Transport de mesures sur un espace d'Alexandrov

Jérôme Bertrand

L'objet de cette note est de présenter quelques résultats concernant le transport optimal de mesures sur un espace métrique à courbure minorée (au sens d'Alexandrov). La principale motivation de cette étude repose sur des travaux récents (et indépendants) de Lott-Villani et de Sturm [4, 5, 9, 10]. L'objet de ces articles est de définir une « bonne » notion de courbure de Ricci minorée sur un espace métrique mesuré. Ces auteurs proposent une définition (en fait, les définitions sont légèrement différentes mais l'idée est la même) s'appuyant sur la théorie du transport optimal de mesures. Il se trouve qu'un espace métrique d'Alexandrov de dimension (de Hausdorff) finie admet une mesure qui lui est naturellement associée, la mesure de Hausdorff correspondante. La notion d'espace d'Alexandrov étant une extension naturelle aux espaces métriques des variétés riemanniennes de courbure sectionnelle minorée, il est légitime de se demander s'ils vérifient cette notion de courbure de Ricci minorée généralisée.

La présentation des résultats dans cet article est volontairement informelle et donc parfois approximative. Le lecteur sourcilleux pourra consulter l'article [2] pour les résultats concernant le transport optimal sur un espace Alexandrov, les articles [4, 5, 9, 10] ainsi que l'article de synthèse de J. Lott [6] pour la notion de courbure de Ricci minorée sur un espace métrique mesuré et ses propriétés. Nous renvoyons au livre de C. Villani [11] pour tout ce qui concerne le transport de mesures ; signalons également l'ouvrage [12] du même auteur qui traite de multiples aspects du transport de mesures, y compris la notion généralisée de courbure de Ricci minorée. Les résultats de [2] dans le cas d'une variété riemannienne ont été démontrés par R.J. McCann [7].

1 Transport optimal de mesures

La théorie du transport optimal de mesures englobe (au moins) deux questions distinctes : le problème de Monge (historiquement le premier) et le problème de Kantorovich.

Definition 1.1 *Soit (X, d) un espace métrique et $\mu, \nu \in \mathcal{P}(X)$ deux mesures de probabilités boréliennes sur X. Soit $c : X \times X \longrightarrow \mathbb{R}^+$ une fonction continue.*

Le problème de Monge consiste à étudier les minimiseurs éventuels de la fonctionnelle :

$$\int_{X\times X} c(x, T(x))d\mu(x)$$

parmi toutes les applications mesurables de X dans X telles que la mesure image de μ par T vérifie $T_\sharp\mu = \nu$. Un telle application est appelée « application de transport » de μ sur ν.

Le problème de Kantorovich consiste lui à étudier les minimiseurs éventuels de :

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

parmi toutes les mesures π appartenant à $\Pi(\mu, \nu)$ où $\Pi(\mu, \nu)$ désigne l'ensembles des mesures de probabilités boréliennes sur $X \times X$ dont les marginales sont μ et ν. Un élément de cet ensemble est appelé « plan de transport » entre μ et ν.

Les deux problèmes ont en commun de chercher à minimiser un coût moyen pour des mesures μ et ν données. La différence est que dans le cas du problème de Monge, on s'interdit de « diviser la masse » localisée en un point alors que cela est permis dans le cas du problème de Kantorovich. A une application de transport T correspond un plan de transport donné par la relation $(Id, T)_\sharp \mu$. L'infimum de la fonctionnelle du problème de Monge est par conséquent toujours supérieur ou égal à celui du problème de Kantorovich.

Remarque 1.2 *Les deux problèmes sont en réalité de nature assez différente. Par exemple, il existe toujours au moins un plan de transport (la mesure produit) alors qu'il n'y a pas, en général, d'application de transport entre deux mesures données. Signalons également qu'il est possible prouver l'existence de plans de transport optimaux sous des hypothèses peu contraignantes, voir par exemple, le début du paragraphe 3.*

2 Notion de courbure de Ricci minorée sur un espace métrique mesuré

De nombreux invariants riemanniens peuvent être bornés à l'aide, entre autres, d'un minorant de la courbure de Ricci de la variété, voir, par exemple, les articles de revue [14, 13]. Dans la plupart de ces résultats, la dimension de la variété intervient également dans les estimations et celles-ci sont uniformes lorsque l'on suppose la dimension majorée par une constante donnée. Cette propriété a été étendue au cas des variétés riemanniennes munies d'une forme volume à densité, on parle alors d'une condition de courbure-dimension, celle-ci porte sur un tenseur de Ricci modifié (tenseur de Bakry-Emery) [1].

C'est évidemment cette condition de courbure-dimension que l'on peut espérer étendre à des espaces métriques mesurés généraux. Cependant sur un espace d'Alexandrov, le problème de la dimension ne se pose pas puisqu'il a été démontré que la dimension de Hausdorff est constante (*i.e.* cette dimension est constante sur tout ouvert de l'espace et est entière, dès que l'on suppose que la dimension de Hausdorff de tout l'espace est finie [3]). Nous nous bornerons donc à donner une définition n'incluant pas de majorant de la dimension (condition du type $CD(K, \infty)$), celle-ci étant notablement plus simple à énoncer que la définition générale.

Definition 2.1 (Espace de Wassertein) *Soit (X, d) un espace métrique complet séparable et $\mathcal{P}_2(X)$ l'espace des mesures de probabilités sur X admettant un moment d'ordre deux. Muni de*

$$d_W(\mu, \nu) = \left(\min_{\pi \in \Pi(\mu,\nu))} \int_{X \times X} d^2(x, y) d\pi(x, y) \right)^{\frac{1}{2}},$$

$(\mathcal{P}_2(X), d_W)$ est un espace métrique de longueurs complet et séparable, appelé espace de Wassertein (quadratique).

À l'aide de cette définition/proposition, nous pouvons maintenant définir la notion d'espace métrique mesuré à courbure de Ricci minorée par un réel K.

Definition 2.2 (Condition $CD(K, \infty)$) *Soit (X, d, μ) un espace métrique mesuré et $P_2^{ac}(X)$ l'espace des mesures de probabilités admettant un moment d'ordre deux et absolument continues par rapport à la mesure μ. L'espace métrique mesuré (X, d, μ) vérifie $CD(K, \infty)$ si, pour toute mesure ν_0, ν_1 appartenant à $P_2^{ac}(X)$, il existe une géodésique $(\nu_t)_{t \in [0,1]}$ pour la distance de Wassertein, incluse dans $P_2^{ac}(X)$ et telle que, pour tout $t \in [0, 1]$:*

$$Ent_\mu(\nu_t) \leq (1 - t) Ent_\mu(\nu_0) + t\, Ent_\mu(\nu_1) - \frac{K}{2} t(1 - t) d_W^2(\nu_0, \nu_1),$$

où Ent_μ désigne l'entropie relative à μ, définie pour $\nu \in P_2^{ac}(X)$ par la formule

$$Ent_\mu(\nu) = \int_X h \log h d\mu$$

avec h la densité de ν par rapport à μ.

Remarque 2.3 *Notez que la définition requiert que la condition soit satisfaite pour UNE géodésique de l'espace de Wassertein. Cette condition semble importante (sinon nécessaire) pour disposer de la propriété de stabilité évoquée ci-dessous. En d'autres termes, il s'agit d'une condition de K-convexité géodésique (faible) de l'entropie sur l'espace de Wassertein.*

Il est possible de montrer que cette définition coïncide avec la définition usuelle dans le cas d'une variété riemannienne munie de la distance riemannienne et du volume riemannien (éventuellement couplé à une densité, elle est alors équivalente à une hypothèse de courbure-dimension). La seconde propriété importante de cette notion est qu'elle est stable par passage à la limite lorsque l'on considère une suite d'espaces métriques mesurés convergeant pour la distance de Gromov-Hausdorff mesurée.

Afin d'établir un lien entre l'existence d'application de transport optimale et la propriété $CD(K, \infty)$, esquissons la preuve du fait qu'une variété riemannienne de courbure de Ricci minorée par K vérifie la condition $CD(K, \infty)$. La première étape est justement de montrer que le transport d'une mesure absolument continue par rapport au volume riemannien sur une autre mesure est donnée par une application de transport optimal $T(x) = exp\,(-\nabla\phi(x))$. On considère ensuite les applications interpolées $T_t(x) = exp\,(-t\nabla\phi(x))$; il s'agit de montrer qu'elles sont assez régulières pour pouvoir leur appliquer la formule du changement de variables. Le lien avec la courbure de Ricci se fait alors par l'intermédiaire du déterminant de ces applications. En effet, par définition de T et pour tout vecteur tangent u, l'application $t \longrightarrow dT_t(u)$ est un champ de Jacobi le long de la géodésique reliant x et $T(x)$. Le jacobien de ces applications (et par conséquent la densité des mesures interpolées) vérifie donc une équation différentielle (du type Sturm-Liouville) faisant intervenir la courbure de Ricci. L'objet du paragraphe suivant est d'expliquer comment on peut étendre aux espaces d'Alexandrov le résultat d'existence (et d'unicité) d'une application de transport optimale ainsi que l'expression explicite de cette application.

3 Existence d'une application de transport optimale

La première étape pour montrer l'existence et l'unicité d'une application de transport optimale est l'utilisation de « la dualité de Kantorovich ». Pour le coût quadratique, le résultat s'énonce comme suit :

Théorème 3.1 (Dualité de Kantorovich) *Soit (X, d) un espace métrique complet et séparable, μ, ν deux mesures de probabilités sur X telles que*

$$\int_{X\times X} d^2(x,y)\ d\mu(x)d\nu(y) < +\infty.$$

Alors,

$$\max_{(\psi,\phi)\in E_c} \int_X \phi(x)d\mu(x) + \int_X \psi(y)d\nu(y) = \min_{\pi\in\Pi(\mu,\nu)} \int_{X\times X} d^2(x,y)d\pi(x,y)$$

où $E_c = \{(\phi,\psi) \in L^1(\mu) \times L^1(\nu);\ \phi(x) + \psi(y) \leq d^2(x,y)\ \mu\otimes\nu\, p.p.\}$.

La fonctionnelle apparaîssant dans le membre de gauche ci-dessus est appelé problème dual de Kantorovich. Des fonctions (ϕ, ψ) qui maximisent cette fonctionnelle sont appelés potentiels de Kantorovich. Il est facile de constater que la masse de tout plan de transport optimal est incluse dans l'ensemble :

$$\partial_c \phi := \{(x, y); \phi(x) + \psi(y) = d^2(x, y)\}$$

où (ϕ, ψ) sont des potentiels de Kantorovich.

La stratégie pour en déduire qu'il existe une unique application de transport optimale est de montrer que l'ensemble $\partial_c \phi$ vérifie la propriété :

$$\text{Pour } \mu \text{ presque tout } x, \text{ il existe un unique } y \text{ tel que } \phi(x) + \psi(y) = d^2(x, y). \tag{1}$$

En effet, en notant $T(x)$ l'unique (μ p.p.) y vérifiant la condition ci-dessus, il est assez facile de voir que T est mesurable et vérifie $T_\sharp \mu = \nu$. Autrement dit, tout plan de transport optimal est induit par une application de transport optimale et celle-ci est unique, définie par la propriété (1).

Formellement, la preuve de la proprété (1) est très simple.
Notons $H_y(z) = d^2(z, y) - \phi(z)$ et considérons $(x, y) \in \partial_c \phi$. L'hypothèse $(\phi, \psi) \in E_c$ entraîne que H_y admet un minimum en x ; par conséquent le gradient de H est nul en ce point et cela donne

$$2d(x, y)\nabla d_y(x) = \nabla \phi(x).$$

Cela détermine y de manière unique.

La difficulté (au moins du point de vue de l'auteur) est de rendre cet argument rigoureux. Cela est possible lorsque la mesure μ est absolument continue par rapport à la mesure de Hausdorff. Le principal outil technique est l'existence d'une structure de variétés sur une partie dense et de mesure pleine de X, notée $Reg(X)$. Signalons que cet ensemble ne possède pas, en général, de propriétés topologiques sympathiques. Dans [8], les auteurs fournissent un exemple où son complémentaire est dense. L'énoncé, découlant de résultats de [3] et de [8], est le suivant.

Théorème 3.2 (Cartes sur $Reg(X)$) *Il existe un atlas $(\phi, U_\phi)_{\phi\in\Phi}$ recouvrant $Reg(X)$, une partie dense et de mesure pleine de X. Plus précisément, pour tout $\phi \in \Phi$, $\phi : U_\phi \mapsto \mathbb{R}^n$ est un homéomorphisme bilispchitzien défini sur un ouvert U_ϕ de X. De plus, il existe une partie dense et de mesure pleine V_ϕ de U_ϕ telle que :*

$$\bigcup_{\phi\in\Phi} V_\phi \supset Reg(X).$$

Enfin, si $\phi, \psi \in \Phi$ sont tels que $U_\phi \cap U_\psi \neq \emptyset$, $\psi \circ \phi^{-1}$ est continûment différentiable sur $\phi(V_\phi \cap V_\psi \cap Reg(X))$.

Ce théorème permet de définir la notion de fonction différentiable en un point de $Reg(X)$ (on choisit une carte ϕ telle que $x \in V_\phi$ et on utilise cette carte pour se ramener à la définition usuelle), donc preque partout. Cet énoncé donne également, à peu de choses près, l'existence d'une métrique riemannienne continue sur $Reg(X)$ (voir [8] ou [2] pour plus de détails). Par conséquent, la notion de gradient d'une fonction est bien définie presque partout.

Le point suivant est de prouver qu'un potentiel de Kantorovich est une fonction différentiable presque partout. Il est facile de voir que c'est une fonction lipschitzienne (car la fonction coût est lipschitzienne). Autrement dit, il faut vérifier que le théorème de Rademacher s'applique aux espaces d'Alexandrov.

Corollaire 3.3 (Théorème de Rademacher) *Une fonction lipschitzienne sur un espace d'Alexandrov de dimension finie est différentiable presque partout.*

Preuve : Soit f une fonction lispchitzienne et ϕ une carte. Par construction, l'application $f \circ \phi^{-1}$ est une application lischitzienne définie sur un ouvert de $\mathbb{R}^n$ donc le théorème de Rademacher usuel s'applique. On conclut en utilisant de nouveau le caractère bilipschitz de ϕ et le fait que ϕ soit différentiable sur V_ϕ donc presque partout. ∎

Passons à la preuve du lemme. La démonstration originale de McCann dans le cadre d'une variété riemannienne utilise de manière cruciale la régularité de l'application exponentielle. Pour pouvoir s'affranchir de cette régularité, il convient de remarquer que l'image de presque tout point x par une application optimale T est un point « régulier » de la fonction distance au point x au sens où, même si $T(x)$ appartient au Cut Locus de x (défini comme l'ensemble des points n'appartenant pas à l'intérieur d'une géodésique minimisante issue de x), il existe une unique géodésique entre x et $T(x)$. Ce résultat s'obtient à partir de la formule de la variation première.

Remerciements : Je tiens à remercier chaleureusement G. Besson et G. Rappa pour leur soutien respectif.

Références

[1] Dominique Bakry and Michel Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

[2] Jérôme Bertrand. Existence and uniqueness of optimal maps on Alexandrov spaces. *Preprint 2007.*

[3] Yuri Burago, Mikhail Gromov, and Gregory Perel′man. A. D. Aleksandrov spaces with curvatures bounded below. *Uspekhi Mat. Nauk*, 47(2(284)) :3–51, 222, 1992.

[4] John Lott and Cédric Villani. Ricci curvature for metric-measure spaces via optimal mass transport. *To appear in Annals of math.*

[5] John Lott and Cédric Villani. Weak curvature conditions and poincaré inequalities. *J. Funct. Anal.* 245(1) : 311–333, 2007.

[6] John Lott. Optimal transport and Ricci curvature for metric-measure spaces. *Surveys in Differential Geometry XI (2007).*

[7] Robert J. McCann. Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.*, 11(3) :589–608, 2001.

[8] Yukio Otsu and Takashi Shioya. The Riemannian structure of Alexandrov spaces. *J. Differential Geom.*, 39(3) :629–658, 1994.

[9] Karl-Theodor Sturm. On the geometry of metric measure spaces. I. *Acta Math.*, 196(1) :65–131, 2006.

[10] Karl-Theodor Sturm. On the geometry of metric measure spaces. II. *Acta Math.*, 196(1) :133–177, 2006.

[11] Cédric Villani. *Topics in optimal transportation.*, volume 58 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2003.

[12] Cédric Villani. *Optimal transport, old and new (to appear).*

[13] Guofang Wei. Manifolds with A Lower Ricci Curvature Bound. *Surveys in Differential Geometry XI (2007).*

[14] Shunhui Zhu. The comparison geometry of Ricci curvature. In *Comparison geometry (Berkeley, CA, 1993–94)*, volume 30 of *Math. Sci. Res. Inst. Publ.*, pages 221–262. Cambridge Univ. Press, Cambridge, 1997.

Jérôme Bertrand
Institut de Mathématiques
Université de Toulouse
UMR 5219 (UPS-CNRS)
118, route de Narbonne
31062 Cedex 4 Toulouse
mél : bertrand@math.ups-tlse.fr