

Mémoire de stage
École Normale Supérieure de Cachan

Théorie générale du transport et applications

Cécile CARRÈRE, Didier LESESVRE, Paul PEGON

Filippo SANTAMBROGIO — Directeur de stage

Table des matières

1	Récit d'un périple dans le monde du transport	4
1.1	La théorie du transport, de MONGE à KANTOROVICH	4
1.2	Les grandes étapes de l'aventure	5
2	Résumé du cours de Filippo SANTAMBROGIO	7
2.1	Dualité de MONGE – KANTOROVICH	7
2.2	Quelques cas particuliers de coûts	7
2.3	Retour sur le problème de MONGE	8
2.4	Le problème de BECKMAN	8
2.5	Cadre topologique et métrique	8
2.6	Application à des problèmes d'équations aux dérivées partielles	9
2.7	Introduction aux flots gradients	10
3	Groupe de lecture : Haïm Brézis, <i>Analyse Fonctionnelle</i>	11
3.1	Théorème de HAHN – BANACH, théorie des fonctions convexes	11
3.2	Théorèmes de BAIRE et de BANACH – STEINHAUS	11
3.3	Topologies faibles, espaces réflexifs et séparables	11
3.4	Les espaces L^p	12
3.5	Les espaces de HILBERT	13
3.6	Espaces de SOBOLEV et problèmes aux limites	13
4	Étude d'articles de recherche	14
4.1	G. CARLIER, I. EKELAND, <i>Equilibrium structure of a bidimensional asymmetric city</i>	14
4.2	Y. BRENIER <i>et al.</i> , <i>Reconstruction of the early Universe as a convex optimization problem</i>	15
5	Simulations numériques : l'algorithme des enchères	17
6	Colloque de recherche : <i>Optimal Transport : algorithms and applications</i>	20
7	Un problème ouvert : Minimisation d'un critère de type oscillation	21
7.1	Motivation	21
7.2	Le problème d'optimisation	21
7.3	Reformulation du problème	21
7.4	Existence d'un optimum	22
7.5	Recherche d'un transport dans le cas réel et compact	23
7.5.1	Paysage de travail	23
7.5.2	Premières idées	23
7.5.3	Formulation duale	23
7.5.4	Un théorème d'existence	24
8	Références	28

Résumé

Le stage d'initiation à la recherche, que nous avons effectué sous la direction de Filippo SANTAMBROGIO et coordonné par l'École Normale Supérieure de Cachan, a été l'occasion de découvrir le passionnant domaine de la théorie du transport, à travers un cours sur le sujet que nous avons suivi à l'école doctorale de la faculté d'Orsay, donné par notre encadrant. Le sujet est vaste et repose sur des acquis solides et variés en analyse fonctionnelle et en théorie de la mesure, et les premiers mois du stage ont, partant, été consacrés à la compréhension et à l'étude du cours ainsi qu'à la rédaction d'un mémoire sur la théorie du transport, reprenant et détaillant le cours que nous avons suivi, et précisant les résultats connexes nécessaires à l'abordage efficace de la théorie. Nous avons en parallèle constitué un groupe de lecture, orchestré par Adina CIOMAGA, dans le but de lire et d'assimiler dans le détail la majeure partie de *l'Analyse Fonctionnelle* de Haim BRÉZIS.

Une fois les cours à Orsay terminés, nous nous sommes attachés à l'étude de deux articles récents de recherche en théorie du transport, l'un sur une application économique du transport à la modélisation de l'évolution dans une ville, l'autre sur une application physique à la modélisation de l'expansion de l'Univers. Ce dernier a été l'occasion d'étudier et de coder l'algorithme des enchères, proposé pour la première fois par Dimitri BERTSEKAS dans les années 1970 pour résoudre le problème de l'allocation optimale de ressources dans le cas discret, et d'en tester la rapidité et la correction. Nous avons par la suite assisté aux conférences du colloque *Optimal Transport : algorithms and applications*, organisé pendant trois jours à l'Institut Henri Poincaré par l'ANR OTARIE. Pour finir, nous avons réfléchi et partiellement résolu un sujet de recherche ouvert, celui de la recherche de plans de transfert et de transports optimaux pour un critère de type L^∞ . De bons résultats ont été obtenus en dimension 1 dans le cas compact et pour des mesures sans atomes, et la réflexion sur l'extension de nos résultats et de nos idées en dimension supérieure est pour le moment inachevée.

Chapitre 1

Récit d'un périple dans le monde du transport

1.1 La théorie du transport, de MONGE à KANTOROVICH

Le problème du transport optimal est introduit par Gaspard MONGE en 1781, alors professeur l'école du génie militaire de Mézières, dans son *Mémoire sur la théorie des déblais et des remblais*, alliant ainsi comme son habitude un sens aigu du concret et des problèmes pratiques avec une grande capacité d'abstraction dans la formulation et l'étude de ces problèmes.

Il s'agit de combler un fossé avec le sable d'un talus, en minimisant l'énergie nécessaire à ce déplacement. MONGE cherchait à minimiser le coût

$$c(T) = \int_X |T(x) - x| dx$$

parmi les transport envoyant la masse initiale sur la masse finale. Le talus et le fossé peuvent être modélisés par des mesures respectives μ et ν , de densités f et g . Si chaque unité de sable en x est envoyée en $T(x)$, la fonction de transport T doit vérifier, pour toute fonction ϕ ν -mesurable,

$$\int_{\Omega} \phi(y) d\nu(y) = \int_{\Omega} \phi(T(x)) d\mu(x) \quad i.e. \quad T\#\mu = \nu$$

Le coût pour déplacer une unité de sable de la position x à la position y peut être plus généralement une fonction $c(x, y)$, et l'on cherche à minimiser le coût total du transport parmi les T vérifiant cette condition la fonctionnelle

$$c(T) = \int_{\Omega} c(x, T(x)) d\mu(x)$$

Ce problème est « très peu linéaire », et des preuves d'existence semblent difficiles à trouver. C'est pourquoi ce problème ne donna lieu qu'à très peu de publications jusqu'au milieu du vingtième siècle. Dans les années 1940, l'économiste et mathématicien russe Leonid KANTOROVITCH pose un regard nouveau sur ce problème séculaire, rendant fertile une terre depuis bien longtemps désertique : nous ne cherchons plus à trouver une application de transport, mais une mesure γ sur Ω^2 vérifiant $\pi_1\#\gamma = \mu$ et $\pi_2\#\gamma = \nu$, où π_1 et π_2 sont les projections respectivement sur l'axe des x et des y . On dit que γ est un plan de transfert. Intuitivement, on n'impose plus qu'un élément de masse initial soit envoyé sur un unique élément de masse à l'arrivée, mais on autorise le *mass splitting*. La quantité à minimaliser est alors

$$c(\gamma) = \int_{\Omega^2} c(x, y) d\gamma(x, y)$$

Ce problème est un problème de minimisation linéaire sous contraintes convexes, il admet donc un problème dual : il s'agit de minimiser, pour ϕ et ψ deux fonctions de Ω vérifiant $\phi(x) + \psi(y) \leq c(x, y)$ pour tout $x, y \in \Omega$, le critère

$$\int_{\Omega} \phi(x) d\mu(x) + \int_{\Omega} \psi(y) d\nu(y)$$

Cette formulation permet de trouver des propriétés intéressantes sur le transport, par exemple qu'il en existe un optimum si la fonction c est convexe, et qu'il est unique si elle est strictement convexe. Cette formulation marque une véritable renaissance de la théorie du transport, et ouvre la porte à tous les outils d'analyse convexe et de théorie de la mesure, qui permettent de prouver de nombreux résultats très généraux, et de relier la théorie du transport à de nombreux autres problèmes, notamment d'équations aux dérivées partielles. Un aperçu des principaux résultats de la théorie est donné dans la suite du présent rapport, et un mémoire détaillé sur la théorie du transport est joint en annexe.

1.2 Les grandes étapes de l'aventure

L'arrivée dans le monde du transport optimal a été une véritable entrée en terrain inconnu, et nous avons été initié aux arts de cette théorie par Filippo SANTAMBROGIO, en suivant le cours qu'il a donné à l'École Doctorale de la faculté d'Orsay sur la théorie du transport. Durant huit semaines, avec des séances hebdomadaires de trois heures, nous avons eu un panorama relativement complet des résultats fondamentaux et des grandes applications de la théorie du transport, et avons consacré ces semaines à la compréhension de ce riche cours et à l'étude des résultats théoriques annexes nécessaires à la théorie.

La théorie du transport optimal requiert de bonnes connaissances en analyse fonctionnelle, c'est pourquoi nous avons étudié les principaux chapitres du livre de Haïm BREZIS, *Analyse fonctionnelle, Théorie et applications* avec Adina CIOMAGA. Pendant des sessions de deux à trois heures, nous avons présenté oralement et dans le détail les chapitres sur le théorème de HAHN-BANACH et ses applications, sur la topologie faible et les espaces réflexifs, séparables et uniformément convexes, sur les espaces L^p , sur les espaces de HILBERT et sur les espaces de SOBOLEV. Les résultats sur la topologie faible et les espaces de SOBOLEV sont particulièrement utiles dans la compréhension de la théorie du transport optimal, et donnent toute sa puissance à la formulation de KANTOROVICH du problème.

Nous avons aussi étudié des applications plus directes de la théorie du transport, à travers la lecture de deux articles de recherche. Premièrement, l'article de G. CARLIER et I. EKELAND présente une modélisation nouvelle de l'évolution des villes considérées comme des systèmes travailleurs – employeurs – propriétaires terriens. Il s'agit du premier modèle très général d'évolution, et la difficulté est concentrée dans la modélisation des comportements et des sens d'évolution de chacun. Une fois un modèle relativement élaboré établi, on est simplement ramenés à un problème de transport, et les résultats de la théorie permettent d'affirmer l'existence d'un équilibre dans ces modèles.

Puis, la lecture de l'article de BRENIER *et al.*, sur l'utilisation du transport dans un modèle physique de l'Univers pour en déduire une reconstruction de l'état de l'Univers dans un passé proche, a été l'occasion d'étudier un algorithme pour déterminer un transport optimal dans le cas discret, nous avons décidé de le coder sous Matlab et d'en évaluer quelques résultats, qui était utilisé lors des simulations dont les résultats étaient présentés dans l'article. Il s'agit ici de transporter, par exemple, des sacs de farine de différentes usines vers différentes boulangeries ; en terme de mesures, la mesure de départ et la mesure d'arrivée sont des sommes de masses de DIRAC. On peut montrer qu'une masse de DIRAC ne peut être envoyée par un transport que vers une masse de DIRAC portant la même masse : on étudiera donc le cas où toutes les masses sont unitaires. Quitte à supposer que les masses sont toutes rationnelles, qui approchent arbitrairement bien toutes les valeurs des masses possibles, on peut toujours se ramener à ce cas, en disant que certaines masses sont plusieurs masses unitaires superposées. Ce programme – l'algorithme des enchères de Dimitri BERTSEKAS – est basé sur des idées économiques d'offre et de demande.

Le colloque scientifique *Optimal transport : applications and algorithms*, organisé durant trois jours au mois de juin à l'Institut Henri POINCARÉ par l'ANR OTARIE, a été l'occasion de découvrir comment se déroulent ce genre de rassemblement de scientifiques spécialistes d'un même sujet. Chacun présentant ses sujets de recherche actuels ou ses passions anciennes sur la théorie du transport au cours de conférences d'une heure environ, suivies de questions et de commentaires de l'auditoire, nous avons pris conscience des multiples directions dans lesquelles peuvent aller les applications de la théorie du transport, et les problèmes et questions en suspens sont l'occasion de discussions constructives entre les chercheurs, de naissance de nouvelles idées, d'apparition de nouvelles connexions entre des sujets distincts.

Nous avons finalement abordé un problème ouvert de minimisation de fonctionnelle de type oscillant, autrement dit de recherche de transfert ou de transports optimaux ou quasi-optimaux pour un coût particulier. Ce coût représente par exemple des enjeux liés au bon brouillage des positions exactes dans une situation où l'on doit transmettre une position approximative, suffisamment représentative de la position exacte pour que les personnes qui étaient proches, dans une limite à préciser, avant transformation de leur position restent proche après. Le coût n'est autre que la proximité maximale après transformation de points « proches ».

Ces recherches peuvent s'appliquer dans le domaine de respect de la vie privée : les téléphones portables proposent parfois des applications qui permettent de savoir si certains contacts sont à proximité. Pour ce faire, un calculateur externe détermine les distances entre ces téléphones. Mais pour que ce calculateur ne puisse pas remonter à la position du téléphone, il faut modifier la carte. Des transformations rigides, comme des rotations, ne suffisent pas, car en comparant les densités de téléphones et les positions des villes, on peut retrouver la bonne orientation de la carte. Il faut donc transformer la distribution actuelle en une distribution uniforme par exemple, mais pour limiter les calculs inutiles, on exige que des points éloignés l'un de l'autre dans la nouvelle distribution soient aussi éloignés dans la distribution réelle. Cela revient à minimiser un « module de continuité partiel » de T , évalué en une valeur δ , où T est un transport de la distribution réelle des téléphones vers la distribution uniforme, pour laquelle toute isométrie simple, telle une rotation ou une symétrie, est intraquable.

Chapitre 2

Résumé du cours de Filippo SANTAMBROGIO

2.1 Dualité de MONGE – KANTOROVICH

La relaxation (semi-continue inférieurement) du problème de MONGE mène exactement au problème de KANTOROVICH, ce qui prouve que ce dernier n'a pas trop élargi le problème initial, et que la nouvelle formulation du problème en est la meilleure généralisation possible. Cependant la contrainte d'appartenance à $\pi(\mu, \nu)$ n'est pas très aisée à manier et à caractériser, et on cherche naturellement à se ramener à un problème de minimisation sur des mesures quelconques, nous laissant plus de liberté. Pour cela, on se sert de la fonction caractéristique de $\pi(\mu, \nu)$:

$$\chi_{\pi(\mu, \nu)} = \sup_{\phi, \psi} \left(\int \phi d\mu + \int \psi d\nu - \int (\phi(x) + \psi(y)) d\gamma \right) = \begin{cases} 0 & \text{si } \gamma \in \pi(\mu, \nu) \\ \infty & \text{sinon} \end{cases}$$

Le problème revient donc à minimiser le critère de KANTOROVICH auquel on ajoute cette fonction caractéristique, sur *toutes* les mesures. Le théorème de FENCHEL – ROCKAFELLAR permet alors de permuter l'infimum et le supremum, et d'obtenir le problème dual

$$\inf_{\gamma \in \pi(\mu, \nu)} \int c d\gamma = \sup_{\phi \oplus \psi \leq c} \left(\int \phi d\mu + \int \psi d\nu \right) = \max_{\phi \oplus \psi \leq c} \left(\int \phi d\mu + \int \psi d\nu \right)$$

Deux fonctions ϕ et ψ réalisant ce maximum sont des *potentiels de KANTOROVICH* du problème, et ils sont au centre de nombre de démonstrations en théorie du transport. On peut les supposer continus avec le même module de continuité que le coût, et même les supposer c -transformés l'un de l'autre.

2.2 Quelques cas particuliers de coûts

Le cas du transport discret est représenté par des mesures μ et ν qui sont des sommes de masses unité de DIRAC. Un plan de transfert est alors représenté par la matrice $\gamma = (\gamma_{i,j})$ de la répartition de masse d'un x_i entre les y_j . Les théorèmes de CHOQUET et de BIRKHOFF prouvent que les matrices minimisant le critère sont les matrices de permutation. Dans ce cas, les transferts optimaux sont donc tous des transports. Ce cas s'étend au cas de masses rationnelles, en normalisant la masse de manière à obtenir des entiers, et en scindant les masses supérieures à un en des masses unité. On obtient ainsi tous les plans de transfert optimaux, qui ne sont plus nécessairement des transports.

Le cas d'un coût découplé est celui qui correspond intuitivement à des efforts fournis essentiellement au départ et à l'arrivée, et non concentrés sur le mouvement. Il représente des systèmes physiquement asymétriques et irréversibles. L'étude des potentiels de KANTOROVICH dans ce cas simple mène à l'obtention de tous les potentiels optimaux, et prouve le peu d'importance des termes découplés dans le coût : ils n'influent pas sur les contraintes de l'optimisation, et on peut désormais se borner à ne considérer que des coûts où les variables sont réellement couplées.

Le cas d'un coût strictement convexe en la distance $h(x - y)$ est très riche et nous y prouvons un résultat puissant : il existe un unique transport optimal, qui est de la forme $y \mapsto x - \nabla h^*(\nabla \phi(x))$, où ϕ est un

potentiel de KANTOROVICH. Le théorème de BRENIER est un cas particulier de ce résultat dans le cas du coût c_2 , pour lequel il existe un unique transport qui est le gradient d'une fonction convexe, à condition que les mesures soient absolument continues par rapport à la mesure de LEBESGUE.

2.3 Retour sur le problème de MONGE

Le problème initialement posé par MONGE est celui de la recherche de transport dans le cas de la distance euclidienne. Une condition intuitivement – et effectivement – nécessaire à un transport optimal est que les points qui portent de la masse ne peuvent être permutés entre eux et donner ainsi un transport moins coûteux. C'est la condition de *c-monotonie cyclique*, valable pour un coût continu et pour des points du support :

$$\forall ((x_i, y_i))_i \in \text{Supp}(\gamma), \forall \sigma \in \mathfrak{S}_n, \sum_{i=1}^n c(x_i, y_i) < \sum_{i=1}^n c(x_i, y_{\sigma_i})$$

Cette propriété permet de prouver un résultat dans le cas d'un coût convexe en la distance : il existe un unique transport optimal croissant, et il s'agit de l'unique transport optimal dans le cas où le coût est strictement convexe en la distance.

Le cas de la distance est plus délicat que les précédents, mais possède des particularités permettant une étude efficace du problème. Le support des plans de transfert optimaux sont partitionnés en des segments, appelés *rayons de transport*, sur lesquels les potentiels de KANTOROVICH sont linéaires et différentiables. L'étude est alors séparée sur chaque rayon, et aboutit à un recollement final des transports obtenus. Ce partitionnement en rayons de transport caractérise entièrement les plans de transfert optimaux pour le coût distance. Il n'y a plus unicité du transport optimal pour ce coût, cependant il est unique si l'on impose la minimisation de c_2 également.

Le cas du coût c_∞ est un cas où la contrainte imposée par le coût est très faible, et tous les plans de transferts dont le support est situé dans des bandes obliques $\{|x - y| \leq L_\infty\}$ sont optimaux. Cependant, comme précédemment, il existe un unique transport optimal optimisant également c_2 .

2.4 Le problème de BECKMAN

Le problème de BECKMANN est celui de la minimisation de la quantité de mouvement du déplacement d'une distribution de fluide vers une autre distribution, représentées somme à l'usuel par deux mesures μ et ν . Autrement dit, on recherche un champ de vecteurs $v : \Omega \rightarrow \mathbb{R}^d$ minimisant le problème

$$\int_{\Omega} |v(x)| dx \quad \text{où} \quad \nabla \cdot v = \mu - \nu$$

Le problème de BECKMANN, bien que d'apparence difficile du fait du manque de propriété de compacité sur l'espace des champs de vecteurs, est en fait équivalent aux problèmes primal et dual de MONGE – KANTOROVICH.

2.5 Cadre topologique et métrique

Les résultats obtenus en théorie du transport permettent de munir l'espace $P(\Omega)$ des mesures de probabilités d'une structure métrique riche. Nous définissons sur cet espace la distance de WASSERSTEIN d'ordre p par

$$W_p(\mu, \nu) = \min \left\{ \left(\int_{\Omega} |x - y|^p d\gamma \right)^{\frac{1}{p}} \mid \gamma \in \pi(\mu, \nu) \right\}$$

Qui se généralise naturellement au cas $p = \infty$ par

$$W_\infty(\mu, \nu) = \min \{ \max |x - y| \mid \gamma \in \pi(\mu, \nu), (x, y) \in \text{Supp} \gamma \}$$

Les distances W_p définissent des distances sur l'espace $P(\Omega)$. Pour des mesures d'ordre p fini, la convergence au sens de la topologie engendrée par W_p est caractérisée par

$$W_p(\mu_n, \mu) \xrightarrow[n_\infty]{} 0 \iff \begin{cases} \mu_n \xrightarrow[n_\infty]{} \mu \\ \int |x|^p d\mu_n \xrightarrow[n_\infty]{} \int |x|^p d\mu \end{cases}$$

Munis de ces « bonnes » distances, il est naturel d'étudier des problèmes plus physiques faisant intervenir la vitesse. Nous introduisons notamment, pour une courbe de mesures $(\mu_t)_t$ lipschitzienne dans $P(\Omega)$, l'équation de continuité pour les champs de vitesse $(v_t)_t$, où v_t est μ_t -intégrable presque partout :

$$\partial_t v_t + \nabla \cdot (v_t \cdot \mu_t) = 0$$

Il y a une unique solution faible $(\mu_t)_t$ de l'équation de continuité pour v uniformément bornée et uniformément lipschitzienne et pour μ_0 fixés. Cette équation permet notamment de caractériser les courbes W_p lipschitziennes : toute courbe lipschitzienne admet presque partout un champ de vitesses L^p , borné par $|\mu'|$, vérifiant l'équation de continuité; et toute solution à l'équation de continuité dont le champ de vitesse est borné au sens L^p est lipschitzienne.

Un autre problème surgissant naturellement lors de l'étude d'une distance est celui de l'étude des géodésiques, *i.e.* des courbes minimisant la longueur parmi les courbes existant entre deux extrémités fixées. Dans le cas des espaces de WASSERSTEIN, nous connaissons explicitement ces géodésiques. En effet, l'existence d'un transport optimal de μ à ν correspond à l'idée intuitive de minimiser le « trajet » qu'il y a à faire pour passer de μ à ν , ce qui correspond aussi à l'idée intuitive que l'on se fait d'une géodésique entre μ et ν , intuition qui est justifiée. Pour tout transport optimal T dans un convexe, l'image de μ par $(1-t)id + tT$ est une géodésique de μ à ν . Plus généralement, on peut toujours trouver une géodésique qui est la projection d'un plan de transfert optimal sur un parcours de segment π_t .

Les géodésiques à vitesses constantes sont parfaitement connues par rapport aux plans de transferts optimaux : les géodésiques à vitesses constantes passent par des mesures μ_t pour lesquelles il existe un unique transport optimal de μ à μ_t et un unique transport optimal de μ_t à ν ; et les géodésiques à vitesse constante sont exactement les projections des plans de transfert optimaux sur les π_t , pour $p > 1$.

2.6 Application à des problèmes d'équations aux dérivées partielles

L'étude précédente des courbes des espaces de WASSERSTEIN mène au problème de la minimisation de la longueur des courbes parmi celles vérifiant l'équation de continuité. Mais la formulation du problème en terme de minimisation de la norme de la vitesse ne semble guère adaptée, et le problème se reformule en posant judicieusement $E_t = v_t \mu_t$, par la minimisation du critère convexe

$$\frac{1}{p} \iint \left| \frac{E_t}{\mu_t} \right|^p d\mu_t \quad \text{où} \quad \partial_t \mu_t + \nabla \cdot E_t = 0$$

Quelques manipulations mènent alors à la forme de BÉNAMOÛ – BRENIER du problème :

$$\sup \left\{ \int (a\psi_t + b \cdot E_t) \mid a + \frac{1}{q}|b|^q \leq 0, \partial_t \mu_t + \nabla \cdot E_t = 0 \right\}$$

On absorbe les condition par les fonctions caractéristiques correspondantes, et le nouveau critère de minimisation devient alors, après l'interversion formelle de l'infimum et du supremum justifiée par un théorème d'analyse convexe dû à ROCKAFELLAR :

$$\min_{\substack{\mu \in P_p(\Omega), E_t = v_t \mu_t \\ \partial_t \mu_t + \nabla \cdot E_t = 0}} \sup_{\substack{a(t,x), b(t,x), \phi \in C^1(\Omega) \\ a + \frac{1}{q}|b|^q \leq 0}} \int \int a d\mu + \int \int b dE - \int \int \partial_t \phi d\mu - \int \nabla \phi \cdot E + G(\phi)$$

Pour simplifier les notations, on pose $m = (\mu, E)$ et $\xi = (a, b)$, et on utilise une méthode du lagrangien augmenté en ajoutant au critère un terme positif et nul à l'optimum, aboutissant à la formulation

$$\min_m \sup_{\xi, \phi} m \cdot \xi - \nabla_{t,x} \phi \cdot m + G(\phi) - I_k(\xi) + \|\xi - \nabla_{t,x} \phi\|^2$$

On approche l'optimum grace à l'algorithme de BRENIER, de type gradient :

$$\begin{cases} F(m) = \sup_x f(m, x) \\ \nabla F(m) = \nabla_m f(m, x^*) \text{ avec } x^* \text{ réalisant le supremum} \end{cases}$$

Le problème est que cet algorithme ne s'applique que pour des fonctions suffisamment régulières pour pouvoir être différentiées presque partout. Dans le cas général, on effectue une discrétisation du problème et on utilise un procédé récursif pour l'approximation, constituant l'algorithme de BRENIER :

$$\forall k \in \mathbb{N}, \begin{cases} \phi_{k+1} = \phi \text{ maximisant le critère avec } (m_k, \xi_k) \\ \xi_{k+1} = \xi \text{ minimisant le critère avec } (m_k, \phi_{k+1}) \\ m_{k+1} = m_k - r(\xi_{k+1} - \nabla_{x,t} \phi_{k+1}) \end{cases}$$

2.7 Introduction aux flots gradients

On se place dans un espace de dimension finie que l'on assimile à \mathbb{R}^n . On considère le problème

$$(E) \begin{cases} x'(t) = -\nabla F(x(t)) \\ x(t_0) = x_0 \end{cases}, \quad t \in \mathbb{R}, \quad F \text{ différentiable}$$

et on cherche à minimiser $F : \mathbb{R}^n \rightarrow \mathbb{R}$ parmi les solutions du problème (E). Autrement dit, il s'agit d'optimiser le mouvement de F à partir de la seule condition initiale, donc de suivre en permanence la direction donnée par la plus grande pente, *i.e.* la direction du gradient.

On fixe un pas temporel $\tau > 0$ et on cherche une suite $(x_k)_k$ représentant les positions successives de la courbe de la solution optimale aux temps $k\tau$. Le pas τ est appelé à tendre vers 0, ce qui fait que la distance entre les x_k tendra également vers 0 dès que F est suffisamment régulière. Donc pour la recherche d'une solution minimisante, on peut imposer que x_{k+1} minimise le critère

$$F(x) + \frac{1}{2\tau}|x - x_k|^2$$

ce qui ne nuit pas à la généralité *ad infinitum*. Le x_{k+1} optimisant le critère doit annuler le gradient, autrement dit il doit satisfaire

$$\frac{1}{\tau}(x_{k+1} - x) + \nabla F(x_{k+1}) = 0 \text{ d'où } \frac{x_{k+1} - x_k}{\tau} = \frac{\Delta x}{\Delta t} = -\nabla F(x_{k+1})$$

On définit donc la suite $(x_k)_k$ récursivement par

$$\begin{aligned} & x_0 \\ \forall k \in \mathbb{N}, & x_{k+1} = x_k - \tau \nabla F(x_{k+1}) \end{aligned}$$

Il s'agit du *schéma d'EULER* implicite, qui correspond intuitivement bien à ce que l'on cherche : on choisit de suivre à intervalle de temps régulier la plus grande pente possible au point courant, donnée par l'opposé du gradient. On « retrouve » alors F en interpolant les x_k lorsque le pas tend vers zéro. En fait tout ceci se généralise dans le cas plus général des espaces métriques, et non plus nécessairement euclidiens, en imposant à x_{k+1} de minimiser $F(x) + \frac{1}{2\tau}d(x, x_k)^2$.

Quelques développements se servant des outils précédemment développés en théorie du transport permettent d'affirmer que les minimiseurs de $F(\mu) + \frac{1}{2\tau}W_2^2(\mu_k, \mu)$ vérifient nécessairement, si ϕ est un potentiel de KANTOROVICH associé :

$$\nabla \frac{\delta F}{\delta \mu} + \frac{\nabla \phi}{\tau} = 0 \text{ donc } T(x) = x - \nabla \phi(x) \text{ soit } \nabla \phi = id - T$$

Le théorème de BRENIER donne, pour $h : x \mapsto \frac{1}{2}|z|^2$, $T(x) = x - (\nabla h)^{-1}(\nabla \phi(x))$. On peut alors bien connaître la trajectoire de la particule en introduisant

$$v_k^\tau : x \mapsto \frac{1}{\tau}(x - T(x)) = -\nabla \frac{\delta F}{\delta \mu}(x) \text{ et } \mu_k = T\#\mu \text{ optimisant le déplacement de la particule}$$

Chapitre 3

Groupe de lecture : Haïm Brézis, *Analyse Fonctionnelle*

3.1 Théorème de HAHN – BANACH, théorie des fonctions convexes

E est un \mathbb{R} -espace vectoriel.

Théorème. (Hahn – Banach, analytique) *Si $p : E \rightarrow \mathbb{R}$ est une semi-norme, i.e. positive-homogène et sous-additive, et si $g : G \rightarrow \mathbb{R}$ est une forme linéaire d'un sous-espace G de E majorée par p , alors il existe une forme linéaire f qui prolonge g sur E et qui est majorée par p .*

Théorème. (Hahn – Banach, géométrique) *A et B ouverts convexes, disjoints non vides de E . Si A est ouvert, alors il existe un hyperplan fermé qui sépare A et B au sens large. Si A est fermé et B est compact, alors il existe un hyperplan fermé qui sépare A et B au sens strict.*

Théorème. (Fenchel – Moreau) *Si φ est convexe, s.c.i. et $\varphi \neq +\infty$, alors $\varphi^{**} = \varphi$.*

Théorème. (Fenchel – Rockafellar) *Si φ et ψ sont deux fonctions convexes, finies en un même point x_0 , et si φ est continue en x_0 , alors la formulation duale de FENCHEL – ROCKAFELLAR est valide*

$$\inf_{x \in E} (\varphi(x) + \psi(x)) = \sup_{f \in E'} (-\varphi^*(-f) - \psi^*(f)) = \max_{f \in E'} (-\varphi^*(-f) - \psi^*(f))$$

3.2 Théorèmes de BAIRE et de BANACH – STEINHAUS

Théorème. (Baire) *Dans un espace métrique complet, une union de fermés d'intérieurs vides est d'intérieur vide, i.e. une intersection d'ouverts denses est dense.*

Théorème. (Banach – Steinhaus) *Une famille d'opérateurs linéaires et continus entre deux espaces de BANACH vérifiant $\sup_i \|T_i(x)\|$ fini pour tout x est uniformément bornée.*

3.3 Topologies faibles, espaces réflexifs et séparables

E est un espace de BANACH.

Définition. (Topologie faible) *La topologie faible $\sigma(E, E')$ est la topologie la moins fine sur E , i.e. avec le minimum d'ouverts, qui rende continues toutes les formes linéaires sur E .*

Théorème. (Propriétés de $\sigma(E, E')$)

* $\sigma(E, E')$ est séparée

* $(x_n \rightarrow x) \iff (\forall f \in E', f(x_n) \rightarrow f(x))$

* La convergence forte implique la convergence faible : $x_n \rightarrow x \implies x_n \rightharpoonup x$

Théorème. (Dimension finie) En dimension finie, les topologies faible et forte sont identiques.

Définition. (Topologie faible $*$) La topologie faible $*$ $\sigma(E', E)$ est la topologie la moins fine sur E' , i.e. avec le minimum d'ouverts, qui rende continues toutes les formes $f \mapsto \langle f, x \rangle$.

Théorème. (Propriétés de $\sigma(E', E)$)

* $\sigma(E', E)$ est séparée

* $(f_n \xrightarrow{*} f) \iff (\forall x \in E, f_n(x) \rightarrow f(x))$

* La convergence forte implique la convergence faible : $f_n \rightarrow f \implies f_n \xrightarrow{*} f$

Théorème. (Banach – Alaoglu – Bourbaki) $\mathcal{B}_{E'}$ est compact pour la topologie faible $*$ $\sigma(E', E)$.

Définition. (Réflexivité) E est réflexif si l'injection canonique de E dans E'' est une bijection, autrement dit si les formes linéaires de E' sont les $x \mapsto \langle f, x \rangle$.

Théorème. (Kakutani) E est réflexif si et seulement si \mathcal{B}_E est compact pour la topologie $\sigma(E, E')$.

Définition. (Séparabilité) E est séparable s'il existe une partie dénombrable dense dans E .

Théorème. (Dual) E est réflexif (et séparable) si et seulement si E' est réflexif (et séparable).

3.4 Les espaces L^p

Définition. (Espaces L^p)

L^p est l'espace des fonctions mesurables de moment d'ordre p fini. On note $\|f\|_p = (\int |f|^p)^{1/p}$.
 L^∞ est l'espace des fonctions mesurables presque partout bornées. On note $\|f\|_\infty = \sup f$.

Théorème. (Inégalité de Holder) Pour $f \in L^p$ et $g \in L^q$, $fg \in L^1$ et $\|fg\|_1 \leq \|f\|_p \|g\|_q$.

Théorème. (Fischer-Riesz) L^p est un espace de BANACH pour la norme $\|\cdot\|_p$.

Théorème. (Réflexivité, séparabilité, dualité)

L^p est réflexif, séparable et de dual $L^{p'}$ pour $1 \leq p \leq \infty$, où $\frac{1}{p} + \frac{1}{p'} = 1$.

L^1 est séparable, non réflexif et de dual L^∞ .

L^∞ n'est ni réflexif, ni séparable, et son dual contient strictement L^1 .

Théorème. (Densité) $C_c(\Omega)$ est dense dans $L^p(\Omega)$ pour $1 \leq p \leq \infty$.

Théorème. (Régularisation) Si $f \in C_c^k(\Omega)$ et $g \in L_{loc}^1(\Omega)$, alors $f \star g$ est C_c^k , et $D^\alpha(f \star g) = (D^\alpha f) \star g$.
 Une suite régularisante est une suite de fonctions $\rho_n \in C_c^\infty$, positives, de masse 1, et de support dans $B(0, \frac{1}{n})$.

Si $f \in C(\mathbb{R}^n)$, alors $\rho_n \star f$ converge uniformément vers f sur tout compact.

Si $f \in L^p(\mathbb{R}^n)$, $1 \leq p < \infty$, alors $\rho_n \star f$ converge vers f dans L^p .

Théorème. (Ascoli) Si K est un métrique compact et si H est un borné de $C(K)$ uniformément continu, i.e. vérifiant

$$\forall \varepsilon > 0, \exists \delta > 0, \forall (x_1, x_2), \forall f \in H, (d(x_1, x_2) < \delta \implies d(f(x_1), f(x_2)) < \varepsilon)$$

alors H est relativement compact dans $C(K)$.

Théorème. (Riesz – Fréchet – Kolmogorov) Si Ω est un ouvert de \mathbb{R}^n , si $\omega \subset\subset \Omega$ et si F est un sous-ensemble borné de $L^p(\Omega)$, $1 \leq p < \infty$, vérifiant

$$\forall \varepsilon > 0, \exists \delta > 0, \delta < d(\omega, {}^c\Omega), \forall f \in F, \forall h \in \mathbb{R}^n, (\|h\| < \delta \implies \|\tau_h f - f\|_{L^p(\omega)} < \varepsilon)$$

alors F_ω est relativement compact dans $L^p(\omega)$.

3.5 Les espaces de HILBERT

Définition. (Espace de Hilbert) Un espace de HILBERT est un espace vectoriel muni d'un produit scalaire et complet pour la norme engendrée.

Théorème. (Réflexivité) Un espace de HILBERT est uniformément convexe, donc réflexif.

Théorème. (Représentation de Riesz – Fréchet) Toute forme linéaire $\varphi \in H'$ se représente de manière unique sous la forme (f, \cdot) , avec $f \in H$.

Théorème. (Lax – Milgram) Si $a(u, v)$ est une forme bilinéaire, continue et coercive, alors toute forme linéaire $\varphi \in H'$ se représente de manière unique sous la forme $a(u, \cdot)$, avec $u \in H$.

3.6 Espaces de SOBOLEV et problèmes aux limites

Définition. (Espace de Sobolev) $W^{1,p}(I) = \left\{ u \in L^p(I) \mid \exists g \in L^p(I), \forall \phi \in C_c^1(I), \int_I u\phi' = \int_I g\phi \right\}$

Théorème. (Structure des $W^{1,p}$)

$W^{1,p}$ est complet pour $1 \leq p \leq \infty$.

$W^{1,p}$ est séparable pour $1 < p < \infty$.

$W^{1,p}$ est réflexif pour $1 \leq p < \infty$.

$H^1 = W^{1,2}$ est un espace de HILBERT séparable.

Théorème. (Représentation continue) Pour tout $u \in W^{1,p}$, il existe une fonction $\tilde{u} \in C(\bar{I})$, presque partout égale à u et vérifiant

$$\forall x, y \in I, \tilde{u}(x) - \tilde{u}(y) = \int_y^x u'(t) dt$$

Théorème. (Caractérisation des $W^{1,p}$) Pour $1 < p \leq \infty$ et $u \in L^p$,

$$\begin{aligned} & (i) \quad u \in W^{1,p} \\ \iff & (ii) \quad \exists C, \forall \varphi \in C_c^\infty(I), \left| \int u\varphi' \right| \leq C\|\varphi\|_{p'} \\ \iff & (iii) \quad \exists C, \forall \omega \subset\subset I, \forall h \in \mathbb{R}, (|h| < d(\omega, {}^c I) \implies \|\tau_h u - u\|_p \leq C|h|) \end{aligned}$$

Théorème. (Caractérisation de $W^{1,1}$) Dans le cas où I est borné, les fonctions de $W^{1,1}$ sont les fonctions absolument continues, autrement dit les fonctions telles que

$$\forall \varepsilon > 0, \exists \delta > 0, \forall N \in \mathbb{N}, \forall (]a_n, b_n])_n \in P(I)^N, \left(\sum_{n=1}^N |b_n - a_n| < \delta \implies \sum_{n=1}^N |u(b_n) - u(a_n)| < \varepsilon \right)$$

Théorème. (Densité) Pour $1 \leq p < \infty$, toute fonction u de $W^{1,p}$ est approchable au sens de $W^{1,p}$ par des restrictions de fonctions de $C_c^\infty(\mathbb{R})$ à I .

Théorème. (Injections) Il existe une constante C telle que $\|u\|_\infty \leq C\|u\|_{W^{1,p}}$ pour tout $1 \leq p \leq \infty$ et tout $u \in W^{1,p}$. Autrement dit, on a une injection de $W^{1,p}$ dans L^∞ .

$W^{1,p} \subset L^\infty$ est continue pour $1 \leq p \leq \infty$.

$W^{1,p} \subset C(\bar{I})$ est compacte pour $1 < p \leq \infty$.

$W^{1,1} \subset L^q(I)$ est compacte pour $1 \leq q < \infty$.

Chapitre 4

Étude d'articles de recherche

4.1 G. CARLIER, I. EKELAND, *Equilibrium structure of a bidimensional asymmetric city*

Présentation

L'article a pour but d'expliquer la structure interne d'une ville bidimensionnelle dont l'espace urbain est divisé en deux types : des zones résidentielles et des zones d'activité. Ces deux zones peuvent être amenées à évoluer lorsque les divers acteurs du modèles tentent d'optimiser certains critères. La question est de savoir si un équilibre, qui devra être défini, est possible.

La ville compte trois types d'acteurs :

- Les entreprises
Elles veulent attirer beaucoup d'ouvriers, ce qui augmente leur productivité (externalité positive pour le travail), et la production en général.
- Les habitants
Ils doivent se déplacer chaque jour de leur habitation jusqu'à leur lieu de travail, ce qui a un coût, et cherchent donc un travail bien payé qui ne soit pas trop éloigné. Ils veulent précisément optimiser leurs revenus, sachant que leur niveau de bien-être ou *utilité*, qui est fonction concave croissante de l'espace et de leur consommation, est fixée et constante pour tous.
- Les propriétaires terriens
Les deux populations précédentes louent leur terrain à des propriétaires qui ne sont pas présents et observent la loi du plus offrant.

Modélisation mathématique

La ville est située dans un ouvert borné $\Omega \subset \mathbb{R}^2$. On suppose que tous les habitants et toutes les firmes sont identiques : mêmes fonctions de productivité, d'utilité, et de production.

On utilise les notations suivantes :

$\mu(x), \nu(y)$	densités d'habitations et d'entreprises
$z(y)$	productivité des ouvriers travaillant en y
$Z_\nu(y)$	$Z_\nu(y) = g(\int \rho(y, y')\nu(y')dy')$, $g \in C_b$ croissante, $\rho \in C_+(\bar{\Omega} \times \bar{\Omega})$
$U(c, S)$	fonction utilité où c est l'argent disponible pour consommer, S la surface de l'habitation
$N(x), n(y)$	densités d'habitants et d'ouvriers dans les zones résidentielles et d'activité respectivement
$f(z, n)$	fonction de production
$\phi(x), \psi(y)$	revenus des habitants en x , salaires des employés par les firmes en y
$Q(\phi), q(z, \psi)$	prix du m^2 pour les habitants et les firmes (dépendances en ces variables <i>a posteriori</i>)

En liant les quantités précédentes à l'aide notamment de considérations économiques et en traduisant les conditions d'optimalité pour chaque acteur du modèle, on obtient des relations entre ces mêmes quantités, ce qui pousse les auteurs à définir un équilibre comme suit :

Définition. c, Ω, f, U, g, ρ sont donnés avec les hypothèses de 4.1. Un équilibre est un 7-uplet $(\mu, \nu, \gamma, \phi, \psi, z, \theta)$

$$\text{tel que } \left\{ \begin{array}{l} \mu, \nu \in \mathcal{L}^1(\Omega, \mathbb{R}_+), \gamma \text{ mesure Borélienne} \\ \phi, \psi \in \mathcal{C}^0(\Omega, \mathbb{R}_+^*) \\ z \in \mathcal{C}^0(\Omega, \mathbb{R}) \\ \theta \text{ fonction Borélienne à valeurs dans } [0, 1] \end{array} \right. \quad \text{et satisfaisant :}$$

1. $\int \mu = \int \nu$
2. $z = Z_\nu$
3. $\mu(x) = (1 - \theta(x))N(\phi(x))$ et $\nu(y) = \theta(x)n(z(y), \psi(y))$
4. $q(z, \psi) > Q(\phi) \Rightarrow \theta = 1$ et $q(z, \psi) < Q(\phi) \Rightarrow \theta = 0$
5. ϕ et ψ sont conjuguées
6. $\gamma \in \Pi(\mu, \nu)$ et $\phi(x) - \psi(y) = c(x, y) - \gamma \cdot p \cdot p$.

L'existence d'un tel équilibre peut se traiter à l'aide de la théorie du transport optimal. D'où l'un des principaux résultats de l'article :

Théorème (cas de Cobb-Douglas). On suppose Ω convexe, $\gamma, \eta \geq 0$, $(\alpha, \beta) \in [0, 1]^2$ et $\left\{ \begin{array}{l} f(z, n) = z^\gamma n^\alpha \\ U(c, S) = c^\beta S^{1-\beta} \\ c(x, y) = M \cdot d(x, y)^\eta \end{array} \right.$

Alors

1. Cas strictement convexe : si $\eta > 1$ et $\alpha \geq 1/2$ alors il existe au moins un équilibre, et ils sont tous purs (c'est un transport)
2. Cas sous-linéaire : si $2(1 - \alpha) \leq \eta \leq 1$ alors il existe au moins un équilibre.

Les auteurs ont montré que le résultat reste vrai pour des fonctions f, U, c plus générales et sous des hypothèses raisonnables. Ainsi il existe une situation d'équilibre économique de la ville, et il est intéressant de noter que la situation où chacun travaille depuis chez lui n'est pas en général une situation d'équilibre.

4.2 Y. BRENIER *et al.*, *Reconstruction of the early Universe as a convex optimization problem*

La connaissance de l'état de l'Univers à différentes périodes et son évolution sont des problèmes fondamentaux en cosmologie, et il convient de ne pas se limiter uniquement aux informations obtenues à travers l'observation de l'Univers lointain et la seule exploitation des équations d'expansion, qui sont basées sur la connaissance des vitesses d'éloignement, difficilement accessibles. Yann BRENIER *et al.* choisissent dans cet article la voie de la reconstruction de l'histoire et de l'évolution partielle de l'Univers, en partant de la connaissance de l'état actuel de l'Univers et de son état initial, il y a 14 milliards d'années, où la masse était quasi-uniformément répartie. La formulation de l'évolution se trouve être très adaptée à l'application de la théorie du transport à cette reconstruction, mais rien ne garantit son unicité en général, et il faut se placer dans des zones suffisamment petites pour que l'on puisse obtenir des résultats exploitables. L'autre donnée fondamentale qui est soumise à beaucoup d'approximations est la forme initiale de l'univers, que l'on doit connaître relativement précisément pour espérer une reconstruction fiable.

Il s'agit premièrement de formuler le problème de la reconstruction de l'Univers en termes de théorie du transport. Pour cela, nous partons du modèle usuel de l'expansion, régis par les équations du mouvement d'EULER – POISSON dans le cadre du modèle de l'Univers d'EINSTEIN – DE SITTER :

$$\begin{aligned} \partial_\tau v + (v \cdot \nabla_x) v &= -\frac{3}{2\tau}(v + \nabla_x \phi_g) \\ \partial_\tau \rho + \nabla_x \cdot (\rho v) &= 0 \\ \nabla_x^2 \phi_g &= \frac{\rho - 1}{\tau} \end{aligned}$$

Ce modèle impose notamment les conditions initiales de *slaving* :

$$v_i(x) + \nabla_x \phi_{g,i} = 0 \quad \text{et} \quad \rho_i(x) = 1$$

Ces conditions de *slaving* sont fondamentales pour plusieurs points. Elles permettent de ne faire dépendre l'évolution que d'un seul paramètre, par exemple le potentiel gravitationnel ϕ_g , et empêchent également le

mass splitting. Nous simplifions le modèle en nous plaçant dans l'approximation de ZEL'DOVICH, qui est relativement bien vérifiée dans des régions de quelques mégaparsecs de diamètre.

En supposant que la distribution de masse de l'Univers en l'état présent ne possède pas de singularité, la conservation de la masse peut être reformulée, après quelques technicités, par une équation de MONGE – AMPÈRE :

$$\det(\nabla_{x_i} \nabla_{x_j} \Theta(x)) = \rho_0(x)$$

et il est classique de faire le lien entre l'équation de MONGE – AMPÈRE et les problèmes de la théorie du transport, notamment le problème du transport optimal de masse, qui n'est autre que le problème de MONGE – KANTOROVICH.

La discrétisation du problème mène à la recherche d'un transport optimal, problème qui est résolu par des algorithmes efficaces, notamment développés par Dimitri BERTSEKAS, qui ont une complexité en $O(N^3)$. Ces algorithmes sont alors appliqués à la reformulation du problème pour obtenir des reconstructions des états antérieurs proches de l'Univers, avec des résultats satisfaisants et cohérents avec les observations.

Chapitre 5

Simulations numériques : l'algorithme des enchères

Le programme reçoit en entrée les positions des points de départ et d'arrivée, la fonction de coût de transport, et le nombre maximal d'itérations.

À chaque point de départ est placé un acheteur, et à chaque point d'arrivée, un vendeur. On suppose que tous les vendeurs possèdent une unité d'un bien que les acheteurs demandent. Les vendeurs proposent un premier prix (le même pour tous), et les acheteurs déduisent quel est le point de vente le plus avantageux pour eux, coût de transport compris. Les vendeurs qui ont plusieurs clients peuvent alors se permettre d'augmenter leurs prix l'un après l'autre, et les acheteurs concernés suivent les enchères jusqu'à ce qu'un autre point de vente soit plus avantageux pour eux.

L'algorithme s'arrête lorsque chaque vendeur a un unique client : cette correspondance est celle d'un transport optimal.

L'utilisateur indique un nombre maximal d'itérations, mais l'algorithme converge toujours : en effet, si α_i est le prix que l'acheteur i va payer (frais de transport inclus) et β_j le prix annoncé par le vendeur j , alors la quantité $U = \sum_i \alpha_i - \sum_j \beta_j$ est toujours inférieure à la somme de tous les coûts de transport possibles, et croît à chaque fois qu'un vendeur augmente son prix, car il ne le fait que si plusieurs acheteurs peuvent le suivre, et ils augmentent tous leurs enchères d'autant.

Afin d'assurer que U est constant à partir d'un certain rang, on modifie légèrement les positions des points, de façon que les coûts de transport soient tous des multiples entiers d'une même quantité ϵ . A chaque étape, U diminue alors d'au moins ϵ .

La correspondance entre points de départ et points d'arrivée trouvée par l'algorithme est bien celle d'un transport optimal. En effet, les prix payés par les vendeurs et ceux proposés par les acheteurs peuvent être identifiés aux potentiels de KANTOROVICH (à un signe près), donc ils vérifient l'égalité

$$\alpha_i - \beta_j = c_{i,j}$$

sur le support du plan de transfert optimal, où $c_{i,j}$ est le coût de transport entre le point de départ i et le point d'arrivée j .

On observe sur la figure 5.1 le transport trouvé par l'algorithme pour le coût c_1 et pour 40 points de départ et d'arrivée tirés au hasard. Les points de départ sont figurés en vert, et les points d'arrivée en rouge. On observe, comme prévu par la théorie, que les rayons de transport ne se croisent pas.

Sur la figure Fig. 5.2 sont présentés cette fois les résultats trouvés par l'algorithme pour trois coûts différents : les correspondances ne sont pas identiques. Par exemple, les rayons de transport peuvent se croiser dans les cas des coûts c_2 et $c_{1/2}$, où

$$c_2((x, y), (x', y')) = (x - x')^2 + (y - y')^2 \quad c_{1/2}((x, y), (x', y')) = \sqrt{\sqrt{(x - x')^2 + (y - y')^2}}$$

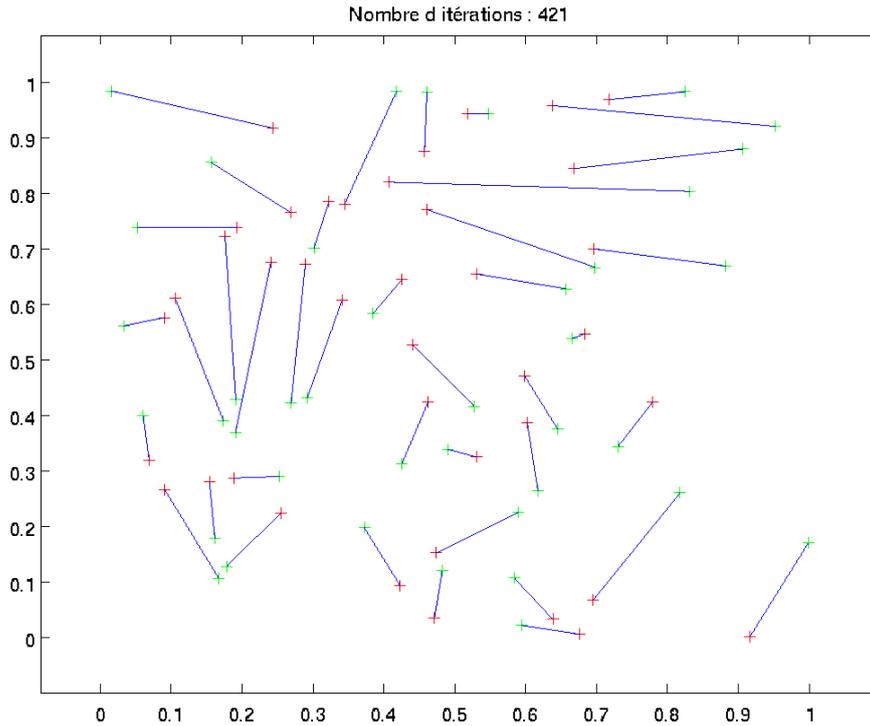


FIGURE 5.1 – 1

Par ailleurs, le coût c_2 est fortement convexe : afin d'éviter les trop grands déplacements, il favorise des déplacements de taille à peu près équivalente pour tous les points. Par contre, le coût $c_{1/2}$ favorise les déplacements très courts, car les trajets longs ne sont pas trop défavorables.

L'utilisateur indique un nombre maximal d'itérations, car le nombre effectif d'itérations nécessaires peut beaucoup varier.

Sur la figure Fig. 5.3 est représenté, sur un échantillon de 200 problèmes d'optimisation tirés au hasard par Matlab, le nombre d'itérations qui lui est nécessaire pour résoudre le problème. Le nombre maximal d'itérations est 800. Pourtant, 30 configurations nécessitent encore plus de 800 itérations pour être résolues.

D'après l'article *Reconstruction of the early Universe*, la complexité de cet algorithme est en $O(n^3)$, où n est le nombre de points à déplacer. Pour cette figure, nous avons étudié des problèmes à 20 points.

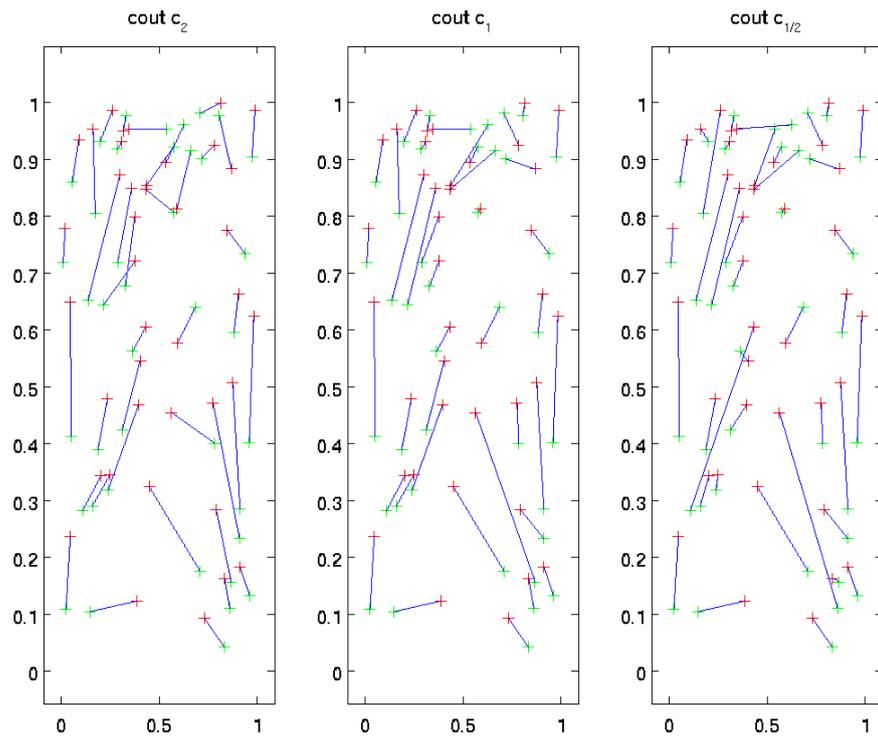


FIGURE 5.2 – 2

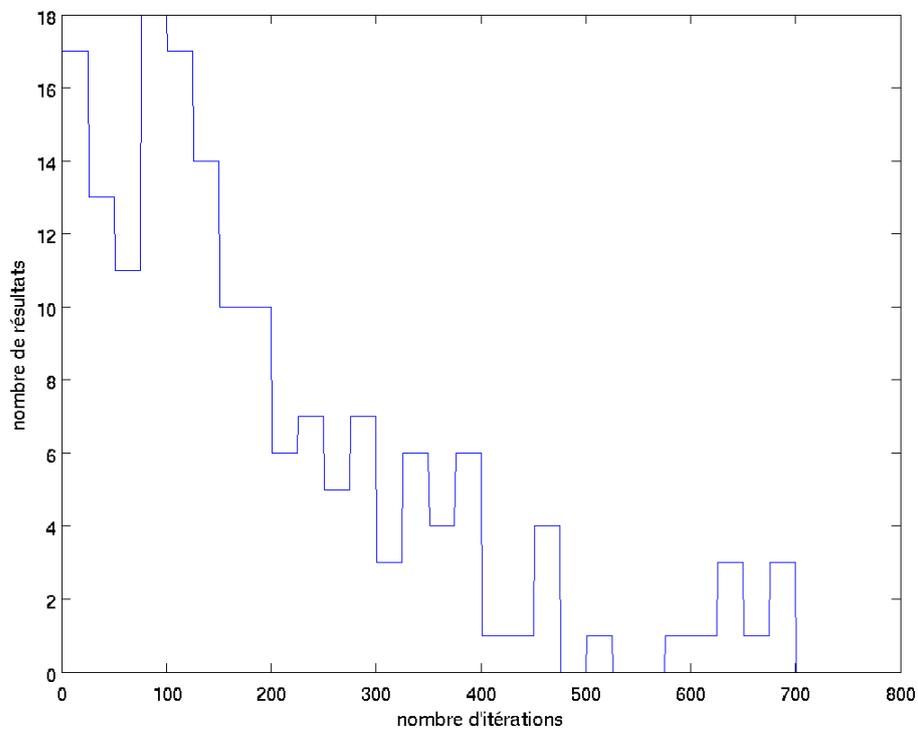


FIGURE 5.3 – 3

Chapitre 6

Colloque de recherche : *Optimal Transport : algorithms and applications*

Dans le cadre de notre découverte des contrées du transport optimal et de ses applications ainsi que dans l'objectif de nous familiariser avec le monde de la recherche, nous avons suivi les conférences du colloque *Optimal Transport : algorithms and applications* à l'Institut Henri POINCARÉ, organisé par l'ANR OTARIE les 7, 8 et 9 juin. Ce fut l'occasion de découvrir le déroulement et l'esprit de ce genre de réunions, et d'écouter des présentations de nombreux sujets de recherche actuels dans la théorie du transport et ses applications en mathématiques, en physique ou en économie. Les thèmes de recherche se révèlent être extrêmement variés, et les points de vue également, comme le révèle la liste des conférences et des conférenciers :

Jean-David BÉNAMOÛ – Projet POEMS, INRIA, Rocquencourt
Unbalanced data and mixed L^2 and Wasserstein distances

Damien BOSC – Department of Economics, Ecole Polytechnique, Palaiseau
Numerical approximation of optimal transport maps

Yann BRENIER – Laboratoire J. A. Dieudonné, Université de Nice, Nice
Optimal transport with concentrations and reconstruction of the early universe

Luigi DE PASCALE – Università di Pisa, Pisa
Existence and duality for L^∞ -optimal transport

Ivar EKELAND – CEREMADE, Université Paris Dauphine, Paris
Problèmes variationnels sous contrainte de convexité : un algorithme et des conjectures

Alfred GALICHON – Department of Economics, Ecole Polytechnique, Palaiseau
Revealed preferences and optimal transport

Wilfrid GANGBO – Georgia Institute of Technology, Atlanta
Particules dynamics with interactions on the real line

Xin Yang LU – Scuola Normale Superiore Pisa, Pisa
Bifurcation behavior for minimizers of average distance functional in irreversible evolutionary processes

Bertrand MAURY – Laboratoire de Mathématiques, Université Paris-Sud 11, Orsay
Crowd motion modeling under density constraints – Part 1

Bruno NAZARET – CEREMADE, Université Paris Dauphine, Paris
Diffusion equations and entropies : a gradient flow approach

Edouard OUDET – Lama, Université de Savoie, Chambéry
Concentration effects in mass transportation

Filippo SANTAMBROGIO – Laboratoire de Mathématiques, Université Paris-Sud 11, Orsay
Crowd motion modeling under density constraints – Part 2

Andrei SOBOLEVSKI – IITP, Moscow & Laboratoire J.-V Poncelet, Moscow
From Bellman to Kardar–Parisi–Zhang : optimization and randomness in stochastic growth models

Eugene STEPANOV – St. Petersburg State University & CVGMT, Pisa
Transportation in metric spaces : transport plans or currents ?

Chapitre 7

Un problème ouvert : Minimisation d'un critère de type oscillation

7.1 Motivation

Nous avons aussi abordé un problème nouveau lié au transport. Au lieu de minimiser le coût du transport, on cherche à minimiser son module de continuité. Ces recherches peuvent s'appliquer dans le domaine de la sécurité et du respect de la vie privée. En effet, les téléphones portables proposent parfois des applications qui permettent de savoir si certains contacts sont à proximité. Pour ce faire, un calculateur externe détermine les distances entre ces téléphones. Mais pour que ce calculateur ne puisse pas remonter à la position du téléphone, il faut modifier la carte. Des transformations rigides, comme des rotations, ne suffisent pas, car en comparant les densités de téléphones et les positions des villes, on peut retrouver la bonne orientation de la carte. Il faut donc transformer la distribution actuelle en une distribution uniforme par exemple, mais pour limiter les calculs inutiles, on exige que des points éloignés l'un de l'autre dans la nouvelle distribution soient aussi éloignés dans la distribution réelle.

Cela revient à minimiser le module de continuité de T évalué en une valeur δ , où T est un transport de la distribution réelle des téléphones vers la distribution uniforme.

7.2 Le problème d'optimisation

Nous nous intéressons ici au problème de minimisation du critère

$$\theta_\delta : T \in \mathcal{T}(\mu, \nu) \longmapsto \sup_{\|x-y\| < \delta} \|T(x) - T(y)\|$$

où $\mathcal{T}(\mu, \nu) = \{T \in \mathcal{B}(\Omega, \Omega) : T\#\mu = \nu\}$

7.3 Reformulation du problème

On considère naturellement une relaxation de ce problème en n'imposant plus nécessairement la minimisation sur les *transports* T , mais sur les *plans de transfert* $\gamma \in \mathcal{P}(\mu, \nu)$. Le critère devient alors

$$\vartheta_\delta : \gamma \in \Pi(\mu, \nu) \longmapsto \sup_{\substack{\|x-x'\| \leq \delta \\ (x,y), (x',y') \in \text{Supp}(\gamma)}} \|x' - y'\|$$

où $\Pi(\mu, \nu) = \{\gamma \in \mathcal{P}(\Omega) : \pi_1\#\gamma = \mu, \pi_2\#\gamma = \nu\}$

Une telle formulation en terme de supremum et d'éléments astreints à être dans le support de γ semble peu adaptée à des manipulations aisées, c'est pourquoi nous préférons la formulation suivante :

$$\vartheta_\delta(\gamma) = \left\| \|\pi_2(a) - \pi_2(b)\| \mathbf{1}_{\|\pi_1(a) - \pi_1(b)\| \leq \delta} \right\|_{L^\infty(\gamma \otimes \gamma)}$$

Dans la suite, toutes les notions topologiques sur $\mathcal{P}(K)$ où K est compact concernent implicitement la topologie faible \star sur le dual des fonctions continues.

7.4 Existence d'un optimum

Théorème 1. Si $\mu, \nu \in \mathcal{P}(\Omega)$, $\delta > 0$, il existe un plan de transfert $\gamma \in \Pi(\mu, \nu)$ minimisant le coût ϑ_δ .

Démonstration. On pose

$$\eta : \rho \in \mathcal{P}(\Omega^4) \mapsto \left\| |\pi_2(a) - \pi_2(b)| \mathbf{1}_{\|\pi_1(a) - \pi_1(b)\| \leq \delta} \right\|_{L^\infty(\rho)}$$

de sorte que $\vartheta_\delta(\gamma) = \eta(\gamma \otimes \gamma)$. Montrons que η est semi-continue inférieurement. On pose

$$\begin{aligned} \eta_p : \rho \in \mathcal{P}(\Omega^4) &\mapsto \left\| |\pi_2(a) - \pi_2(b)| \mathbf{1}_{\|\pi_1(a) - \pi_1(b)\| \leq \delta} \right\|_{L^p(\rho)} \\ \text{et } \lambda : x \in \Omega &\mapsto \mathbf{1}_{\|x\| < \delta} \end{aligned}$$

On peut alors réécrire le critère sous la forme

$$\eta_p^p(\rho) = \int_{\Omega^2 \times \Omega^2} \|\pi_2(a) - \pi_2(b)\|^p \lambda(\pi_1(a) - \pi_1(b))^p d\rho(a, b)$$

λ étant semi-continue inférieurement à valeurs réelles, elle s'écrit comme limite croissante de fonctions continues, donc η_p^p est également limite croissante de fonctions continues, ce qui fait de η_p une fonction semi-continue inférieurement. Par suite η est limite croissante des η_p qui sont semi-continues inférieurement, donc η est aussi semi-continue inférieurement.

On montre maintenant que $\mathcal{P}(\Omega^2) \otimes \mathcal{P}(\Omega^2)$ est un fermé de $\mathcal{P}(\Omega^4)$, donc un compact. Soit $\gamma_k \otimes \gamma'_k \rightharpoonup \rho$. En intégrant des fonctions continues d'une seule variable, on obtient que $\gamma_k \rightharpoonup \pi_1 \# \rho = \gamma$ et $\gamma'_k \rightharpoonup \pi_2 \# \rho = \gamma'$. Montrons que $\rho = \gamma \otimes \gamma'$. Pour toutes fonctions $f, g \in \mathcal{C}^0(\Omega^2, \mathbb{R})$, on a :

$$\int (f \cdot g)(a, b) d(\gamma_k \otimes \gamma'_k)(a, b) = \int f d\gamma_k \cdot \int g d\gamma'_k$$

ce qui donne en faisant tendre k vers l'infini :

$$\int (f \cdot g)(a, b) d\rho(a, b) = \int f d\gamma \cdot \int g d\gamma' = \int (f \cdot g)(a, b) d(\gamma \otimes \gamma')(a, b)$$

On étend ce résultat des fonctions continues à variables séparables sur $\Omega^2 \times \Omega^2$ à toute fonctions de $\mathcal{C}^0(\Omega^2 \times \Omega^2, \mathbb{R})$ à l'aide du

Théorème de STONE – WEIERSTRASS. Si A est une sous-algèbre de $\mathcal{C}^0(X, \mathbb{R})$ telle que la topologie engendrée par A est séparée, i.e. $\forall x \neq y \in X, \exists \alpha \in A \quad \alpha(x) \neq \alpha(y)$, alors A est dense dans $\mathcal{C}^0(X, \mathbb{R})$ pour la topologie uniforme.

Ici, on prends A l'espace vectoriel engendrée par les fonctions continues à variables séparables, qui est bien une sous-algèbre de $\mathcal{C}^0(\Omega^2 \times \Omega^2, \mathbb{R})$ séparant tout couple de points distincts. Or si $f_n \cdot g_n$ converge uniformément vers h , pour tout $\varrho \in \mathcal{P}(\Omega^4)$ on a $\int_{\Omega^2 \times \Omega^2} |f_n \cdot g_n - h| d\varrho \leq \|f_n \cdot g_n - h\|_\infty \rightarrow 0$ donc on obtient $\rho = \gamma \otimes \gamma'$.

Soit alors γ_n minimisant ϑ_δ sur $\Pi(\mu, \nu)$. Quitte à extraire, puisque $\Pi(\mu, \nu)$ est compact, on peut supposer que $\gamma_n \otimes \gamma_n$ converge vers une mesure, qui est de la forme $\gamma \otimes \gamma$. Quitte à extraire, on peut supposer que les γ_n convergent également. Partant, la continuité des projections permet d'affirmer que γ_n converge vers γ , d'où

$$\begin{aligned} \vartheta_\delta(\gamma) &= \eta(\gamma \otimes \gamma) \leq \underline{\lim} \eta(\gamma_n \otimes \gamma_n) = \underline{\lim} \vartheta_\delta(\gamma_n) \\ \text{i.e. } \vartheta_\delta(\gamma) &= \inf_{\Pi(\mu, \nu)} \vartheta_\delta = \min_{\Pi(\mu, \nu)} \vartheta_\delta \end{aligned}$$

□

7.5 Recherche d'un transport dans le cas réel et compact

7.5.1 Paysage de travail

On pose :

$$\begin{aligned} g(x) &= \sup \{y : (x, y) \in \text{Supp}(\gamma)\} \\ f(x) &= \inf \{y : (x, y) \in \text{Supp}(\gamma)\} \end{aligned}$$

Le support de γ est alors contenu dans la *bande* F délimitée dans le plan par les fonctions f et g . Nous cherchons à extraire – ou à construire à partir – de γ un transport optimal, il n'importe guère de conserver exactement le même support, mais il est fondamental de veiller à conserver le coût – qui est minimal pour γ – et les marginales. On suppose donc que F n'a pas de « trous » verticaux : tous les éléments y entre un $f(x)$ et un $g(x)$ appartiennent à F .

7.5.2 Premières idées

La contrainte imposée par le coût, qui est de type L^∞ , porte uniquement sur le support des mesures. On cherche donc, à partir du support d'un γ optimal, à construire un fermé F le contenant et suffisamment grand pour espérer y trouver le graphe d'un transport, mais suffisamment proche de $\text{Supp}(\gamma)$ pour garantir la minimalité du coût. Etant donné un tel fermé F (égal à $\text{Supp}(\gamma)$ à défaut de mieux), on cherche ensuite à trouver un transport respectant les marges et de graphe inclus dans F . Or le théorème de Brenier donne l'existence et l'unicité d'un transport optimisant le coût c_2 , sans cette dernière contrainte. D'où l'idée de minimiser le coût c_2 parmi les plans de transfert dont le support est inclus dans F , dont γ fait partie :

$$c_2 : \gamma \in \Gamma \longmapsto \int_{\Omega \times \Omega} |x - y|^2 d\gamma(x, y)$$

où $\Gamma = \{\gamma \in \Pi(\mu, \nu) : \text{Supp}(\gamma) \subset F\}$

Notons tout d'abord que Γ est un fermé. En effet, si γ_n est une suite convergeant vers γ , cette dernière est dans $\Pi(\mu, \nu)$ qui est déjà connu pour être un fermé, et on a de plus :

$$\forall n \in \mathbb{N}, \quad \gamma_n(F) = 1 \quad \text{i.e.} \quad \int_{\Omega \times \Omega} \mathbf{1}_F d\gamma_n = 1$$

Or, puisque la fonction caractéristique d'un fermé est semi-continue supérieurement, on a

$$1 = \overline{\lim} \int_{\Omega^2} \mathbf{1}_F d\gamma_n \leq \int_{\Omega^2} \mathbf{1}_F d\gamma \leq 1$$

ce qui fait que γ a également sa masse concentrée sur F , soit $\text{Supp}(\gamma) \subset F$, et cela achève la preuve de la fermeture de Γ , qui donc est également compact.

7.5.3 Formulation duale

La semi-continuité de c_2 permet d'affirmer l'existence d'un minimum sur Γ :

$$\inf_{\Gamma} c_2(\gamma) = \min_{\Gamma} c_2(\gamma)$$

Nous cherchons à obtenir une formulation duale à la MONGE–KANTOROVICH. Pour cela, nous explicitons deux fonctions caractéristiques :

$$\begin{aligned} \alpha(\gamma) &= \sup_{\phi, \psi} \left(\int \phi d\mu + \int \psi d\nu - \int \phi \oplus \psi d\gamma \right) = \begin{cases} 0 & \text{si } \gamma \in \Pi(\mu, \nu) \\ +\infty & \text{sinon} \end{cases} \\ \beta(\gamma) &= \sup_{\xi | \xi(F) = \{0\}} \int \xi d\gamma = \begin{cases} 0 & \text{si } \text{Supp}(\gamma) \subset F \\ +\infty & \text{sinon} \end{cases} \end{aligned}$$

Il vient alors, en admettant que le principe de dualité de FENCHEL – ROCKAFELLAR s'applique, *i.e.* que les deux bornes inférieure et supérieure intervenant dans le problème permutent :

$$\begin{aligned}
\inf_{\gamma \in \Gamma} \int c_2 d\gamma &= \inf_{\gamma \in \mathcal{P}(\Omega^2)} \left(\int c_2 d\gamma + \alpha(\gamma) + \beta(\gamma) \right) \\
&= \inf_{\gamma \in \mathcal{P}(\Omega^2)} \sup_{\substack{\phi, \psi, \xi \\ \xi \mathbf{1}_F = 0}} \left(\int (\phi \oplus \psi - c_2 + \xi) d\gamma + \int \phi d\mu + \int \psi d\nu \right) \\
&= \sup_{\substack{\phi, \psi, \xi \\ \xi \mathbf{1}_F = 0}} \left(\int \phi d\mu + \int \psi d\nu + \inf_{\gamma \in \mathcal{P}(\Omega^2)} \left(\int (\phi \oplus \psi - c_2 + \xi) d\gamma \right) \right) \\
&= \sup_{\substack{\phi, \psi \\ \xi | \xi \mathbf{1}_F = 0 \\ \xi \geq \phi \oplus \psi - c_2}} \left(\int \phi d\mu + \int \psi d\nu \right) \\
&= \sup_{(\phi \oplus \psi) \mathbf{1}_F \leq c_2} \left(\int \phi d\mu + \int \psi d\nu \right)
\end{aligned}$$

Nous cherchons maintenant à préciser nos connaissances et les hypothèses que l'on peut imposer sur ϕ et ψ . Prenons (ϕ, ψ) un couple de fonctions admissibles, autrement dit vérifiant

$$\forall (x, y) \in F, \quad \phi(x) + \psi(y) \leq c(x, y)$$

Notons F_x l'ensemble des y tels que (x, y) soit dans F ; et de manière analogue F^y l'ensemble des x tels que (x, y) soit dans F . On considère les transformées

$$\begin{aligned}
\forall (x, y) \in F, \quad \phi(x) &\leq \inf_{y \in F_x} (c(x, y) - \psi(y)) = \phi^c(x) \\
\psi(y) &\leq \inf_{x \in F_y} (c(x, y) - \phi(x)) = \psi^{cc}(y)
\end{aligned}$$

Le nouveau couple (ϕ^c, ψ^{cc}) est également admissible, mais l'on n'obtient pas nécessairement des fonctions plus régulières.

7.5.4 Un théorème d'existence

Théorème 2. Soient Ω un compact de \mathbb{R} , μ et ν des mesures de probabilité sur Ω . Si μ est sans atome, alors il existe un transport optimal $T \in T(\mu, \nu)$ pour le coût θ_δ , où

$$\theta_\delta : T \mapsto \sup_{|x-y| < \delta} |T(x) - T(y)|$$

Preuve. Considérons un plan de transfert $\gamma \in \Pi(\mu, \nu)$ optimal pour le coût ϑ_δ :

$$\vartheta_\delta(\gamma) = \sup_{a, b \in \text{Supp}(\gamma)} \{ |\pi_2(a) - \pi_2(b)| : |\pi_1(a) - \pi_1(b)| < \delta \}$$

On note $K = \vartheta_\delta(\gamma)$ le coût optimal, et on pose f et g les deux fonctions délimitant la plus petite bande convexe selon y contenant le support de γ :

$$\begin{aligned}
g(x) &= \sup \{ y : (x, y) \in \text{Supp}(\gamma) \} \\
f(x) &= \inf \{ y : (x, y) \in \text{Supp}(\gamma) \}
\end{aligned}$$

Ainsi, la bande $[f, g] = \{(x, y) : f(x) \leq y \leq g(x)\}$ vérifie les propriétés suivantes :

$$\text{Supp}(\gamma) \subset [f, g] \tag{7.1a}$$

$$\forall (x, y), (x', y') \in [f, g], \quad (|x - x'| < \delta \implies |y - y'| \leq K) \tag{7.1b}$$

Les fonctions f et g sont *a priori* seulement mesurables, et nous cherchons naturellement à les remplacer par des fonctions plus régulières, définies sur un domaine plus commode et vérifiant les propriétés (1). Il sera plus aisé de travailler avec de telles fonctions, qui définissent une nouvelle bande, et nous prouvons qu'il existe un transport de $T(\mu, \nu)$ dans cette bande, qui est optimal pour θ_δ .

Puisque $|f(x) - g(y)| \leq K$ pour $|x - y| < \delta$, on a nécessairement

$$f(x) \leq \sup_{|y-x|<\delta} g(y) - K$$

$$g(x) \geq \inf_{|y-x|<\delta} f(y) + K$$

On définit alors naturellement les transformées d'une fonction $\phi : X \rightarrow \Omega$, définies sur une partie X de Ω , comme les fonctions délimitant la bande la plus large possible pour garder un coût inférieur à K si ϕ est l'un des bords de la bande :

$$\phi^\downarrow(x) = \sup_{y \in X, |y-x|<\delta} \phi(y) - K \quad (7.2a)$$

$$\phi^\uparrow(x) = \inf_{y \in X, |y-x|<\delta} \phi(y) + K \quad (7.2b)$$

Par construction, la nouvelle bande $[f^{\uparrow\downarrow}, f^\uparrow]$ délimite un espace pour le support de mesure n'ayant pas un coût excédant K . Il est aisé de constater que f^\uparrow est supérieure à g et que $f^{\uparrow\downarrow}$ est inférieure à f , ce qui prouve que l'objectif d'élargir la bande initiale est atteint, et donc que le support de γ est également inclus dans cette bande. On remplace alors $[f, g]$ par $[f^{\uparrow\downarrow}, f^\uparrow]$, que l'on note toujours $[f, g]$. Afin d'obtenir des fonctions définies sur un « meilleur » ensemble que X , on remplace préalablement f par la transformée g^\downarrow de g , définie par la formule (7.2b) sur tout $Y = X + [-\delta/2, +\delta/2]$, qui est une réunion finie de segment $(S_i)_{i=1\dots N}$.

L'existence de transports $T_i : S_i \rightarrow \Omega$ dans $T(\pi_1 \# \gamma|_{S_i \times \Omega}, \pi_2 \# \gamma|_{S_i \times \Omega})$ dont le graphe est inclus dans la bande $[f, g]$ permet de construire un transport $T \in T(\mu, \nu)$ restant dans la bande, de sorte qu'on se limite au cas où Y est un segment non trivial S .

On étudie maintenant les propriétés des transformées définies précédemment, pour vérifier que la transformation de la bande a les effets régularisants désirés.

Définition 1. Soit $\phi, \psi : S \rightarrow \Omega$. On dit que (ϕ, ψ) est un couple conjugué si $\phi^\uparrow = \psi$ et $\psi^\downarrow = \phi$. On vérifie facilement que $\phi^{\uparrow\downarrow} = \phi^\uparrow$, i.e. le couple $(\phi^{\uparrow\downarrow}, \phi^\uparrow)$ est conjugué.

Lemme 1. Soit $f : S \rightarrow \Omega$ mesurable. Alors f^\uparrow et f^\downarrow sont réglées, donc continues \mathcal{L} -presque partout, et respectivement semi-continue supérieurement et semi-continue inférieurement.

Preuve. Puisque f est mesurable et bornée, elle s'écrit comme limite uniforme de fonctions étagées f_n , que l'on peut choisir croissante. Il convient alors d'étudier l'effet de la transformation sur les fonctions étagées. Soit une fonction étagée

$$\phi = \sum_{i=1}^N \alpha_i \mathbf{1}_{A_i} \quad \text{où} \quad \alpha_1 < \dots < \alpha_N$$

Nous construisons algorithmiquement sa transformée ϕ^\uparrow . Celle-ci étant définie par l'infimum de la fonction sur la boule de rayon δ autour du point considéré, elle va valoir α_1 dès que ϕ vaut α_1 en un point proche de δ , donc elle vaut α_1 sur la trace sur Ω de l'ensemble $A_1 +]-\delta, +\delta[=: A_1^\uparrow$ des tels points, qui est de ce fait une réunion finie d'intervalles de longueur supérieure à 2δ , sauf éventuellement pour les deux intervalles situés aux bords. De la même manière, elle vaut α_2 sur la trace de $A_2^\uparrow := (A_2 +]-\delta, +\delta[) \setminus A_1^\uparrow$ sur Ω , etc. Plus généralement, ϕ^\uparrow vaut α_k sur la trace sur Ω de $A_k^\uparrow := (A_k +]-\delta, +\delta[) \setminus \left(\bigcup_{i=1}^{k-1} A_i^\uparrow \right)$, qui sont tous des réunions d'intervalles.

De plus, $(f_n)_n$ étant supposée croître vers f et la transformée étant croissante, si f_n est proche de f à $\frac{1}{n}$ près, on a :

$$f - 1/n \leq f_n \leq f \quad \text{donc} \quad (f - 1/n)^\uparrow = f^\uparrow - 1/n \leq f_n^\uparrow \leq f^\uparrow$$

Ce qui prouve que f^\uparrow est limite croissante et uniforme des f_n^\uparrow , qui sont en escalier : c'est donc une fonction réglée, donc une fonction continue \mathcal{L} -presque partout. Le même raisonnement s'applique à f^\downarrow . Notons que le lemme 3 qui suit, dont la preuve est indépendante du résultat qui vient d'être faite, prouve que le nombre de points de discontinuité des transformées des f_n est uniformément borné, et cela prouve en particulier que le nombre de points de discontinuité de la limite f est non seulement négligeable, mais dénombrable.

Montrons maintenant que f^\uparrow est semi-continue supérieurement. Considérons une suite $(x_n)_n$ convergeant vers x . Puisque $f^\uparrow(x) = \inf_{]x-\delta, x+\delta[} f + K$, si $a \in]x - \delta, x + \delta[$, pour n assez grand $a \in]x_n - \delta, x_n + \delta[$, donc $f^\uparrow(x_n) \leq f(a) + K$, d'où

$$\overline{\lim} f^\uparrow(x_n) \leq f(a) + K \quad \text{donc} \quad \overline{\lim} f^\uparrow(x_n) \leq \inf_{a \in]x-\delta, x+\delta[} f(a) + K = f^\uparrow(x)$$

Ce qui prouve la semi-continuité supérieure de f^\uparrow . Comme $f^\downarrow = -(-f)^\uparrow$, elle est semi-continue inférieurement. \square

Lemme 2. Soient $f, g : S \rightarrow \Omega$ en escalier et conjuguées, g s'écrivant $\sum_{j=1 \dots N} \alpha_j \mathbf{1}_{I_j}$ où les I_j sont des intervalles partitionnant I rangés par ordre croissant, et les α_j sont consécutivement distincts. Alors pour tout triplet $(\alpha_{k-1}, \alpha_k, \alpha_{k+1})$ tel que $\alpha_{k-1}, \alpha_{k+1} > \alpha_k$, I_k a une longueur au moins égale à 2δ .

Preuve. La fonction g étant semi-continue supérieurement, les intervalles sont de la forme $I_{k-1} =]a, b[$, $I_k =]b, c[$ et $I_{k+1} =]c, d[$, où la notation $]a, b[$ signifie que la borne en a peut être ouverte ou fermée. Tout d'abord, c doit être à une distance supérieure ou égale à δ du bord gauche de I , sinon $g(x)$, qui est l'infimum de f sur $]x - \delta, x + \delta[$, serait monotone sur $]a, d[$. Soit alors α la valeur prise par g sur I_k , et $\varepsilon > 0$ assez petit pour que $c - \varepsilon$ soit dans $]b, c[$. On a $g(c) > g(c - \varepsilon) = \alpha$, donc $\alpha = \inf_{]c-\varepsilon-\delta, c-\delta[} f + K$, ce qui donne par semi-continuité inférieure de f , en faisant tendre ε vers 0, $\alpha \geq f(c - \delta) + K$. D'où $g(x) \leq f(c - \delta) + K = \alpha$ sur $S \cap]c - 2\delta, c[$, ce qui implique $]b, c[\supset]c - 2\delta, c[$. \square

Ce résultat permet de contrôler *uniformément* le nombre de changements de monotonie de g . Le même résultat pour les maxima tient pour f .

Lemme 3. Pour f, g vérifiant les hypothèses du lemme 2, il existe une subdivision de taille bornée indépendamment de f et g , telle que sur chacun de ses intervalles, f et g aient la même monotonie.

Preuve. Conservons les notations du lemme 2 :

$$g = \sum_{j=1 \dots N} \alpha_j \mathbf{1}_{I_j}$$

On regroupe les intervalles I_j en un nombre fini de réunions connexes U_k construites comme suit : deux intervalles S et T appartiennent à la même réunion si et seulement s'ils sont adjacents et de même forme $]a, b[$ ou $]a, b[$, à l'exception des deux intervalles aux bords. Cela revient à regrouper les intervalles en portions respectivement croissantes et décroissantes. Reste les cas des intervalles du type $]a, b[$ et $]a, b[$, qui correspondent à des plateaux extrémaux. On place les intervalles du type $]a, b[$ – les plateaux maximaux – dans la réunion U_k croissante en contact avec a . Les intervalles $]a, b[$ – les plateaux minimaux –, de longueur au moins égale à 2δ , sont découpés en deux parties $]a, b - \delta[$ et $]b - \delta, b[$ qu'on place respectivement dans le U_k décroissant contenant a et le U_l croissant contenant b . Quant aux intervalles aux bords, ou bien ils respectent la monotonie de la réunion U_k qui leur est adjacente et on les y ajoute, ou ils constituent chacun un nouveau U_k .

On obtient naturellement un ensemble fini d'intervalles U_k sur lesquels g est monotone. Leur nombre M est borné indépendamment de g grâce à la longueur minimale de 2δ des plateaux minimaux :

$$M \leq 1 + \lfloor \frac{l}{\delta} \rfloor \quad \text{où} \quad l = |S|$$

Montrons maintenant que sur chacun de ces U_k , $k \in 1 \dots M$, la fonction $f : x \mapsto \inf_{]x-\delta, x+\delta[\cap S} g$ est monotone de même sens que g . Si g est croissante sur un $U_k =]a, b[$ qui n'est pas en contact avec le bord, alors g est croissante sur $]a + \delta, b[$, puisque U_k contient d'une part une portion $]a, a + \delta[$ obtenue par découpage d'un plateau minimal et est d'autre part suivi de U_{k+1} décroissant, qui ou bien contient $]b, b + \delta[$ (inclus dans la portion gauche du plateau minimal suivant) ou atteint le bord droit, ce qui garantit encore la croissance de f jusque b compris. Le cas des bords et des U_k décroissants se traitent de manière analogue grâce au bon découpage formé par les U_k . \square

Lemme 4. Si $f, g : S \rightarrow \Omega$ sont mesurables et conjuguées, alors il existe une subdivision de S en nombre uniformément borné d'intervalles telle que sur chacun de ses intervalles, f et g aient la même monotonie.

Preuve. Soit f_n une suite de fonctions étagées convergeant uniformément vers f , on a vu qu'alors f_n^\uparrow converge uniformément vers $f^\uparrow = g$ et $f_n^{\uparrow\downarrow}$ converge uniformément vers $f^{\uparrow\downarrow} = f$. Quitte à renommer $f_n^{\uparrow\downarrow}$ en f_n et f_n^\uparrow en g_n on a donc deux suites f_n et g_n conjuguées et continues par morceaux qui convergent uniformément

vers f et g respectivement. D'après le lemme 3, pour tout n , il existe une partition $(U_k^n)_{k=1\dots M_n}$ de S (rangés par ordre croissant) telle que f_n et g_n sont simultanément croissantes ou décroissantes sur chaque U_k^n . Quitte à extraire, M_n étant une suite d'entiers bornée par M , on peut supposer qu'elle est constante. Pour chaque $k = 1 \dots M$ on extrait à nouveau pour que les bornes de U_k^n convergent, puis on extrait une dernière fois de sorte que pour chaque k , la monotonie reste constante sur la suite d'intervalles U_k^n lorsque n varie. On vérifie que sur chaque $U_k =]a_k, a_{k+1}[$, f et g ont même monotonie. En effet, pour tous x et y dans U_k tels que $x < y$, pour n assez grand, $x, y \in U_k^n$ et en supposant par exemple que f_n et g_n sont croissantes sur U_k^n , donc que $f_n(x) \leq f_n(y)$ et $g_n(x) \leq g_n(y)$ on obtient naturellement que f et g sont croissantes sur U_k par passage à la limite. \square

Lemme 5. *Soit $f, g : X =]a, b[\rightarrow Y =]c, d[$ sont deux fonctions mesurables, monotones de même monotonie et telles que $f \leq g$. On suppose de plus qu'il existe $\gamma \in \Pi(\mu, \nu)$ où $\mu \in \mathcal{P}(X), \nu \in \mathcal{P}(Y)$ tel que $\text{Supp}(\gamma) \in [f, g]$ et $T \in T(\mu, \nu)$ de même monotonie que f . Alors le graphe de T est μ -presque partout inclus dans $[f, g]$.*

Preuve. On traite par exemple le cas où f, g sont croissantes. Notons γ_x la désintégration de γ selon π_1 . Pour tout t , on a $\gamma_x(]c, t]) \leq \delta_{f(x)}(]c, t])$ par définition de f , donc :

$$\begin{aligned} \forall t \in X \quad \int_X \gamma_x(]c, t]) d\mu(x) &\leq \int_X \delta_{f(x)}(]c, t]) d\mu(x) \\ \text{i.e.} \quad \nu(]c, t]) &\leq \mu\{x : f(x) \leq t\} \\ \text{i.e.} \quad \mu\{x : T(x) \leq t\} &\leq \mu\{x : f(x) \leq t\} \end{aligned}$$

Ceci signifie en particulier que pour tout $t \in Y$, $\mu\{x : T(x) \leq t < f(x)\} = 0$. En considérant une suite t_n dense dans Y , il vient

$$\mu\{x : T(x) < f(x)\} = \mu\left(\bigcup_{n \in \mathbb{N}} \{x : T(x) \leq t_n < f(x)\}\right) = 0$$

Donc $f \leq T$ μ -presque partout. En appliquant un raisonnement analogue à g on obtient que T est un transport dans $T(\mu, \nu)$ dont le graphe est μ -presque partout inclus dans $[f, g]$. \square

Le lemme 4 prouve que l'on a une subdivision de S en M intervalles $X_1 = [a_1, a_2], \dots, X_M = [a_M, a_{M+1}]$ sur lesquels f et g sont simultanément de même monotonie. On pose $Y_j = [f(a_j), g(a_{j+1})]$, $\mu_j = \mu|_{X_j}$ et $\nu_j = \nu|_{Y_j}$ et $\gamma_j = \gamma|_{X_j \times Y_j}$ de sorte que $\gamma_j \in \Pi(\mu_j, \nu_j)$ ait son support inclus dans la bande $[f, g]$ restreinte à $X_j \times Y_j$. Les μ_j étant absolument continues par rapport à la mesure de Lebesgue, il existe des transports $T_j \in T(\mu_j, \nu_j)$ de même monotonie que f, g . Le lemme 5 prouve l'existence de transports dans $T(\mu_j, \nu_j)$ dont les graphes sont inclus dans $[f, g]$, que l'on notera toujours T_j . En posant $T(x) = T_j(x)$ sur $[a_j, a_{j+1}[$, $T(a_M) = T_M(a_{M+1})$, on vérifie qu'on obtient un transport dans $T(\mu, \nu)$ dont le graphe est inclus dans la bande $[f, g]$ ce qui implique $\theta_\delta(T) = K$. Il s'agit donc d'un transport optimal. \square

Chapitre 8

Références

Théorie générale du transport et applications

- [1] L. Ambrosio, *Lecture notes*, « Mathematical aspects of evolving interfaces », 2000.
- [2] Ambrosio, Gligli, Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Birkhäuser, 2005.
- [3] C. Villani, *Topics in optimal transportation*, AMS, 2003.
- [4] C. Villani, *Optimal transport : old and new*, Springer, 2009.

Articles de recherche en théorie du transport

- [5] Benamou, Brenier, *A computational fluid mechanics solutions to the Monge – Kantorovich mass transfer problem*, 2001.
- [6] Brenier *et al.*, *Reconstruction of the early Universe as a convex optimization problem*, 2008.
- [7] Carlier, Ekeland, *Equilibrium structure of a bidimensional asymmetric city*, 2005.
- [8] F. Santambrogio, *Absolute continuity and summability of transport densities : simpler proofs and new estimates*, 2009.

Fondements théoriques et outils

- [9] H. Brézis, *Analyse Fonctionnelle*, McGraw-Hill, Inc., 1987.
- [10] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.
- [11] Rockafellar, Wets, *Variational Analysis*, Springer, 2009.
- [12] W. Rudin, *Real and Complex Analysis*, Dunod, 1999.