Le théorème de factorisation polaire

Yoan Tardy et Nicolas Masson
sous la direction de Bertrand Maury

Résumé

Le but de nombreux théorèmes classiques en mathématiques est de décomposer un objet en produit d'objets afin de simplifier l'étude du premier. Certains d'entre eux, connus sous le nom de "théorèmes de factorisation polaire", traitent de problématiques en apparence très différentes mais semblent profondément liés ; liens que nous nous proposons d'étudier dans ce mémoire. Le travail ici présenté est essentiellement dû à Yann Brenier, qui est parvenu à extraire l'essence commune de ces résultats pour les unifier en un théorème unique : le théorème de Brenier. Nous commencerons par énoncer les théorèmes classiques dits de "factorisation polaire" afin de relever ce qui les unit, pour ensuite proposer une définition abstraite de la notion de factorisation polaire, accompagnée d'un énoncé général de décomposition. Cependant, l'approche choisie en premier lieu se heurte à des difficultés techniques et échoue à produire un énoncé suffisamment général, nous contraignant à adopter une approche différente, mobilisant des outils de transport optimal. Le lien entre le problème de factorisation et le transport optimal étant subtil et difficile d'accès, il nous a semblé pertinent de sensibiliser le lecteur aux questions que posent le problème de Monge-Kantorovich. Nous étudierons ensuite un exemple précis pour permettre au lecteur de comprendre pourquoi le théorème de Brenier peut se formuler comme un problème de Monge-Kantorovich, avant de terminer par la démonstration du théorème de Brenier.

Table des matières

1 Introduction

L'objectif de ce mémoire est d'unifier un certain nombre de résultats classiques en apparence décorrélés, on commencera donc par lister ces théorèmes avant d'essayer de comprendre ce qui les lie. La majeure partie des résultats ici présentés proviennent de [1] et [4].

1.1 Les théorèmes de factorisation polaire classiques

Théorème 1.1.1 (Coordonnées polaires dans le plan complexe)**.** *Tout nombre complexe z s'écrit $z = re^{i\theta}$, où $r \geq 0$ et $\theta \in \mathbb{R}/2\pi\mathbb{Z}$. Si $z \neq 0$, la décomposition est unique.*

Théorème 1.1.2 (Décomposition polaire des matrices réelles)**.** *Toute matrice $M \in \mathcal{M}_n(\mathbb{R})$ s'écrit $M = RU$, avec $R \in \mathcal{S}_n^+(\mathbb{R})$ et $U \in \mathcal{O}_n(\mathbb{R})$. Si de plus M est inversible, la décomposition est unique.*

Théorème 1.1.3 (Décomposition de Helmholtz des champs de vecteurs)**.** *Soit Ω un ouvert de $\mathbb{R}^d$, supposé borné, lisse et connexe. Alors tout champ de vecteurs z sur Ω peut s'écrire $z = \nabla \times A + \nabla p$, où p est une fonction réelle lisse, définie sur Ω et A un potentiel vecteur.*

Avant d'énoncé le dernier théorème de cette classe, introduisons quelques notations et définitions nécessaires à sa compréhension.

Definition 1.1.1. Si (X, μ), (Y, ν) sont deux espaces de probabilité, on dit qu'une application $s : X \to Y$ préserve la mesure si elle est mesurable et que pour tout $A \subset Y$ ν-mesurable :

$$\mu\left(s^{-1}(A)\right) = \nu(A),$$

ou de manière équivalente, si pour toute fonction ν-intégrable f, $f \circ s$ est μ-intégrable et

$$\int_X (f \circ s)\mathrm{d}\mu = \int_Y f\mathrm{d}\nu.$$

S'il existe une bijection qui préserve la mesure entre (X, μ), (Y, ν), ces derniers sont dits isomorphes.

Definition 1.1.2. Si $p \in [1, +\infty[$ et $d \geq 1$, on appelle réarrangement de $u \in L^p(X, \mu, \mathbb{R}^d)$ sur (Y, ν) tout $v \in L^p(Y, \nu, \mathbb{R}^d)$ tel que :

$$\int_X f(u(x))\mathrm{d}\mu(x) = \int_X f(v(y))\mathrm{d}\nu(y),$$

et ce pour tout $f \in C(\mathbb{R}^d)$ tel que $|f(\xi)| \leq C(1 + \|\xi\|^p)$.

Definition 1.1.3. On dit que $u : (X, \mu) \to (Y, \nu)$ satisfait la condition de non-dégénérescence si l'image réciproque de tout ensemble ν-négligeable par u est de μ-mesure nulle.

Avec ces définitions, nous sommes en mesure d'énoncer le dernier résultat classique de décomposition polaire, qui concerne les réarrangements de fonctions à valeurs réelles.

Théorème 1.1.4. *Soit (X, μ) un espace de probabilité isomorphe à $([0, 1], \lambda)$ où λ est la mesure de Lebesgue. Soit $u \in L^p(X, \mu, \mathbb{R}^d)$ vérifiant la condition de non-dégénérescence (conformément à la définition 1.1.3). Alors il existe un unique réarrangement $u^\# \in L^p([0, 1])$ et une unique application préservant la mesure $s : (X, \mu) \to ([0, 1], \lambda)$ tels que*

$$u = u^\# \circ s.$$

Remarque. Il est possible de montrer que s est l'unique application préservant la mesure qui maximise $\int_X u(x)s(x)\mathrm{d}x$, ce qui en donne une caractérisation.

Tous ces résultats peuvent être unifiés par un seul et même théorème : le théorème de Brenier, énoncé ci-dessous, et dont la démonstration sera l'objet de la section 4. On considère Ω un ouvert borné connexe de $\mathbb{R}^d$, et β une mesure sur $\bar{\Omega}$ telle que $\beta(\partial\Omega) = 0$. On suppose que $d\beta(z) = \beta(z)dz$, où β est une fonction minorée par une constante strictement positive sur tout compact de Ω. L'espace de Sobolev :

$$W^{1,p}(\Omega, \beta) = \{f \in L^p(\Omega, \beta),\ \nabla f \in L^p(\Omega, \beta; \mathbb{R}^d)\}$$

jouera un rôle important dans la suite, ainsi que l'ensemble :

$$K = \{\nabla\psi;\ \psi \in W^{1,p}(\Omega, \beta);\ \psi \text{ convexe}\}.$$

Théorème 1.1.5 (Théorème de Brenier)**.** *Soit N l'ensemble des $u \in L^p(X, \mu; \mathbb{R}^d)$ contrevenant à la condition de non-dégénérescence (conformément à la définition 1.1.3). Alors, $\forall u \in L^p(X, \mu; \mathbb{R}^d) \setminus N$, il existe un unique couple $(u^\#, s)$ tel que $u^\# \in K$, s est une application de (X, μ) dans $(\bar{\Omega}, \beta)$ préservant la mesure, et :*

$$u = u^\# \circ s.$$

On a de plus :

1. *$u^\#$ est l'unique réarrangement de u dans K.*
2. *s est l'unique application préservant la mesure préservant la mesure maximisant $\int_X u(x) \cdot s(x) d\mu(x)$.*
3. *l'application $u \mapsto (u^\#, s)$ est continue de $L^p(X, \mu); \mathbb{R}^d) \backslash N$ dans $L^p(\Omega, \beta; \mathbb{R}^d) \times L^q(X, \mu, \mathbb{R}^d)$ où q est tel que $\frac{1}{p} + \frac{1}{q} = 1$.*

1.2 Le problème de projection

Le lien entre ces théorèmes étant assez difficile à mettre en évidence au premier abord, nous nous proposons d'exposer en premier lieu l'origine du théorème de Brenier. En effet, pour bien comprendre comment Yann Brenier a réussi à extraire l'essence commune de ces résultats, il nous apparaît nécessaire d'exposer le problème qu'il s'était initialement posé : le problème de projection.

Soit Ω un ouvert de $\mathbb{R}^d$ supposé borné, lisse et connexe, l'espace de Hilbert $H = L^2(\Omega, \mathbb{R}^d)$ muni de son produit hilbertien canonique $(.,.)$. Notons S l'ensemble des applications préservant la mesure de (Ω, λ) dans lui-même.

Remarque. S est fermé, borné, et vérifie, $\forall s \in S,\ f \in H,\ \|f \circ s\|_H = \|f\|$.

Problème 1.1 (Projection)**.** Trouver, pour $u \in H$, $s \in S$ qui minimise $\|u - s\|_H^2$, ou de façon équivalente qui maximise $\int_\Omega u(x) \cdot s(x)$.

En se souvenant de la remarque faite dans le théorème 1.1.4, on note que pour $d = 1$ la solution au problème de projection n'est autre que la partie "angulaire" de la décomposition, ce qui pousse à penser qu'il doit y avoir un lien assez fort entre le problème de projection et le théorème 1.1.4, qui pourrait donner lieu à une généralisation de ce dernier. Si on choisit de partir de ce problème pour aborder celui de la factorisation polaire, c'est-à-dire sans *a priori* sur la partie "radiale" de la décomposition, il est légitime de se demander comment retrouver $u^\#$ à partir de s, et quelles contraintes cela impose sur $u^\#$.
La première remarque à faire est que S a une structure de semi-groupe, ce qui naturellement incite à étudier le cas où s est inversible : si cette hypothèse est vérifiée, il suffit alors de définir $u^\# = u \circ s^{-1}$ pour avoir la partie "radiale". Le problème 1.1 étant formulé en des termes métriques, il est naturel d'étudier les propriétés de $u^\#$ imposées par s aussi d'un point de vue métrique. Or si on regarde $\|u^\# - \sigma\|$ pour $\sigma \in S$, on a :

$$\begin{aligned} \|u^\# - \sigma\| &= \|u \circ s^{-1} - \sigma\| \\ &= \|u - \sigma \circ s\| \end{aligned}$$

grâce à la remarque

$$\geq \|u - s\| = \|u^{\#} - id\|.$$

Ainsi, $u^{\#}$ vérifie

$$\forall \sigma \in S, \quad \|u^{\#} - id| \leq \|u^{\#} - \sigma\|,$$

ce qui se traduit plus simplement dans un Hilbert par

$$\forall \sigma \in S, \quad (u^{\#}, id - \sigma) \geq 0.$$

Or l'ensemble

$$K = \{u \in H, \quad (u, id - s) \geq 0\}$$

est bien connu de ceux qui ont fait de l'analyse convexe : il s'agit du cône polaire de S. Cette interprétation géométrique nous donne de bonnes pistes de réflexion pour comprendre le lien entre nos résultats initiaux.

1.3 Lien entre les différents résultats

C'est ici que l'approche précédente se révèle fructueuse, dans la mesure où elle nous permet de bien comprendre le rôle des deux parties de la décomposition. En effet, on remarque que tous les résultats classiques exposés en introduction se placent sur un espace de Hilbert (sauf le dernier, pour lequel il faut supposer de plus $p = 2$), et traitent d'une décomposition dont une partie appartient à un semi-groupe fermé, borné, et l'autre au cône polaire de ce semi-groupe. On comprend donc un peu mieux la terminologie utilisée : on a une partie "angulaire" qui correspond au semi-groupe, et une partie "radiale" qui correspond au cône polaire.

On notera également que pour bénéficier des résultats sous leur forme la plus satisfaisante (par exemple avec unicité de la décomposition), on est contraint de se restreindre à un sous-ensemble de l'espace de Hilbert. Cependant, on remarque que l'ensemble des éléments pour lesquels la décomposition n'est pas complètement satisfaisante est toujours relativement petit : pour le théorème 1.1.1, la décomposition est unique pour tout $z \in \mathbb{C} \setminus \{0\}$, pour le théorème 1.1.2, elle est unique sur $GL_n(\mathbb{C})$, soit sur un ouvert dense de $\mathcal{M}_n(\mathbb{C})$, et enfin pour le théorème 1.1.4, elle est unique pour tout u vérifiant la condition de non-dégénérescence, et on montrera que l'ensemble des u contrevenant à cette condition est négligeable au sens de Baire.

2 Une version abstraite du théorème de factorisation polaire

2.1 Le concept abstrait de factorisation polaire

Le paragraphe précédent nous a aidé à mieux comprendre ce qui fait l'essence de ces théorèmes, et nous permet de proposer quelques définitions abstraites du concept de factorisation polaire. Dans cette section, la notion de "factorisation polaire" est donc définie dans un cadre général qui mobilise :

(a) un espace de Hilbert H, muni de son produit $(.,.)$ et de la norme $\|.\|$ qui en dérive,

(b) un sous-ensemble fermé borné S de H,

(c) une loi de composition $* : S \times H \to H$ telle que $(S, *)$ soit un semi-groupe, avec un élément neutre e, ce qui veut dire que :

$$s_1 * s_2 \in S \quad \text{dès que} \quad s_1, s_2 \in S, \quad s * e = e * s = s, \quad \forall s \in S$$

On suppose de plus que :

$$s * (\alpha u + \beta v) = \alpha s * u + \beta s * v, \quad \forall s \in S, \quad \forall u, v \in H, \quad \forall \alpha, \beta \in \mathbb{R}$$

que $\forall u \in H$, $s \in S \to s * u$ est continue, et que $\|s * u\| = \|u\|$

(d) Le cône polaire de S, noté K, défini par :

$$K = \{u \in H, \quad (u, e - s) \geq 0, \quad \forall s \in S\}.$$

On dit alors que H admet une factorisation polaire par le semi groupe S et son cône polaire K s'il existe un sous-ensemble $N \subset H$ négligeable au sens de Baire tel que :

(i) Pour tout $u \in H \setminus N$, il existe une unique factorisation $u = k * s$, où $(k, s) \in K \times S$,

(ii) $u \to (s, k)$ est continue sur $H \setminus N$.

Afin de montrer la pertinence de ces définitions, revenons désormais à nos résultats de départ.

Exemple 2.1 (Décomposition polaire des nombres complexes)**.** Ici, $H = \mathbb{C}$ (vu comme un espace de Hilbert réel), S est le cercle unité, $N = \{0\}$. Il n'est pas difficile de voir que le cône polaire du cercle unité est $K = \mathbb{R}_+$

Exemple 2.2 (Décomposition polaire des matrices réelles)**.** Ici, $H = \mathcal{M}_n(\mathbb{R})$ muni du produit scalaire euclidien canonique, N est l'ensemble des matrices non inversibles, qui est donc un fermé d'intérieur vide et par conséquent négligeable au sens de Baire ; et $S = \mathcal{O}_n(\mathbb{R})$. On va montrer que le cône polaire de S est $\mathcal{S}_n^+(\mathbb{R})$. Soit donc une matrice M telle que :

$$\forall U \in \mathcal{O}_n(\mathbb{R}), \quad \mathrm{Tr}(M^t(I-U)) \geq 0,$$

ce qui équivaut à

$$\forall U \in \mathcal{O}_n(\mathbb{R}), \quad \mathrm{Tr}\, M \geq \mathrm{Tr}(MU).$$

$\mathrm{Tr}(M)$ apparaît donc naturellement comme le maximum de l'application infiniment différentiable $\Phi : U \in \mathcal{O}_n(\mathbb{R}) \to \mathrm{Tr}(MU)$. Il est donc logique d'exploiter d'un point de vue différentiel la propriété de maximum, en notant que :

$$T_I(\mathcal{O}_n(\mathbb{R})) = \mathcal{A}_n(\mathbb{R}) \subset \mathrm{Ker}(\mathrm{d}\Phi(I))$$

(où $T_I(\mathcal{O}_n(\mathbb{R}))$ désigne l'espace tangent à $\mathcal{O}_n(\mathbb{R})$ en I, et $\mathcal{A}_n(\mathbb{R})$ l'espace des matrices antisymétriques de taille n). Or $\mathrm{d}\Phi(I).H = \mathrm{Tr}(MH)$, on a donc :

$$\forall A \in \mathcal{A}_n(\mathbb{R}), \quad \mathrm{Tr}(MA) = 0,$$

ce qui veut dire que $M \in \mathcal{A}_n(\mathbb{R})^\perp = S_n(\mathbb{R})$.

2.2 Un théorème abstrait de factorisation polaire

Avant d'énoncer le théorème abstrait de factorisation polaire, introduisons quelques définitions d'analyse convexe :

Théorème 2.2.1 (Théorème de factorisation polaire abstrait)**.** *Si H est un espace de Hilbert, $(S,*)$ est un groupe fermé borné de H, alors il existe N un ensemble négligeable au sens de Baire tel que $\forall u \in H \setminus N$, on peut écrire de manière unique $u = k * s$ avec :*

1. *il y a un unique $\pi(u)$ projection de Hilbert sur S de u.*
2. *$s = \pi(u)$.*
3. *$\partial J(u) = \{\pi(u)\}$ où $J(u) = \sup_{s\in S}\langle u, s\rangle$ et $\partial J(u)$ désigne le sous-gradient de J en u, c'est-à-dire l'ensemble des vecteurs $v \in H$ tels que $\forall h \in H$, $J(h) \geq J(u) + \langle v, h-u\rangle$.*

Pour cela, nous allons commencer par montrer que cette décomposition concerne effectivement une large partie des éléments de H :

Théorème 2.2.2 (Edelstein)**.** *Soit S fermé borné de H. Alors si l'on pose*

$$H \setminus N = \{\ u \in H, \quad \forall \varepsilon > 0, \quad \exists \delta > 0, \quad \forall s_1, s_2 \in S,\ \forall i = 1, 2,$$
$$\|s_i - u\| \leq d(u, S) + \delta, \Rightarrow \|s_1 - s_2\| \leq \varepsilon\},$$

alors tout élément de $H \setminus N$ admet une unique projection de Hilbert $\pi(u)$ sur S. De plus, $\pi : H \setminus N \to H$ est continue.

Moralement, l'ensemble $H \setminus N$ est l'ensemble des points qui "ne sont pas uniformément entourés par S", ils sont plus proches d'une partie de S en particulier et c'est en cela que l'on peut assurer l'unicité de la projection.

Preuve. On va montrer que $H \setminus N$ est une intersection dénombrable d'ouverts denses, à savoir des

$$T_\varepsilon = \{u \in H \mid \exists \delta > 0 \mid \forall s_1, s_2 \in S, \forall i = 1, 2,$$
$$\|s_i - u\| \leq d(u, S) + \delta \Rightarrow \|s_1 - s_2\| \leq \varepsilon\},$$

pour ε rationnel strictement positif. Montrons d'abord que les T_ε sont ouverts : si $u \in T_\varepsilon$, on peut se munir d'un δ comme dans la définition. Montrons que $B(u, \frac{\delta}{3}) \subset T_\varepsilon$: Soit u' tel que $\|u - u'\| \leq \frac{\delta}{3}$. Supposons que :

$$\forall i \in \{1, 2\}, \|s_i - u'\| \leq d(u', S) + \frac{\delta}{3}.$$

Alors :

$$\begin{aligned}
|s_i - u\| &\leq \|s_i - u'\| + \frac{\delta}{3} \quad \text{par inégalité triangulaire,} \\
&\leq d(u', S) + 2\frac{\delta}{3} \quad \text{par hypothèse,} \\
&\leq d(u, S) + 3\frac{\delta}{3} \quad \text{par inégalité triangulaire.}
\end{aligned}$$

Donc, $\|s_1 - s_2\| \leq \varepsilon$ car $u \in T_\varepsilon$. On conclut alors que $u' \in T_\varepsilon$ et donc le fait que T_ε est ouvert.
Montrons maintenant que T_ε est dense :
Prenons $u \in H$, on peut supposer $d(u, S) > 0$, car dans le cas contraire, $u \in S \subset T_\varepsilon$

(on peut prendre $\delta = \frac{\varepsilon}{2}$ et conclure avec une inégalité triangulaire). Puis, on peut alors de plus supposer que $u = 0$ et $d = d(u,S) = 1$ quitte à translater et à dilater S. Alors, on peut trouver $s_0 \in S$ tel que $\|s_0\| = R$ où R assez proche de 1 en en restant strictement supérieur. On pose alors, pour $r \in]0,1[$, $u_r = r\frac{s_0}{R}$. Le but est de trouver un bon r tel que $u_r \in T_\varepsilon$.
On prend $\varepsilon > 0$, on cherche r tel que $\|u_r - u\| \leq \varepsilon$, or, $\|u_r\| = r$, fixons alors un $r \leq \varepsilon$. Posons :

$$A = \{s \in S / \|s - u_r\| \leq d(u_r, S) + \delta\}.$$

Le but est de trouver δ tel que $diam(A) \leq \varepsilon$. Faisons d'abord deux remarques simples :

1. $A \subset B(u_r, R - r + \delta)$ car $d(u_r, S) \leq \|u_r - s_0\| = \|\frac{(r-R)s_0}{R}\| = R - r$ car $\|s_0\| = R$.
2. $S \subset \{v \in H / \|v\| \geq 1\}$ car $d(0,S) = 1$.

Fixons $v \in A$ et estimons $\|v - s_0\|^2$:

$$\begin{aligned}
\|v - s_0\|^2 &= \|v\|^2 + R^2 - 2\langle s_0, v\rangle \\
&= \frac{R}{r}\left(\|v\|^2 + r^2 - 2\frac{r}{R}\langle s_0, v\rangle\right) + \left(1 - \frac{R}{r}\right)\|v\|^2 - Rr + R^2 \\
&\leq \frac{R}{r}(R - r + \delta)^2 + (1 - \frac{R}{r})\|v\|^2 - Rr + R^2 \\
&= (1 - \frac{R}{r})\|v\|^2 + R^2 + \frac{R}{r}\left[(R - r + \delta)^2 - r^2\right] \\
&\leq 1 - \frac{R}{r} + R^2 + \frac{R}{r}[(R - r + \delta)^2 - r^2] \text{car } R > r, \|v\| \geq 1 \\
&= 1 - \frac{R}{r} + R^2 + \frac{R}{r}[(R - r)^2 + 2(R - r)\delta + \delta^2 - r^2] \\
&= 1 - \frac{R}{r} + R^2 + Rr\left(\frac{R}{r} - 1\right) = Rr\left(\frac{R}{r} - 1\right)^2 + 2R\left(\frac{R}{r} - 1\right)\delta + \delta^2\frac{R}{r} \\
&= \left(1 - \frac{R}{r}\right)\left[-1 + Rr + Rr\left(\frac{R}{r} - 1\right) + 2R\delta\right] + \frac{R}{r}\delta^2 \\
&= \left(1 - \frac{R}{r}\right)[Rr - 1 + R^2 - Rr + 2R\delta] + \frac{R}{r}\delta^2 \\
&= \left(1 - \frac{R}{r}\right)[R^2 - 1 - 2R\delta] + \frac{R}{r}\delta^2 \\
&\leq C_1[\varepsilon + 2C_2\delta] + C_3\delta^2
\end{aligned}$$

si l'on fixe R assez proche de 1, et r comme indiqué précédemment.

$$\leq C_4\varepsilon \quad \text{si } \delta \text{ assez petit.}$$

Donc, pour δ assez petit, on bien $diam(A) \leq \varepsilon$ via l'inégalité triangulaire. Ainsi, pour les r choisis, on a $u_r \in T_\varepsilon$ par définition de A et de T_ε, et u_r assez petit. On peut trouver un élément de T_ε aussi proche de 0 que souhaité, ce qui nous permet de conclure quant à la densité de T_ε d'après les hypothèses simplificatrices du début. De là, on a bien N négligeable au sens de Baire.
Prenons maintenant un élément u de $H \setminus N$ et montrons qu'il admet une unique projection de Hilbert sur S. Supposons un instant qu'il en possède en fait 2 (la distance à un fermé est atteinte, on a au moins une projection), notons-les l_1 et l_2. Alors, par définition, $\|l_1 - u\| = d(u, S)$ et de même pour l_2, d'où, comme $u \in H/N$, on a $\|l_1 - l_2\| \leq \varepsilon$ pour tout $\varepsilon > 0$, d'où l'égalité des deux projections.
Il reste à montrer que $\pi : H \setminus N \to H$ est continue. Prenons $\varepsilon > 0$ et $u, v \in H \setminus N$, alors il existe δ_u de la définition de $H \setminus N$. Alors, si on prend $v \in H \setminus N$ tel que $\|u - v\| \leq \frac{\delta_u}{2}$, on a :

1. $\|u - \pi(u)\| \leq d(u, S)$.
2. $\|u - \pi(v)\| \leq \frac{\delta_u}{2} + d(v, S) \leq d(u, S) + \delta_u$ par inégalité triangulaire.
   Donc, par définition de δ, $\|\pi(u) - \pi(v)\| \leq \varepsilon$. D'où la continuité désirée.

Notons pour la suite que $\forall s \in S, \|s\| = \|e * s\| = \|e\| = cste$. S vérifie donc les hypothèses de la proposition ci-dessous :

Proposition 2.2.3 (Caractérisation de π)**.** *Soit S fermé borné inclus dans une sphère centrée en l'origine d'un Hilbert H. Alors, la projection $\pi :\to S$ est caractérisée par le gradient de*

$$J : u \to sup_{s\in S} \langle u|s\rangle .$$

Plus précisément, $\partial J(u) = \{\pi(u)\}$ pour $u \in H/N$.

Remarque. On a : $J(u) = \langle u, \pi(u)\rangle$.

Preuve. Remarquons tout d'abord que maximiser $\langle u|s\rangle, s \in S$ revient à minimiser $\|u-s\|,\ s \in S$ car $\|s\|$ est une constante sur S, et donc $J(u) = \langle u|\pi(u)\rangle, \forall u \in H\setminus N$ par définition de $\pi(u)$. Le fait que π est le gradient de J découle de deux lemmes :

Lemme 2.2.4. $\forall u \in H, \partial J(u) \subset \overline{Conv(S)}$.

Lemme 2.2.5. $\forall u \in H, \forall p \in \partial J(u), \langle p|u\rangle = J(u)$.

Commençons par déduire de ces deux lemmes la proposition. Fixons $u \in H\backslash N$. Tout d'abord, $\pi(u) \in \partial J(u)$. En effet, si $v \in H$:

$$\begin{aligned} J(v) \geq J(u) + \langle \pi(u) | v - u \rangle &\Longleftrightarrow J(v) \geq \langle u | \pi(u) \rangle + \langle \pi(u) | v - u \rangle \\ &\Longleftrightarrow J(v) \geq \langle \pi(u) | v \rangle, \end{aligned}$$

ce qui est vrai étant donnée la définition de J. Puis, si $p \in \partial J(u)$, on a que $p = \pi(u)$. En effet, par les lemmes précédents, on a $\langle p | u \rangle = J(u)$ et $p \in \overline{Conv(S)}$. Par définition de $H \setminus N$:

$$\forall \varepsilon > 0,\ \exists \delta > 0 \ /\ \forall s \in S,\ \|s - u\|^2 \leq \|\pi(u) - u\|^2 + \delta \Rightarrow \|\pi(u) - s\| \leq \frac{\varepsilon}{2}.$$

Prenons $\gamma > 0$ tel que $\gamma\|u\| \leq \frac{\delta^2}{2}$. Tout d'abord, comme $p \in \overline{Conv(S)}$, $\exists (\theta_i)_{i \in I}$, $(s_i)_{i \in I}$ tels que $\sum_{i \in I} \theta_i = 1$, et $\|p - \sum_{i \in I} \theta_i s_i\| \leq \gamma$, où I ensemble fini. On a :

$$\begin{aligned} \sum_{i \in I} \theta_i \langle s_i | u \rangle &= \left\langle \sum_{i \in I} \theta_i s_i | u \right\rangle \\ &= \left\langle \sum_{i \in I} \theta_i s_i - p | u \right\rangle + \langle p | u \rangle \\ &\geq \langle p | u \rangle - \|u\| \|\sum_{i \in I} \theta_i s_i - p\| \quad \text{par Cauchy-Schwarz,} \\ &\geq \langle p | u \rangle - \gamma \|u\| \\ &\geq J(u) - \frac{\delta^2}{2} \text{ car } p \in \partial J(u), \text{ d'où en } v = 0,\ 0 \geq J(u) - \langle p | u \rangle. \end{aligned}$$

On pose alors pour $i \in I$:

$$\begin{aligned} a_i &= 2(J(u) - \langle s_i | u \rangle) \\ &= 2 \langle \pi(u) - s_i | u \rangle \\ &= \|u - s_i\|^2 - \|u - \pi(u)\|^2 \geq 0. \end{aligned}$$

On a :

$$\sum_{i \in I} \theta_i a_i \leq \delta^2.$$

En effet, par l'inégalité obtenue précédemment, comme $\sum_{i \in I} \theta_i = 1$, on a :

$$\delta^2 \geq 2 \sum_{i \in I} \theta_i (J(u) - \langle s_i | u \rangle) = \sum_{i \in I} \theta_i a_i.$$

De cette inégalité, on déduit que :

$$\exists I' \subset I \text{ tel que } a_i \leq \delta \ \forall i \in I' \text{ et } \ \sum_{i \in I'} \theta_i \geq 1 - \delta.$$

En effet, si on pose $I' = \{i \mid a_i \leq \delta\}$, on a :

$$\begin{aligned}
\delta^2 &\geq \sum_{i \notin I'} \theta_i a_i \geq \delta \sum_{i \notin I} \theta_i \\
\Rightarrow \delta &\geq \sum_{i \notin I'} \theta_i \\
\Rightarrow \sum_{i \in I'} \theta_i &\geq 1 - \delta.
\end{aligned}$$

Ainsi, si $i \in I'$, $\|s_i - u\|^2 = \|u - \pi(u)\| + a_i \leq \|u - \pi(u)\|^2 + \delta$. Donc, $\|\pi(u) - s_i\| \leq \varepsilon$ par définition de δ, car $u \in H \setminus N$. De là :

$$\begin{aligned}
\|\pi(u) - p\| &\leq \|\pi(u) - \sum_{i \in I'} \theta_i s_i\| + \|\sum_{i \notin I} \theta_i s_i\| + \|\sum_{i \in I} \theta_i s_i - p\| \\
&\leq \varepsilon + \delta \|e\| + \gamma \\
&\leq 2\varepsilon \quad \text{quitte à prendre } \gamma, \delta \text{ plus petits.}
\end{aligned}$$

Or, cela vaut quelque soit $\varepsilon > 0$, donc $p = \pi(u)$.

Il ne reste plus qu'à prouver les deux lemmes et l'on aura alors démontré la proposition.

Preuve (lemme 2.2.4). Supposons qu'il existe $x \in \partial J(u) \setminus \overline{Conv(S)}$, alors comme $\{x\}$ compact convexe, $\overline{Conv(S)}$ fermé convexe et H localement convexe, on a par Hahn-Banach l'existence de ϕ une forme linéaire représentée par un vecteur w tel que :

$$\forall s \in S, \ \langle w | x \rangle > \langle w | s \rangle.$$

Or, $x \in \partial J(u)$, d'où

$$\forall v \in H, J(v) \geq J(u) + \langle x | v - u \rangle.$$

On prend alors $v = w + u$, et donc :

$$\begin{aligned} & J(v) > J(u) + \langle s|w\rangle \quad \forall s \in S \\ \Leftrightarrow \quad & \langle v|\pi(v) - s\rangle > \langle u|\pi(u) - s\rangle \quad \forall s \in S. \end{aligned}$$

On particularise en $s = \pi(v) \in S$, et il vient : $\langle u|\pi(v)\rangle > \langle u|\pi(u)\rangle$, ce qui est absurde car $\langle u|\pi(u)\rangle = \sup_{s\in S}\langle u|s\rangle$.

Preuve (lemme 2.2.5). Soit $p \in \partial J(u)$. On a $J(v) \geq J(u) + \langle p|v - u\rangle$ $\forall v \in H$. En appliquant en $v = 0$, il vient, $\langle p|u\rangle \geq J(u)$. Puis, par un argument de convexité et de continuité/passage à la limite, on a l'inégalité inverse car $J(u) = \sup_{s\in S}$ et $p \in \overline{Conv(S)}$.

On a donc démontré la proposition 3.2.1, qui nous permet de démontrer le théorème 2.2.1 :

Preuve (Théorème de factorisation abstraite). On séparera la preuve en trois parties :

1. Existence : Edelstein nous donne pour chaque $u \in H \setminus N$ une unique projection $s = \pi(u)$ sur S. Or S est un groupe, donc, on peut poser $k = s^{-1} * u$ et alors, $u = s * k$. Il suffit alors de montrer que $k \in K$. Soit $\sigma \in S$. On a :

$$\begin{aligned} \langle k|e - \sigma\rangle &= \langle s * k|s - s * \sigma\rangle \quad \text{par l'identité de polarisation,} \\ &= \langle u|s - s * \sigma\rangle \\ &= \frac{1}{2}(\|u - s * \sigma\|^2 - \|u - s\|^2) \geq 0 \quad \text{par définition de s.} \end{aligned}$$

   Donc k appartient effectivement à K.

2. Unicité : Il suffit de montrer l'unicité de s étant donné que cet élément est inversible, l'unicité de k en découlera automatiquement. On a :

$$\begin{aligned} s = \pi(u) &\Longleftrightarrow \forall \sigma \in S \quad \langle u|s\rangle \geq \langle u|\sigma\rangle \\ &\Longleftrightarrow \forall \sigma \in S \quad \langle s * k|s\rangle \geq \langle s * k|\sigma\rangle \\ &\Longleftrightarrow \forall \sigma \in S \quad \langle k|e\rangle \geq \langle k|s^{-1} * \sigma\rangle \\ &\Longleftrightarrow \forall \sigma \in S \quad \langle k|e - s^{-1} * \sigma\rangle \geq 0 \quad \text{car } k \in K. \end{aligned}$$

3. Continuité de la factorisation polaire : Prenons $u_n = s_n * k_n \in H \setminus N$ où $u_n \to u$ quand $n \to +\infty$. Par Edelstein, on a $s_n = \pi(u_n) \to \pi(u) = s$ par

continuité de π.
Enfin :

$$\begin{aligned}\|k_n - k\| &= \|s_n^{-1} * u_n - s^{-1} * u\| \\ &= \|\|\|u_n - s_n * s^{-1} * u\| \quad \text{car la norme est invariante par l'action de } S, \\ &\leq \|u_n - u\| + \|u - s_n * s^{-1} * u\|,\end{aligned}$$

tendant tout deux vers 0 à l'infini car $\forall u \in H, s \to s * u$ est continue par hypothèse. D'où la continuité désirée.

Il ne reste plus qu'à montrer que $K = \{\nabla\psi, \psi \in W_{1,2}(\Omega), \psi \text{ } convexe\}$.

Preuve. montrons tout d'abord que $\{\nabla\psi, \psi \in W_{1,2}(\Omega), \psi \text{ } convexe\} \subset K$, si l'on a ψ convexe, alors,

$$\forall s \in S, \forall x \in \Omega,\ \psi(s(x)) \geq \psi(x) + \langle \nabla\psi(x) | s(x) - x \rangle$$

et donc,

$$\int_\Omega \langle \nabla\psi(x) | x - s(x) \rangle \, dx \geq \int_\Omega \psi(x) - \psi(s(x)) dx = 0$$

car s préserve la mesure. D'où $\nabla\psi \in K$. Montrons maintenant l'inclusion inverse. Nous admettons tout d'abord, tout élément de K est gradient d'une fonction de $W_{1,2}(\Omega)$. Puis, u est monotone, et donc ψ est convexe. En effet, il suffit pour cela de montrer que

$$\langle u(x_1) - u(x_2) | x_1 - x_2 \rangle \geq 0$$

presque partout sur Ω^2. On fixe alors x_1, x_2 dans Ω et on pose pour ε tel que les boules centrées en x_1 et en x_2 de rayon ε ne se rencontrent pas l'application s tel que :

— $s(x) = x - x_1 + x_2$ si $x \in B(x_1, \varepsilon)$
— $s(x) = x - x_2 + x_1$ si $x \in B(x_2, \varepsilon)$
— $s(x) = x$ sinon.

C'est une application préservant la mesure qui ne fait qu'échanger les boules données. Comme $u \in K$, on a :

$$\int_\Omega \langle u(x) | x - s(x) \rangle \, dx \geq 0.$$

Avec la définition de s, cela revient à :

$$\int_{B(x_1,\varepsilon)} \langle u(x)|x_1 - x_2\rangle + \int_{B(x_2,\varepsilon)} \langle u(x)|x_1 - x_2\rangle$$

$$= \left\langle x_1 - x_2 \middle| \int_{B(0,1)} u(x_1 + \varepsilon y) - \int_{B(0,1)} u(x_2 + \varepsilon y) \right\rangle \geq 0.$$

D'où, en faisant $\varepsilon \to 0$, comme u est intégrable, on a l'inégalité voulue.

3 Digression sur le transport optimal

Si l'approche géométrique du problème nous a permis d'identifier ce qui constituait l'essence de ces résultats classiques, elle échoue malheureusement à produire une preuve dans le cas où S est seulement muni d'une structure de semi-groupe. Cependant, il est tout de même possible d'unifier ces théorèmes en exhibant un lien fort entre le problème et la théorie du transport optimal. L'objectif de cette section sera simplement d'exposer quelques éléments de cette théorie pour mieux comprendre sur quoi repose ce lien. Les résultats présentés dans cette section proviennent de [3].

3.1 Optimisation sous contrainte

De nombreuses questions mathématiques se formulent comme des problèmes de minimisation ou de maximisation de fonctions. La détermination directe des *extrema* d'une fonction étant en générale trop complexe, l'objectif est alors de trouver ce qui caractérise ces *extrema*. La situation se complique sensiblement lorsque le problème d'optimisation possède des contraintes, cas que nous allons explorer pour mieux comprendre quelques idées essentielles en transport optimal. Commençons par une situation simple :

Proposition 3.1.1. *Soit U un ouvert d'un espace de Hilbert H, et J une fonctionnelle différentiable. Si u est un minimum local de J sur U, alors $\nabla J(u) = 0$.*

L'argument essentiel ici est que dans un ouvert, toutes les directions d'approche de u sont autorisées, ce qui n'est évidemment pas le cas si l'on impose des contraintes. Nous nous intéresserons par la suite à des contraintes d'un type bien particulier : les contraintes d'inégalité.

3.1.1 Mise en place du problème et premiers résultats

Soit N un entier supérieur à 1 et $\varphi_1, ..., \varphi_N$ une collection de N applications différentiables de H dans $\mathbb{R}$. On s'intéresse ici à la minimisation de fonctionnelles sur des ensembles du type

$$K = \{v \in H,\ \varphi_i(v) \leq 0,\ i = 1, \cdots, N\}.$$

Definition 3.1.1 (Contraintes actives)**.** On dit que la contrainte i est active en $u \in H$ dès que $\varphi(u) = 0$. On note I_u l'ensemble des i tels que la contrainte est active en u.

Definition 3.1.2 (Qualification des contraintes)**.** Soit $u \in H$. On dit que les contraintes $[\varphi_i \leq 0]$ sont qualifiées en u s'il existe un vecteur $h \in H$ tel que :

$$\nabla \varphi_i(u) \cdot h < 0$$

pour tout $i \in I_u$

Proposition 3.1.2. *Soit J une fonctionnelle C^1 définie sur un ouvert U de H, et u un minimiseur local de J sur $U \cap K$. On suppose que les contraintes sont qualifiées en u. Il existe alors $\lambda_1, ..., \lambda_N \geq 0$ tels que*

$$\nabla J(u) + \sum_{i=1}^{N} \lambda_i \nabla \varphi_i = 0.$$

Si ce résultat a déjà un intérêt en soi, il laisse en général un peu sur sa faim. Cependant, continuer à aborder le problème frontalement s'avère improductif, on a donc besoin d'une nouvelle approche en "dualisant" le problème. L'idée, extrêmement fertile, consiste à reformuler le problème d'une manière différente en "dualisant" les contraintes, ce que nous nous proposons de faire dans le paragraphe suivant.

3.1.2 Dualité

Lemme 3.1.3. *Soient V et Λ deux ensembles, et $L(.,.)$ une application de $V \times \Lambda$ dans $\mathbb{R}$. On définit :*

$$G(q) = \inf_{v \in V} L(v, q),\ F(v) = \sup_{q \in \Lambda} L(v, q).$$

On a alors :

$$G(q) \leq F(v) \quad \forall q \in \Lambda,\ v \in V.$$

Par suite, s'il existe u et p tels que $G(p) = F(u)$, alors

$$G(p) = \max G = \min F = F(u).$$

Definition 3.1.3. Dans le contexte, et avec les notations du lemme précédent, on appellera :

— problème **primal** le problème de minimisation de F sur V, et

— problème **dual** le problème de maximisation de G sur Λ.

Definition 3.1.4 (Point-selle)**.** En reprenant les notations du lemme, on dit que (u,p) est un point-selle de L sur $V \times \Lambda$ si

$$L(u,q) \leq L(u,p) \leq L(v,p) \quad \forall q \in \Lambda,\ v \in V.$$

Proposition 3.1.4. *Les deux assertions suivantes sont équivalentes :*

(i) *$L(.,.)$ admet un point-selle (u,p).*

(ii) *Le sup de G est atteint en un point $p \in \Lambda$, l'inf de F est atteint en un point $u \in v$, et ces deux quantités sont égales.*

Le lien entre les problèmes de minimisation sous contrainte et la notion de point-selle passe par la définition d'une fonctionnelle appelée Lagrangien :

Definition 3.1.5 (Lagrangien)**.** Soit J une fonctionnelle d'un ensemble X dans $\mathbb{R}$, et K un ensemble défini par N_u contraintes d'inégalité et N_e contraintes d'égalité :

$$K = \{v \in X,\ \varphi_i(v) \leq 0,\ \psi_j(v) = 0 \quad \forall i,j,\ 1 \leq i \leq N_u,\ 1 \leq j \leq N_e\}.$$

Le Lagrangien associé au problème de minimisation de J sur K est défini par :

$$(u,p,q) \in X \times \mathbb{R}_+^{N_u} \times \mathbb{R}^{N_e} \mapsto L(u,p,q) = J(u) + \sum_{i=1}^{N_u} p_i \varphi_i(u) + \sum_{i=1}^{N_e} q_j \psi_j(u).$$

Proposition 3.1.5. *On considère une fonctionnelle d'un ensemble X dans $\mathbb{R}$, et l'on suppose que le Lagrangien associé au problème de minimisation de J sur K (défini comme dans la définition précédente) admet un point selle $(u,p,q) \in X \times \mathbb{R}_+^{N_u} \times \mathbb{R}^{N_e}$. Alors u minimise J sur K, et l'on a $p_i \varphi_i(u) = 0$ pour tout i.*

On a donc réussi à faire le lien entre le problème primal et le problème dual, permettant ainsi à qui doit résoudre un problème de minimisation de choisir la formulation qui lui semble plus confortable.

3.2 Transport optimal discret

3.2.1 Le problème de Monge

La théorie du transport optimal prend sa source dans un problème très simple et très concret d'optimisation de coût de transport de marchandises entre leur lieu de production et un lieu d'exploitation. Mathématiquement parlant, il se formule comme suit :

Problème 3.1 (Problème de Monge, ou **MP**)**.** Soient N entier et $(c_{i,j})_{1\leq i,j\leq N} \in \mathbb{R}^{N^2}$ un collection de coûts. Le problème consiste à trouver une bijection φ qui minimise la quantité

$$\sum_{i=1}^{N} c_{i,\varphi(i)}.$$

Le problème ci-dessus ne présente pas d'intérêt théorique particulier, l'existence est évidemment assurée et la recherche d'un solution peut-être extrêmement pénible. Nous allons voir que l'on a intérêt, comme Kantorovich l'a fait, de poser une version relaxée du problème, le rendant immédiatement beaucoup plus intéressant.

3.2.2 Le problème de Monge-Kantorovich

L'idée consiste à partir du même problème mais de relaxer certaines contraintes. Concrètement, au lieu d'imposer un transfert de matière total d'un point de production vers un seul point d'exploitation, on se laisse la possibilité de répartir le transfert entre plusieurs points d'exploitation. Dans ce qui suit on notera $\gamma_{i,j}$ la quantité de matière allant de i vers j. On appellera $\gamma = (\gamma_{i,j})$ un *plan de transport.*

Problème 3.2 (Problème de Monge-Kantorovich, ou **MKP**)**.** On considère deux ensembles finis X et Y de cardinaux respectifs N et $M \in \mathbb{N}$, et l'on se donne une collection de coûts $c_{i,j} \in \mathbb{R}$. On considère deux mesures de probabilité discrètes μ et ν sur X et Y, respectivement. On cherche à minimiser le coût :

$$C(\gamma) = \sum_{i,j} c_{i,j}\gamma_{i,j}$$

sous la contrainte que γ transporte μ vers ν, c'est-à-dire :

$$\gamma_{i,j} \geq 0, \quad \sum_{j} \gamma_{i,j} = \mu_i,\ \forall i \quad \sum_{i} \gamma_{i,j} = \nu_j\ \forall j,$$

ce que l'on écrira $\gamma \in \Pi(\mu, \nu)$.

Remarque. Les contraintes imposées sur les plans de transports peuvent se formuler en des termes probabilistes, en considérant γ comme une loi de probabilité sur $X \times Y$ dont les marginales sur X et Y sont respectivement μ et ν.

Remarque. L'existence d'une solution au **MKP** est claire dans le cadre ici présenté, dans la mesure, où le coût est continu en γ et que Π est fermé borné, donc compact car en dimension finie. Cependant, on peut formuler les **MP** et **MKP** de manière générale, et la situation sera alors beaucoup plus délicate.

Dans le cas où les cardinaux sont les mêmes et les mesures équidistribuées, on peut préciser le lien entre le modèle relaxé fondé sur les plans de transports et le problème de Monge.

Proposition 3.2.1. *On se place dans le cas $N = M$, et $\mu_i = \mu_j = 1/N$, et on note Π_S l'ensemble des plans de transport associés à une affectation, c'est-à-dire tels que $\gamma_{i,j} = \delta_{i,\varphi(i)}$ où φ est une permutation. L'ensemble Π est l'enveloppe convexe de Π_S.*

Preuve. C'est une conséquence du théorème de Krein-Milman en dimension finie.

Corollaire 3.2.2. *On se place dans les mêmes conditions que la propriété 3.2.1. Le problème de Monge-Kantorovich admet au moins une solution dans Π_S, i.e une solution au problème de Monge.*

Ainsi, la relaxation du problème n'aboutit pas à un problème complètement différent dans la mesure où les solutions de **MP** et de **MKP** sont liées. Mais nous y avons véritablement gagné quelque chose puisque le **MKP** a l'avantage de se présenter comme un problème d'optimisation sous contrainte, nous permettant ainsi de nous placer dans un cadre avec des résultats, des méthodes et des outils connus et utiles. Cela va entre autres nous offrir la possibilité d'étudier un problème potentiellement plus confortable : le problème dual.

3.2.3 Formulation duale du problème de Monge-Kantorovich

La formulation duale du **MKP** se fonde sur l'expression duale des contraintes marginales :

$$\sum_j \gamma_{i,j} = \mu_i \Longleftrightarrow \sum_{i=1}^{N} p_i \left(\mu_i - \sum_j \gamma_{i,j} \right) = 0 \quad \forall p \in \mathbb{R}^N,$$

$$\sum_i \gamma_{i,j} = \nu_i \Longleftrightarrow \sum_{j=1}^{M} q_i \left(\nu_j - \sum_i \gamma_{i,j} \right) = 0 \quad \forall q \in \mathbb{R}^M.$$

Conformément à la définition 3.1.5, on introduit le Lagrangien :

$$L : (\gamma, p, q) \in V \times \Lambda \mapsto \sum_{i,j} c_{i,j}\gamma_{i,j} + \sum_{i=1}^{N} p_i \left(\mu_i - \sum_j \gamma_{i,j}\right) + \sum_{j=1}^{M} q_i \left(\nu_j - \sum_i \gamma_{i,j}\right)$$

avec $V = \mathbb{R}_+^{NM}$ et $\Lambda = \mathbb{R}^N \times \mathbb{R}^M$. Le problème primal est le problème consistant minimiser la fonctionnelle :

$$F(\gamma) = \sup_{p,q} L(\gamma, p, q) = \begin{cases} \sum_{i,j} c_{i,j}\gamma_{i,j} & \text{si } \gamma \in \Pi \\ +\infty & \text{sinon} \end{cases}$$

ce qui revient bien à résoudre le **MKP**. Le problème dual consiste à maximiser la fonctionnelle duale $G(p,q) = \inf_\gamma L(\gamma, p, q)$. Cette fonctionnelle s'exprime de la manière suivante :

$$\begin{aligned} G(p,q) &= \inf_{\gamma \in V} \left(\sum_{i,j} (c_{i,j} - p_i - q_j)\gamma_{i,j} + \sum_{i=1}^{N} p_i\mu_i + \sum_{j=1}^{M} q_j\nu_j \right) \\ &= \sum_{i=1}^{N} p_i\mu_i + \sum_{j=1}^{M} q_j\nu_j + \inf_{\gamma \in V} \left(\sum_{i,j} (c_{i,j} - p_i - q_j)\gamma_{i,j} \right). \end{aligned}$$

Comme γ parcourt $V = \mathbb{R}_+^{NM}$, l'inf ci-dessus vaut $-\infty$ à moins que l'on ait $p_i + q_j \leq c_{i,j}$ pour tous i, j et 0 dans ce dernier cas. On a donc :

$$G(p,q) = \begin{cases} \displaystyle\sum_{i=1}^{N} p_i\mu_i + \sum_{j=1}^{M} q_j\nu_j & \text{si } \; p_i + q_j \leq c_{i,j} \quad \forall i, j \\ -\infty & \text{sinon} \end{cases}$$

3.3 Notes sur le transport optimal général

Si le transport optimal discret possède déjà un intérêt en soi, il peut se formuler dans un cadre plus général, donnant lieu à des complications notables mais permettant la résolution de problèmes très riches. On considère par la suite deux espaces mesurés (X, μ) et (Y, ν) et une fonction "coût" continue $c : X \times Y \to \mathbb{R}_+$.

Problème 3.3 (Problème de Monge général)**.** Trouver une application bijective préservant la mesure entre (X, μ) et (Y, ν) minimisant la quantité :

$$\int_X c(x, s(x)) d\mu(x).$$

Problème 3.4 (Problème de Monge-Kantorovich général)**.** Trouver une mesure de probabilité γ sur $X \times Y$ avec pour marginales μ et ν sur X et Y, c'est-à-dire que, pour tout $f \in C_c^0(X)$ et $g \in C_c^0(Y)$:

$$\int_X f(x) d\mu(x) = \int_{X \times Y} f(x) d\gamma(x, y),$$
$$\int_Y g(y) d\nu(y) = \int_{X \times Y} g(y) d\gamma(x, y)$$

minimisant le coût :

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

On généralise la notion de problème dual exposée dans la section 3.1 par

Problème 3.5 (Problème dual)**.** Minimiser la quantité :

$$\int_X f d\mu + \int_Y g d\nu$$

où f et g vérifient :

$$f(x) + g(x) \leq c(x, y).$$

Les questions d'existence se révèlent bien plus compliquées que dans le cas discret, et l'obtention de résultats satisfaisants nécessitent généralement comme des hypothèses de compacité.

4 Le théorème de Brenier

Comme on l'a vu en section 2, l'approche géométrique du théorème de factorisation polaire s'est révélée partiellement infructueuse, dans la mesure où celle-ci introduisait une difficulté qui semble difficile à dépasser : celle de la structure de semi-groupe. Cependant, il est possible d'opter pour une autre approche qui est en mesure de produire un résultat unifiant les théorèmes énoncés en introduction. Brenier a en effet identifié un lien fort entre la théorie du transport optimal et le problème de factorisation polaire, en le formulant comme un problème de Monge. Cependant, s'il est tout à fait possible de présenter directement cette formulation ainsi que la démonstration qu'elle permet de produire, il nous a paru intéressant d'étudier plus en profondeur le lien entre ces problématiques à travers un cas plus simple : le cas discret.

4.1 Un exemple concret pour comprendre le lien entre transport optimal et factorisation polaire

Soient $X = \{x_1, \cdots, x_N\}$ et $Y = \{y_1, \cdots, y_M\}$ deux ensembles finis, et μ, ν deux mesures discrètes sur $\mathbb{R}^d$ définies par :

$$\mu = \sum_{i=1}^{N} \delta_{x_i}, \quad \nu = \sum_{j=1}^{M} \delta_{y_j}.$$

On s'intéresse au problème de transport entre μ et ν avec un coût quadratique $c_{i,j} = |y_j - x_i|^2$. On note γ un minimiseur du problème de transport, et Γ l'application multivaluée qui à x_i fait correspondre l'ensemble des y_j tels que $\gamma_{i,j} > 0$, c'est-à-dire l'ensemble des points vers lesquels le transport est effectif.

Si on fait un dessin, on se doute bien qu'avec un tel coût, l'optimalité devrait être liée à une forme de monotonie du plan de transport, au sens où celui-ci doit éviter des croisements. Il est alors aisé de montrer le résultat suivant :

Proposition 4.1.1. Γ *est monotone au sens suivant :*

$$(y' - y) \cdot (x' - x) \geq 0 \quad \forall x, x' \in X, \quad y \in \Gamma(x),\ y' \in \Gamma(x'). \tag{1}$$

Preuve. Supposons l'existence de $x, x' \in X$ et de $(y, y') \in \Gamma(x) \times \Gamma(x')$ contrevenant à (1). On va construire de manière très intuitive un plan de transport moins coûteux, contredisant ainsi l'optimalité de γ. On définit le plan $\tilde{\gamma}$ par :

$$\tilde{\gamma}_{i,j} = \begin{cases} \gamma_{i,j} - \varepsilon & \text{si } x, y = x_i, y_j \\ \gamma_{i,j} + \varepsilon & \text{si } x, y' = x_i, y_j \\ \gamma_{i,j} - \varepsilon & \text{si } x', y' = x_i, y_j \\ \gamma_{i,j} + \varepsilon & \text{si } x', y = x_i, y_j \\ \gamma_{i,j} & \text{sinon} \end{cases}$$

où $\varepsilon > 0$ est choisi de sorte à ce que les quantités définies ci-dessus soient positives, ce qui est possible car, par hypothèse $y \in \Gamma(x)$ et $y' \in \Gamma(x')$. Il est aisé de montrer que $\tilde{\gamma}$ respecte bien les contraintes marginales, il s'agit donc bien d'un plan de transport. De plus :

$$
\begin{aligned}
C(\tilde{\gamma}) &= \sum_{(z,t)\in X\times Y} \tilde{\gamma}_{z,t} c_{z,t} \\
&= \underbrace{\sum_{z,t\neq x,x',y,y'} c_{z,t}\gamma_{z,t}}_{\alpha} + \sum_{z,t=x,x',y,y'} c_{z,t}\gamma_{z,t} \\
&= \alpha + (\gamma_{x,y}+\varepsilon)\, c_{x,y} + (\gamma_{x,y'}-\varepsilon)\, c_{x,y'} + (\gamma_{x',y'}+\varepsilon)\, c_{x',y'} + (\gamma_{x',y}-\varepsilon)\, c_{x',y} \\
&= \underbrace{\alpha + \gamma_{x,y}c_{x,y} + \gamma_{x,y'}c_{x,y'} + \gamma_{x',y}c_{x',y} + \gamma_{x',y'}c_{x',y'}}_{C(\gamma)} + \varepsilon(c_{x,y} + c_{x',y'} - c_{x',y} - c_{x,y'}) \\
&= C(\gamma) + \varepsilon(|y-x|^2 + |x'-y'|^2 - |y-x'|^2 - |y'-x|^2) \\
&= C(\gamma) + 2\varepsilon(x'\cdot y + x\cdot y' - x\cdot y - x'\cdot y') \\
&= C(\gamma) + 2\varepsilon\underbrace{(y'-y)\cdot(x'-x)}_{<0} < C(\gamma),
\end{aligned}
$$

ce qui contredit l'optimalité de γ.

Ce dernier résultat ne constitue malheureusement pas une caractérisation de l'optimalité, il nous faut donc creuser un peu l'idée de monotonie pour aboutir à une réelle caractérisation. Celle-ci est donnée par la proposition suivante :

Proposition 4.1.2. Γ *est cycliquement monotone, c'est-à-dire que :*

$$\sum_{k=1}^{p} x_{i_k}\cdot(y_{j_k} - y_{j_{k-1}}) \geq 0 \tag{2}$$

pour tous (i_1, j_1),...,(i_p, j_p) *dans le support de* γ*, avec la convention* $j_0 = j_p$

Preuve. On reprend les mêmes idées que précédemment : supposons l'existence de (i_1, j_1),...,(i_p, j_p) dans le support de γ, avec la convention $j_0 = j_p$, contrevenant à (2). On définit alors $\tilde{\gamma}$ par

$$\tilde{\gamma}_{i,j} = \begin{cases} \gamma_{i,j} - \varepsilon & \text{si } i,j = x_{i_k}, y_{j_k} \\ \gamma_{i,j} + \varepsilon & \text{si } i,j = x_{i_k}, y_{j_{k-1}} \\ \gamma_{i,j} & \text{sinon} \end{cases}$$

De la même manière que précédemment, on montre que $\tilde{\gamma}$ est bien un plan de transport pour ε bien choisi. Notons I l'ensemble des couples (x_{i_k}, y_{j_k}) et $x_{i_k}, y_{j_{k-1}}$. On a :

$$
\begin{aligned}
C(\tilde{\gamma}) &= \sum_{i,j} \tilde{\gamma}_{i,j} c_{i,j} \\
&= \sum_{(i,j)\notin I} \gamma_{i,j} c_{i,j} + \sum_{k=1}^{p} \tilde{\gamma}_{x_{i_k}, y_{j_k}} c_{x_{i_k}, y_{j_k}} + \sum_{k=1}^{p} \tilde{\gamma}_{x_{i_k}, y_{j_{k-1}}} c_{x_{i_k}, y_{j_{k-1}}} \\
&= \sum_{(i,j)\notin I} \gamma_{i,j} c_{i,j} + \sum_{k=1}^{p} (\gamma_{x_{i_k}, y_{j_k}} - \varepsilon) c_{x_{i_k}, y_{j_k}} + \sum_{k=1}^{p} (\gamma_{x_{i_k}, y_{j_{k-1}}} + \varepsilon) c_{x_{i_k}, y_{j_{k-1}}} \\
&= C(\gamma) + \varepsilon \sum_{k=1}^{p} (c_{x_{i_k}, y_{j_{k-1}}} - c_{x_{i_k}, y_{j_k}}) \\
&= C(\gamma) + \varepsilon \sum_{k=1}^{p} \left(|x_{i_k} - y_{j_{k-1}}|^2 - |x_{i_k} - y_{j_k}|^2 \right) \\
&= C(\gamma) + \varepsilon \underbrace{\sum_{k=1}^{p} \left(|y_{j_{k-1}}|^2 - |y_{j_k}|^2 \right)}_{=0} + 2\varepsilon \underbrace{\sum_{k=1}^{p} x_{i_k} \cdot (y_{j_k} - y_{j_{k-1}})}_{<0} \\
&< C(\gamma),
\end{aligned}
$$

ce qui contredit l'optimalité de γ.

Cette condition de monotonie cyclique est plus forte qu'une simple monotonie, et il s'avère qu'elle permet une caractérisation de l'optimalité :

Proposition 4.1.3. *On se place dans le cas de mesure équidistribuées, avec γ un plan de transport. Si Γ est cycliquement monotone, alors γ est optimal*

Preuve. On se restreindra ici à montrer la proposition dans le cas où le plan de transport est de type permutation. Quitte à renuméroter, on peut supposer que γ envoie x_1 sur y_1, etc. On va commencer par montrer que le coût de tout plan de transport associé à une permutation de type cycle est supérieur à $C(\gamma)$, avant de généraliser. Soit donc $\sigma \in S_N$ de type permutation circulaire. On effectue la construction suivante :

$$
\begin{aligned}
j_N = N, &\quad i_N = N \\
j_{N-1} = \sigma(i_N), &\quad i_{N-1} = j_{N-1} \\
j_{N-2} = \sigma(i_{N-1}), &\quad i_{N-2} = j_{N-2} \\
&\vdots \\
j_1 = \sigma(i_2) &\quad i_1 = j_1.
\end{aligned}
$$

Vu que σ est une permutation, $\sigma(i_1) = \sigma^2(i_2) = ... = \sigma^N(i_N) = N$, et comme de plus σ est une permutation circulaire, on a une bijection entre $\{i_1, ..., i_N\}$ et $\{1, ...N\}$. On en déduit :

$$\begin{aligned}\sum_{i=1}^{N} |y_{\sigma(i)} - x_i|^2 - C(\gamma) &= \sum_{i=1}^{N} |y_{\sigma(i)} - x_i|^2 - \sum_{i=1}^{N} |y_i - x_i|^2 \\ &= 2\sum_{i=1}^{N} x_i \cdot (y_i - y_{\sigma(i)}) \\ &= 2\sum_{k=1}^{N} x_{i_k} \cdot (y_{i_k} - y_{\sigma(i_k)}) \\ &= 2\sum_{k=1}^{N} x_{i_k} \cdot (y_{j_k} - y_{j_{k-1}}) \geq 0.\end{aligned}$$

On généralise à tout plan de transport associé à une permutation quelconque ϕ en décomposant cette dernière en produit de cycles à supports disjoints. Enfin, on montre que γ est optimal en se souvenant de la propriété 3.2.1, qui nous permet de dire que, si $\tilde{\gamma}$ est un plan de transport, alors $\tilde{\gamma}$ est un barycentre de plans de transports associés à des permutations, ce qui permet de conclure aisément.

On a donc réussi à trouver une bonne définition de monotonie telle que nous le suggérait l'intuition. Une autre idée toute aussi naturelle aurait été de définir la monotonie par analogie avec le cas réel à l'aide de fonctions convexes. En effet, il paraît légitime de penser une application monotone comme le gradient d'une fonction convexe, et nous verrons que ces deux approches sont équivalentes, comme en témoigne la proposition suivante.

Proposition 4.1.4. *Soit Φ une fonction convexe régulière. Alors l'application $\Gamma = \nabla\Phi$ est cycliquement monotone.*

Preuve. Soient (i_1, j_1),...,(i_p, j_p) tels que $j_0 = j_p$. On a :

$$\begin{aligned}\sum_{i=1}^{p} x_{i_k} \cdot (\nabla\Phi(x_{i_k}) - \nabla\Phi(x_{ik-1})) &= \sum_{k=1}^{p} x_{i_k} \cdot \nabla\Phi(x_{i_k}) - \sum_{k=1}^{p} x_{i_k} \cdot \nabla\Phi(x_{i_{k-1}}) \\ &= \sum_{k=1}^{p} x_{i_k} \cdot \nabla\Phi(x_{i_k}) - \sum_{k=1}^{p} x_{i_{k+1}} \cdot \nabla\Phi(x_{i_k})\end{aligned}$$

avec la convention $x_{p+1} = x_1$

$$= \sum_{i=1}^{p} \nabla\Phi(x_{i_k}) \cdot (x_{i_k} - x_{i_{k+1}}).$$

Or, par convexité de Φ :

$$\nabla\Phi(x)\cdot(x-y) \geq \Phi(x)-\Phi(y) \quad \forall x,y$$

ce qui nous permet d'avoir :

$$\begin{aligned}\sum_{i=1}^{p} x_{i_k}\cdot(\nabla\Phi(x_{i_k})-\nabla\Phi(x_{ik-1})) &\geq \sum_{k=1}^{p}\big(\Phi(x_{i_k})-\Phi(x_{i_{k+1}})\big)\\ &=0,\end{aligned}$$

d'où la monotonie cyclique.

Corollaire 4.1.5. *Si γ est un plan de transport dont le Γ associé peut s'écrire comme le gradient d'une fonction convexe, alors γ est optimal.*

Il s'avère qu'une forme de réciproque du corollaire est vraie, nous permettant ainsi de voir la notion précédente comme une caractérisation de l'optimalité.

Proposition 4.1.6. *On se place toujours dans le cas où les cardinaux sont égaux et les mesures équidistribuées. Soit γ un plan de transport optimal de type permutation. Quitte à renuméroter, on suppose que γ envoie x_i sur y_i. Alors il existe Φ convexe différentiable telle que*

$$y_i \in \partial\Phi(x_i),$$

où conformément à la définition donnée dans l'énoncé du théorème 2.2.1, $\partial J(u)$ désigne le sous-gradient de J en u, c'est-à-dire l'ensemble des vecteurs $v \in H$ tels que $\forall h \in H,\ J(h) \geq J(u) + \langle v, h-u\rangle$.

Preuve. Nous nous contenterons ici de démontrer cette proposition dans le cas plus simple où Pour de simples questions de commodité, on considérera plutôt un coût $c_{i,j} = \frac{1}{2}|y_j - x_i|^2$. Soit (p,q) un couple de solutions au problème dual.On se contentera ici de démontrer cette proposition dans le cas où l'égalité $p_i + q_j = c_{i,j}$ n'est réalisée que sur le support de γ. On définit alors :

$$\Phi(x) = \max_j \underbrace{\left(x\cdot y_j - \frac{1}{2}|y_j|^2 + q_j\right)}_{\phi_j}.$$

Φ est évidemment convexe en tant que maximum de fonctions affines. De plus, on note que $\forall i,j$:

$$\begin{aligned}
\phi_i(x_i) - \phi_j(x_i) &= x_i \cdot y_i - \frac{1}{2}|y_i|^2 + q_i - x_i \cdot y_j + \frac{1}{2}|y_j|^2 - q_j \\
&= x_i \cdot y_i - \frac{1}{2}|y_i|^2 + c_{i,i} - p_i - x_i \cdot y_j + \frac{1}{2}|y_j|^2 - q_j \\
&= \underbrace{\frac{1}{2}|x_i|^2 - x_i \cdot y_j + \frac{1}{2}|y_j|^2}_{=c_{i,j}} - \underbrace{(p_i + q_j)}_{<c_{i,j}} \\
&> 0.
\end{aligned}$$

Ainsi, on a qu'il existe un ouvert contenant x_i tel que $\Phi(x) = \phi_i(x)$ sur cet ouvert, ce qui nous permet de voir que Φ est différentiable en x_i, avec :

$$\nabla\Phi(x_i) = y_i.$$

Le travail qui a été fait ci-dessus permet d'exhiber un lien fort entre les notions de factorisation polaire et de transport optimal. En effet, on vient de voir par les propositions 4.1.3 et 4.1.6 que l'optimalité d'un plan de transport pouvait se caractériser par une certaine forme de monotonie, l'idée principale étant que l'optimalité interdit les "croisements". Ainsi, si on se donne un ensemble fini de vecteurs X et $u : X = \{x_1, ..., x_N\} \to \mathbb{R}^d$ injective, alors il existe s bijection de X dans $u(X)$ (qui joue donc le rôle d'une application préservant la mesure) et Φ convexe telle que :

$$u = \nabla\Phi \circ s$$

en choisissant simplement s comme un plan de transport optimal de type permutation.

4.2 Démonstration du théorème

La section précédente nous ayant permis de nous forger une intuition sur le sujet, nous pouvons revenir au problème avec un regard neuf et une nouvelle approche : celle du transport optimal. On rappelle le problème de projection 1.1 que Brenier s'était initialement posé :

Problème 4.1. Trouver, pour $u \in L^2(\Omega, \mathbb{R}^d)$, s un réarrangement qui minimise $\int_\Omega \|u(x) - s(x)\|^2 dx$, ou de façon équivalente qui maximise $\int_\Omega u(x) \cdot s(x) dx$.

On remarque que ce problème peut se formuler comme un problème de Monge pour le coût $c(x, y) = \|u(x) - y\|^2$, que l'on peut relaxer en le problème de Monge-Kantorovich associé. Néanmoins, le choix du coût $c(x, y) = \|u(x) - y\|^2$ pose quelques difficultés techniques, car u ne dispose pas nécessairement de la régularité suffisante. Il serait donc préférable d'intégrer u aux contraintes marginales, en formulant le **MKP** comme suit :

Problème 4.2. Trouver γ mesure de probabilité sur $\mathbb{R}^d \times \bar{\Omega}$ minimisant le coût :

$$C(\gamma) = \int_{\mathbb{R}^d \times \bar{\Omega}} \|x - y\|^2 d\gamma(x, y)$$

avec γ ayant pour marginale sur $\bar{\Omega}$ la mesure de Lebesgue (notée β) et pour marginale sur $\mathbb{R}^d$ la mesure image de u (notée α).

Remarque. On note qu'on présente ici des mesures considérées comme des mesures de probabilité, alors que le problème que l'on pose ne l'impose pas forcément. Cependant, les mesures mobilisées dans le problème sont finies, on peut donc s'y ramener en divisant simplement par la mesure de l'espace entier.

Pour des raisons qui apparaîtront plus clairement par la suite, nous nous intéresserons à un problème de Monge-Kantorovich plus général, où α est une mesure de probabilité sur $\mathbb{R}^d$ telle que $\int (1 + \|y\|) d\alpha(y) < +\infty$, et β une mesure de probabilité sur $\bar{\Omega}$ telle que $\beta(\partial\Omega) = 0$ et telle que $d\beta(z) = \beta(z)$ où β est une fonction plus grande qu'une constante strictement positive sur tout compact de Ω, ce qui permet à la mesure β d'avoir les mêmes ensembles négligeables que la mesure de Lebesgue. On étudiera donc les problèmes suivants :

Problème 4.3 (**MKP** primal)**.** Trouver une mesure de probabilité p sur $\mathbb{R}^d \times \bar{\Omega}$ qui maximise $\int y \cdot z \, dp(y, z)$ sous les conditions suivantes : $\int \|y\| dp(y, z) < +\infty$, α et β sont les marginales de p sur $\mathbb{R}^d$ et $\bar{\Omega}$, c'est-à-dire :

$$\int f(y) dp(y, z) = \int f(y) d\alpha(y)$$

pour tout $f \in C(\mathbb{R}^d)$ tel que $|f(y)| \leq cst(1 + \|y\|)$, et

$$\int g(z) dp(y, z) = \int g(z)\beta(z) dz, \quad \forall g \in C(\bar{\Omega}).$$

Le problème dual du problème ci-dessus consiste à rechercher ϕ, ψ telles que $\phi(y) + \psi(z) \leq \|z - y\|^2$ et minimisant $\int \phi d\alpha + \int \psi d\beta$. On préférera cependant la formulation suivante, toujours pour des raisons qui s'éclairciront dans les pages à venir.

Problème 4.4 (**MKP** dual)**.** Trouver $\phi \in C(\mathbb{R})^d \cap L^1(\mathbb{R}^d, \alpha)$ et $\psi \in L^1(\Omega, \beta)$ qui minimisent $\int \phi d\alpha$ et vérifient :

$$\int \psi d\beta = 0$$

$$\phi(y) + \psi(z) \geq y \cdot z, \quad \forall (y, z) \in \mathbb{R}^d \times \Omega.$$

Remarque. Le lecteur aura noté que les formulations ne s'inscrivent pas exactement dans le cadre des énoncés de la section 3.1, on prêtera bien entendu attention à expliciter les liens entre les différents problèmes.

Il s'avère que le problème le plus commode à étudier est une formulation mixte des problèmes primal et dual.

Problème 4.5 (**MKP** mixte)**.** Trouver $\phi \in C(\mathbb{R}^d) \cap L^1(\mathbb{R}^d, \alpha)$, $\psi \in C(\Omega) \cap L^1(\Omega, \beta)$ et une mesure de probabilité p sur $\mathbb{R}^d$ et $\bar{\Omega}$ tels que :

$$\phi(y) + \psi(z) \geq y \cdot z, \quad \forall (y, z) \in \mathbb{R}^d \times \Omega,$$

$\int \|y\| dp(y, z) < +\infty$, α et β sont les marginales de p sur $\mathbb{R}^d$ et $\bar{\Omega}$, $\int \psi d\beta = 0$ et $\int \phi d\alpha \leq \int y \cdot z \, dp(y, z)$

Remarque. Notons que, dans les **MKP** dual et mixte, p est nécessairement une mesure de probabilité tendue et $p(\mathbb{R}^d \times \partial\Omega) = \beta(\partial\Omega) = 0$.

4.2.1 Lemmes techniques

Nous allons nous appuyer sur deux propositions pour la démonstration du théorème de factorisation polaire. Ces propositions sont les résultats techniques les plus importants de la preuve ; ils donnent respectivement les différents liens entre les composantes de la solution du problème de Monge-Kantorovich et de bonnes propriétés de convergence des solutions dudit problème. Mais avant cela, démontrons un lemme technique préliminaire :

On pose :

$$K_0 = \{\psi \in W^{1,2}(\Omega, \beta) \cap C(\Omega) \textit{ tel que } \int \psi d\beta = 0 \textit{ et } \exists \tilde{\psi} : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}, \\ \text{convexe, semi-continue inférieurement, et} \quad \psi = \tilde{\psi} \textit{ sur } \Omega\}$$

Lemme 4.2.1 (Propriétés de K_0)**.** *On a :*

1. *On pose $M = \int \|\nabla\psi\| d\beta$. $\forall \psi \in K_0$, on a*
   (a) $-2rM \leq \psi(z) \; \forall z \in \Omega$
   (b) $\psi(z) \leq 2r\|\nabla\psi(z)\| \; \beta - p.p \;\; sur \;\; \Omega$
   (c) $\psi(z) \leq C(\delta) rM \; \forall z \in \Omega$ *où* $\delta = d(z, \partial\Omega)$
   (d) $\int |\psi| d\beta \leq 4rM$

   (e) *La transformée de Legendre ψ^* est r-lipschitzienne, et , $-r\|y\| \leq \psi^*(y) \leq r\|y\| + 2rM \; \forall y \in \mathbb{R}^d$ et, $\psi = \psi^{**} \;\; sur \;\; \Omega$*

2. *Soit $(\psi_n)_{n\in\mathbb{N}} \in K_0^{\mathbb{N}}$ tel que $\int \|\nabla\psi_n\| d\beta \leq M$, alors quitte à extraire on peut trouver $\psi \in C(\Omega) \cap L^1(\Omega,\beta), \phi \in C(\mathbb{R}^d)$ tel que :*
   (a) $\phi(y) + \psi(z) \geq \langle y|z\rangle \quad \forall y \in \mathbb{R}^d, z \in \Omega$
   (b) *$\psi_n \to \psi$ au sens $L^1(\Omega,\beta)$ et uniformément sur tout compact de Ω*
   (c) *$\psi_n^* \to \phi$ uniformément sur tout compact de $\mathbb{R}^d$*
   (d) $|\psi_n^*(y)| \leq r(2M + \|y\|)$

Preuve. Première étape : les estimations qui sont les 4 premières assertions de 1 : Prenons $\psi \in K_0$, qui est continue sur Ω et telle que $\int \psi d\beta = 0$ par définition. On a l'existence de $z_0 \in \Omega$ tel que $\psi(z_0) = 0$ (car sinon l'intégrale serait strictement positive ou strictement négative). Puis $\psi \in W^{1,1}(\Omega,\beta)$ et donc $\nabla\psi$ est défini $\beta{-}p.p$ (et aussi $\mu{-}p.p$ car ces deux mesures ont les mêmes ensembles négligeables comme vu précédemment). Par convexité de ψ :

$$\psi(z_0) \geq \psi(z) + \langle \nabla\psi(z)|z_0 - z\rangle \, \beta - p.p \ sur \ \Omega.$$

Ainsi,

$$\psi(z) \leq \langle \nabla\psi(z)|z - z_0\rangle \leq \|\nabla\psi(z)\| 2r.$$

Par Cauchy-Schwarz. De plus :

$$\psi(z) \geq \psi(z') + \langle \nabla\psi(z')|z - z'\rangle \, .$$

En intégrant par rapport à z' sur Ω :

$$\psi(z) = \psi(z)\int d\beta \geq \int \psi d\beta + \int \langle \nabla\psi(z)|z - z'\rangle \, d\beta(z').$$

D'où,

$$\psi(z) \geq -2r \int \|\nabla\psi\| d\beta = -2Mr \ \beta - p.p.$$

De là,

$$-2rM \leq \psi(z) \leq 2r\|\nabla\psi(z)\| \ \beta - p.p.$$

De ces premières estimations, on déduit $|\psi(z)| \leq 2rM + 2r\|\nabla\psi(z)\|$ et en intégrant :

$$\int |\psi| d\beta \leq 2rM + 2rM = 4rM.$$

Il ne reste plus qu'une inégalité à démontrer :
On fixe $z \in \Omega$, on pose $\delta = d(z, \partial\Omega) > 0$. On a $z + \frac{\delta}{2}B(0,1) \subset \Omega$ et on définit

$$\overline{\omega} = \left\{ z' \in \Omega / d(z', \partial\Omega) \geq \frac{\delta}{2} \right\} \subset \Omega.$$

Par hypothèse, $\beta \geq \eta(\delta) > 0$ sur $\overline{\omega}$. Ainsi, ψ est Lebesgue intégrable sur $\overline{\omega}$, et,

$$\begin{aligned} \int_{\overline{\omega}} |\psi(z')| dz' &\leq \int_{\overline{\omega}} |\psi(z')| \frac{1}{\eta(\delta)} \beta(z') dz' \\ &\leq \frac{4rM}{\eta(\delta)} \end{aligned}$$

par les estimations précédentes. La convexité de ψ nous donne le droit d'utiliser l'inégalité de Jensen car le domaine d'intégration est de mesure finie non nulle :

$$\begin{aligned} \psi(z) &\leq \frac{1}{\mu(z + \frac{\delta}{2} B(0,1))} \int_{z + \frac{\delta}{2} B(0,1)} \psi(z') dz' \\ &\leq \frac{1}{\mu(z + \frac{\delta}{2} B(0,1))} \int_{\overline{\omega}} \psi(z') dz' \end{aligned}$$

car la boule centrée en z est incluse dans $\overline{\omega}$

$$\begin{aligned} &\leq \frac{4rM}{\eta(\delta) \mu(\frac{\delta}{2} B(0,1))} \\ &= \frac{4rM}{\eta(\delta) (\frac{\delta}{2})^d \mu(B(0,1))} \\ &= C(\delta) rM. \end{aligned}$$

D'où,

$$\psi(z) \leq C(\delta) rM \quad \forall z \in \Omega.$$

Deuxième étape : résultats sur la transformée de Legendre :
ψ^* est bien définie sur $\mathbb{R}^d$ car ψ minorée sur Ω*.
Puis, par définition,

$$\begin{aligned} \psi^*(y) &\geq \langle z_0 | y \rangle - \psi(z_0) \\ &\geq -r \|y\| \end{aligned}$$

par Cauchy-Schwarz et parce que $\psi(z_0) = 0$. Puis,

$$\begin{aligned} \psi^*(y) &\leq r\|y\| - \inf_{\Omega} \psi \\ &= r\|y\| + 2rM = r(\|y\| + 2M) \end{aligned}$$

ψ^* est r-lipschitzienne : si $y, y' \in \mathbb{R}^d$, si $\varepsilon > 0$,

$$\psi^*(y) - \psi^*(y') \leq \langle z|y\rangle - \psi(z) + \varepsilon - (\langle z|y'\rangle - \psi(z))$$

car on prend un élément z tel qu'il approche le premier sup par rapport à y, puis on peut minorer $\psi(y')$ par rapport à la valeur prise en z.

$$\begin{aligned} &= \langle z|y - y'\rangle + \varepsilon \\ &\leq r\|y - y'\| + \varepsilon \end{aligned}$$

et ce, pour tout $\varepsilon > 0$. Puis, par symétrie des rôles joués par y et y', on a la même majoration en les échangeant, d'où finalement :

$$|\psi^*(y) - \psi^*(y')| \leq r\|y - y'\|.$$

Il ne reste plus qu'à montrer que $\psi^{**} = \psi$. Pour cela, nous devons faire un petit détour par l'analyse convexe (tirés de [2]), et nous admettrons le lemme suivant :

Lemme 4.2.2. *Si $f : \mathbb{R}^d \to$ convexe, semi-continue inférieurement et n'étant pas identiquement égale à $+\infty$, jamais égale à $-\infty$ (on dit alors que f est propre), alors $\forall x \in \mathbb{R}^d, f(x) = E_f(x)$ où*

$$E_f(x) = \sup\{h(x), h \text{ affine et minorant } f\}.$$

On en déduit la proposition suivante :

Proposition 4.2.3. *si $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ convexe, semi-continue inférieurement et propre, alors $f^{**} = f$*

Preuve (Proposition 4.2.3). Tout d'abord, on pose

$$\forall \phi \in \mathbb{R}^d, h_\phi(x) = \langle \phi|x\rangle - f^*(\phi) \in \{\text{minorants affines de} f\} = AM(f).$$

On pose de plus $AM^* = \left\{h_\phi, \phi \in \mathbb{R}^d\right\} \subset AM(f)$. Alors,

$$f^{**}(x) = \sup_{h \in AM^*} h(x) \leq \sup_{h \in AM(f)} h(x) = E_f(x) \quad \text{pour } x \in \mathbb{R}^d.$$

Inversement, si $h_{\phi,b} \in AM(f)$, alors

$$\forall x \in \mathbb{R}^d, \langle \phi|x\rangle - f(x) \leq -b.$$

D'où, $f^*(\phi) \leq -b$. Ainsi,

$$\forall x \in \mathbb{R}^d \quad \langle \phi|x\rangle - f^*(\phi) \geq \langle \phi|x\rangle + b = h_{\phi,b}(x)$$

donc,

$$f^{**}(x) \geq h(x) \; \forall h \in AM(f).$$

Donc, par passage au sup :

$$f^{**}(x) \geq E_f(x) \quad \forall x \in \mathbb{R}^d.$$

Des deux inégalités, on déduit l'égalité désirée.

Troisième étape : quelques propriétés séquentielles.
Soit $(\psi_n)_{n\in\mathbb{N}}$ tel que $\int \|\nabla\psi_n\| d\beta \leq M$. Alors par les estimations de l'étape 1, on déduit que $(\psi_n)_{n\in\mathbb{N}}$ est bornée en $\|.\|_1$, et si on ajoute l'hypothèse, on a $(\psi_n)_{n\in\mathbb{N}}$ bornée dans $W^{1,1}(\Omega)$. Or, $W^{1,1}(\Omega)$ s'injecte de manière compacte dans $L^1(\Omega)$, d'où quitte à extraire, ψ_n converge dans L^1 vers ψ. Les estimations précédentes nous donnent :

$$\forall n \in \mathbb{N}, \forall z \in \Omega, -2rM \leq \psi_n(z) \leq C(\delta)rM.$$

On a donc une borne uniforme sur chaque compact, et ceci plus le fait que ψ soit convexe nous donne l'existence d'une constante de lipschitz uniforme sur chaque compact. Aussi, par le même argument que précédemment, (ψ_n^*) est r-lipschitzienne, et de plus bornée :

$$\psi^*(y) = \sup_{z\in\Omega} \langle y|z\rangle - \psi_n(z) \leq \|y\|r - \inf_{\psi_n(z)} \leq \|y\|r + 2rM.$$

Puis,

$$\psi_n^*(y) \geq -r\|y\|.$$

En minorant le sup par la valeur en $z_{0,n}$, un point tel que $\psi_n(z_{0,n}) = 0$ d'où

$$|\psi_n^*(y)| \leq r(\|y\| + 2M). \tag{3}$$

Les hypothèses du théorème d'Ascoli sont donc réunies, ont peut donc conclure l'existence de ϕ, ψ tel que quitte à extraire, $\psi_n^* \to \phi$ uniformément sur tout compact et $\psi_n \to \psi$ uniformément sur tout compact (et donc on a gratuitement la convergence L^1 étant donné que ψ_n intégrable pour tout n). Si l'on fait tendre n vers l'infini dans (3) , on obtient :

$$\forall y \in \mathbb{R}^d, |\phi(y)| \leq r(2M + \|y\|).$$

De plus, il est clair, par la définition de la transformée de Legendre, que

$$\forall n \in \mathbb{N}, \; \forall y \in \mathbb{R}^d, \; \forall z \in \Omega, \quad \psi_n^*(y) + \psi_n(z) \geq \langle y|z\rangle .$$

En faisant tendre n vers l'infini, on obtient :

$$\phi(y) + \psi(z) \geq \langle y|z\rangle \quad \forall y, z.$$

On va maintenant montrer les deux propositions suggérées plus haut, sous réserve d'existence d'une solution au **MKP** mixte, ce qui sera démontré en fin d'exposé.

Proposition 4.2.4 (Propriétés de la solution du MKP)**.** *Notons (ϕ, ψ, p) une solution du MKP associée à une mesure α, on a les propriétés suivantes :*

1. $\psi \in K_0$ *et,* $\|\nabla\psi\|_{L^1(\Omega,\beta)} = \int \|y\| d\alpha(y)$
2. $\phi = \psi^*$ $\alpha - p.p$ *sur* $\mathbb{R}^d$*, où* ψ^* *est la transformée de Legendre :*
$$\psi^*(y) = \sup_{z\in\Omega} \{\langle y|z\rangle - \psi(z)\} \, \forall y \in \mathbb{R}^d.$$
3. $dp(y,z) = \delta(y - \nabla\psi(z))\beta(z)dz.$
4. $\int \phi d\alpha = \int \langle y|z\rangle \, dp(y,z)$ *et,* $\int \psi d\beta = 0.$
5. (ψ, ϕ) *est l'unique solution du* ***MKP*** *primal et* p *est l'unique solution du* ***MKP*** *dual. De plus, si* $\alpha \ll \mu$ *:*
6. $z = \nabla\psi^*(y), \; y = \nabla\psi(z) \quad p - p.p$ *sur* $\mathbb{R}^d \times \overline{\Omega}.$
7. $\nabla\psi^*(\nabla\psi(z)) = z \quad \beta - p.p$ *sur* $\overline{\Omega}$ *et,* $\nabla\psi(\nabla\psi^*(y)) = y$ $\alpha - p.p$ *sur* $\mathbb{R}^d.$
8. $dp(y,z) = \delta(y - \nabla\psi^*(y))d\alpha(y).$

Preuve. Etape 1 :On montre tout d'abord les assertions de 1 à 4. Prenons (ϕ, ψ, p) solution du **MKP** mixte, On prend $\tilde{\phi}$ la transformée de Legendre de ψ. Par hypothèse du **MKP** mixte, on a :

$$\phi(y) + \psi(z) \geq \langle y|z\rangle \quad \forall y, z.$$

D'où, en passant au sup après avoir passé le $\psi(z)$ de l'autre coté de l'inégalité :

$$\phi(y) \geq \tilde{\phi}(y) \quad \forall y \in \mathbb{R}^d.$$

On fixe $z_0 \in \Omega$, on a par définition $\tilde{\phi}(y) \geq \langle y|z_0\rangle - \psi(z_0)$. On en déduit que $\tilde{\phi}$ est finie en tout point. De plus elle est r-lipschitzienne comme transformée de Legendre sur Ω. On pose maintenant $\tilde{\psi}$ la transformée de Legendre de $\tilde{\phi}$. Alors, par la proposition 4.2.3, car $\tilde{\phi}$ est propre (ce que l'on vient de montrer) :

$$\tilde{\psi} = \tilde{\phi}^* \text{ et } \tilde{\psi}^* = \tilde{\phi}. \tag{4}$$

De là,

$$\tilde{\psi}(z) \leq \psi(z) \quad \forall z \in \Omega \tag{5}$$

car,

$$\tilde{\psi}(z) \leq \langle z|y\rangle - \tilde{\phi}(y) + \varepsilon$$

où y est choisi de telle manière à approcher le sup à ε près.

$$\leq \langle z|y\rangle - (\langle z|y\rangle - \psi(z)) + \varepsilon = \psi(z) + \varepsilon$$

et ce pour tout $\varepsilon > 0$. De plus,

$$\tilde{\psi}(z) \geq -\tilde{\phi}(0) > -\infty$$

en regardant en $y = 0$. On a donc montré que $\tilde{\psi}$ est minorée par une constante et majorée par une application intégrable, d'où, $\tilde{\psi}$ intégrable sur β car cette mesure est finie. De même les estimations de $\tilde{\phi}$ donne so intégrabilité car α est une mesure finie. On a donc le droit d'écrire :

$$\int \tilde{\psi} d\beta \leq \int \psi d\beta, \int \tilde{\phi} d\alpha \leq \int \phi d\alpha \tag{6}$$

Par définition de $\tilde{\psi}$, on a

$$\forall z, y, \quad \tilde{\psi}(z) + \tilde{\phi}(y) \geq \langle y|z\rangle . \tag{7}$$

On intègre cette inégalité par rapport à p :

$$0 \leq \int \left(\tilde{\psi}(z) + \tilde{\phi}(y) - \langle y|z\rangle\right) dp(y, z)$$

$$= \int \tilde{\psi}(z) d\beta(z) + \int \tilde{\phi}(y) d\alpha(y) - \int \langle y|z\rangle \, dp(y, z)$$

$$\leq \int \psi(z) d\beta(z) + \int \phi(y) d\alpha(y) - \int \langle y|z\rangle \, dp(y, z) \leq 0$$

car $\int \psi(z) d\beta(z) = 0$ et $\int \tilde{\phi}(y) d\alpha(y) \leq \int \langle y|z\rangle \, dp(y, z)$ par hypothèse du **MKP** mixte. Donc, les inégalités dans cet enchaînement ne sont finalement que des égalités, on en déduit :

$$\int \psi(z) d\beta(z) = \int \tilde{\psi}(z) d\beta(z) \tag{8}$$

$$\int \tilde{\phi}(y) d\alpha(y) = \int \phi(y) d\alpha(y). \tag{9}$$

Or, $\tilde{\psi} \leq \psi$, et $\tilde{\phi} \leq \phi$, d'où

$$\tilde{\psi} = \psi \ \alpha - p.p \tag{10}$$

$$\tilde{\phi} = \phi \ \beta - p.p \tag{11}$$

et,

$$\tilde{\psi}(z) + \tilde{\phi}(y) = \langle y|z \rangle \ \ p - p.p. \tag{12}$$

Or, $\psi \in C(\Omega)$ et $\tilde{\psi}$ est r-lipschitzienne, leur égalité presque partout devient une simple égalité sur Ω. Donc, $\nabla\psi$ est bien défini (il semble que cela vienne du théorème de Rademacher étant donné qu'on a une fonction convexe continue) et donc presque partout sur Ω, $\partial\psi(z) = \{\nabla\psi(z)\}$. Alors, de (12), on déduit par de bonnes propriétés d'analyse convexe que

$$z \in \partial\tilde{\phi}, \ y \in \partial\psi(z) \ p - p.p. \tag{13}$$

Puis,

$$p\left(\{(y,z) \in \mathbb{R}^d \times \overline{\Omega}, \partial\psi(z) \neq \{\nabla\psi(z)\}\}\right) = \int_\Omega \mathbb{1}_{\partial\psi(z)\neq\{\nabla\psi(z)\}} d\beta + \int_{\partial\Omega} \mathbb{1}_{\partial\psi(z)\neq\{\nabla\psi(z)\}} d\beta$$

car β marginale de p

$$= \int_\Omega \mathbb{1}_{\partial\psi(z)\neq\{\nabla\psi(z)\}} d\beta = 0$$

car $\beta(\partial\Omega) = 0$. Or,

$$\{(y,z) \in \mathbb{R}^d \times \overline{\Omega}, \ \partial\psi(z) = \{\nabla\psi(z)\}\} \cap \left\{(y,z) \in \mathbb{R}^d \times \overline{\Omega}, \ y \in \partial\psi(z), z \in \partial\tilde{\phi}(y)\right\}$$

$$\subset \left\{(y,z) \in \mathbb{R}^d \times \overline{\Omega}, y = \nabla\psi(z)\right\}.$$

Avec les deux premiers ensembles de probabilité 1, donc leur intersection l'est aussi. Puis, le dernier ensemble est bien un borélien car ensemble des antécédents de 0 d'une fonction mesurable. Donc, le dernier ensemble est de mesure 1, et donc

$$y = \nabla\psi(z) \ p - p.p. \tag{14}$$

Prenons maintenant f continue telle que $|f(y,z)| \leq cste(1 + \|y\|)$. f est alors p-intégrable et par ce que l'on vient de montrer,

$$\int f(y,z) dp(y,z) = \int f(\nabla\psi(z), z) dp(y,z)$$

$$\int f(y,z)dp(y,z) = \int f(\nabla\psi(z),z)d\beta(z),$$

car β est la marginale de p. La généralité sur f nous donne que

$$dp(y,z) = \delta(y - \nabla\psi(z))\beta(z)dz. \tag{15}$$

Regardons le cas particulier où $f(y,z) = \|y\|$:

$$\int \|y\|dp(y,z) = \int \|\nabla\psi(z)\|d\beta(z) < +\infty$$

Donc, $\nabla\psi$ est intégrable. De là, $\psi \in W^{1,1}(\Omega,\beta)$. De plus, $\psi = \tilde{\psi}$, ce qui donne immédiatement $\psi \in K_0$.

Etape 2 : unicité de la solution du **MKP** mixte
On prend (ψ,ϕ,p) une solution du **MKP** mixte associé à α. Prenons de plus (ψ_1,ϕ_1) l'unique solution du **MKP** primal et p_1 l'unique solution du **MKP** dual. On voit directement que (ψ_1,ϕ_1,p) et (ψ,ϕ,p_1) sont solutions du **MKP** mixte. En effet, comme p_1 maximise $\int \langle y|z\rangle\, dp_1$:

$$\begin{aligned}
\int \langle y|z\rangle\, dp &\leq \int \langle y|z\rangle\, dp_1 \\
&\leq \int \psi_1 + \phi_1 dp_1 \\
&= \int \psi_1 d\beta + \int \phi_1 d\alpha \\
&= \int \phi_1 d\alpha \\
&\leq \int \phi d\alpha
\end{aligned}$$

car ϕ_1 minimise l'intégrale par hypothèse du **MKP** primal

$$= \int \langle y|z\rangle\, dp(y,z),$$

car ϕ résout le **MKP** mixte. On en conclut qu'il ne s'agit dans cet enchaînement que d'égalités, d'où les deux triplets exhibés vérifient en effet les hypothèses du **MKP** mixte. Ainsi, par ce que nous a montré l'étape 1, on a :

$$\psi_1^* = \phi_1\ \alpha - p.p, dp_1(y,z) = \delta(y - \nabla\psi(z))\beta(z)dz = dp(y,z).$$

Ainsi que

$$y = \nabla\psi_1(z)\ p - p.p, \quad y = \nabla\psi(z)\ p - p.p.$$

On en déduit que

$$p_1 = p, \quad \nabla\psi_1 = \nabla\psi\ p - p.p.$$

Or, $\psi, \psi_1 \in K_0$, Ω est connexe, d'où $\psi_1 = \psi + cste$, et $cste = 0$ car l'intégrale sur la mesure finie β est nulle. On en déduit que (ϕ, ψ) est l'unique solution du **MKP** primal, et p celle du **MKP** dual car définie à un ensemble négligeable près.

Etape 3 : si $\alpha \ll Leb$. Tout d'abord, $\nabla\psi^*$ est bien définie à un ensemble négligeable près, car ψ^* est r-lipschitzienne, convexe. Or,

$$\left\{y \in \mathbb{R}^d,\ \partial\psi^*(y) \neq \nabla\psi^*(y)\right\}$$

est α négligeable d'où, comme α est marginale de p, alors,

$$\partial\psi^*(y) \neq \nabla\psi^*(y)\ p - p.p.$$

On en déduit immédiatement, parce que $z \in \partial\psi^*(y)\ et\ y \in \partial\psi(z)\ p - p.p$, que

$$z = \nabla\psi^*(y)\ et\ y = \nabla\psi(z)\ p - p.p. \tag{16}$$

Et comme α et β sont les marginales de p, on obtient :

$$z = \nabla\psi^*(\nabla\psi(z))\ \beta - p.p\ et\ y = \nabla\psi(\nabla\psi^*(y))\ \alpha - p.p. \tag{17}$$

On en déduit aussi que :

$$dp(y, z) = \delta(z - \nabla\psi^*(y))d\alpha(y), \tag{18}$$

d'où le résultat final.

Proposition 4.2.5 (Continuité des solutions du MKP par rapport à α)**.** *Soit $(\alpha_n)_{n\in\mathbb{N}}$ une suite de mesures de probabilité sur $\mathbb{R}^d$ telle que $\int f d\alpha_n \to \int f d\alpha$ pour n tendant vers $+\infty$, et ce, $\forall f \in C(\mathbb{R}^d)$, tel que $|f(y)| \leq cste(1 + \|y\|)$.*
Si (ψ_n, ϕ_n, p_n) est une solution du ***MKP*** *mixte associé à α_n, alors, le* ***MKP*** *mixte associé à α a une unique solution (ψ, ϕ, p), et*

1. *$\phi_n \to \phi$ pour n tendant vers l'infini, uniformément sur tout compact de $\mathbb{R}^d$*
2. *$\psi_n \to \psi$ dans $W^{1,1}(\Omega, \beta)$*
3. *$\int f dp_n \to \int f dp \quad \forall f \in C(\mathbb{R}^d \times \overline{\Omega})$ tel que $|f(y, z)| \leq cste(1 + \|y\|)$*

Preuve. Prenons (α_n) une suite de mesure de probabilité sur $\mathbb{R}^d$ tel que

$$\int f d\alpha_n \to \int f d\alpha, \ \forall f \in C(\mathbb{R}^d) \quad \text{tel que } |f(y)| \leq cste(1 + \|y\|).$$

On peut alors se munir de (ϕ_n, ψ_n, p_n) une solution du **MKP** mixte associée. On a par la propriété 4.2.4 :

(i) $\psi_n \in K_0, \phi_n = \psi_n^*, \ \alpha_n - p.p$

(ii) $dp_n(y,z) = \delta(y - \nabla\psi_n(z))\beta(z)dz$

(iii) $\int \|\nabla\psi_n\| d\beta = \int \|y\| d\alpha_n(y) \to \int \|y\| d\alpha(y)$.

On déduit alors du lemme 4.2.1 que quitte à extraire, (ψ_n) converge uniformément sur tout compact et au sens L^1 vers ψ, et (ϕ_n) convergent uniformément sur tout compact vers ϕ, où ϕ et ψ vérifient :

$$\psi \in C(\Omega) \cap L^1(\Omega, \beta), \quad \phi \in C(\mathbb{R}^d)$$
$$\phi(y) + \psi(z) \geq \langle y|z \rangle$$
$$|\phi_n(y)| \leq cste(1 + \|y\|).$$

On a, du fait qu'il s'agisse de solution du **MKP** mixte :

$$\int \phi_n d\alpha_n \leq \int \langle y|z \rangle \, dp_n(y,z), et \int \psi_n d\beta = 0.$$

Ainsi,

$$\int \psi d\beta \ et, \ \int \phi d\alpha \leq \lim_{n\to\infty} \int \phi_n d\alpha_n$$

et,

$$\int |\phi_n(y)| d\alpha_n(y) \leq cste \int 1 + \|y\| d\alpha_n(y) \leq cste.$$

Si on regarde p_n, ses marginales sont α_n et β. Alors,

$$\int 1 + \|y\| dp_n = \int 1 + \|y\| d\alpha_n \to \int 1 + \|y\| d\alpha.$$

Ainsi,

$$\forall f \in C_c^\infty, \int f(y,z) dp_n(y,z) \to \int f(y,z) dp(y,z) \tag{19}$$

et donc p est une mesure finie de marginale α, β, et l'assertion (19) reste valable pour les $f$tels que $|f(y)| \leq cste(1 + \|y\|)$ (on regarde la différence et on conclut grâce au théorème de convergence dominée). On applique en particulier à l'application produit scalaire :

$$\int \langle y|z\rangle \, dp_n(y,z) \to \int \langle y|z\rangle \, dp(y,z)$$

et donc, comme (ϕ_n, ψ_n, p_n) solution du MKP pour chaque n, en passant à la limite dans

$$\int \langle y|z\rangle \, dp_n(y,z) \geq \int \phi_n d\alpha,$$

on obtient :

$$\int \langle y|z\rangle \, dp(y,z) \geq \int \phi d\alpha.$$

Ainsi, (ϕ, ψ, p) est bien solution du **MKP** mixte. Par la proposition 4.2.4, on a (ϕ, ψ) l'unique solution du **MKP** primal, et p est l'unique solution du **MKP** dual, et,

$$dp(y,z) = \delta(y - \nabla\psi(z))\beta(z)dz. \tag{20}$$

Il nous reste à montrer que

$$\psi_n \to \psi \textit{ au sens de } W^{1,1}(\Omega, \beta) \tag{21}$$

au lieu de la convergence L^1 que l'on a déjà obtenue. La convergence uniforme donne clairement la convergence L^1 du gradient, et donc, si l'on a la convergence L^1 de la fonction et de son gradient, on a la convergence au sens de $W^{1,1}$. On a alors une valeur d'adhérence pour la convergence au sens de $W^{1,1}$, et cette valeur d'adhérence est unique par les théorèmes d'unicité susmentionnés, d'où la convergence de la suite de départ. Les convergences des autres composantes s'obtiennent facilement. On précise qu'on a en particulier :

$$\int f(\nabla\psi_n(z), z)\beta(z)dz \to \int f(\nabla\psi(z), z)\beta(z)dz,$$

et ce quelle que soit f continue telle que $|f(y,z)| \leq cste(1 + \|y\|)$.

4.2.2 Résultats préliminaires

Des propositions 4.2.4, 4.2.5 et du théorème d'existence, on montre le théorème suivant :

Théorème 4.2.6 (Solution du MKP)**.** *Le* ***MKP*** *mixte associé à* α *comme déjà décrit a une unique solution* (ϕ, ψ, p)*, où* (ϕ, ψ) *est l'unique solution du* ***MKP*** *primal, et* p *est l'unique solution du* ***MKP*** *dual. De plus, toutes les hypothèses des propositions 4.2.4 et 4.2.5 sont vérifiées*

Nous allons maintenant nous attaquer au dernier résultat technique avant la démonstration du résultat principal, il nous sera pratique dans ce dernier de pouvoir nous donner un autre point de vue ce qu'est $\nabla\psi$, et c'est l'objet de la proposition suivante :

Proposition 4.2.7 (caractérisation de $\nabla\psi$)**.** *Soit* $\psi \in K_0$ *tel que* $\int_\Omega f(\nabla\psi(z))\beta(z)dz = \int_{\mathbb{R}^d} f(y)d\alpha(y)$ $\forall f$ *continue telle que* $|f(y)| \leq cste(1 + \|y\|)$*. Alors,* (ϕ, ψ, p) *est l'unique solution du* ***MKP*** *mixte associé à* α *où*

$$dp(y, z) = \delta(y - \nabla\psi(z))\beta(z)dz.$$

Preuve. Soit $\psi \in K_0$ vérifiant l'hypothèse du théorème. On pose p la mesure tel qu'on l'attend :

$$dp(y, z) = \delta(y - \nabla\psi(z))\beta(z)dz$$

ce qui est légitime car l'appartenance à K_0 assure la définition de $\nabla\psi$. On a alors pour f continue telle que $|f(y)| \leq cste(1 + \|y\|)$:

$$\int f(y, z)dp(y, z) = \int f(y, z)\delta(y - \nabla\psi(z))\beta(z)dz \quad = \int f(\nabla\psi(z), z)\beta(z)dz$$

En particulier,

$$\int 1 + \|y\|dp(y, z) = \int (1 + \|\nabla\psi(z)\|)\beta(z)dz < +\infty.$$

Or, par hypothèse,

$$\int f(\nabla\psi(z))\beta(z)dz = \int f(y)d\alpha(y).$$

Donc, α est une marginale de p, puis

$$\int f(z)dp(y, z) = \int f(z)\delta(y - \nabla\psi(z))\beta(z)dz = \int f(z)\beta(z)dz.$$

D'où, β est une marginale de p elle aussi. Puis, comme $\psi \in K_0$, on a ψ convexe, localement lipschitz sur Ω (car $\nabla\psi$ intégrable), et donc,

$$\psi(z) + \langle\nabla\psi(z)|z' - z\rangle \leq \psi(z'), \ \forall z' \in \Omega, \ \forall z \in \Omega \setminus E$$

où E ensemble de mesure de Lebesgue négligeable. Puis, par définition de p, il existe un ensemble F p-négligeable tel que $y = \nabla\psi(z)$ sur $(\mathbb{R}^d \times \overline{\Omega}) \setminus F$. En effet, on pose

$$X = \{(y, z) \text{ tel que} y = \nabla\psi(z)\ p - p.p\}$$

X est un borélien car ensemble d'antécédents de zéro d'une application mesurable. De plus,

$$\int_X dp(y, z) = \int_{\mathbb{R}^d \times \overline{\Omega}} dp(y, z),$$

car le symbole de Kronecker s'annule justement exactement en dehors de X

$$= p(\mathbb{R}^d \times \overline{\Omega}) = 1.$$

Maintenant, on pose $A = \left(\mathbb{R}^d \times \overline{\Omega} \setminus E\right) \setminus F$, on a

$$\psi(z) + \langle y|z' - z\rangle \leq \psi(z'),\ \textit{et}\ p(A) = 1$$

En effet,

$$1 - p(A) \leq p\left(\mathbb{R}^d \times (\partial\Omega \cup E)\right) + p(F) = \beta(\partial\Omega \cup E) + p(F) = 0.$$

Donc, en posant

$$\phi = \psi^*$$

on a, par l'inégalité presque partout que l'on vient d'obtenir :

$$\psi(z) + \phi(y) \leq \langle y|z\rangle$$

sur A. On intègre alors sur A :

$$\int \phi d\alpha + \int \psi d\beta \leq \int \langle y|z\rangle\, dp(y, z)$$

avec $\int \psi d\beta = 0$. Enfin, par définition de ϕ, on a de plus

$$\phi + \psi \geq \langle y|z\rangle\,.$$

On a donc bien montré que (ϕ, ψ, p) est l'unique solution du **MKP** mixte.

Déduisons immédiatement de la proposition 4.2.7 le théorème suivant :

Théorème 4.2.8. *Pour toute mesure de probabilité α sur $\mathbb{R}^d$ tel que $\int (1 + \|y\|)\, d\alpha < +\infty$, il existe un unique $u^* = \nabla\psi^* \in K$ tel que*

$$\int f(y) d\alpha(y) = \int_\Omega f(\nabla\psi^*(z)) d\beta(z)$$

pour tout f continue telle que $|f(y)| \leq cste(1 + \|y\|)$. De plus, si

$$\int f d\alpha_n \to \int f d\alpha$$

pour ces mêmes f, alors,

$$\psi_n^* \to \psi^*$$

au sens de $W^{1,1}(\Omega, \beta)$.

Preuve. La première partie découle directement des propositions 4.2.4 et 4.2.7. Puis, si, $\int f d\alpha_n \to \int f d\alpha$ pour f continue telle que $|f(y)| \leq cste(1 + \|y\|)$, alors par la proposition 4.2.5, on a directement la convergence souhaitée.

Nous allons alors déduire de ce théorème un Corollaire intéressant en soi :

Théorème 4.2.9 (Réarrangement des fonctions à valeurs vectorielles)**.** *On pose $K = \{\nabla\psi, \psi \in W^{1,1}(\Omega, \beta), \psi \text{ convexe}\}$ $\forall u \in L^1(X, \mu, \mathbb{R}^d)$ il existe un unique $u^* \in K$ réarrangement de u. De plus, l'application $u \to u^*$ est continue.*

Preuve. Soit $u \in L^1(X, \mu, \mathbb{R}^d)$. On pose α la mesure image de μ par u. Alors, par le théorème 4.2.8 , il existe un unique $u^* \in K$ tel que

$$\int f(y) d\alpha(y) = \int_\Omega f(u^*(z)) d\beta(z)$$

pour tout f continue telle que $|f(y)| \leq cste(1 + \|y\|)$. Donc, par définition,

$$\int f(u(x)) d\mu(x) = \int_\Omega f(u^*(z)) d\beta(z).$$

D'où le résultat. Il suffit maintenant de montrer la continuité : Si l'on a $u_n \to u$, on a α_n la mesure image de μ par u_n. On a ainsi si f continue telle que $|f(y)| \leq cste(1 + \|y\|)$:

$$\int f d\alpha_n = \int f(u_n(x)) d\mu(x) \to \int f(u(x)) d\mu(x) = \int f d\alpha$$

par le théorème de convergence dominée. D'où, par le théorème 4.2.8 , $u_n^* \to u^*$.

4.2.3 La preuve du théorème du factorisation

L'heure est maintenant venue de démontrer le résultat principal, énonçons le solennellement :

Théorème 4.2.10 (Factorisation polaire des fonctions à valeurs vectorielles)**.** *Soit* $N = L^p(X, \mu, \mathbb{R}^d) \backslash \big\{ u \in L^p(X, \mu, \mathbb{R}^d) \text{ tel que } \mu(u^{-1}(E)) = 0 \ \forall E \text{ tel que } \mu(E) = 0 \big\}$, *alors,*

$$\forall u \in L^p(X, \mu, \mathbb{R}^d) \setminus N$$

, il existe un unique couple (u^*, s) *tel que* $u^* \in K$, *s est une application préservant la mesure partant de* (X, μ), *allant jusqu'à* $(\overline{\Omega}, \beta)$. *De plus,* $u = u^* \circ s$ *et*

1. *u^* est l'unique fonction donnée par le théorème 4.2.9*
2. *s est l'unique application préservant la mesure qui maximise :* $\int_X \langle u(x) | s(x) \rangle \, d\mu(x)$
3. *l'application* $u \in L^p(X, \mu, \mathbb{R}^d) \setminus N \to (u^*, s) \in L^p(\Omega, \beta, \mathbb{R}^d) \times L^q(X, \mu, \mathbb{R}^d)$ *est continue quelque soit* $q \geq 1$.

Preuve. Etape 1 : existence de la factorisation
Fixons $u \in L^1(X, \mu, \mathbb{R}^d) \setminus N$. Alors, si on pose α la mesure image de μ par u, on a $\alpha \ll \mu$ par définition de $L^1 \setminus N$, et on peut donc écrire $d\alpha = \alpha(y)dy$, où $\alpha \geq 0$ sur $\mathbb{R}^d$. On utilise la proposition 4.2.4 et l'existence d'une solution au **MKP** mixte, on peut se munir de $\psi \in K_0, \phi$ lipschitzienne convexe telle que $\phi = \psi^*$, et vérifiant :

(i) $z = \nabla\phi(\nabla\psi(z)), \quad \beta - p.p$

(ii) $y = \nabla\psi(\nabla\phi(y)), \quad \alpha - p.p$

(iii) $dp(y, z) = \delta(y - \nabla\psi(z))\beta(z)dz = \delta(z - \nabla\phi(y))\alpha(y)dy$

On pose, pour $x \in X, s(x) = \nabla\phi(u(x))$: cette application préserve la mesure et va de (X, μ) vers $(\overline{\Omega}, \beta)$, en effet on a pour $f \in C_c^\infty(\mathbb{R}^d)$,

$$\begin{aligned}
\int f(s(x))d\mu(x) &= \int f(\nabla\phi(u(x))d\mu(x) \\
&= \int f(\nabla\phi(y))\alpha(y)dy \\
&= \int f(z)\delta(z - \nabla\phi(y))\alpha(y)dy \\
&= \int f(z)dp(y, z) \\
&= \int f(z)d\beta(z).
\end{aligned}$$

Puis pas densité, cela reste vrai si $f \in L^1(\Omega, \beta)$. Il nous suffit alors de montrer que $u(x) = \nabla\psi(s(x))\ \mu - p.p.$ On pose

$$\begin{aligned} M &= \{x \in X \ tel\ que\ u(x) \neq \nabla\psi(s(x))\} \\ &= \{x \in X \ tel\ que\ u(x) \neq \nabla\psi(\nabla\phi(u(x)))\} \\ &= u^{-1}\left(\{y \in \mathbb{R}^d \ tel\ que\ y \neq \nabla\psi(\nabla\phi(y))\}\right) \end{aligned}$$

qui est un ensemble négligeable car image réciproque par un élément de $L^1 \subset N$ d'un élément négligeable.

Etape 2 : Unicité de la factorisation.
Supposons que si $u \in L^1 \subset N$, on ai $u = \nabla\psi' \circ s' = \nabla\psi \circ s$. On a, pour f continue telle que $|f(y)| \leq cste(1 + \|y\|)$,

$$\int f(y)dy = \int f(u(x))d\mu(x).$$

Par définition de α

$$= \int f(\nabla\psi' \circ s'(x))d\mu(x).$$

Par hypothèse

$$\int f(\nabla\psi'(z))\beta(z)dz$$

car s' préserve la mesure. Donc, la proposition 4.2.7 nous donne $\psi' = \psi$ car tous deux solutions du **MKP** mixte. Il ne reste plus qu'à montrer que $s' = \nabla\phi \circ u = s\ \mu - p.p.$On a :

$$\begin{aligned} \mu\left(\{x \in X \ tel\ que\ s'(x) \neq \nabla\phi \circ \nabla\psi(s'(x))\}\right) &= \int_X \mathbb{1}_{\{x \in X\ tel\ que\ s'(x) \neq \nabla\phi\circ\nabla\psi(s'(x))\}} d\mu(x) \\ &= \int_{\overline{\Omega}} \mathbb{1}_{\{z \neq \nabla\phi\circ\nabla\psi(z)\}}\beta(z)dz \\ &= \beta\left(\{z \neq \nabla\phi \circ \nabla\psi(z)\}\right) \\ &= p\left(\{(y,z) \neq \nabla\phi \circ \nabla\psi(z)\}\right) = 0 \end{aligned}$$

car ensemble négligeable comme précédemment.

Etape 3 : Continuité de la factorisation polaire :
On sait déjà que l'application $u \to \psi$ est continue par le théorème 4.2.8. Il nous

reste alors à montrer que l'application $u \to s$ est continue. Prenons pour cela une suite (u_n) de $L^1 \setminus N$ tendant vers $u \in L^1 \setminus N$. On sait déjà que $\psi_n \to \psi$ au sens de $W^{1,1}$. Montrons que $s_n \to s$.
On déduit tout d'abord de la proposition 4.2.7 le fait que (ψ_n^*, ψ_n, p_n) est l'unique solution du **MKP** mixte associée à α_n la mesure image de μ par rapport à u_n. Puis la proposition 4.2.5 nous donne les convergences suivantes :

$$\int f(\nabla\psi_n(z), z)\beta(z)dz \to \int f(\nabla\psi(z), z)\beta(z)dz = \int f(y, z)dp(y, z)$$

pour f continue à support compact, et on en déduit :

$$\int f(u_n(x), s_n(x)d\mu(x) = \int f(\nabla\psi_n(z), z)dz \to \int f(\nabla\psi(z), z)\beta(z)dz.$$

On en déduit que :

$$\int f(u(x), s_n(x)d\mu(x) \to \int f(u(x), s(x)d\mu(x).$$

En effet,

$$\begin{aligned} &|\int f(u(x), s_n(x)d\mu(x) - \int f(u(x), s(x)d\mu(x)| \\ \leq &|\int f(u(x), s_n(x)d\mu(x) - \int f(u_n(x), s_n(x)d\mu(x)| \\ + &|\int f(u_n(x), s_n(x)d\mu(x) - \int f(u(x), s(x)d\mu(x)|. \end{aligned}$$

Avec le deuxième terme qui tend vers 0 quand n tend vers l'infini, et pour le premier, on a convergence vers 0 car u_n converge au sens L^1.

On peut étendre cela par densité à toute les fonctions f tel que $f(y, z) = g(y)h(z)$, où $g \in L^1(\mathbb{R}^d, \alpha)$ et, h continue. En effet, si l'on approche g par g_ε au sens $L^1(\alpha)$ à η près :

$$\begin{aligned} |\int f_\varepsilon(u(x), s(x))d\mu(x) - \int f(u(x), s(x))d\mu(x)| &\leq \|h\|_\infty \int |g_\varepsilon(y) - g(y)|\alpha(y)dy \\ &\leq \eta\|h\|_\infty. \end{aligned}$$

De plus, par le cas déjà traité,

$$|\int f_\varepsilon(u(x), s_n(x))d\mu(x) - \int f_\varepsilon(u(x), s(x))d\mu(x)| \leq \eta$$

si n assez grand, puis,

$$|\int f_\varepsilon(u(x),s_n(x))d\mu(x)-\int f(u(x),s_n(x))d\mu(x)|\leq \|h\|_\infty\int |g_\varepsilon(y)-g(y)|\alpha(y)dy$$

$$\leq \eta\|h\|_\infty.$$

Donc, si n assez grand, en recollant les morceaux :

$$|\int f(u_n(x),s_n(x)d\mu(x)-\int f(u(x),s(x)d\mu(x)|\leq \eta(1+2\|h\|_\infty).$$

D'où,

$$\int f(u(x),s_n(x)d\mu(x)\to\int f(u(x),s(x)d\mu(x).$$

Appliquons cela à $f(y,z)=\langle\nabla\phi(y)|z\rangle$, on a :

$$\int \langle\nabla\phi(u(x))|s_n(x)\rangle\, d\mu(x)\to\int \langle\nabla\phi((u(x))|s(x)\rangle\, d\mu(x)=\int \langle s(x)|s(x)\rangle\,.$$

D'où,

$$\int \|s_n(x)\|^2 d\mu(x)=\int \|z\|^2\beta(z)dz=\int \|s(x)\|^2 d\mu(x).$$

Ainsi,

$$\int \|s_n(x)-s(x)\|^2 d\mu(x)\to 0.$$

Donc, $s_n\to s$ au sens L^2. Or, Ω borné, donc cette convergence implique la convergence L^1 car les s_n sont L^1, en effet :

$$\int_X \|s(x)\|d\mu(x)=\int_\Omega \|z\|\beta(z)dz\leq r.$$

Ainsi, comme les normes L^p sont inférieures à la norme L^1 dans le cas où la quantité tend vers 0, on a la convergente L^p pour $p\geq 1$.

Etape 4 : caractérisation des différents facteurs de la décomposition :
On sait déjà, par le théorème 4.2.9 , que $\nabla\psi$ est l'unique réarrangement de u appartenant à $\{\nabla\psi,\psi\in K_0\}$. Montrons que s est l'unique élément de S qui maximise :

$$\int \langle s(x)|u(x)\rangle\, d\mu(x).$$

Tout d'abord, s maximise effectivement cette quantité sur S : si ψ convexe, localement lipschitzienne, on a :

$$\forall s' \in S, \psi(s'(x)) \geq \psi(s(x)) + \langle \nabla\psi(s(x)) | s'(x) - s(x) \rangle \;\; \mu - p.p.$$

On intègre alors :

$$\int \psi \circ s' d\mu \geq \int \psi \circ s d\mu + \int \langle \nabla\psi \circ s | s' - s \rangle \, d\mu.$$

D'où,

$$\int \langle \nabla\psi \circ s | s \rangle \, d\mu \geq \int \langle \nabla\psi \circ s | s' \rangle \, d\mu.$$

D'où le résultat. Montrons maintenant c'est le seul qui maximise cette quantité : on prend s' un autre élément de S qui maximise aussi la dite quantité. On pose p' mesure de probabilité sur $\mathbb{R}^d \times \overline{\Omega}$ tel que

$$\int f(y,z) dp(y,z) = \int f(u(x), s'(x)) d\mu(x) \;\forall f \text{continue telle que } |f| \leq cste(1 + \|y\|).$$

Alors, p' est l'unique mesure solution du **MKP** dual associée à α. En effet :

$$\int f(y) dp'(y,z) = \int f(u(x)) d\mu(x) = \int f(y)\alpha(y)$$

$$\int f(z) dp'(y,z) = \int f(s'(x)) d\mu(x)) \int f(z)\beta(z) dz.$$

On a donc les bonnes marginales. Puis p' maximise la quantité donnée par le **MKP** dual :

$$\int \langle y | z \rangle \, dp'(y,z) = \int \langle u(x) | s'(x) \rangle \, d\mu(x)$$

par définition

$$= \int \langle u(x) | s(x) \rangle \, d\mu(x)$$

car s aussi maximise

$$\int \langle \nabla\psi(s(x)) | s(x) \rangle \, d\mu(x)$$

$$\int \langle \nabla\psi(z)|z\rangle\, \beta(z)dz$$

$$\int \langle y|z\rangle\, dp(y,z)$$

par définition de p. De plus, par sa définition,

$$\int \|y\| dp'(y,z) = \int \|u(x)\| d\mu(x) < +\infty$$

car u intégrable. Donc, p' est solution du **MKP** dual, et donc par unicité, $p' = p$
De là, on a pour f continue telle que $|f| \leq cste(1 + \|y\|)$,

$$\int f(u(x), s'(x))d\mu(x) = \int f(y,z)dp(y,z)$$

par ce qui précède

$$= \int f(u(x), s(x))d\mu(x).$$

Par définition de p. Alors, en particularisant une nouvelle fois en $f(y,z) = \langle \nabla\phi(y)|z\rangle$, on a :

$$\int \langle \nabla\phi(u(x))|s'(x)\rangle\, d\mu(x) = \int \langle \nabla\phi(u(x))|s(x)\rangle\, d\mu(x)$$

Donc,

$$\int \langle s(x)|s'(x)\rangle\, d\mu(x) = \int \|s(x)\|^2 d\mu(x) = \int \|s'(x)\|^2 d\mu(x)$$

car s, s' préservent la mesure. On conclut alors :

$$\int \|s - s'\|^2 d\mu = 0.$$

On déduit donc, $s = s'$ $\mu - p.p$, d'où le résultat.

4.2.4 Existence d'une solution au MKP mixte

Il nous faut prouver l'existence d'une solution au MKP pour justifier les démonstrations faites jusqu'à présent. La preuve de l'existence se fera en deux parties : une première où l'on ne considère que le cas où α est une mesure à support compact, et une deuxième où l'on étendra ce cas au cas général.

I - Le cas compact : On suppose ici que le support de α est inclus dans $B(0,R)$. Nous allons admettre la proposition suivante :

Proposition 4.2.11 (Principe fort de dualité)**.** *Il existe une mesure de probabilité p sur $B(0,R)\times\overline{\Omega}$ avec pour marginales α, β qui satisfont $\int \langle y|z\rangle\, dp(y,z) = I$ où,*

$$I = \inf\left\{\int \phi d\alpha + \int \psi d\beta, \phi \in C(B(0,R)), \psi \in C(\overline{\Omega})\right\}.$$

Prenons maintenant (ψ_n, ϕ_n) une suite fournie par la proposition 4.2.11 telle que

$$\phi_n(y) + \psi_n(z) \geq \langle y|z\rangle\,,\ \forall (y,z) \in B(0,R)\times\overline{\Omega} tel\ que\ \int \phi_n d\alpha + \int \psi_n d\beta \to I.$$

On peut, quitte à retrancher une constante à chaque ϕ_n et l'ajouter à ψ_n, que $\min_{\overline{\Omega}} \phi_n = 0$ car ces applications sont continues sur un compact, donc atteignent leur minimum. On note alors $\tilde{\psi_n}$ la transformée de Legendre de ϕ_n (on prend le sup sur $B(0,R)$)et $\tilde{\phi_n}$ la transformée de Legendre de $\tilde{\psi_n}$ (on prend le sup sur $\overline{\Omega}$). On sait alors que $\tilde{\psi_n}$ est R-lipschitzienne avec $\tilde{\psi_n}(0) = 0$. On en déduit alors :

$$0 = -\tilde{\psi_n}(0) \leq \tilde{\phi_n}(y) \leq r\|y\| - \inf \tilde{\psi_n} \leq r\|y\| + Rr,$$

car espace de départ contenu dans $B(0,r)$ et R-lipschitzien. D'où,

$$0 \leq \tilde{\phi_n}(y) \leq r(\|y\| + R), \forall y \in \mathbb{R}^d$$

et, $\tilde{\phi_n}$ est r-lipschitzienne. On déduit alors par le théorème d'Ascoli l'existence de ϕ, ψ tel que :

$$\tilde{\psi_n} \to \psi, \tilde{\phi_n} \to \phi$$

uniformément sur $B(0,R)$ et $\overline{\Omega}$. Alors, les fonctions limites héritent des propriétés de continuité et d'intégrabilité car la convergence est uniforme et parce que les mesures d'intégrations sont des mesures finies. De plus,un passage à la limite dans les inégalités donne :

$$\phi(y) + \psi(z) \geq \langle y|z\rangle\,, \forall y, z.$$

Enfin, comme $\tilde{\phi_n} = \phi_n^{**}$ on a, par de bonnes propriétés d'analyse convexe :

$$\tilde{\phi_n} \leq \phi_n$$

et de même,

$$\tilde{\psi_n} \leq \psi_n.$$

Ainsi,

$$\int \phi + \int \psi \leq I.$$

De là, les suites $\tilde{\psi_n}$ et $\tilde{\phi_n}$ sont bien des suites qui tendent vers l'inf I. On peut alors supposer $\int \psi = 0$ quitte faire basculer une constante de ψ vers ϕ. Ainsi, par rapport à la contrainte

$$\phi(y) + \psi(z) \geq \langle y|z \rangle\,,$$

on a bien $\int \phi$ qui est minimisé. Si l'on pose alors p comme on l'imagine, on a bien l'existence d'une solution au MKP.

II-Cas non compact
Ce paragraphe repose sur la proposition 4.2.5 : on approche α par des mesures à support compact α_n :

$$\int f d\alpha_n = C_n \int_{\|y\| \leq n} f(y) d\alpha(y), \forall f \in C_c(\mathbb{R}^d)$$

où C_n tel que l'on renormalise la mesure α_n. On a alors clairement

$$\int f d\alpha_n \to \int f d\alpha, \forall f \text{ continue telle que } \ |f| \leq cste(1 + \|y\|)$$

La proposition 4.2.5 nous donne alors l'existence d'une solution au MKP.

5 Conclusion

Le théorème de Brenier constitue à la fois une très jolie généralisation de la notion de factorisation polaire et un objet d'intérêt en soi. Il nous semble en effet que son lien avec le transport optimal mériterait d'être approfondi et que certaines hypothèses du problème pourraient être relaxées, comme le note Brenier lui-même dans son article.

Références

[1] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. 1991.

[2] Haïm Brezis. *Analyse Fonctionnelle*. Dunod, 2005.

[3] Bertrand Maury. *Cours de modélisation de M1*. 2016.

[4] Filippo Santambrogio. *Optimal Transport for Applied Mathematicians*. Birkhäuser.