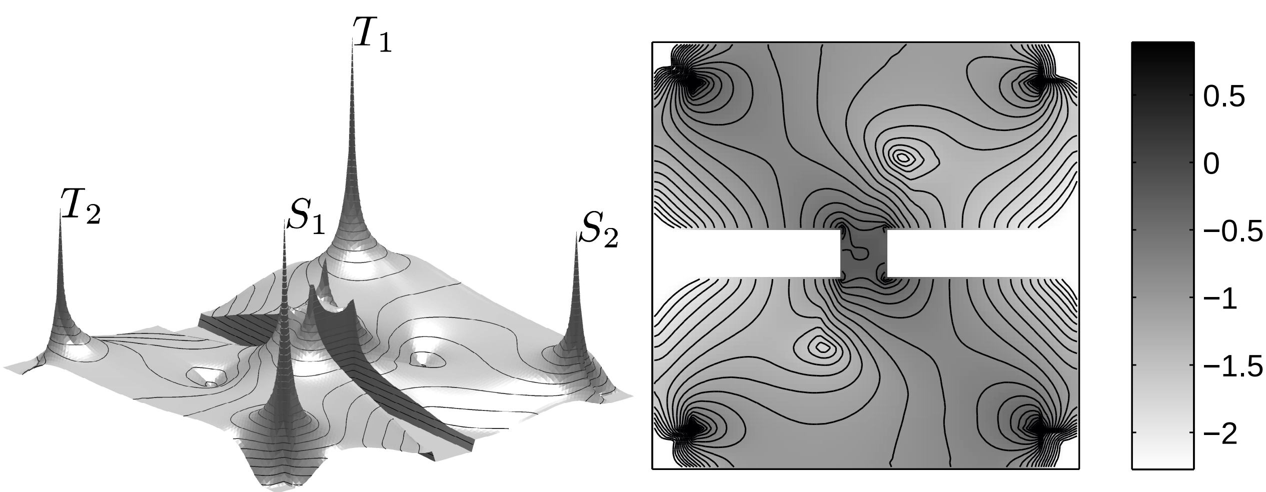
The main results of [7] are summarized in the following theorem

Theorem 2.10. *The functional J^h is a strictly convex functional whose solution may be approximated by standard subgradient methods (since it is non-differentiable, see Section 2.4 for its subdifferential), thus guaranteeing convegence to the minimizer. Moreover, it can be extended to functional, still denoted by J^h, defined on $L^{p'}(\Omega)$, putting it at $+\infty$ on those ξ which are not piecewise constant n the cells of the grid, and in this case we have*

$$J^h \xrightarrow{\Gamma} J$$

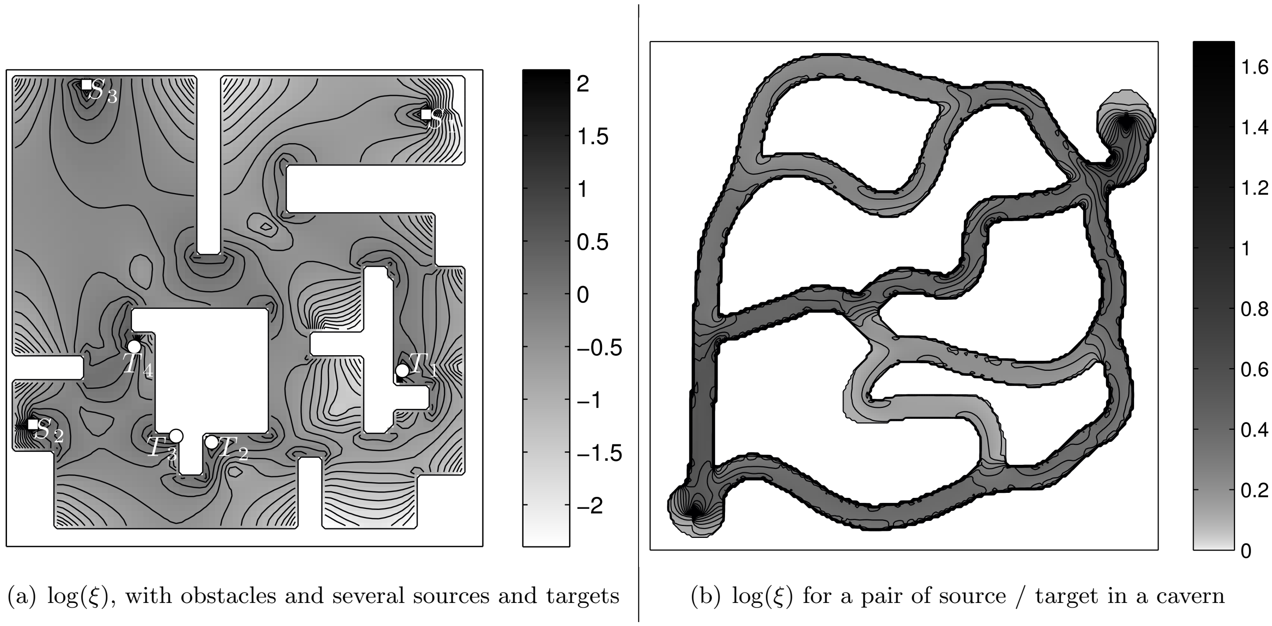
provided $\gamma^h \rightharpoonup \gamma$ (the $\Gamma-$convergence being intended w.r.t. the weak convergence in $L^{p'}$). As a by-product, the minimizers of J^h converge to the unique minimizer $\bar{\xi}$ of J.

The key point for the numerical simulations - some of which are presented here - is the computation of the gradient of the term c^h_ξ with respect to ξ, which has been useful for approaching other problems as well (see next section).



(a) 3D view of ξ (b) Flat view of $\log(\xi)$

Figure 2.1: Two sources and two targets, with a river and a bridge on a symmetric configuration with asymmetric traffic weights.



(a) $\log(\xi)$, with obstacles and several sources and targets

(b) $\log(\xi)$ for a pair of source / target in a cavern

Figure 2.2: Traffic congestion equilibrium metric in the case of several sources and targets with obstacles. On figure (b), one can see that the traffic is higher on the shortest (in the Euclidian sense) pathways.

2.4 Useful tools and perspectives

I want to start this section by a short list of results that appeared during the preparation of the papers composing this part of my researches, which are likely or have already shown to be useful in other fields, but which were not originally meant as the main part of the work they were included in.

First of all, during the preparation of [5] we ended up on the Wardrop equilibrium condition with a metric $\xi \in L^p$. It was necessary to define a distance associated to this metric, and this lead to the definition of $\bar{c}$. Yet, the interpretation in terms of viscosity solution of the Eikonal equation could lead to different definitions. One of them is through the maximal a.e. subsolution, other could come from the L^p-theory of viscosity solutions (see [62] for a definition via $W^{1,q}$ test functions and local a.e. inequality and [63] for some other notions and equivalences in the case $\xi \in L^\infty$).

We can define

$$v_\xi = \sup\{v \,:\, v \in W^{1,q}(\Omega),\, v(S_0) = 0,\, |\nabla v| \leq \xi \text{ a.e. }\},$$

which is the maximal a.e. subsolution of $|\nabla v| = \xi$ Then, we highlighted in [7] the following statement.

Lemma 2.11. *The following equality hold:* $\bar{c}_\xi(S_0, .) = v_\xi = \lim_{\varepsilon\to 0} c_{\xi*\rho_\varepsilon}(S_0, \cdot)$.

I do not believe that this equivalence is novel, but I think that it deserves to be underlined, due to its unexpected applications (for instance in this field, i.e. continuous congestion games).

I will stick a little bit more to the field of PDE while presenting the next result that I want to highlight. During the study of the very degenerate elliptic equation $\nabla \cdot \nabla K_{(p)} = f$, we needed to prove global L^∞ and H^1 bounds for $G(\nabla u)$. Globalness of the bound was crucial.

First, I must say that the proof we gave in [8] contained an error in the boundary estimates and it is not complete. We could fix this error some time later, and the detailed computations will be included in Lorenzo Brasco's PhD thesis. For both the estimates, Sobolev and L^∞, the most easy way for dealing with Neumann boundary conditions is the following: if you are able to prove local interior regularity, then perform reflections across the boundary and local results on the reflected solution will be turned into global ones for the original solution. If, for instance, Ω is a cube this is quite easy and a correct proof is achieved very quickly. If on the contrary $\partial\Omega$ needs to be "rectified" through local diffeomorphisms, the equation turns into a variable-coefficients one. The regularity of the coefficients depends on the regularity of the boundary and this is why, for usual non-degenerate elliptic equations, $C^{1,1}$ at least is needed. Actually, we followed a paper by Carstensen and Müller ([66]), which uses a precise and explicit diffeomorphisms whose Jacobian matrix has some commutation properties with the simmetrization matrix (properties which are needed to deduce a PDE on the simmetrized solution) and this requires to lose one derivative more and leads at least to $C^{2,1}$. For the L^∞ regularity this is enough, and this should not be astonishing: as I already mentioned, this part of the estimate (L^∞) is the most classical one and - in the case of constant coefficients - could be easily deduced from standard results on subsolutions. For the variable coefficients case (i.e. the global result up to the boundary), the technique of DiBenedetto in [76] can be adapted. Hence L^∞ up to the boundary is true for $C^{2,1}$ domain.

As far as the Sobolev proof is concerned it is more complicated: the strategy is that of estimating the differential quotient $(G(\nabla u_h) - G(\nabla u))/|h|$, where u_h denotes a translation of a vector h of the solution u. One needs to prove L^2 bounds for this quotients. As I already underlined in Remark 2.1, the terms that let $u_h - u$ appear need another (discrete) integration by parts since in this degenerate case we cannot control $|u_h - u|$ with $|G(\nabla u_h) - G(\nabla u)|$. In the case of interior regularity (i.e. constant coefficients) the only extra integration by parts is correctly handled once one supposes $f \in W^{1,q}$. For boundary regularity and variable coefficients, one needs extra regularity on the coefficients themselves and $C^{1,1}$ would be enough. Yet, the problem is that when one reflects the equation across the boundary, no more than Lipschitz regularity may be guaranteed for the coefficients, since they are themselves obtained as the reflection of the coefficients inside Ω and reflecting a Lipschitz function gives a Lipschitz function, while reflecting a $C^{1,1}$ function does not give in general a $C^{1,1}$ function !

In our specific case, a term like

$$\int_\Omega \big(c_h(x) - c(x)\big) \cdot \big(\nabla u_h(x) - \nabla u(x)\big)\, dx$$

had to be estimated with something of the order of $|h|^2$, where the vector function c is a general expression for the variable coefficients of the equation and $\Omega = \Omega^+ \cup \Omega^-$ is to be thought as the union of a diffeomorphic image Ω^+ of a part of the original domain near the boundary and of its reflection. If $c \in C^{1,1}$, one can integrate by parts in a discrete forms (i.e. the two parts $\big(c_h(x) - c(x)\big) \cdot \nabla u_h(x)$ and $\big(c_h(x) - c(x)\big) \cdot \nabla u(x)$ are separated, the change of variable $y = x + h$ on the first one is performed, and the two integrals are reconsidered together), thus getting

$$\int_\Omega \Delta_h c(x) \cdot \nabla u(x)\, dx, \quad \text{where } \Delta_h c = c_h - 2c + c_{-h}$$

and, using $\Delta_h c = c_h - 2c + c_{-h} \leq C|h|^2$ and the integrability assumptions on ∇u, one can conclude. Yet, in this case c is separately $C^{1,1}$ on Ω^+ and Ω^- but only Lipschitz on the whole Ω. This means that, if one sets $\Omega_h^\pm := \{x \in \Omega^\pm \,:\, x+h, x-h \in \Omega^\pm\}$ and $\Gamma_h = \Omega \setminus (\Omega_h^+ \cup \Omega_h^-)$, we have

$$\begin{aligned}\int_\Omega \Delta_h c(x) \cdot \nabla u(x)\, dx &\leq C|h|^2\big(||\nabla u||_{L^p(\Omega_h^+)} + ||\nabla u||_{L^p(\Omega_h^-)}\big) + \int_{\Gamma_h} \Delta_h c(x) \cdot \nabla u(x)\, dx \\ &\leq C|h|^2\big(||\nabla u||_{L^p(\Omega_h^+} + ||\nabla u||_{L^p(\Omega_h^-}\big) + |h|\,|\Gamma_h|\ \ ||\nabla u||_{L^\infty}.\end{aligned}$$

Since $|\Gamma_h| \leq C|h|$, this shows that an L^∞ bound on the gradient allows to conclude the Sobolev regularity.

I stressed this part of the computations since I think it is not an usual idea to use L^∞ bounds as a part of the proof of H^1 regularity.

After these two remarks on PDEs, I think I must move to algorithms and numerics. Actually, one of the main side-contributions of what we studied with the goal of giving results for Continuous Traffic Congestion is probably linked to numerical computations. The Fast Marching Method is nowadays universally used as an efficient method for computing geodesic distances and what we did is computing the derivatives of its result with respect to the metric

In the continuous case, if we compute $\mathcal{U}_{x_0,\xi}(y)$ and let ξ vary, if ξ is replaced by $\xi + \varepsilon h$, we have

$$\frac{d}{d\varepsilon}\mathcal{U}_{x_0,\xi+\varepsilon h}(y) = \int_0^1 h(\omega_{x_0,y}(t))|\omega'_{x_0,y}(t)|\, dt$$

where the integral follows a geodesic $\omega_{x_0,y}$ (with respect to ξ).

The discrete analogous of this derivative "concentrated on the geodesics" has been studied in [6]. Here the starting point is the fact that the value $\mathcal{U}(y)$ satisfies

$$\text{either } (\mathcal{U}(y) - \mathcal{U}(y_1))^2 + (\mathcal{U}(y) - \mathcal{U}(y_2))^2 = \xi^2(y) \text{ or } \mathcal{U}(y) - \mathcal{U}(y_1) = h\xi(y),$$

for one or two "parents" y_1, y_2. These parents are the points that the FMM algorithm has visited before $y = (i, j)$ and whose values are active in the computation of $D_x\mathcal{U}_{i,j}$ and $D_y\mathcal{U}_{i,j}$.

Theorem 2.12. *The value $\mathcal{U}_\xi(y)$ is a concave function of ξ and its variations are given by*

$$\delta\mathcal{U}(y) = \frac{2h^2\xi(y)\delta\xi(y) + \delta\mathcal{U}(y_1)(\mathcal{U}(y) - \mathcal{U}(y_1)) + \delta\mathcal{U}(y_2)(\mathcal{U}(y) - \mathcal{U}(y_2))}{2\mathcal{U}(y) - \mathcal{U}(y_1) - \mathcal{U}(y_2)}$$

or

$$\delta\mathcal{U}(y) = h\delta\xi(y) + \delta\mathcal{U}(y_1)$$

in the case of two or one parents, respectively, which allows to compute recursively $\nabla_\xi \mathcal{U}(y)$ visiting the points in the same order of the FMM and during the same routine, with a computational cost of $O(N^2 \ln(N))$ if N is the number of points of the grid.

The vector $\nabla_\xi \mathcal{U}(y)$ which is obtained is either the gradient at differentiability points, or an element of the supergradient in general.

This computation, together with a gradient algorithm, has been applied to different setting than traffic congestion: for instance when maximizing $\xi \mapsto d_\xi(S,T)$ under integral and L^∞ constraints on ξ (the "landscape design" problem, where a metric has to be chosen so as to slow down the movement of a potential enemy who arrives from S and attacks T). If the constraints are $0 \leq \xi \leq \overline{\xi}$ and $\int \xi \leq M$ the solution is known to be composed by two balls around S and T, while there is no explicit solution if the lower bound for ξ is not zero but a positive value. See [61] and [6] and look at Figure 2.3 for an idea of the solution.

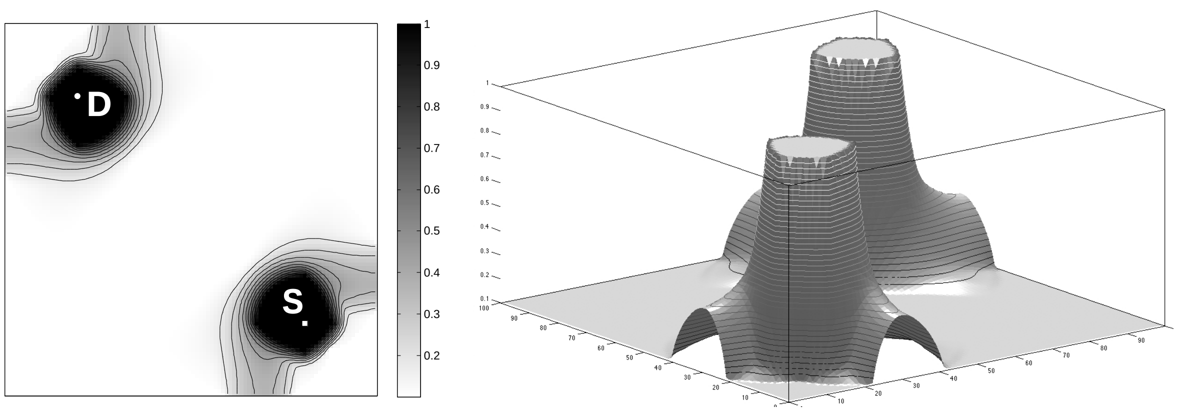


Figure 2.3: 2D and 3D display of the optimal metric $\xi^\star$.

As far as perspective research linked to this subject is concerned, there is still something to be done both in the modeling and in the technical part.

In modeling, it would be interesting to insert anisotropy in the models (it is not true in reality that the speed to pass through a point only depend on how many people are passing through the same point: it also depends on our and their directions). This could be done by means of Finsler metrics. Moreover, one natural question is whether these models have some link with the theory of Mean Field Games. These problems as well consist in the movement of some players whose individual goal is fixed but whose utility also looks at the density of all the other players (i.e. their movement is more expensive if they pass where the density is higher), which is the case of what J.-M. Lasry and P.-L. Lions introduced in [86]. And, finally, due to their game theory aspect, these Wardrop equilibra issues in a continuous setting should undergo the usual analysis of more sophisticated questions that game theorists have already applied to the network case (repeated games,. . .).

From the technical point of view, I think several questions have been opened in the study of very degenerate elliptic equations, and probably the first one should be seeing whether the unnatural

Sobolev assumption on f is necessary.

The applications of the Subgradient Marching Algorithm to other non-convex problems is also a delicate issue. In [6] we did an experiment in travel-time tomography, for the reconstruction of an unknown density of material from the wave propagation (either sound or light) inside it, but one could also think to use such a gradient computation outside variational models, in evolutionary processes, for instance.

Chapter 3

Last contributions to the theory of branched transport

Résumé La théorie du transport branché est la variante du transport usuel où le coût pour une masse m qui e déplace sur une longuer l n'est pas ml ou $mc(l)$ mais $m^\alpha l$, pour un exposant $\alpha \in]0,1[$, ce qui favorise le fait que des masses différentes se mettent ensemble. Cela donne lieu à des structures branchées et les branchement se font d'après certaines lois d'angle. Au moins, cela est ce qui se passe dans le cas d'un graphe fini, mais en continu on a des graphes infinis, voire des ensembles rectifiables. Le problème admet néanmoins une formulation en continu, avec des mesures vectorielles singulières (concentrées sur des rectifiables) à divergence fixée ou avec des mesures sur l'espace des courbes. Le modèle et le propriétés de ces problèmes variationnels ont été étudiées à fond.

Ce qu'on démontre ici est un résultat de régularité pour le réseau optimal, sous l'hypothèse que la mesure irriguée à partir d'une seule source soit équivalente à celle de Lebesgue. On présente en suite un contre-exemple quand cette hypothèse manque.

En suite, grâce à son expression en tant que minimisation d'une fonctionnelle concave parmi les mesures vectorielles, on propose une approximation du problème par une suite des fonctionnelle "elliptiques": on minimise, parmi des vrais champs de vecteurs (non pas des mesures), une fonctionnelle du type $\varepsilon \int |\nabla v|^2 + \varepsilon^{-1} \int |v|^\alpha$ (avec d'autres exposants, trouvés de manière à garantir la convergence). On démontre la $\Gamma-$convergence de cette énergie vers celle du transport branché et ce résultat de convergence est à la base des résultats numériques mentionnés dans la partie "perspectives".

The name "branched transport" is now often used for addressing all the transport problems where the cost for a mass m moving on a distance l is proportional to l but sub-additive w.r.t. m and typically proportional to a power m^α $(0 < \alpha < 1)$. The distributions of sources and destinations are given and one looks for the path followed by each particle, sums up the mass which moves together on each part of the path, and associates to every configuration its total cost $\sum_i l_i m_i^\alpha$. The adjective "branched" in the name stands for one of the main features of the optimal solutions: they gather mass together, masses tend to move jointly as long as possible, and then they branch towards different destinations, thus giving rise to a tree-shaped structure.

We present here the framework of the optimization problem proposed by Xia in [107, 108] and then studied by many authors (see for instance [49] for a whole presentation of the theory).

Let $\Omega \subset \mathbb{R}^d$ be an open set with compact closure $\overline{\Omega}$ and $\mathcal{M}(\Omega)$ the set of finite vector measures on $\overline{\Omega}$ with values in $\mathbb{R}^d$ and such that their divergence is a finite scalar measure. On this space we consider the convergence $u_\varepsilon \to u$ corresponding to $u_\varepsilon \rightharpoonup u$ and $\nabla \cdot u_\varepsilon \rightharpoonup \nabla \cdot u$ as measures. When a function is considered as an element of this space, or a functional space as a subset of it, we always think of absolutely continuous measures (with respect to the Lebesgue measure on Ω) and the functions represent their densities. On the other hand, when we take $u \in \mathcal{M}(\Omega)$ and we write $u = U(M, \theta, \xi)$ we mean that u is a rectifiable vector measure (it is the translation in the language of vector measures of the concept of rectifiable currents) $u = \theta\xi \cdot \mathcal{H}^1_{|M}$ whose density with respect to the $\mathcal{H}^1-$Hausdorff measure on M is given by the real multiplicity $\theta : M \to \mathbb{R}^+$ times the orientation $\xi : M \to \mathbb{R}^d$, ξ being a measurable vector field of unit vectors belonging to the (approximate) tangent space to M at $\mathcal{H}^1-$almost any point.

For $0 < \alpha < 1$, we consider the energy

$$M^\alpha(u) = \begin{cases} \int_M \theta^\alpha d\mathcal{H}^1 & \text{if } u = U(M, \theta, \xi), \\ +\infty & \text{otherwise.} \end{cases} \tag{3.1}$$

The problem of branched transport amounts to minimizing M^α under a divergence constraint:

$$\min \left\{ M^\alpha(u) \ : \ \nabla \cdot u = f := f^+ - f^- \right\}. \tag{3.2}$$

This is not the original definition by Xia of the Energy M^α: Xia proposed it in [107] as a relaxation from the case of finite graphs, and then proved that formula (3.1) can be seen as a representation formula for the relaxed energy

$$M^\alpha(u) = \inf \left\{ \liminf_n E^\alpha(G_n) \ : \ G_n \text{ finite graph}, \ u_{G_n} \to u \right\},$$

where

$$E^\alpha(G) := \sum_h w_h^\alpha \mathcal{H}^1(e_h), \tag{3.3}$$

for a weighted oriented graph $G = (e_h, \hat{e}_h, w_h)_h$ (where e_h are the edges, $\hat{e}_h$ their orientations, w_h the weights), and u_G is the associated vector measure given by

$$u_G := \sum_h w_h \hat{e}_h \mathcal{H}^1_{|e_h},$$

(and the convergence is in the sense of $\mathcal{M}(\Omega)$).

In this way a continuous model has been obtained starting from the widely-studied discrete problem on graphs, as proposed by Gilbert (see [83]), and generalizing it.

In general Problem (3.2) admits a solution with finite energy for any pair of probability measures (f^+, f^-) (or, more generally, for any pair of equal mass finite positive measures), provided $\alpha > 1-1/d$ (this is proven in [107] by means of an explicit construction). this is coherent with the case $\alpha = 1$, corresponding to Monge, where all compactly supported measures may be linked with finite energy.

For the same problem other models have been suggested and a very wide literature uses (as we did in the congestion part) measures on the set of paths.

Let us denote by C the set of $1-$Lipschitz curves $\omega : [0, +\infty[\to \Omega$ that are eventually constant. It means that, if we define the stopping time of a curve ω by

$$T(\omega) = \inf \{s : \omega \text{ is constant on } [s, +\infty[\},$$

these are curves with $T(\omega) < +\infty$.

Given a probability measure Q on the space C, for any point $x \in \mathbb{R}^d$ the $Q-$multiplicity of x is defined by

$$[x]_Q := Q \{\omega \in C : x \in \omega([0, T(\omega)])\}. \tag{3.4}$$

Then we can define $Z_Q(\omega) = \int_0^{T(\omega)} [\omega(t)]_Q^{\alpha-1} dt$ and $J(Q) = \int_C Z_Q \, dQ$. We can also call, as in [47], *traffic plan* any measure Q on the space C with $\int T(\omega) dQ < +\infty$ (and every Q such that $J(Q) < +\infty$ has such a property) and *fiber* of the traffic plan Q any curve ω such that $Z_Q(\omega) < +\infty$ (as in [23]).

Finally, we consider the maps $\pi_0, \pi_\infty : C \to \Omega$, given by $\pi_0(\omega) = \omega(0)$, and $\pi_\infty(\omega) = \omega(T(\omega))$. The two image measures $(\pi_0)_\# Q$ and $(\pi_\infty)_\# Q$, which belong to $\mathcal{P}(\Omega)$, will be called the starting and the terminal measure of Q, respectively, or irrigating and irrigated measures.

The minimization problem proposed in [47] is

$$(P) \quad \min \quad J(Q) \ : \ Q \in \mathcal{P}(C) \, (\pi_\infty)_\# Q = \mu, \, (\pi_0)_\# Q = \nu,$$

where μ and ν are given measures in $\mathcal{P}(\Omega)$.

In this chapter of the Mémoire I want to present the last contributions I gave to this theory after my PhD thesis. They are contributions mainly conceived during a my post-doc à Cachan and they concern a regularity result on the optimal networks obtained in collaboration with J.-M. Morel ([10]), and an approximation result by means of elliptic functionals in view of possible numerical schemes as well as to do a bridge with other theories in Calculus of Variations. This approximation result has already proven to be useful for numerical purposes, and this is the object of work in progress with E. Oudet that I will present in the perspectives section, which will also contain another work-in-progress with L. Brasco on the equivalence with a model using curves of measures, and some other issues from game theory.

3.1 Regularity

In the discrete setting, the irrigated mass and the irrigating mass are finite atomic masses and the optimal graph has no circuits and is therefore a tree, with a finite number of vertices joined by straight edges. In addition, the following equilibrium equation is satified at all vertices:

$$\sum_{i\in I} m_i^\alpha \vec{e}_i = 0 \tag{3.5}$$

where $m_i = w(e_i)$ and $\vec{e}_i$ are the flows and directions of all edges e_i arriving or leaving a given vertex and all $\vec{e}_i$'s are oriented inwards the corresponding edge.

One of the main challenges of the continuous model is to explore the regularity of infinite networks. It is evident that no precise result concerning angle law as above can be stated in a continuous setting if we cannot prove the existence of tangent directions at every point of the graph. In [108, 109] the blow-up limits of an optimal network are studied, far from the support of the two measures and inside them, respectively. Yet, the limit of the blow-up was only considered up to subsequences, which is exactly the main problem in trying to give the angle between tangent directions at a point. Far from the support of the two measures, under some additional assumptions (either the two supports are disjoint or one of the measures is purely atomic), the problem has been solved for $\alpha > 1 - 1/d$ in [48] by Bernot, Caselles and Morel, who proved, among other results, that the network is locally a finite graph on $\mathbb{R}^d \setminus (\mathrm{supp}(\mu) \cup \mathrm{supp}(\nu))$.

In [10] we proved, in collaboration with J.-M. Morel, in the case $\mu = \delta_0$ and $\nu = f \cdot \mathcal{L}$ with $c \le f \le C$, a regularity result saying that "on every curve composing the network, the direction of the tangent is a locally BV function" and we identified its (measure) derivative. In particular this implies the existence of a tangent direction at any point which is not a branching point and the existence of a tangent for any branch at junctions.

Theorem 3.1. *Let $\omega(s)$, $s \in [0, T(\omega)]$ be a fiber of an optimal traffic plan Q irrigating a measure ν equivalent to $\mathcal{L}^d$ from δ_0. Suppose that ω is parametrized by arclength and, for $[a, b] \subset [0, T(\omega)[$, set $\Sigma =: \omega([a, b])$. Let ε_i, $i \in I$, be the masses of all trees branching from Σ. Then $\sum_i \varepsilon_i^\alpha < \infty$ and Σ has a bounded total curvature, i.e. $\omega' \in BV(a, b)$. As a consequence Σ has two half-tangents at all points and a tangent at all points which are not branching points.*

The most important consequence is the following:

Corollary 3.2. *Every branching point x of an optimal traffic plan has a tangent cone made of a finite (and bounded by a constant depending on α and d) number of segments whose directions e_i and masses m_i satisfy the equation* (3.5).

The next step is to strengthen the result, by providing a differential equation which is satisfied by ω. We have just proven that ω' is BV. In addition $[\omega]_Q^\alpha$ is BV since it is monotone decreasing. Thus the product $[\omega]_Q^\alpha \omega'$ is BV and its derivative is a measure. More precisely, we have

Theorem 3.3. *Let ω be a fiber of an optimal traffic plan Q irrigating a measure ν, equivalent to $\mathcal{L}^d$, from δ_0. Then ω satisfies in the sense of distributions the elliptic equation*

$$-([\omega(t)]_Q^\alpha \omega'(t))' = \sum_{i \in I} \varepsilon_i^\alpha \delta_{\omega(t_i)} \hat{n}_i \tag{3.6}$$

where $\hat{n}_i$ is the tangent of the branch stemming from ω at $\omega(t_i)$ with mass ε_i. Notice that this tangent vector exists and that the right hand side is a vector measure with finite mass, thanks to Theorem 3.1.

The reasons for looking only at the case $\mu = \delta_0$ are various and mainly technical, since, due to the monotonicity of $[\omega(t)]_Q$ along fibers, many results are easier to obtain in this case; on the other hand, this case is already quite important in applications, since many branched systems are composed by a network transporting a fluid from one single source to multiple destinations or the other way around (for instance blood vessels, river basins...).

On the contrary, the assumptions of ν for this regularity result could seem unnatural (being equivalent to the Lebesgue measure). However, we provided in [10] an example showing that the existence of the tangent direction may not be true for any optimal network stemming from δ_0. In our counter-example we chose ν to be an infinite combination of Dirac masses, so that we actually have lots of zero-densities (outside the support of ν, which is countable) and of "infinite" densities (on the atoms).

Our counterexample is the following:

$$\mu = \delta_a, \ \nu = (1/2 - \varepsilon)\delta_b + (1/2 - \varepsilon)\delta_c + \sum_{i=1}^{\infty} \varepsilon_i \delta_{z_i},$$

where $\sum_{i=1}^{\infty} \varepsilon_i = 2\varepsilon$ and the points z_i belong to a cone T centered around the segment linking b and a and accumulate close to b, so that the optimal traffic plan from μ to ν has a part contained in T which arrives up to b, and then goes straight up to c, according to the picture.

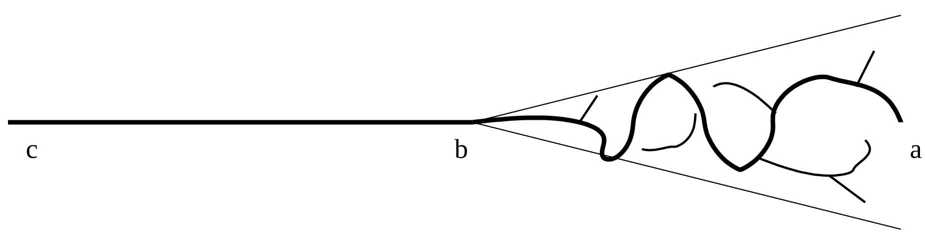


Figure 3.1: A sketch of the counterexample: the main fiber passes through b and then goes straight with no branching

Moreover, we will choose the points z_i so that the traffic plan Q will be forced to follow those points (the network will be consequently composed by the straight line segments $z_i z_{i+1}$, which will converge to b, and by the segment bc). It will be possible to choose the points satisfying the additional criterion that they oscillate from one side of T to the other, thus having as a consequence that the tangent of the traffic plan at b does not exist.

Theorem 3.4. *Let us make the following choices, according to our previous notations (and complex notations for points in the plane): θ is an angle sufficiently small; $f : \mathbb{R} \to [-\frac{\theta}{2}, \frac{\theta}{2}]$ is a Lipschitz periodic function such as $f(t) = \frac{\theta}{2}\sin t$; $a = 1$, $b = 0$, $c = -1$; $z_n = A^{-n}e^{if(n^\omega)}$; $\alpha, \omega > 0$ and $\alpha + \omega < 1$; $\varepsilon_n = cn^{(\omega-1)/\alpha}$. Suppose moreover that A is large enough and c small enough. Under these assumptions there is only one optimal traffic plan from $\mu = \delta_a$ to $\nu = (1/2-\varepsilon)\delta_b + (1/2-\varepsilon)\delta_c + \sum_{i=1}^{\infty}\varepsilon_i\delta_{z_i}$, and it is given by a single simple curve connecting a to z_1, $z_2, \dots, z_n, \dots, b$ and c by straight line segments. In particular, since the argument of z_n oscillates from $-\frac{\theta}{2}$ to $\frac{\theta}{2}$, there is no right hand side tangent at the point $b = 0$.*

3.2 Elliptic approximation *à la* Modica-Mortola

The result I presented in [11] and that I present again here is somehow inspired by, or at least recalls most of the results in the elliptic approximation of free discontinuity problems (Modica-Mortola, Ginzburg-Landau or Aviles-Giga, see for instance [31, 36, 37, 39, 40, 50, 51, 95]). We will only mention in details the following (from [95], see also [55]) because of its simplicity, even if it is probably not the closest one in this two-dimensional setting where Aviles-Giga seems closer. Moreover, I was recently informed of a study by Bouchitté, Seppecher and collaborators on the same subject, nased on the possibility of adapting a previous paper, [54], which looks at a one-dimensional problem but may be the starting point for a slicing procedure.

I start by recalling the main result by Modica and Mortola, which was presented as one of the first exemples of $\Gamma-$convergence ([72]).

Theorem (Modica-Mortola). *Define the functional F_ε on $L^1(\Omega)$ through*

$$F_\varepsilon(u) = \begin{cases} \frac{1}{\varepsilon}\int W(u(x))dx + \varepsilon\int |\nabla u(x)|^2 dx & \text{if } u \in H^1(\Omega); \\ +\infty & \text{otherwise.} \end{cases}$$

Then, if $W(0) = W(1) = 0$ and $W(t) > 0$ for any $t \neq 0, 1$, we have $\Gamma-$convergence of the functionals F_ε towards the functional F given by

$$F(u) = \begin{cases} cPer(S) & \text{if } u = 1 \text{ on } S,\ u = 0 \text{ on } S^c \text{ and } S \text{ is a finite-perimeter set;} \\ +\infty & \text{otherwise,} \end{cases}$$

where the constant c is given by $c = 2\int_0^1 \sqrt{W(t)}dt$.

In order to approximate the branched transport problem, we need to consider functionals of the form

$$M_\varepsilon^\alpha(u) = \varepsilon^{\gamma_1}\int_\Omega |u(x)|^\beta dx + \varepsilon^{\gamma_2}\int_\Omega |\nabla u(x)|^2 dx, \tag{3.7}$$

defined on $u \in H^1(\Omega; \mathbb{R}^2)$ and set to $+\infty$ outside $H^1 \subset \mathcal{M}(\Omega)$, for two exponents $\gamma_1 < 0 < \gamma_2$.

As one can see the functional recalls Modica-Mortola's functional to recover the perimeter as a limit, where the double-well potential is replaced with a concave power. Notice that concave powers, in their minimization, if the average value for u is fixed in a region (which is in some sense the meaning of weak convergence, i.e. the convergence we use on $\mathcal{M}(\Omega)$), prefer either $u = 0$ or $|u|$ being as large as possible, i.e. there is sort of a double well on zero and infinity.

In [11] I gave a heuristics for determining the exponents β, γ_1 and γ_2, in order to prove a $\Gamma-$convergence result in the space $\mathcal{M}(\Omega)$. The correct choice for a possible convergence result towards the energy (3.1) which is proportional to m^α is obtained by imposing that the exponent of m is α and the exponent of ε is zero, i.e.

$$\beta = \frac{2 - 2d + 2\alpha d}{3 - d + \alpha(d-1)}; \quad \frac{\gamma_1}{\gamma_2} = \frac{(d-1)(\alpha-1)}{3 - d + \alpha(d-1)}.$$

Notice that γ_1 and γ_2 may not both be determined since one can always replace ε with a power of ε, thus changing the single exponents but not their ratio. Notice also that the exponent β is positive and less than 1 as soon as $\alpha \in]1 - 1/d, 1[$, which is the usual condition.

The main result proved in [11] is

Theorem 3.5. *Suppose $d = 2$ and $\alpha \in]1/2, 1[$: then we have $\Gamma-$convergence of the functionals M_ε^α to cM^α, with respect to the convergence of $\mathcal{M}(\Omega)$, as $\varepsilon \to 0$, where c is a finite and positive constant (the value of c is actually $c = \alpha^{-1}\left(4c_0\alpha/(1-\alpha)\right)^{1-\alpha}$, being $c_0 = \int_0^1 \sqrt{t^\beta - t}dt$).*

Besides the interesting comparison aspects of this result with the similar ones in the approximation of free discontinuity problems, one of the main goal of this study concerned possible numerical applications, an in general approximation of the Xia minimization problem (3.2).

We would like to replace the problem of minimizing M^α under divergence constraints with a simpler problem, i.e. minimizing M_ε^α.

The idea would be solving

$$\min\left\{M_\varepsilon^\alpha(u)\ :\ \nabla\cdot u = f_\varepsilon\right\}, \tag{3.8}$$

being f_ε a suitable approximation of $f = \mu - \nu$., and proving that the minimizers of (3.8) converge to the minimizers of (3.2).

Theorem 3.5 proves a $\Gamma-$convergence result, which should give the convergence of the minimizers, but the problem is that we did not address the condition $\nabla\cdot u = f_\varepsilon$, nor we discussed the choice of f_ε.

What we can consider more easily are penalization methods, which are quite natural as far as numerics is concerned. One could decide to replace the (quite severe) condition $\nabla\cdot u = f$ at the limit with a weaker one, concerning a distance between $\nabla\cdot u$ and f. We can consider problems of the form

$$\min \quad M_\varepsilon^\alpha(u) + G(\nabla\cdot u; f), \tag{3.9}$$

where G is a functional defined on pairs of finite measures, which is continuous w.r.t. weak convergence, and such that $G = 0$ implies $\nabla\cdot u = f$. This would play the game, since $\Gamma-$convergence is stable when we add continuous perturbation.

Yet, this would converge to

$$\min \quad M^\alpha(u) + G(\nabla\cdot u; f), \tag{3.10}$$

which in general is not exactly the same as imposing $\nabla \cdot u = f$.

Actually, it is possible to obtain an equivalent problem if one properly chooses the penalization G. We start from the following consideration: as explained in [107], the quantity $d_\alpha(\mu,\nu) := \min\{M^\alpha(u) \ : \ \nabla\cdot u = \mu - \nu\}$ is a distance on the set of probability measures that metrizes weak convergence.

As a consequence of d_α being a distance, due to triangular inequality, solving

$$\min_{\rho^+,\rho^-\in\mathcal{P}(\overline{\Omega})} 2d_\alpha(\mu,\rho^+) + \min\{M^\alpha(u) \ : \ \nabla\cdot u = \rho^+ - \rho^-\} + 2d_\alpha(\nu,\rho^-)$$

amount to choosing $\mu = \rho^+$, $\nu = \rho^-$ and solving $\min\{M^\alpha(u) \ : \ \nabla\cdot u = \mu - \nu\}$.

This means that choosing $2d_\alpha$ as a penalization is a clever strategy (even if it only allows for considering probabilities or, in general, fixed mass measures).

To describe a precise problem where we can insert this penalization we need to introduce a slightly different space. Consider the space $Y(\Omega) \subset \mathcal{M}(\Omega) \times \mathcal{P}(\overline{\Omega}) \times \mathcal{P}(\overline{\Omega})$ defined by

$$Y(\Omega) := \{(u,\mu,\rho^-) \in \mathcal{M}(\Omega) \times \mathcal{P}(\overline{\Omega}) \times \mathcal{P}(\overline{\Omega}) \ : \ \nabla\cdot u = \mu - \rho^-\},$$

endowed with the obvious topology of componentwise weak convergence.

On this space consider the functionals

$$(u,\rho^+,\rho^-) \mapsto M_\varepsilon^\alpha(u) + 2d_\alpha(\mu,\rho^+) + 2d_\alpha(\nu,\rho^+). \tag{3.11}$$

In this way, as an easy consequence of our previous $\Gamma-$convergence result, one has obtained a useful approximation of any branched transport problem.

Nevertheless, the approximation is not satisfactory yet, especially for numerical purposes: we are trying to suggest methods to approximate the minimization of M^α and we propose to use the distance d_α itself!! For computing it one should probably solve a problem of the same kind and no progress would have been done.

A possible escape we suggest is replacing d_α with other quantities, which are larger but still vanish when the two measures coincide. This is possible for instance thanks to [21], where the inequality $d_\alpha \le CW_1^{2\alpha-1}$ is proven, W_1 being the usual Wasserstein distance on $\mathcal{P}(\overline{\Omega})$. In this way one can replace the d_α distance with a more convincing power of the W_1 distance.

It is anyway important to stress that the approach on $\mathcal{M}(\Omega)$ without penalization stays useful for a lot of problems where the divergence is not prescribed but enters the optimization (think at $\min_\mu d_\alpha(\mu,\nu) + F(\mu)$). Some of this problems are addressed in [23], for instance for urban planning or biological shape optimization.

The second issue we want to address, after the one concerning divergence constraints, deals with the convergence of the minimizers. $\Gamma-$convergence is quite useless if we cannot deduce that the minimizers u_ε converge, at least up to subsequences, to a minimizer u. Yet, this requires a little bit of compactness. The compactness we need is compactness in $\mathcal{M}(\Omega)$, i.e. we want bounds on the mass of $\nabla \cdot u_\varepsilon$ and of u_ε. The first bound, has been guaranteed by the fact that we decided to stick to the case of difference of probability measures. On the contrary, the bound on $|u_\varepsilon|(\Omega)$ has to be proven.

Notice that $M_\varepsilon^\alpha(u_\varepsilon) \leq C$ will not be sufficient for such a bound, as one can guess looking at the limit functional. Think of a finite graph with a circle of length l and mass m on it: its energy is $m^\alpha l$ which provides no bound on ml (its mass), if m is allowed to be large. Actually, what happens on the limit functional is "bounded energy configuration have not necessarily bounded mass, but optimal configuration do have". This is due to the fact that, if μ and ν are probabilities, then $m \leq 1$ on optimal configurations (and no cycle are possible, by the way). Notice that this statement does not depend on $m \mapsto m^\alpha$ being concave, but simply increasing in m. The same kind of behavior is likely to be true on the approximating problems, but after months of work a proof of this fact did not appear.

Hence I will conclude with two points: I will give a final - *naïve* - suggestion for getting a useful Γ-convergence result and I will list this question of the L^1 estimate on u_ε in the next section.

The suggestion is, as I mentioned, quite *naïve* : just take a sufficiently large number K so that every minimizer u of the limit problem satisfies $|u|(\overline{\Omega}) \leq K$ and then restrict the analysis to the compact subset $Y_K(\Omega) := \{(u, \mu, \nu) \in Y(\Omega) \ : \ |u|(\overline{\Omega}) \leq K\}$.

The problem one obtains is

$$\min \left\{ M_\varepsilon^\alpha(u) + CW_1^{2\alpha-1}(\mu, \rho^+) + CW_1^{2\alpha-1}(\nu, \rho^-) \ (u, \rho^+, \rho^-) \in Y_K(\Omega) \right\}.$$

It approximates (with $\Gamma-$convergence and convergence of the minimizers) the limit problem given by $\min_{Y_K(\Omega)} \ M^\alpha(u) + CW_1^{2\alpha-1}(\mu, \rho^+) + CW_1^{2\alpha-1}(\nu, \rho^-)$, which is equivalent to $\min\{M^\alpha(u) \ : \ \nabla \cdot u = \mu - \nu\}$.

3.3 Perspectives in numerics, modeling and game theory

3.3.1 Elliptic approximation

One of the main goal when I started discussing about the elliptic approximation of branched transport problems with J.-M. Morel was exploiting this techniques for numerical purposes. It happened that, as soon as I posted the paper on a preprint server, Edouard Oudet contacted me for working on the numerical side. He actually had already used some Modica-Mortola approach for problems involving the perimeter (like in Kelvin's conjecture, see [96]) and turning the codes to the branched transport was easy. I just show the first preliminary results

Since the first numerical attempts are very satisfactory (see Figure 3.2), there is a list of questions and tasks to be approached in the near future:

- give answers to the pending questions of the theoretical $\Gamma-$convergence result of [11]: imposing the divergence constraint and proving a bound on $||u_\varepsilon||_{L^1}$;
- generalize the same results to dimensions higher than 2: it seems that also Bouchitté and collaborators conjecture this result, but the proof is far from being achieved;
- generalizing similar results to the case $\alpha \leq 1 - 1/d$ and in particular to $\alpha = 0$, in order to retrieve the Steiner problem. This could probably be done, at least in dimension 2, by looking

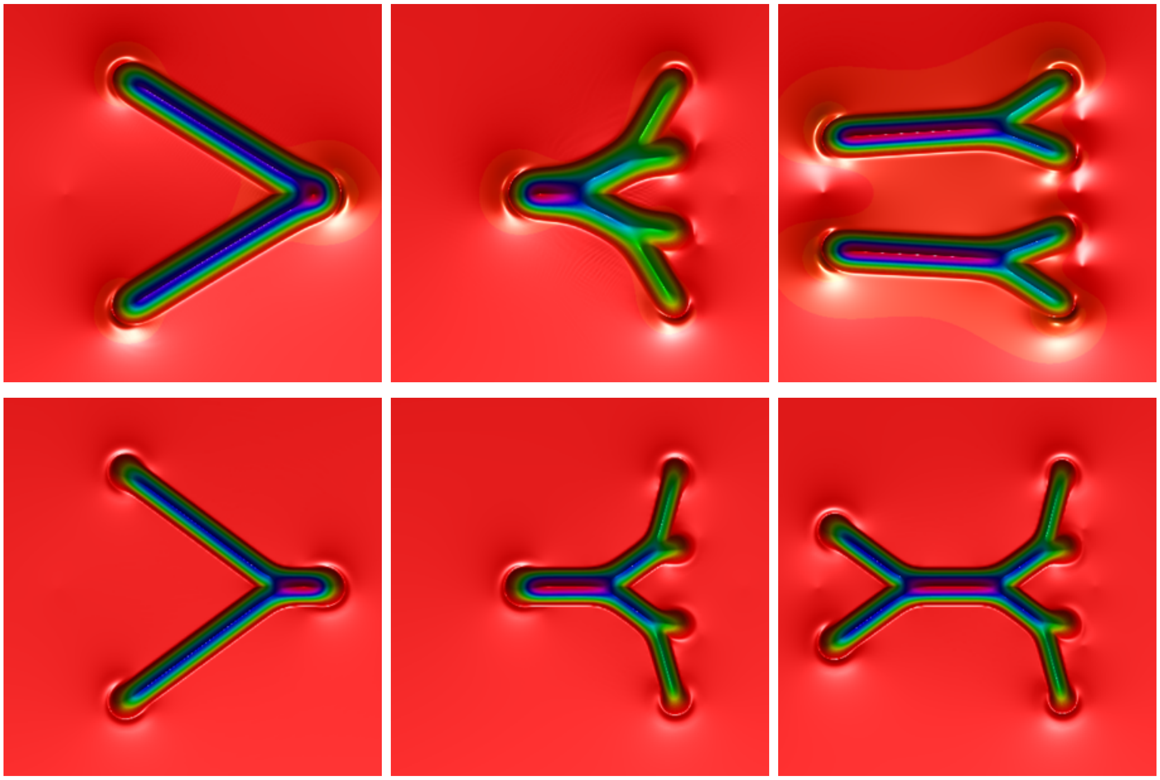


Figure 3.2: Three examples of branched transport: the parameter α equals 0.75 for the three configurations above and 0.6 below; the thickness of the branches exactly comes from the parameter ε which has been added.

at the techniques of [54] and it would have important applications in numerics; more generally, it would be interesting, as G. Alberti pointed out to me after a talk where I presented the result in Pisa, to look at which kind of lower-dimensional energies may be approximated through this approach;

- prove rigorously the convergence of the numerical method which is used.

3.3.2 Curves of measures

When the problem of branched transport arrived to the calculus of variations community, one attempt was done to look at it in terms of curves of measures rather than measures on curves. This meant minimizing a functional

$$\int_0^1 G_\alpha\big(\mu(t)\big)|\mu'|(t)dt$$

among possible curves μ connecting μ_0 and μ_1 in the space of probability measures endowed with the W_p Wasserstein distance (and the speed $|\mu'|(t)$ is to be intended as a metric derivative in such a space), with

$$G_\alpha(\mu) = \begin{cases} \sum_i a_i^\alpha & \text{if } \mu = \sum_i a_i \delta_{x_i}, \\ +\infty & \text{otherwise.} \end{cases}$$

In this way the problem becomes a geodesic problem in the Wasserstein space with a non-uniform metric. It has been proposed and studied in [15].

Yet, this model turned out not to be equivalent. There are two main problems in the model of [15]. One is in the way masses and length are mixed to compute the cost (say, one pays $(\sum m_i^\alpha)(\sum m_i l_i^p)^{1/p}$, instead of $\sum m_i^\alpha l_i$), so that the result is not at all what is expected in Branched Transport Theory. The other is in the fact that the model actually let the still mass pay as well (i.e. masses that have already reached their final destination are taken into account in G_α). The first is not that difficult to solve, if one accepts to work in the (less investigated) space W_∞, i.e. the space of probability measures with metric

$$W_\infty(\mu,\nu) = \min\{\gamma{-}ess\sup |x-y| \,:\, \gamma \in \Pi(\mu,\nu)\} = \min\big\{\max\{|x-y|,\, (x,y)\in \operatorname{supp}(\gamma) \,:\, \gamma\in\Pi(\mu,\nu)\big\}$$

(see section 1.3 where some L^∞ problems in optimal transport appear).

In this case one can see that it arrives to a situation which is very similar to the formulation by Bernot, Morel and Caselles, up to the fact that particles which have already reached their destination are still counted in the moving measure μ_t.

A solution for the case $\mu_0 = \delta_0$ (this case is always simpler and has lots of extra features) is in study in a joint work with L. Brasco.

The idea is considering pair of measures (μ_t, ν_t) with $\nu_t \le \mu_t \in \mathcal{P}(\Omega)$ and $\mu_t - \nu_t$ increasing. This reproduces the fact that ν_t is the moving mass and hence the functional we consider is

$$J(\mu,\nu) = \int_0^1 G_\alpha\big(\nu(t)\big)|\mu'|(t)dt.$$

After several reductions of the problem one can find the correct formulation for replacing measures n curves with curves of measures.

It is a work in progress whose interest lies both in the fact that it gives a better understanding of the interactions between the models and in the intermediate lemmas one has to prove, which ask for a better comprehension of the curves in the space W_∞.

3.3.3 Cooperative games

One question that arose with J.-M. Morel after the completion of [23] was the following: the landscape function z, which is roughly defined by associating to every point x the integral of $\theta^{\alpha-1}$ along the fiber of the network from the origin (if $\mu = \delta_0$) to x, is a way of partitioning the total cost of the optimal network among the agents, which are located at every point of Ω according to the density ν. Not only, it is maybe the most reasonable repartition, where every common segment is paid by all its user proportionally to the actual use they make of it. Is it a stable price repartition, according to the *core* theory for *cooperative games*? This means, does every subset of agents accept this repartition, or is there a subset who prefers to build its own optimal network to the point 0, thus paying a lower price? Notice that, in real life, problems of this kind often appear: if two or more people need to share a taxi, so as to save money in total with respect to different taxis, how should they pay the total fare?

It happens that also in discrete and simple cases this is not true and that there also exist cases where the core is empty (i.e., no cost allocation is possible). I supervised a short internship of a student in Dauphine, Lambert Piozin, who looked at this question, finding interesting examples (see [97]), adapting a paper by N. Megiddo, [94], who looked at the Steiner case. Yet, the question is open in the continuous case since it seems like the discrete and symmetric structure of these examples played an important role and several questions may be of interest in the discrete case as well.

This cooperative counterpart of the equilibrium problems on traffic network that appeared in the congestion part of my researches (Chapter 2) sounds very interesting to me and I would like to spend more time on these and on other more economics-oriented applications of the optimal transport theory and its variants. As I said in the introduction, approaching these problems has been one of the interesting points of my years in Dauphine, but I needed some time to start having a better view on them.

Chapter 4

Pressures and transport evolutions with density constraints

Résumé. Ce chapitre est le seul qui traite des modèles d'évolution. On en approche deux: le modèle variationnel pour les fluides incompressibles proposé par Y. Brenier et un modèle macroscopique de mouvement de foules introduit par B. Maury. Dans les deux cas, des contraintes sur les densités induisent une dynamique influencée par les effets d'une pression.

Le modèle de Brenier peut se décrire rapidement comme la minimisation de l'énergie cinétique moyenne d'une mesure sur l'espace des courbes, sous la contrainte que la marge au temps t de cette mesure soit toujours la mesure de Lebesgue (incompressibilité) et que le déplacement entre $t = 0$ et $t = 1$ soit prescrit. Sur ce problème, on donne des nouveaux résultats en dimension $d = 1, 2$, notamment un résultat d'unicité pour $d = 1$ et une caractérisation des solutions dans un cas précis (le déplacement de x en $-x$ dans un disc) pour $d = 2$ qui a fait apparaître des nouvelles solutions inconnues à Brenier.

Le modèle de mouvement des foules est présenté sous la forme d'un flot-gradient dans l'espace de Wasserstein. La fonctionnelle est la distance moyenne à la porte, plus la contrainte que la densité ne dépasse pas un certain seuil. Comme équation, cela donne lieu à l'équation de continuité où la vitesse est la projection de la vitesse souhaitée (l'opposée du gradient de la distance à la porte) sur l'ensemble des vitesses admissibles (celles qui ont divergence positive sur l'ensemble où la densité sature la contrainte).

In this chapter I want to present my researches on some transport-related models for evolution PDE mainly linked to fluid mechanics, which have this point in common: there are particles in evolution and their density must satisfy some incompressibility constraints. These constraints give rise to a pressure field affecting the dynamics.

There is an analogy with the theory presented in Chapter 2, but in congested traffic there is not a true evolution (time only plays a fictitious role, as if the flow was cyclical and continuous) and the density is not constrained but penalized.

Two joint works are presented, the first on incompressible Euler Equation and the second on crowd motion. The first looks at a purely incompressible situation, so that the density ρ is constantly equal to a uniform density at every time. It is natural that in such a case the model cannot be only eulerian, but that it is more interesting to look at the motion of every single particle, through a measure Q on the space of curves, which is a common tool that is present throughout the researches that I presented in almost all of this mémoire. On the contrary, in the second there is only a one-sided constraint $\rho \leq 1$, which leads to the study of a curve of measures. Moreover, this second model will be seen as a gradient flow in the space of measures where the initial configuration is given and then the density evolves, so that it is a first-order model with intial data only. On the contrary, in the first what we prescribe is somehow the initial *and final* configurations, and the velocity is an unknown too, satisfying its own PDE, so that it is actually a second-order model.

4.1 On the Brenier variational formulation for the Euler Equation of incompressible fluids

4.1.1 The Euler Equation and its variational formulation

Classically, the velocity field of an incompressible fluid moving inside a smooth domain $\Omega \subset \mathbb{R}^d$ is represented by a time-dependent and divergence-free vector field $\boldsymbol{u}(t,x)$ which is parallel to the boundary $\partial\Omega$. The Euler equations for incompressible fluids describing the evolution of such a velocity field $\boldsymbol{u}$ and of the pressure field p are

$$\begin{cases} \partial_t \boldsymbol{u} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} = -\nabla p & \text{in } [0,T]\times\Omega, \\ \nabla\cdot\boldsymbol{u} = 0 & \text{in } [0,T]\times\Omega, \\ \boldsymbol{u}\cdot\hat{n} = 0 & \text{on } [0,T]\times\partial\Omega. \end{cases} \tag{4.1}$$

If we assume that $\boldsymbol{u}$ is smooth, the trajectory of a particle initially at position x is obtained by solving

$$\begin{cases} \dot g(t,x) = \boldsymbol{u}(t,g(t,x)), \\ g(0,x) = x. \end{cases}$$

Since $\boldsymbol{u}$ is divergence free, for each time t the map $g(t,\cdot):\Omega\to\Omega$ is a measure-preserving diffeomorphism of Ω (we will say that $g(t,\cdot)\in \mathrm{SDiff}(\Omega)$), which means

$$g(t,\cdot)_\#\mathcal{L}^d_{\lfloor\Omega} = \mathcal{L}^d_{\lfloor\Omega}$$

(here and in the sequel $f_\#\mu$ is the push-forward of a measure μ through a map f, and $\mathcal{L}^d_{\lfloor\Omega}$ is the Lebesgue measure inside Ω). Writing Euler equations in terms of g, we get

$$\begin{cases} \ddot{g}(t,x) = -\nabla p\,(t, g(t,x)) & \text{in } [0,T]\times\Omega, \\ g(0,x) = x & \text{in } \Omega, \\ g(t,\cdot) \in \mathrm{SDiff}(D) & \text{for } t\in[0,T]. \end{cases} \tag{4.2}$$

In [38], Arnold interpreted the equation above, and therefore (4.1), as a *geodesic* equation on the space $\mathrm{SDiff}(\Omega)$, viewed as an infinite-dimensional manifold with the metric inherited from the embedding in $L^2(\Omega)$ and with tangent space corresponding to the divergence-free vector fields. According to this interpretation, one can look for solutions of (4.2) by minimizing

$$\int_0^T \int_\Omega \frac{1}{2}|\dot{g}(t,x)|^2\,d\mathcal{L}^d_{\lfloor\Omega}(x)\,dt \tag{4.3}$$

among all paths $g(t,\cdot) : [0,T] \to \mathrm{SDiff}(\Omega)$ with $g(0,\cdot) = f$ and $g(T,\cdot) = h$ prescribed (typically, by right invariance, f is taken as the identity map $\boldsymbol{i}$). In this way, the pressure field arises as a Lagrange multiplier from the incompressibility constraint.

Shnirelman proved in [100, 101] that when $d \geq 3$ the infimum is not attained in general, and that when $d = 2$ there exists $h \in \mathrm{SDiff}(\Omega)$ which cannot be connected to $\boldsymbol{i}$ by a path with finite action. These "negative" results motivate the study of relaxed versions of Arnold's problem.

The first relaxed version of Arnold's minimization problem was introduced by Brenier in [56]: he considered probability measures Q in C, the space of continuous paths $\omega : [0,T] \to \Omega$, and solved the variational problem

$$\text{minimize} \qquad \mathcal{A}_T(Q) := \int_C \int_0^T \frac{1}{2}|\dot{\omega}(\tau)|^2\,d\tau\,dQ(\omega), \tag{4.4}$$

with the constraints

$$(\pi_0, \pi_T)_\# Q = (\boldsymbol{i}, h)_\# \mathcal{L}^d_{\lfloor\Omega}, \qquad (\pi_t)_\# Q = \mathcal{L}^d_{\lfloor\Omega} \qquad \forall\, t \in [0,T] \tag{4.5}$$

(here and in the sequel $\pi_t(\omega) := \omega(t)$ are the evaluation maps at time t). According to Brenier, we shall call these Q *generalized incompressible flows* in $[0,T]$ between $\boldsymbol{i}$ and h. The existence of a minimizing Q is a consequence of the coercivity and lower semicontinuity of the action, provided that there exists at least a generalized flow Q with finite action (see [56]). This is the case for instance if $\Omega = [0,1]^d$ or if it is a bilipschitz image of the cube (this follows from the results in [56, 59]).

Here as well the pressure p will appear as a Lagrange multiplier and it will happen that Q is optimal if and only if $(\pi_t)_\# Q = \mathcal{L}^d_{\lfloor\Omega}$ and Q is concentrated on curves which are minimizing the action

$$\omega \mapsto \int_0^T \big(\frac{1}{2}|\dot{\omega}(\tau)|^2 - p(\omega(\tau))\big)d\tau \tag{4.6}$$

among curves with the same starting and arrival points.

4.1.2 Uniqueness and multiplicity of solutions in dimension 1 and 2

In [12], we considered Problem (4.4)-(4.5) in the particular cases where Ω is either the ball $B_1(0)$ or an annulus $B_R(0) \setminus B_r(0)$, in dimension 1 and 2. We were mainly concerned with uniqueness and characterization issues, as existence always holds in these cases.

If $\Omega = B_1(0) \subset \mathbb{R}^2$ is the unit ball, the following situation arises: an explicit solution of Euler equations is given by the transformation

$$g(t,x) = R_t x,$$

where $R_t : \mathbb{R}^2 \to \mathbb{R}^2$ denotes the counterclockwise rotation of an angle t. Indeed the maps $g(t,\cdot) : D \to D$ are clearly measure preserving, and moreover we have

$$\ddot{g}(t,x) = -g(t,x),$$

so that $\boldsymbol{u}(t,x) = \dot{g}(t,y)|_{y=g^{-1}(t,x)}$ is a solution to the Euler equations with the pressure field given by $p(x) = \frac{1}{2}|x|^2$ (so that $\nabla p(x) = x$). It is not difficult to prove that the generalized incompressible flow induced by g is optimal if $T \le \pi$, and is the unique one if $T < \pi$ (since the corresponding curves optimize the functional in (4.6)).

This implies in particular that there exists a unique minimizing geodesic from $\boldsymbol{i}$ to the rotation R_T if $0 < T < \pi$. On the contrary, for $T = \pi$ more than one optimal solution exists, as both the clockwise and the counterclockwise rotation of an angle π are optimal. Moreover, Brenier found in [56, Section 6] an example of action-minimizing path Q connecting $\boldsymbol{i}$ to $-\boldsymbol{i}$ in time π which is not induced by a classical solution of the Euler equations (and it cannot be simply constructed using the two opposite rotations). The idea of such a solution is that from each point x there is a multiplicity of curves going to $-x$ solving $\omega''(t) = -\omega(t)$ but with different initial velocity. It happens that, if the distribution of initial velocities at x is uniform among the vectors of modulus equal to $\sqrt{1-|x|^2}$, then the corresponding flow is incompressible and optimal.

A similar construction may be done in dimension one, where there is no possible rotation and the only known solution was obtained in a similar way: the pressure is still of the form $p(x) = \frac{1}{2}x^2$, every point at x splits into different trajectories with initial velocity given by $\sqrt{1-x^2}\cos\theta$, with a uniform distribution on $\theta \in [0, 2\pi]$. All of these trajectories arrive at $-x$ at time $T = \pi$ and the corresponding flow is incompressible and optimal.

By the way, the one-dimensional case is a bit particular since, if $\Omega = [-1,1]$, the space of measure-preserving diffeomorphisms consist of $\{\boldsymbol{i}, -\boldsymbol{i}\}$, and so Arnold's problem is trivial (there are only two continuous curves belonging to $\mathrm{SDiff}([-1,1])$). However the relaxed problem was non-trivial, and the solution mentioned above was found by Brenier in [56].

In [12] we proved that

Theorem 4.1. *If $\Omega = [-1,1] \subset \mathbb{R}$, Problem* (4.4)-(4.5) *for $h = -\boldsymbol{i}$ has a unique minimizer.*

and then we strengthened the result:

Theorem 4.2. *Let $\Omega = [-1,1]$ and suppose that the pressure p is of class C^∞. Then there exists a unique minimizer Q to Problem* (4.4)-(4.5).

In dimension two, on the other hand, we found a whole family of solutions which was not known to Brenier, mainly based on the following lemma:

Theorem 4.3. *If $\Omega = B(0,R) \setminus B(0,r)$ is an annulus, the rotations are still solutions and the pressure is still $p(x) = \frac{1}{2}|x|^2$. Moreover, there also exists an incompressible Q which is constructed in the following way: from every point x, several trajectories stem with initial velocities v which are distributed on a circle $\{v : |v| = C(R, r, |x|)\}$, with radius of the circle and densities on it depending on $|x|, R$ and r (in the case $r = 0$ the uniform distribution on the whole circle works). Moreover, this solution is unique if one adds the assumption that all the curves are "clockwise" (i.e. the sign of $\omega'(t) \wedge \omega(t)$ is prescribed).*

This allows to find lots of different solutions, for instance separating the disc into several annuli and using one such solution in each annulus. Another characterization result is the following, concerning stationary solutions (i.e. solutions such that the image of Q through the map (π_t, v_t) does not depend on t, where $v_t(\omega) = \omega'(t)$).

Theorem 4.4. *If $\Omega \subset \mathbb{R}^2$ is either the disc or the annulus, then there exists only one stationary clockwise minimizer, and it is rotationally invariant.*

4.2 Crowd motion as a Gradient Flow in W_2

This model, opposite to the previous one, is expressed in terms of evolving densities. It looks at the evolution of a mass of particles (say, people in a room) who move according to their desired velocity (say, the maximal speed they can afford in the direction of the door, as if a fire had burst in the room), but they are forced to adapt their movement due to the presence of other particles, so that their density does not exceed a fixed value. It is a quite simple model for crowd motion in emergency evacuation situations : the behaviour of individuals is based on optimizing their very own trajectory, regardless of others, but the fulfillment of individual strategies is made impossible because of congestion. It has been studied in a *microscopic* framework, where the agents are represented by incompressible disks and the constraint is that they can not overlap, by B. Maury and J. Venel in [91, 92, 102, 103]. Here we present a *macroscopic* framework, where the agents are represented through their density ρ and the constraint is $\rho \leq 1$. It is the subject of a joint work with Bertrand Maury and Aude Roudneff-Chupin and the main topic of the PhD thesis (in progress) of Aude herself.

4.2.1 Eulerian model

The model takes the following form: given a domain Ω (the building), whose boundary Γ is composed of Γ_{out} (the exit) and Γ_w (the walls), we describe the current distribution of people by a measure ρ of given mass (say 1 without loss of generality) supported within $\overline{\Omega}$. To model the fact that people getting through the door are out of danger, yet keeping a constant total mass without having to model the exterior of the building, we shall assume that ρ may concentrate on Γ_{out}. The idea is that people are safe once they have exited Ω, and that we do not take into account what they do once they have passed the door. This is why we let the mass stay on the door Γ_{out} itself and we do not care about its concentration on Γ_{out}.

In this spirit, we denote by K the set of all those probability measures over $\mathbb{R}^d$ that are supported in $\overline{\Omega}$, and that are the sum of a diffuse part in Ω, with density between 0 and 1, and a singular part carried by Γ_{out}.

We shall denote by $\mathcal{U}$ the spontaneous velocity field: $\mathcal{U}(x)$ represents the velocity that an individual at x would have if he were alone. It is taken equal to 0 outside Ω. The set C_ρ of feasible velocities corresponds to all those fields which do not increase ρ on the saturated zone $\{\rho = 1\}$, and which account for walls (people do not walk through them). Unformally, $\nabla \cdot \boldsymbol{u} \geq 0$ in $\{\rho = 1\}$. As we plan to define C_ρ as a closed convex set in $L^2(\Omega)$, those constraints do not make sense as they are, and we shall favor a dual definition of this set. Let us introduce the "pressure" space

$$H^1_\rho = \big\{q \in H^1(\Omega)\,,\ q \geq 0 \text{ a.e. in } \Omega\,,\ q(x) = 0 \text{ a.e. on } \{\rho < 1\}\,,\ q_{|\Gamma_{out}} = 0\big\}.$$

The proper definition of C_ρ reads

$$C_\rho = \{\boldsymbol{v} \in L^2(\Omega; \mathbb{R}^d)\,,\ \int_\Omega \boldsymbol{v} \cdot \nabla q \leq 0 \quad \forall q \in H^1_\rho\} \tag{4.7}$$

The model is based on the assumption that the actual instantaneous velocity field is the feasible field which is the closest to $\mathcal{U}$ in the least-square sense, i.e. it is defined as the L^2-projection of $\mathcal{U}$ onto the closed convex cone C_ρ. Finally the problem consists in finding a trajectory $t \mapsto \rho(t) \in K$ which is advected by $\boldsymbol{u}$, i.e. such that $(\rho, \boldsymbol{u})$ is a (weak) solution of the transport equation in $\mathbb{R}^d$

$$\partial_t \rho + \nabla \cdot (\rho \boldsymbol{u}) = 0, \tag{4.8}$$

where $\boldsymbol{u}$ verifies, for almost every t,

$$\boldsymbol{u} = P_{C_\rho} \mathcal{U}. \tag{4.9}$$

4.2.2 Gradient flow formulation

We want to see when the problem of finding a solution to (4.8) and (4.9) may become a gradient flow of a suitable functional in a suitable space. This will only be adapted to the case where $\mathcal{U}$ is a gradient field. In this case the gradient flow formulation was already available for the microscopic setting. Here we will look for a gradient flow formulation in the Wasserstein space $\mathcal{P}_2$ (the set of probability measures over $\mathbb{R}^d$ endowed with the Wasserstein distance W_2, see [32] and [35] for the whole theory on this topic).

We denote by K the set of feasible densities

$$K = \{\rho \in \mathcal{P}_2\,,\ \operatorname{supp}(\rho) \subset \overline{\Omega}\,,\ \rho = \rho_{out} + \rho_\Omega\,,\ \rho_\Omega(x) \leq 1 \ \text{ a.e. }\,,\ \operatorname{supp}(\rho_{out}) \subset \Gamma_{out}\}. \tag{4.10}$$

Let an initial density ρ^0 be given, and $\tau > 0$ a time step. We build $\rho^0_\tau = \rho^0$, ρ_1, ... as follows

$$\rho^{k+1}_\tau \in \underset{\mathcal{P}_2(\mathbb{R}^d)}{\operatorname{argmin}} \left\{ J(\rho) + \mathbf{I}_K + \frac{1}{2\tau} W_2^2(\rho, \rho^k_\tau) \right\} \tag{4.11}$$

where W_2 is the Wasserstein distance, and J is the dissatisfaction functional defined as :

$$J(\rho) := \int_\Omega D(x)\rho(x)\, dx. \tag{4.12}$$

The function D is typically the distance to the door Γ_{out}, and to D we associate the vector field $\mathcal{U} = -\nabla D$. It is important in order to have vanishing velocities on the door that D is minimal and constant on Γ_{out}.

We admit here that under reasonable assumptions this process is indeed an algorithm (i.e. we can always find a measure ρ_τ^{k+1} and that, if possible, it is uniquely defined as the minimizer of the functionnal above), and we denote by ρ_τ the piecewise constant interpolate of ρ_τ^0, ρ_1^τ, ...

It is quite easy to see, thanks to general methods in [27], that ρ_τ actually converges, up to subsequences, to an absolutely continuous curve ρ, but the goal is to identify such a curve, together with its velocity field, and to prove that it is the curve we are searching.

The general theory of Gradient flows in Wasserstein spaces would say that, as τ goes to 0, ρ_τ converges to some trajectory $t \mapsto \rho$ in K, which is a (weak) solution to

$$\partial_t \rho + \nabla \cdot (\rho \boldsymbol{u}) = 0,$$

where $\boldsymbol{u}$ is such that, for almost every t,

$$\boldsymbol{u} \in -\partial\left(J + I_K\right)$$

where $\partial\Psi$ denotes the strong subdifferential of Ψ, defined in [32, 35]. Furthermore $\boldsymbol{u}$ minimizes the L^2 norm among all those fields in the subdifferential above.

We can prove unformally that this characterizes the instantaneous velocity as the projection of $\mathcal{U} = -\nabla D$ onto C_ρ. This subdifferential of a function Ψ at ρ in the Wasserstein setting is defined as the set of fields $\boldsymbol{u}$ such that,

$$\Psi(\rho) + \int_\Omega \langle \boldsymbol{u}, \mathbf{t}(x) - x\rangle\, d\rho(x) \leq \Psi(\mathbf{t} \# \rho) + o(||\mathbf{t} - \mathbf{i}||)$$

where $\mathbf{t}$ denotes a transport map acting on ρ. Note that the previous inequality does not provide any information as soon as $\mathbf{t} \# \rho$ is not feasible (in that case the right-hand side is $+\infty$). Let us consider a feasible field $\boldsymbol{v} \in C_\rho$, and let us assume that, for ε small enough, $\mathbf{t}_\varepsilon = \mathbf{i} + \varepsilon \boldsymbol{v}$ pushes forward ρ onto a measure in K. Notice that $\mathbf{t}_\varepsilon$ is defined ρ-almost everywhere, with Γ_{out} carrying some mass, but as it vanishes on Γ_{out}, the singular part of ρ remains unchanged. Letting ε go to 0 in the subdifferential inequality, we obtain

$$\int \nabla D \cdot \boldsymbol{v}\, d\rho(x) + \int \boldsymbol{u} \cdot \boldsymbol{v}\, d\rho(x) \leq 0,$$

so that $\boldsymbol{u} + \nabla D = \boldsymbol{u} - \mathcal{U}$ belongs to C_ρ°, the polar cone to C_ρ. As $\boldsymbol{u}$ minimizes the L^2 norm over $\mathcal{U} + C_\rho$, $\boldsymbol{u}$ identifies with the projection of $\mathcal{U}$ onto C_ρ, which ends this formal proof.

Yet, there are at least two problems in order to apply the general theory of [32]. The first may be seen in the following remark, and it also refers to the simpler case where $\Gamma_{out} = \emptyset$ (i.e. where no concentration on the boundary is allowed, even if D is still the distance to a subset of the boundary):

Remark 4.1. In general, there exist feasible densities $\rho \in K$ (defined by (4.10)) and fields $\boldsymbol{v} \in C_\rho$ (defined by (4.7)) such that $(\mathbf{i}+\varepsilon\boldsymbol{v})\,\#\,\rho$ exits K for any $\varepsilon > 0$, this is why the considerations above do not make a rigorous proof. Consider for example ω a dense open subset in Ω, with a small measure, and define ρ as $\mathbf{1}_{\omega^c}$. The pressure space is $\{0\}$, and C_ρ is $L^2(\Omega)$: any field is feasible. If one considers now a strictly contractant field (with negative divergence), it is clear that $(\mathbf{i}+\varepsilon\boldsymbol{v})\,\#\,\rho \notin K$ for any $\varepsilon > 0$. By the way, the same problem would occur if instead of $(\mathbf{i}+\varepsilon\boldsymbol{v})$ one would follow for a time ε the flow of the vector field v (which is more natural and usually solves some of the problems linked to the difference between "infinitesimally admissible" velocity and "truly admissible", i.e. the difference between the Jacobian and its first-order development).

The second problem is peculiar to the case where there is an exit door Γ_{out} where concentration is admitted: in this case the functional $J + I_K$ is not geodesically convex! This depends on the fact that K is not a geodesically convex set (see next section). This rules out the possibility of using a certain part of the general theory, even if the previous problem could be solved, at least in the most interesting case, which is the one with an exit.

We now go back to the recursive minimization scheme given in (4.11).

We define on $\mathring{\Omega}$ the discrete velocities : $\mathbf{v}_\tau^{\mathbf{k}} = \dfrac{\mathbf{i} - \mathbf{t}_\tau^{\mathbf{k}}}{\tau}$, where $\mathbf{t}_\tau^{\mathbf{k}}$ is the unique optimal transport function from ρ_τ^k to ρ_τ^{k-1}, which is well defined on $\mathring{\Omega}$ (but not necessarily on Γ_{out}, due to the singular part of ρ_τ^k). We also define $\mathbf{E}_\tau^{\mathbf{k}} = \rho_\tau^k \mathbf{v}_\tau^{\mathbf{k}}$ on $\mathring{\Omega}$. We can interpolate these discrete values $(\rho_\tau^k, \mathbf{v}_\tau^{\mathbf{k}}, \mathbf{E}_\tau^{\mathbf{k}})_{k\geq 0}$ by the piecewise constant functions defined by :

$$\left\{\begin{array}{l} \rho_\tau(t,.) = \rho_\tau^k \\ \mathbf{v}_\tau(t,.) = \mathbf{v}_\tau^{\mathbf{k}} \qquad \text{if } t \in \,](k-1)\tau, k\tau] \\ \mathbf{E}_\tau(t,.) = \mathbf{E}_\tau^{\mathbf{k}} \end{array}\right. \tag{4.13}$$

Our goal is to prove that ρ_τ converges when $\tau \to 0$ to a solution of the continuity equation (4.8)-(4.9). Here is our main result :

Theorem 4.5. *Let Ω be a convex bounded set of $\mathbb{R}^d$, $D : Ov \to \mathbb{R}$ a continuous function with $D_{|\Gamma_{out}} = \min D$, ρ^0 a probability density, and (ρ_τ^k) constructed following the recursive scheme (4.11). Then there exists a family of probability densities $(\rho(t,.))_{t>0}$, and a family of velocities $(\mathbf{u}(t,.))_{t>0}$ such that $(\rho_\tau(t,.), \mathbf{E}_\tau(t,.))$ converge up to subsequences - to $(\rho(t,.), \rho(t,.)\,\mathbf{u}(t,.))$ for a.e. t. Moreover, $(\rho, \mathbf{u})$ satisfies the continuity equation :*

$$\left\{\begin{array}{rcl} \partial_t \rho + \nabla.(\rho \mathbf{u}) & = & 0 \\ \mathbf{u}(t,.) & = & P_{\mathcal{C}_{\rho(t,.)}} \mathbf{U} \quad \textit{for a.e. } t \\ \rho(0,.) & = & \rho^0 \end{array}\right. \tag{4.14}$$

where $\mathbf{U} = -\nabla D$, and $\mathcal{C}_{\rho(t,.)}$ is defined in (4.7).

Unfortunately, the theorem above only addresses existence matters, and uniqueness and stability are not solved yet. This is quite typical of gradient flow problems for non-$\lambda-$convex functionals. Actually, stability estimates (which imply uniqueness) seem easy to get if D is $\lambda-$convex and $\Gamma_{out} = \emptyset$

but are completely open when the exit door Γ_{out} is considered. The same difficulties we find in approaching these questions have been found by the authors of [82] and they seem to be interesting matters of investigation for the future.

4.3 Useful tools

After some discussions with experts of the field of gradient flows in Wasserstein spaces, I thought that it would have been useful to stress the peculiar strategy we used in [13] to obtain optimality conditions for the discrete problems 4.11. The main particular feature of the strategy is the fact that we used "vertical" perturbations $\rho_\varepsilon = (1-\varepsilon)\rho_\tau^k + \varepsilon\rho$ rather than "horizontal" perturbations $(id + \varepsilon T)_\# \rho_\tau^k$. I summarize the result in the following lemma and I give an idea of the proof:

Lemma 4.6. *If ρ_τ^k, v_τ^k and H^1_ρ are defined as in previous session, then one has the following decomposition of the spontaneous velocity $\mathcal{U} = -\nabla D$*

$$\mathcal{U} = \mathbf{v}_\tau^{\mathbf{k}} + \nabla p_\tau^k \quad \textit{with}\; p_\tau^k \in H^1_{\rho_\tau^k}, \tag{4.15}$$

which is the correct equation to be passed to the limi tas $\tau \to 0$ so as to obtain that v is the projection of $\mathcal{U}$.

To see how one could prove this statement (see [13] for details), we start from the optimality conditions on the minimality ρ_τ^k, under perturbations like $\rho_\varepsilon = (1-\varepsilon)\rho_\tau^k + \varepsilon\rho$, for $\rho \in K$. This method is less used in transport theory than the perturbations by transports which are close to the identity, but it has been widely used in [18, 22, 19, 1]. From this kind of variations one can deduce

$$\int_\Omega \left(D + \frac{\bar{\varphi}}{\tau}\right)\rho \;\geq\; \int_\Omega \left(D + \frac{\bar{\varphi}}{\tau}\right)\rho_\tau^k \quad \text{for all } \rho \in K, \tag{4.16}$$

where $\bar{\varphi}$ is a Kantorovitch potential for the quadratic transport from ρ_τ^k to ρ_τ^{k-1}. This essentially comes from the fact that the derivative (or subgradient) of the functional $\frac{1}{2}W_2^2$ is given by the Kantorovitch potential itself.

Equation (4.16) means that ρ_τ^k optimizes the integral of a certain given function f (which is in this case $D + \bar{\varphi}/\tau$) among densities in K. In the easiest case where we do not consider the door Γ_{out} (i.e. K is the set of all densities smaller than 1, but the other case is not much more difficult) the solution of this problem is easy to detect: the optimal measure must be concentrated on a sublevel $\{f < t\}$, and t has to be chosen so that this sublevel has measure 1. More precisely,

$$\begin{cases} \rho_\tau^k = 1 & \text{on} \quad \{f < l\} \\ \rho_\tau^k \leq 1 & \text{on} \quad \{f = l\} \\ \rho_\tau^k = 0 & \text{on} \quad \{f > l\} \end{cases}$$

We can then define a pressure like function :

$$p_\tau^k(x) := (l - F(x))_+ = \left(l - D(x) - \frac{\bar{\varphi}(x)}{\tau}\right)_+ \tag{4.17}$$

which satisfies : $p_\tau^k \geq 0$, and $p_\tau^k = 0$ on $\{\rho_\tau^k < 1\}$. Then, it is not difficult to check, thanks to the relation between the Kantorovitch potential and the displacement (which gives $v_\tau^k = \nabla\bar{\varphi}/\tau$) that we the desired decomposition for the spontaneous velocity holds true ρ_τ^k-a.e.

After the previous part which was mainly devoted to techniques, I want to go on in this section with an easy but useful inequality in Wasserstein spaces that appeared in the preparation of [13]. This inequality leads, for its generalization to the case where K takes into account the existence of a door Γ_{out}, to some problem involving incompressibility and measure-preserving bilipschitz homeomorphisms.

We start from the following inequality, which is well known,

$$\int_\Omega f\, d(\mu-\nu) \leq ||\nabla f||_{L^\infty} W_1(\mu,\nu),$$

and it comes exactly from the dual formulation of the problem of Kantorovitch when the cost equals distance, which involves Lipschitz functions.

One could wonder whether the exponents 1 and ∞ on the right hand side could be replaced by any pair p and p'. In particular one could wonder whether he can estimate the integral on the left with the H^1 norm of f (or the L^2 norm of its gradient) and the W_2 distance between the two measures. This was necessary during the proofs of [13], so as to estimate some error terms involving the pressure, which was naturally an H^1 function in this framework.

It is easy to see that the inequality

$$\int_\Omega f\, d(\mu-\nu) \leq ||\nabla f||_{L^2} W_2(\mu,\nu)$$

cannot be true, since H^1 functions in dimension higher than one are not continuous nor bounded, and this result would give an estimate of their oscillation if one takes μ and ν to be Dirac masses.

Yet, is is not difficult to prove

Lemma 4.7. *Assume that μ and ν are absolutely continuous measures, whose densities are bounded by a same constant C. Then, for all function $f \in H^1(\Omega)$, we have the following inequality :*

$$\int_\Omega f\, d(\mu-\nu) \;\leq\; \sqrt{C}\, ||\nabla f||_{L^2(\Omega)} W_2(\mu,\nu)$$

The idea of the proof is simple: let μ_t be a geodesic for the distance W_2 connecting μ to ν, and v_t the corresponding velocity vector field: then one has

$$\int_\Omega f\, d(\mu-\nu) = \int_0^1 \left(\int_\Omega \nabla f\cdot v_t\, d\mu_t\right) dt \leq \int_0^1 ||\nabla f||_{L^2(\mu_t)} ||v_t||_{L^2(\mu_t)} dt.$$

Then one uses the fact that $||v_t||_{L^2(\mu_t)} = |\mu'|_{W_2}(t)$, so that $\int_0^1 ||v_t||_{L^2(\mu_t)} dt = W_2(\mu,\nu)$ and the fact that all the interpolating measures will satisfy the same bound on the density (by geodesic convexity of the set of densities bounded by the same constant C), so that $||\nabla f||_{L^2(\mu_t)} \leq \sqrt{C}||\nabla f||_{L^2}$.

Lemma 4.7 was useful for the - simpler - case where there was no exit Γ_{out}. For the general case, when $\mu,\nu \in K$, it is easy to prove the following:

Lemma 4.8. *For $\mu, \nu \in K$, we define the length*

$$L(\mu, \nu) = \inf \left\{ \int_0^1 |\sigma'|_{W_2}(t)dt \ : \ \sigma(t, .) \in K, \ \sigma(0) = \mu, \ \sigma(1) = \nu \right\}.$$

Then the following inequality holds for all functions $f \in H^1(\Omega)$ with $f = 0$ on Γ_{out}:

$$\int_\Omega f \, d(\mu - \nu) \ \leq \ ||\nabla f||_{L^2(\Omega)} L(\mu, \nu).$$

To be able to use Lamma 4.8 as we used Lemma 4.7 in the case with no exit, it is convenient to prove the following:

Lemma 4.9. *The minimal length L, defined in Lemma 4.8, is a distance on K which is continuous (and hence topologically equivalent) with respect to the weak convergence. More precisely we have*

$$L(\mu, \nu) \leq C(\Omega, \Gamma_{out}) W_2^{1/(4d)}(\mu, \nu).$$

It is clear that if K was geodesically convex this minimal length L would have been the same as the distance W_2 and the statement of the above Lemma would have been straightforward. The construction to prove the general case is quite involved and it is first performed, in [13], in the case where $\Omega =]0, 1[^d$ and Γ_{out} is one of the faces of the cube.

The generalization to more general domains is done thanks to the following result, which is useful in the Brenier's variational model for the incompressible Euler equation, in the estimates for traffic congestion of Chapter 2, and in the study of crowd motion of Section 4.2 in the case of an exit Γ_{out}. For this result, I refer to [70]

Theorem. *For any sufficiently good domain $\Omega \subset \mathbb{R}^d$ which is homeomorphic to the cube, there exists a bi-lipschitz homeomorphism $\phi : \overline{\Omega} \to [0, 1]^d$ such that $\phi_\#(\mathcal{L}^d_{|\Omega}) = c\mathcal{L}^d_{|Q}$ and $\phi_{|\partial\Omega} = \phi_0$, where ϕ_0 is an arbitrary bi-lipschitz diffeomorphism of $\partial\Omega$ onto $\partial[0, 1]^d$.*

I also want to present here a more-precisely-stated result in dimension two, which is actually quite easy to prove by hand.

Lemma 4.10. *For any star-shaped domain $\Omega \subset \mathbb{R}^2$ with Lipschitz boundary and any closed connected subset $\Gamma \subset \partial\Omega$ which is neither a point nor the whole boundary, there exists a bi-lipschitz homeomorphism $\phi : \overline{\Omega} \to \overline{Q}$ such that $\phi_\#(\mathcal{L}^2_{|\Omega}) = c\mathcal{L}^2_{|Q}$ and $\phi(\Gamma) = S$.*

This final remark on the Dacorogna-Moser technique for passing from a domain to another suggests that some of the technical lemmas of the two topics included in this chapter are in common between the two of them, even if I tried to stress the mathematical differences that they have at the beginning of the chapter.

Bibliography

[1] F. Santambrogio, *Variational problems in transport theory with mass concentration*, PhD Thesis, Edizioni della Normale, Birkhäuser, 2007

Chapter 1

[2] G. Carlier, A. Galichon and F. Santambrogio, From Knothe's transport to Brenier's map and a continuation method for optimal transport, *SIAM J. Math. An.*, to appear.

[3] F. Santambrogio, Absolute continuity and summability of transport densities: simpler proofs and new estimates, published on line in *Calc. Var. Par. Diff. Eq.*.

[4] G. Carlier, L. De Pascale and F. Santambrogio, A strategy for non-strictly convex transport costs and the example of $||x-y||^p$ in $\mathbb{R}^2$, submitted, available online at `http://cvgmt.sns.it`

Chapter 2

[5] G. Carlier, C. Jimenez and F. Santambrogio, Optimal transportation with traffic congestion and Wardrop equilibria, *SIAM J. Control Optim.* (47), 2008, 1330-1350.

[6] F. Benmansour, G. Carlier, G. Peyré and F. Santambrogio, Fast Marching Derivatives with Respect to Metrics and Applications, submitted, available online at `http://cvgmt.sns.it`

[7] F. Benmansour, G. Carlier, G. Peyré and F. Santambrogio, Numerical Approximation of Continuous Traffic Congestion Equilibria, *Net. Het. Media*, 4 (3), 2009, 605–623.

[8] L. Brasco, G. Carlier and F. Santambrogio, Congested traffic dynamics, weak flows and very degenerate elliptic equations, published online by *J. Math. Pures et Appl.*

[9] F. Santambrogio and V. Vespri, Continuity for a very degenerate elliptic equation in two dimensions, submitted, available online at `http://cvgmt.sns.it`

Chapter 3

[10] J-M. Morel and F. Santambrogio, The regularity of optimal irrigation patterns, published on line in *Arch. Rati. Mech. An.*.

[11] F. Santambrogio, A Modica-Mortola approximation for branched transport, submitted, available online at `http://cvgmt.sns.it`

Chapter 4

[12] M. Bernot, A. Figalli and F. Santambrogio, Generalized solutions for the Euler equations in one and two dimensions, *J. Math. Pures et Appl.* 91, no. 2, 2009, 137–155

[13] B. Maury, A. Roudneff-Chupin and F. Santambrogio A macroscopic crowd motion model of gradient flow type, *Math. Models Meth. Appl. Sci.*, to appear.

Previous and parallel works

I only list those works which are related to the the subject of this Mémoire. For the complete list of publications, see `http://www.ceremade.dauphine.fr/~filippo` or `http://cvgmt.sns.it`.

[14] L. Ambrosio and F. Santambrogio, Necessary optimality conditions for geodesics in weighted Wasserstein spaces. *Rend. Lincei Mat. Appl.*, (18), no. 1, 23–37, 2007.

[15] A. Brancolini, G. Buttazzo and F. Santambrogio, Path Functionals over Wasserstein spaces, *Journal of the European Mathematical Society*, vol. 8, No. 3, 2006, 415–434;

[16] A. Brancolini, G. Buttazzo, F. Santambrogio and E. Stepanov Long-term planning versus short-term planning in the asymptotical location problem, *Contr. Opt. Calc. Var.* 15 (2009), 509–524

[17] L. Brasco and F. Santambrogio, On the equivalence of different models of branched transport, in preparation.

[18] G. Buttazzo and F. Santambrogio, A model for the optimal planning of an urban area. *SIAM J. Math. Anal.* (37), no. 2, 514–530, 2005.

[19] G. Buttazzo and F. Santambrogio, A Mass Transportation Model for the Optimal Planning of an Urban Region. *SIAM Rev.* (51), no. 3, 593-610, 2009

[20] G. Carlier and F. Santambrogio, A variational model for urban planning with traffic congestion, *ESAIM Control Optim. Calc. Var.* **11** (2007), 595-613.

[21] J.-M. Morel and F. Santambrogio, Comparison of distances between measures, *Appl. Math. Lett.*, (20), no. 4, 2007, 427–432.

[22] F. Santambrogio, Transport and concentration problems with interaction effects. *J. Global Optim.* 38 (2007), no. 1, 129–141.

[23] F. Santambrogio, Optimal Channel Networks, Landscape Function and Branched Transport, *Interfaces and Free Boundaries*, **9** (2007), 149–169.

[24] F. Santambrogio, Introduction to Optimal Transport Theory, lecture notes for the Summer School *Optimal transportation : Theory and applications*, Grenoble, 2009, available at http://www.ceremade.dauphine.fr/∼filippo/teaching.html, to be published.

[25] F. Santambrogio, Models and applications of Optimal Transport Theory, lecture notes for the Summer School *Optimal transportation : Theory and applications*, Grenoble, 2009, available at http://www.ceremade.dauphine.fr/∼filippo/teaching.html, to be published.

Other references

[26] R. A. Adams, *Sobolev spaces*, Academic press, 1975.

[27] L. Ambrosio, Movimenti minimizzanti, *Rend. Accad. Naz. Sci. XL Mem.Mat. Sci. Fis. Natur.*, (113), 191-246, 1995.

[28] L. Ambrosio, Lecture Notes on Optimal Transport Problems, in *Mathematical Aspects of Evolving Interfaces*, 1–52, Springer Verlag, Berlin, 2003.

[29] L. Ambrosio, Transport equation and Cauchy problem for BV vector fields, Invent. Math. **158** (2004), 227-260.

[30] L. Ambrosio and G. Crippa, Existence, uniqueness, stability and differentiability properties of the flow associated to weakly differentiable vector fields, *Transport Equations and Multi-D Hyperbolic Conservation Laws, Lecture Notes of the Unione Matematica Italiana*, Vol. 5.

[31] L. Ambrosio, C. De Lellis and C. Mantegazza, Line energies for gradient vector fields in the plane, *Calc. Var. Partial Differential Equations* (9) no. 4, 1999, 327–255.

[32] L. Ambrosio, N. Gigli and G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Lectures in Mathematics, ETH Zürich (2005)

[33] L. Ambrosio, B. Kirchheim and A. Pratelli, Existence of optimal transport maps for crystalline norms, *Duke Math. J.* , 125 (2), 2004, 207-241.

[34] L. Ambrosio and A. Pratelli. Existence and stability results in the L^1 theory of optimal transportation, in *Optimal transportation and applications*, Lecture Notes in Mathematics (CIME Series, Martina Franca, 2001) 1813, L.A. Caffarelli and S. Salsa Eds., 123-160, 2003.

[35] L. Ambrosio and G. Savaré, *Gradient flows in spaces of probability measures.* Handbook of Differential Equations. Evolutionary equations III, North Holland 2007.

[36] L. Ambrosio and V. M. Tortorelli, Approximation of functionals depending on jumps by elliptic functionals via Γ-convergence. *Comm. Pure Appl. Math.* (43), no. 8, 1990, 999–1036.

[37] L. Ambrosio and V. M. Tortorelli, On the approximation of free discontinuity problems. *Boll. Un. Mat. Ital. B (7)*, 6, no. 1, 1992, 105–123.

[38] V.Arnold: Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. (French) *Ann. Inst. Fourier (Grenoble)*, **16** (1966), fasc. 1, 319–361.

[39] P. Aviles and Y. Giga, A mathematical problem related to the physical theory of liquid crystal configurations. *Miniconference on geometry and partial differential equations, 2* (Canberra, 1986), *Proc. Centre Math. Anal. Austral. Nat. Univ., 12,* ANU, Canberra, 1–16, 1987.

[40] P. Aviles and Y. Giga, Singularities and rank one properties of Hessian measures. *Duke Math. J.* 58 (1989), no. 2, 441–467.

[41] M. Beckmann, A continuous model of transportation, *Econometrica* 20, 643–660, 1952.

[42] M. Beckmann, C. McGuire and C. Winsten, C., *Studies in Economics of Transportation.* Yale University Press, New Haven, 1956.

[43] M. Beckmann and T. Puu, *Spatial Economics: Density, Potential and Flow*, North-Holland, Amsterdam, 1985.

[44] J.-D. Benamou and Y. Brenier, A numerical method for the optimal time-continuous mass transport problem and related problems, *Contemp. Math.*, 226, pp. 1–11, 1999.

[45] J.-D. Benamou and Y. Brenier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem, *Numer. Math.*, 84, pp. 375–393, 2000.

[46] P. Bernard and B. Buffoni, The Monge Problem for supercritical Mañé Potentials on compact Manifolds, *Advances in Mathematics*, 207 n. 2, 691-706, 2006.

[47] M. Bernot, V. Caselles and J.-M. Morel, Traffic plans, *Publicacions Matemàtiques* (49), no. 2, 417–451, 2005.

[48] M. Bernot, V. Caselles and J.-M. Morel, The structure of branched transportation networks, *Calc. Var. Par. Diff. Eq.*, 32 (2008), no. 3, 279–317.

[49] M. Bernot, V. Caselles and J.-M. Morel, *Optimal transportation networks, Models and Theory*, Lecture Notes in Mathematics, Springer, Vol. 1955 (2008).

[50] F. Bethuel, H. Brezis and F. Hélein, Asymptotics for the minimization of a Ginzburg-Landau functional, *Calc. Var. and Par. Diff. Eq.*, (1), no. 2, 1993, 123–148

[51] F. Bethuel, H. Brezis and F. Hélein, *Ginzburg-Landau Vortices*, Birkhäuser, 1994.

[52] S. Bianchini and M. Gloyer, On the Euler-Lagrange equation for a variational problem: the general case II, preprint available online at the page `https://digitallibrary.sissa.it/handle/1963/2551`, 2008.

[53] G. Bouchitté, G. Buttazzo and P. Seppecher, Shape optimization solutions via Monge-Kantorovich equation. *C. R. Acad. Sci. Paris Sér. I Math.* 324, n. 10, 1185–1191, 1997.

[54] G. Bouchitté, C. Dubs and P. Seppecher, Transitions de phases avec un potentiel dégénéré à l'infini, application à l'équilibre de petites gouttes. *C. R. Acad. Sci. Paris Sér. I Math.* 323, 1103–1108, 1996.

[55] A. Braides, *Approximation of free-discontinuity problems.* Lecture Notes in Mathematics, 1694, Springer-Verlag, Berlin, 1998.

[56] Y.Brenier: The least action principle and the related concept of generalized flows for incompressible perfect fluids. *J. Amer. Mat. Soc.*, **2** (1989), 225–255.

[57] Y.Brenier: The dual least action problem for an ideal, incompressible fluid. *Arch. Ration. Mech. Anal.*, **122** (1993), 323–351.

[58] Y.Brenier: Homogeneous hydrostatic flows with convex velocity profiles. *Nonlinearity* **12** (1999), 495–512.

[59] Y.Brenier: Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations. *Comm. Pure Appl. Math.*, **52** (1999), 411–452.

[60] Y.Brenier: Generalized solutions and hydrostatic approximation of the Euler equations. Physica D, to appear.

[61] G. Buttazzo, A. Davini, I. Fragalà and F. Maciá, Optimal Riemannian distances preventing mass tranfer, *J. Reine Angew. Math.*, 575: 157–171, 2004.

[62] L. Caffarelli, M. G. Crandall, M. Kocan and A. Swiech, On Viscosity Solutions of Fully Nonlinear Equations with Measurable Ingredient, *Comm. on Pure and Appl. Math.*, **49** (4): 365–398, 1999.

[63] F. Camilli and E. Prados, Shape-from-Shading with discontinuous image brightness, *Appl. Num. Math.*, **56** (9): 1225–1237, 2006.

[64] L. Caravenna , A proof of Sudakov theorem with strictly convex norms, preprint, 2009, available at `http://cvgmt.sns.it`

[65] L. Caravenna and S. Daneri, The disintegration of the Lebesgue measure on the faces of a convex function, preprint, 2009, available at `http://cvgmt.sns.it`

[66] C. Carstensen and S. Müller, *Local stress regularity in scalar non-convex variational problems.* SIAM J. Math. Anal. 34, 495-509 (2001)

[67] T. Champion and L. De Pascale, The Monge problem for strictly convex norms in $\mathbb{R}^d$, to appear in *J. Eur. Math. Soc.*, 2009, available at `http://cvgmt.sns.it`

[68] T. Champion and L. De Pascale The Monge problem in R^d, preprint, 2009, available at `http://cvgmt.sns.it`

[69] T. Champion, L. De Pascale and P. Juutinen, The ∞-Wasserstein distance: local solutions and existence of optimal transport maps, *SIAM Journal of Mathematical Analysis,* 40 (1), 2008, 1–20.

[70] B. Dacorogna and J. Moser, On a partial differential equation involving the Jacobian determinant. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 7 (1990), no. 1, 1–26.

[71] G. Dal Maso: *An Introduction to* $\Gamma-$*convergence*. Birkhauser, Basel, 1992.

[72] E. De Giorgi and T. Franzoni, Su un tipo di convergenza variazionale, *Atti Acc. Naz. Lincei Rend.*, (58) no. 8, 1975, pp. 842-850.

[73] L. De Pascale and A. Pratelli, Regularity properties for Monge Transport Density and for Solutions of some Shape Optimization Problem, *Calc. Var. Par. Diff. Eq.* 14, n. 3, pp. 249–274, 2002.

[74] L. De Pascale, L. C. Evans and A. Pratelli, Integral estimates for transport densities, *Bull. of the London Math. Soc.*. 36, n. 3, pp. 383–385, 2004.

[75] L. De Pascale and A. Pratelli, Sharp summability for Monge Transport density via Interpolation, *ESAIM Control Optim. Calc. Var.* 10, n. 4, pp. 549–552, 2004.

[76] E. DiBenedetto, $C^{1+\alpha}$ *local regularity of weak solutions of degenerate elliptic equations*, Nonlinear Analysis. Theory, Methods & Applications **7** (1983), 827-850.

[77] E. DiBenedetto, *Degenerate Parabolic Equations.* Universitext. Springer-Verlag, New York, 1993

[78] E. DiBenedetto and V. Vespri, *On the singular equation* $\beta(u)_t = \Delta u$. Arch. Rational Mech. Anal. 132 (1995), no. 3, 247–309.

[79] L. C. Evans and W. Gangbo, Differential equations methods for the Monge-Kantorovich mass transfer problem, Mem. Amer. Math. Soc. **137**, 653 (1999)

[80] M. Feldman and R. McCann, Uniqueness and transport density in Monge's mass transportation problem, *Calc. Var. Par. Diff. Eq.* 15, n. 1, pp. 81–113, 2002.

[81] A. Figalli and N. Juillet, Absolute continuity of Wasserstein geodesics in the Heisenberg group, *J. Funct. Anal.*, 255, n. 1, pp. 133-141, 2008.

[82] A. Figalli and N. Gigli, A new transportation distance between non-negative measures, with applications to gradients flows with Dirichlet boundary conditions, preprint.

[83] E. N. Gilbert, Minimum cost communication networks, *Bell System Tech. J.* (46), 2209–2227, 1967.

[84] C. Jimenez, Optimisation de Problèmes de Transport, PhD thesis of Université du Sud-Toulon-Var, 2005.

[85] L. V. Kantorovich, *On the transfer of masses*, Dokl. Akad. Nauk. SSSR, **37**, 227-229 (1942)

[86] Jean-Michel Lasry and Pierre-Louis Lions, Mean-Field Games, *Japan. J. Math.* 2, 229–260(2007)

[87] P.L.Lions: *Mathematical topics in fluid mechanics, Vol. I: incompressible models.* Oxford Lecture Series in Mathematics and its applications, **3**, Oxford University Press, 1996.

[88] S. Lisini, Characterization of absolutely continuous curves in Wasserstein spaces. *Calc. Var. Par. Diff. Eq.*, Vol. 28, no. 1, 2007, 85–120

[89] F. Maddalena and S. Solimini, Transport distances and irrigation models, *J. Conv. An.*, 2009.

[90] F. Maddalena, S. Solimini and J.-M. Morel, A variational model of irrigation patterns, *Interfaces and Free Boundaries*, **5** (2003), 391–416.

[91] B. Maury and J. Venel, Handling of contacts in crowd motion simulations, *Traffic and Granular Flow*, Springer (2007)

[92] B. Maury and J. Venel, A mathematical framework for a crowd motion model, *C.R. Acad. Sci. Paris, Ser I* **346** 1245-1250 (2008)

[93] R. J. McCann, A convexity principle for interacting gases. *Adv. Math.* 128, no. 1, 153–159, 1997.

[94] N. Megiddo, Cost allocation for Steiner trees, *Networks* 8 (1978), No. 1, 1–6

[95] L. Modica and S. Mortola, Un esempio di Γ-convergenza. (Italian) *Boll. Un. Mat. Ital. B* (14), no. 1, 285–299, 1977.

[96] E. Oudet, Approximation of partitions of least perimeter by $\Gamma-$convergence : around Kelvin's conjecture, preprint, 2009, available at `http://www.lama.univ-savoie.fr/~oudet`

[97] L. Piozin, Réseaux branchés de transport, optimisation et équilibre, mémoire de stage de licence, MIDO, Université Paris-Dauphine, 2009

[98] E. Rouy and A. Tourin. A viscosity solution approach to shape from shading. *SIAM Journal on Numerical Analysis*, **29**:867–884, 1992.

[99] J. A. Sethian. *Level Set Methods and Fast Marching Methods.* Cambridge University Press, 1999.

[100] A.I. Shnirelman: The geometry of the group of diffeomorphisms and the dynamics of an ideal incompressible fluid. (Russian) *Mat. Sb. (N.S.)*, **128 (170)** (1985), no. 1, 82–109.

[101] A.I. Shnirelman: Generalized fluid flows, their approximation and applications. *Geom. Funct. Anal.*, **4** (1994), no. 5, 586–620.

[102] J. Venel, Integrating strategies in numerical modelling of crowd motion, *Pedestrian and Evacuation Dynamics*, Springer (2008)

[103] J. Venel, A discrete contact model for crowd motion, soumis

[104] C. Villani, *Topics in optimal transportation*, Grad. Stud. Math. 58, AMS, Providence (2003)

[105] C. Villani, *Optimal Transport, Old and New*, Grundlehren der mathematischen Wissenschaften, Vol. 338, Springer, 2009.

[106] J. G. Wardrop, Some theoretical aspects of road traffic research, *Proc. Inst. Civ. Eng.* **2** (1952), 325-378.

[107] Q. Xia, Optimal Paths related to Transport Problems. *Comm. Cont. Math.* (5), no. 2, 251–279, 2003.

[108] Q. Xia, Interior regularity of optimal transport paths. *Calc. Var. and and Par. Diff. Eq.* (20), no. 3, 2004, 283–299.

[109] Q. Xia, Boundary regularity of optimal transport paths, preprint 2004.