

Mémoire de mathématiques

réalisé au cours de l'année scolaire 2014-2015
dans le cadre du M2 EDP&CS de l'université Paris-Sud

Espaces de Sobolev par rapport à des mesures quelconques et application au transport optimal

Écrit par :

Hugo LAVENANT

Sous la direction de :

Filippo SANTAMBROGIO



ÉCOLE NORMALE SUPÉRIEURE



UNIVERSITÉ PARIS-SUD

4 septembre 2015

Résumé

Dans ce mémoire, on présente la définition des espaces de Sobolev par rapport à des mesures quelconques proposée par Bouchitté [2, 3] et on montre qu'elle coïncide avec celle donnée par Zhikov [17] quand l'ordre de dérivation est un.

Quand de plus la dimension de l'espace ambiant est égale à un, on arrive à une description satisfaisante de tels espaces et, de manière similaire à Louet [12], on s'en sert pour montrer un résultat d'existence pour un problème de transport optimal dans lequel on rajoute une pénalisation en gradient.

Table des matières

1	Introduction et organisation du mémoire	2
1.1	Motivation	2
1.2	État de l'art et organisation du mémoire	4
2	La définition dans le cas général	6
2.1	Le contexte et le défaut d'unicité des dérivées	6
2.2	La définition des espaces de Sobolev	8
3	Le cas des dérivées premières : l'espace tangent	14
3.1	Une digression sur les multifonctions mesurables	14
3.2	L'espace tangent de Zhikov et la structure du graphe	19
3.3	Un autre point de vue : la définition de Bouchitté	24
3.4	Une description explicite en dimension un	25
4	Application au transport optimal	37
4.1	Le transport optimal avec pénalisation en gradient	37
4.2	Un résultat d'existence en dimension un	40
	Références	46
A	Relaxée d'une fonction	48

1 Introduction et organisation du mémoire

1.1 Motivation

« *Quelles sont donc les conditions qui s'imposent logiquement au choix des hypothèses sur lesquelles doit reposer la théorie physique ? Ces conditions sont de trois sortes.*

En premier lieu, une hypothèse ne sera pas une proposition contradictoire en soi, car le physicien entend ne pas énoncer des non-sens.

En second lieu, les diverses hypothèses qui doivent porter la Physique ne se contrediront pas les unes les autres [. . .].

En troisième lieu, les hypothèses seront choisies de telle manière que, de leur ensemble, la déduction mathématique puisse tirer des conséquences qui représentent, avec une approximation suffisante, l'ensemble des lois expérimentales. »

Pierre Duhem, *La théorie physique – son objet, sa structure.*

Si les deux premières conditions qu'énonce Duhem et que toute théorie physique devrait respecter *a priori*, avant toute confrontation avec l'expérience, paraissent raisonnables, leur *vérification* peut s'avérer compliquée. Par exemple, si la solution d'une équation est censée capturer le comportement d'une entité physique, ce que l'on peut imposer *a priori* est que l'équation en question possède effectivement une solution, et c'est une propriété qui n'est pas toujours évidente à vérifier. Pour ce faire, il est nécessaire d'utiliser les mathématiques, et de manière plus générale, ce sont elles qui permettent de s'assurer de la non-contradiction et de la cohérence des théories scientifiques ainsi que de la rigueur de leurs déductions.

Le présent travail s'inscrit dans cette démarche, puisqu'il explore une formulation mathématique rigoureuse des espaces de Sobolev définis par rapport à des mesures quelconques, avant de l'utiliser pour étudier des équations issues de la théorie de l'élasticité et du transport optimal. Commençons donc par une présentation succincte de ces deux théories.

Les problèmes d'élasticité Il s'agit de décrire les déformations d'un solide soumis à des contraintes. La méthode est de définir une fonction énergie E prenant pour argument les configurations possibles du solide et de postuler que l'on observe en pratique

celle qui minimise E tout en satisfaisant les contraintes (pour une présentation complète on pourra consulter [11]). Si l'on s'intéresse à la forme que peut prendre un pont, une configuration est codée par une fonction u allant d'un domaine Ω de \mathbb{R}^2 dans \mathbb{R} , la valeur de u en un point x correspondant à l'altitude du pont en x . Si le pont est uniquement soumis à son propre poids ainsi qu'aux contraintes élastiques internes, son énergie possède une expression de la forme

$$E(u) = \int_{\Omega} H(D^2u) + u, \quad (1.1)$$

où H est une fonction convexe croissant de manière quadratique. L'ensemble des configurations possibles est l'espace des fonctions dont la différentielle seconde est de carré intégrable, c'est-à-dire $W^{2,2}$.

Dans le cas où l'objet O dont on étudie la déformation est plus compliqué (le pont correspond au cas où O est un domaine du plan), on le code par une mesure μ qui est concentrée sur O et qui donne (par exemple) la répartition de masse sur O . Une déformation est vue comme une fonction u définie sur O , ou, ce qui revient au même, définie sur l'espace tout entier mais μ -presque partout. L'énergie associée à une configuration u prend toujours la forme

$$E(u) = \int (H(D^2u) + u) \, d\mu, \quad (1.2)$$

et l'espace des configurations devrait être $W^{2,2}$, mais où l'on impose que la différentielle seconde soit de carré intégrable *par rapport* à μ (pour plus de détails on pourra lire l'introduction de [3]). La définition précise de cet espace ainsi que les questions usuelles associées au problème de calcul des variations consistant à minimiser E (existence, unicité et régularité des minimiseurs) nécessitent alors un traitement rigoureux.

Le transport optimal avec pénalisation en gradient Le transport optimal naît lorsque Monge pose le problème suivant : on dispose d'un trou et d'un tas de sable à proximité dont le volume coïncide avec celui du trou ; de quelle manière faut-il déplacer le sable dans le trou si l'on cherche à minimiser l'effort fourni ? Plus généralement, on suppose données deux mesures μ et ν de même masse totale et l'on cherche à *transporter* l'une sur l'autre, c'est-à-dire à trouver une application T transportant la masse de μ sur celle de ν . Mathématiquement parlant, la masse de μ transportée par T sur un borélien B est exactement $\mu(T^{-1}(B))$, et on veut que celle-ci coïncide avec la masse de ν sur B , c'est-à-dire que l'on impose

$$\forall B \text{ borélien, } \mu(T^{-1}(B)) = \nu(B), \quad (1.3)$$

que l'on note de façon condensée $T\#\mu = \nu$. Il faut ensuite trouver un moyen de mesurer l'effort fourni, la théorie classique le fait en introduisant une fonction de coût c ayant deux arguments : $c(x, y)$ désigne le coût nécessaire pour amener une unité de masse de x en y . Le coût total du transport pour une application T est alors

$$\int c(x, T(x)) \, d\mu(x). \quad (1.4)$$

La question qui se pose est l'existence d'une application de transport *optimale*, c'est-à-dire qui minimise le coût (1.4) parmi toutes toutes les applications T vérifiant $T\#\mu = \nu$. La principale difficulté est que la contrainte $T\#\mu = \nu$ portant sur la mesure image est non-linéaire, non-locale et passe mal à la limite, ce qui fait que la méthode directe du calcul des variations ne fonctionne pas (pour une présentation générale du transport optimal et de ses applications, on pourra consulter [16] ou [14]).

Une variante, qui a été proposée par Louet et Santambrogio (c'est l'objet de [12]), est d'ajouter une *pénalisation en gradient*, c'est-à-dire de pénaliser les applications de transport dont le gradient est trop grand. Concrètement, si $p > 1$, on s'intéresse au problème de calcul des variations suivant :

$$\inf \left\{ \int (c(x, T(x)) + |DT(x)|^p) \, d\mu(x); T\#\mu = \nu \right\}. \quad (1.5)$$

Il est alors nécessaire de définir le terme $\int |DT(x)|^p \, d\mu(x)$ et plus généralement l'espace de Sobolev $W^{1,p}$ par rapport à la mesure μ dans lequel on va chercher l'application de transport optimale.

1.2 État de l'art et organisation du mémoire

On trouve dans la littérature plusieurs définitions des espaces de Sobolev par rapport à des mesures quelconques. Les plus générales sont données dans le cadre des espaces métriques (voir [9] ou [13, section 5]), même si l'on ne peut alors « dériver » qu'une seule fois les fonctions. On pourra aussi consulter [10] qui traite le cas où l'espace est euclidien et la mesure à densité par rapport à la mesure de Lebesgue. Dans ce mémoire, on utilise la définition proposée par Bouchitté dans [2] puis [3] et qui coïncide avec celle donnée par Zhikov dans [17].

Plus précisément, la deuxième section s'attache à donner une définition générale des espaces de Sobolev par rapport à des mesures quelconques qui englobe les cas du premier et du second ordre étudiés dans les articles de Bouchitté. Bien que les résultats ne

soient pas nouveaux, la formulation proposée ici ne semble pas avoir été explicitement utilisée auparavant.

La troisième section étudie plus en détail le cas où l'ordre de dérivation est un en introduisant une notion d'espace tangent à une mesure. Pour celle-ci, la définition que l'on utilise est à la base celle donnée par Zhikov mais on montre pourquoi elle coïncide avec celle de Bouchitté. On regarde aussi le cas particulier où l'espace ambiant est de dimension un, puisqu'une description explicite de l'espace tangent est alors possible et a été donnée par Louet [12]. Toujours dans ce cas, la description de l'espace de Sobolev $W^{1,p}$ donnée par Louet n'est en fait pas correcte, on explique pourquoi et on corrige son résultat.

Enfin, la quatrième section applique les résultats de la précédente au transport optimal avec pénalisation en gradient. Après une présentation des questions que l'on peut se poser, on améliore un résultat de la thèse de Louet [12] dans le cas où la dimension de l'espace ambiant est un en obtenant un critère fin, valable pour n'importe quelle fonction de coût, pour prouver l'existence d'une application de transport optimale .

Les notations mathématiques utilisées ainsi que les hypothèses faites tout au long de ce mémoire sont données dans la section 2.1.

2 La définition dans le cas général

Dans cette section, on donne une définition des espaces de Sobolev par rapport à une mesure quelconque pour un ordre de dérivation arbitraire.

2.1 Le contexte et le défaut d'unicité des dérivées

Donnons tout d'abord quelques notations et hypothèses qui seront utilisées dans toute la suite.

- La dimension de l'espace ambiant est notée d et Ω est un domaine (c'est-à-dire un ouvert connexe) *borné* de \mathbf{R}^d .
- La lettre p désigne un réel de l'intervalle $]1, +\infty[$ (on *exclut* les cas critiques $p = 1$ et $p = +\infty$), son exposant conjugué, défini par $\frac{1}{p} + \frac{1}{p'} = 1$, est noté p' .
- Enfin, μ est une mesure de Radon positive et *finie* sur Ω , la mesure de Lebesgue sur Ω est quant à elle notée \mathcal{L} .

Si $(E, \|\cdot\|)$ est un espace de Banach de dimension finie, on note $L_\mu^p(\Omega, E)$ l'ensemble des fonctions (on utilise indifféremment les termes « fonction » et « application ») $u : \Omega \rightarrow E$ mesurables qui vérifient

$$\|u\|_{p,\mu}^p := \int_{\Omega} |u|^p d\mu < +\infty, \quad (2.1)$$

et que l'on quotiente par la relation d'égalité μ -presque partout. Muni de la norme $\|\cdot\|_{p,\mu}$, $L_\mu^p(\Omega, E)$ est un espace de Banach réflexif. Dans le cas $E = \mathbf{R}$, on le note simplement $L_\mu^p(\Omega)$, et si $\mu = \mathcal{L}$, on omet l'indice μ , c'est-à-dire que l'on utilise la notation $L^p(\Omega, E)$, la norme étant notée $\|\cdot\|_p$. Enfin, on note $L_{\text{loc}}^1(\Omega)$ l'ensemble des fonctions u mesurables à valeurs réelles telles que pour tout compact K inclus dans Ω , la fonction $u\mathbb{1}_K$ appartienne à $L^1(\Omega)$, où $\mathbb{1}_K$ désigne la fonction indicatrice de K .

On note $\mathcal{D}(\Omega)$ l'ensemble des fonctions C^∞ à support compact sur Ω , et si φ est un élément de $\mathcal{D}(\Omega)$ et $a \in \mathbf{N}$, $D^a\varphi$ désigne la différentielle a -ième de φ . Lorsque $a = 1$, on utilise la notation abrégée $D\varphi$. Si $a \in \mathbf{N}$, on note $S^a\mathbf{R}^d$ l'ensemble des applications a -linéaires symétriques sur \mathbf{R}^d , il hérite naturellement de la structure euclidienne de \mathbf{R}^d : c'est l'espace auquel appartient $D^a\varphi(x)$ si $x \in \Omega$ et $\varphi \in \mathcal{D}(\Omega)$.

Si $k \in \mathbf{N}$, on a envie de dire qu'une fonction u appartient à l'espace de Sobolev d'ordre k par rapport à μ (que l'on note $W_\mu^{k,p}(\Omega)$) si et seulement si $D^a u$ appartient à $L_\mu^p(\Omega, S^a\mathbf{R}^d)$ pour tout $a \leq k$. Dans le cas où μ est la mesure de Lebesgue, $D^a u$ est définie au sens des distributions, mais ce n'est plus possible pour une mesure μ quelconque. En effet, pour

qu'une fonction u mesurable s'identifie à une distribution d'ordre 0 sur Ω *via* la formule

$$\langle u, \varphi \rangle_{\mathcal{D}(\Omega)' \times \mathcal{D}(\Omega)} := \int_{\Omega} u \varphi \, d\mathcal{L}, \quad (2.2)$$

il faut et il suffit que $u \in L^1_{\text{loc}}(\Omega)$. Mais la condition $u \in L^p_{\mu}(\Omega)$ n'assure pas nécessairement que u appartienne à $L^1_{\text{loc}}(\Omega)$. Puisqu'un élément de $L^p_{\mu}(\Omega)$ est défini μ -presque partout tandis qu'un élément de L^1_{loc} est défini \mathcal{L} -presque partout, il est clair que si μ et \mathcal{L} sont mutuellement singulières, on ne peut pas identifier un élément de $L^p_{\mu}(\Omega)$ à un élément de $L^1_{\text{loc}}(\Omega)$, *a fortiori* à une distribution.

Dans le cas où la mesure μ est absolument continue par rapport à \mathcal{L} , toute fonction mesurable et bornée peut être vue à la fois comme un élément de $L^p_{\mu}(\Omega)$ et de $L^1_{\text{loc}}(\Omega)$. On dit alors que l'injection de $L^p_{\mu}(\Omega)$ dans $L^1_{\text{loc}}(\Omega)$ est continue si, pour tout compact K , il existe une constante C telle que, pour toute fonction u mesurable et bornée, on ait $\|u \mathbb{1}_K\|_1 \leq C \|u\|_{p,\mu}$. On a alors le critère suivant :

Proposition 2.1. *Soit $f \in L^1(\Omega)$ une fonction positive, on note $\mu := f \cdot \mathcal{L}$. L'injection de $L^p_{\mu}(\Omega)$ dans $L^1_{\text{loc}}(\Omega)$ est continue si et seulement si $f^{-\frac{1}{p-1}} \in L^1_{\text{loc}}(\Omega)$.*

Démonstration. Sens réciproque. Supposons que $f^{-\frac{1}{p-1}} \in L^1_{\text{loc}}(\Omega)$, en particulier f est strictement positive \mathcal{L} -presque partout. On se donne un compact K ainsi qu'une fonction u mesurable bornée et l'on utilise l'inégalité de Hölder pour écrire

$$\|u \mathbb{1}_K\|_1 = \int_K |u| \, d\mathcal{L} = \int_K \left(|u| f^{\frac{1}{p}} \right) f^{-\frac{1}{p}} \, d\mathcal{L} \quad (2.3)$$

$$\leq \left(\int_K |u|^p f \, d\mathcal{L} \right)^{\frac{1}{p}} \left(\int_K f^{-\frac{p'}{p}} \, d\mathcal{L} \right)^{\frac{1}{p'}} \quad (2.4)$$

$$= \|u\|_{p,\mu} \left\| f^{-\frac{1}{p-1}} \mathbb{1}_K \right\|_1^{\frac{1}{p'}}, \quad (2.5)$$

ce qui montre que l'injection est continue, puisque $\left\| f^{-\frac{1}{p-1}} \mathbb{1}_K \right\|_1^{\frac{1}{p'}}$ est fini par hypothèse.

Sens direct. Il suffit de considérer le cas d'égalité dans l'inégalité de Hölder. Plus précisément, on raisonne par contraposée et on suppose que $f^{-\frac{1}{p-1}} \notin L^1_{\text{loc}}(\Omega)$: on peut trouver un compact K tel que $\left\| f^{-\frac{1}{p-1}} \mathbb{1}_K \right\|_1 = +\infty$. Pour $\varepsilon > 0$, on considère la fonction u définie par $u := f^{-\frac{1}{p-1}} \mathbb{1}_K \mathbb{1}_{\{f > \varepsilon\}}$, c'est une fonction mesurable et bornée par $\varepsilon^{-\frac{1}{p-1}}$. On vérifie alors que $|u|^p f = f^{-\frac{1}{p-1}} \mathbb{1}_K \mathbb{1}_{\{f > \varepsilon\}}$ de sorte que

$$\frac{\|\mathbb{1}_K u\|_1}{\|u\|_{p,\mu}} = \left(\int_K f^{-\frac{1}{p-1}} \mathbb{1}_{\{f > \varepsilon\}} \, d\mathcal{L} \right)^{1-\frac{1}{p}}, \quad (2.6)$$

et il suffit de faire tendre ε vers 0 pour que le membre de droite tende (par convergence monotone) vers $+\infty$, ce qui montre que l'injection n'est pas continue. \square

On voit donc que dans le cas général, on ne peut pas espérer définir les dérivées des éléments de $L^p_\mu(\Omega)$ à l'aide de la théorie des distributions. En particulier, comme on le verra par la suite, une fonction peut avoir plusieurs dérivées, et un enjeu crucial sera de décrire le défaut d'unicité de ces dernières.

2.2 La définition des espaces de Sobolev

Introduisons tout d'abord une notation supplémentaire : on rappelle que $(S^a\mathbf{R}^d, | \cdot |)$ est l'ensemble des formes a -linéaires symétriques sur \mathbf{R}^d muni de la norme euclidienne. Si $k \in \mathbf{N}$, on note $S^{\leq k}\mathbf{R}^d$ l'ensemble

$$S^{\leq k}\mathbf{R}^d := \bigoplus_{a=0}^k S^a\mathbf{R}^d \simeq \mathbf{R} \times \mathbf{R}^d \times S^2\mathbf{R}^d \times \dots \times S^k\mathbf{R}^d. \quad (2.7)$$

Si $z = (z_0, z_1, \dots, z_k)$ est un élément de $S^{\leq k}\mathbf{R}^d$, sa norme est définie par

$$|z| := (|z_0|^p + |z_1|^p + \dots + |z_k|^p)^{\frac{1}{p}}. \quad (2.8)$$

Muni de cette norme, $(S^{\leq k}\mathbf{R}^d, | \cdot |)$ est un espace de Banach de dimension finie. Un élément de l'espace $L^p_\mu(\Omega, S^{\leq k}\mathbf{R}^d)$ est alors vu comme une collection de $k + 1$ fonctions (u_0, u_1, \dots, u_k) telle que $u_a \in L^p_\mu(\Omega, S^a\mathbf{R}^d)$ pour tout $a \in \{0, 1, \dots, k\}$. On note $\pi_a : L^p_\mu(\Omega, S^{\leq k}\mathbf{R}^d) \rightarrow L^p_\mu(\Omega, S^a\mathbf{R}^d)$ la projection sur la a -ième composante, c'est-à-dire que $\pi_a(u_0, u_1, \dots, u_k) = u_a$. Avec ces notations, si $\varphi \in \mathcal{D}(\Omega)$, la fonction $(\varphi, D\varphi, D^2\varphi, \dots, D^k\varphi)$ est naturellement vue comme un élément de $L^p_\mu(\Omega, S^{\leq k}\mathbf{R}^d)$.

Cela nous conduit à la définition suivante pour le « graphe » G :

Définition 2.2. On note G le sous-ensemble de $L^p_\mu(\Omega, S^{\leq k}\mathbf{R}^d)$ défini par

$$G := \{(\varphi, D\varphi, D^2\varphi, \dots, D^k\varphi); \varphi \in \mathcal{D}(\Omega)\}, \quad (2.9)$$

et \bar{G} son adhérence.

Remarquons que puisque G est un sous-espace vectoriel, \bar{G} est l'adhérence de G à la fois par rapport à la topologie forte et la topologie faible sur $L^p_\mu(\Omega, S^{\leq k}\mathbf{R}^d)$. L'objet \bar{G} encode toute la structure de l'espace de Sobolev que l'on cherche à définir :

Définition 2.3. On note $W_\mu^{k,p}(\Omega)$ le sous-ensemble de $L_\mu^p(\Omega)$ défini par

$$W_\mu^{k,p}(\Omega) := \pi_0(\bar{G}), \quad (2.10)$$

et la norme d'un élément u de $W_\mu^{k,p}(\Omega)$ est définie par

$$\|u\|_{k,p,\mu} := \inf \{ \|v\|_{p,\mu}; v \in \bar{G} \text{ et } \pi_0(v) = u \}. \quad (2.11)$$

Remarque. Notons $\pi_0^{\bar{G}}$ la restriction de π_0 à l'espace de Banach \bar{G} . Il est clair que l'espace $W_\mu^{k,p}(\Omega) = \pi_0(\bar{G})$ est isomorphe à $\bar{G}/\ker(\pi_0^{\bar{G}})$. Un lecteur attentif remarquera que cet isomorphisme vaut aussi pour la norme, c'est-à-dire que la norme $\|\cdot\|_{k,p,\mu}$ correspond à la norme quotient sur $\bar{G}/\ker(\pi_0^{\bar{G}})$.

Une manière de lire cette définition est de dire qu'un élément $u \in L_\mu^p(\Omega)$ est dans $W_\mu^{k,p}(\Omega)$ si et seulement si il existe une suite $(\varphi_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{D}(\Omega)$ telle que $\varphi_n \rightarrow u$ dans $L_\mu^p(\Omega)$ lorsque $n \rightarrow +\infty$, et $(D\varphi_n)_{n \in \mathbb{N}}, (D^2\varphi_n)_{n \in \mathbb{N}}, \dots, (D^k\varphi_n)_{n \in \mathbb{N}}$ soient des suites convergentes dans respectivement $L_\mu^p(\Omega, \mathbf{R}^d), L_\mu^p(\Omega, S^2\mathbf{R}^d), \dots, L_\mu^p(\Omega, S^k\mathbf{R}^d)$. Dans ce cas, si on note v_1, v_2, \dots, v_k les limites respectives des suites $(D\varphi_n)_{n \in \mathbb{N}}, (D^2\varphi_n)_{n \in \mathbb{N}}, \dots, (D^k\varphi_n)_{n \in \mathbb{N}}$, alors $(u, v_1, v_2, \dots, v_k) \in \bar{G}$. Ainsi, on retrouve bien la définition proposée par Zhikov [17] et reprise par Louet [12]. Remarquons aussi que l'on peut supposer que les suites $(D\varphi_n)_{n \in \mathbb{N}}, (D^2\varphi_n)_{n \in \mathbb{N}}, \dots, (D^k\varphi_n)_{n \in \mathbb{N}}$ sont seulement bornées, puisque cela suffit à garantir, à extraction près, qu'elles convergent faiblement, et \bar{G} est aussi l'adhérence de G pour la topologie faible.

Pour expliquer à quoi correspond la norme $\|\cdot\|_{k,p,\mu}$, introduisons la fonctionnelle H définie sur $L_\mu^p(\Omega)$ par

$$H(\varphi) := \begin{cases} \int_\Omega (|\varphi|^p + |D\varphi|^p + \dots + |D^k\varphi|^p) \, d\mu & \text{si } \varphi \in \mathcal{D}(\Omega) \\ +\infty & \text{sinon} \end{cases}, \quad (2.12)$$

et notons \bar{H} la relaxée de H (pour la définition de la relaxée on pourra consulter l'annexe A).

Proposition 2.4. Soit $u \in L_\mu^p(\Omega)$. Alors $\bar{H}(u) < +\infty$ si et seulement si $u \in W_\mu^{k,p}(\Omega)$, et dans ce cas $\bar{H}(u) = \|u\|_{k,p,\mu}^p$.

Démonstration. Pour cela, on utilise la représentation suivante de \bar{H} (cf. proposition

A.3) :

$$\bar{H}(u) = \inf \left\{ \liminf_{n \rightarrow +\infty} H(\varphi_n); \forall n \in \mathbf{N}, \varphi_n \in \mathcal{D}(\Omega) \text{ et } \lim_{n \rightarrow +\infty} \varphi_n = u \text{ dans } L_\mu^p(\Omega) \right\}. \quad (2.13)$$

Ainsi, si $\bar{H}(u)$ est fini, il existe une suite $(\varphi_n)_{n \in \mathbf{N}}$ d'éléments de $\mathcal{D}(\Omega)$ telle que $\varphi_n \rightarrow u$ et $H(\varphi_n) \rightarrow \bar{H}(u)$ lorsque $n \rightarrow +\infty$. En particulier, $(H(\varphi_n))_{n \in \mathbf{N}}$ est bornée. D'après l'expression de H , cela signifie que les suites $(D\varphi_n)_{n \in \mathbf{N}}, (D^2\varphi_n)_{n \in \mathbf{N}}, \dots, (D^k\varphi_n)_{n \in \mathbf{N}}$ sont bornées, et donc que $u \in W_\mu^{k,p}(\Omega)$ comme on l'a fait remarquer ci-dessus. À extraction près, ces suites convergent faiblement vers des limites que l'on note respectivement v_1, v_2, \dots, v_k et $(u, v_1, v_2, \dots, v_k) \in \bar{G}$. En particulier, comme la norme sur $L_\mu^p(\Omega, S^{\leq k}\mathbf{R}^d)$ est semi-continue inférieurement pour la convergence faible,

$$\|u\|_{k,p,\mu}^p \leq \|(u, v_1, v_2, \dots, v_k)\|_{p,\mu}^p \quad (2.14)$$

$$\leq \liminf_{n \rightarrow +\infty} \|(\varphi_n, D\varphi_n, D^2\varphi_n, \dots, D^k\varphi_n)\|_{p,\mu}^p \quad (2.15)$$

$$= \liminf_{n \rightarrow +\infty} H(\varphi_n) \quad (2.16)$$

$$= \bar{H}(u). \quad (2.17)$$

Pour obtenir l'inégalité dans l'autre sens, on se donne $u \in W_\mu^{k,p}(\Omega)$. Par densité de G dans \bar{G} , on peut trouver une suite $(v_n)_{n \in \mathbf{N}}$ d'éléments de G telle que $\|v_n\|_{p,\mu} \rightarrow \|u\|_{k,p,\mu}$ et $\pi_0(v_n) \rightarrow u$ lorsque $n \rightarrow +\infty$. Pour tout $n \in \mathbf{N}$, il existe une fonction $\varphi_n \in \mathcal{D}(\Omega)$ telle que $v_n = (\varphi_n, D\varphi_n, \dots, D^k\varphi_n)$. En particulier, $H(\varphi_n) = \|v_n\|_{p,\mu}^p \rightarrow \|u\|_{k,p,\mu}^p$ et $\varphi_n \rightarrow u$ lorsque $n \rightarrow +\infty$, ce qui permet de voir que $\bar{H}(u) \leq \|u\|_{k,p,\mu}^p$. □

Remarquons que dans le cas où μ est la mesure de Lebesgue, pour tout $u \in W_\mu^{k,p}(\Omega)$, il n'y a qu'un seul élément v dans \bar{G} vérifiant $\pi_0(v) = u$ et c'est $v = (u, Du, D^2u, \dots, D^k u)$, où les dérivées sont prises au sens des distributions. Dans le cas général il peut y avoir plusieurs éléments dans l'ensemble $\pi_0^{-1}(\{u\}) \cap \bar{G}$: c'est ici que se manifeste le fait que les dérivées ne sont pas définies de manière unique.

Exemple. Soit $x_0 \in \Omega$, on considère $\mu := \delta_{x_0}$ une masse de Dirac concentrée en x_0 . On voit que dans ce cas

$$G = \bar{G} = L_\mu^p(\Omega, S^{\leq k}\mathbf{R}^d). \quad (2.18)$$

En effet, si $z = (z_0, z_1, \dots, z_k) \in S^{\leq k}\mathbf{R}^d$, il est facile de construire une fonction $\varphi \in \mathcal{D}(\Omega)$ qui vaut z_0 en x_0 et dont les k premières dérivées en x_0 sont prescrites et valent respectivement z_1, z_2, \dots, z_k .

En conséquence, pour tout $u \in W_\mu^{k,p}(\Omega)$, $\|u\|_{k,p,\mu} = |u(x_0)|$, de sorte que l'espace $W_\mu^{k,p}(\Omega)$ coïncide avec $L_\mu^p(\Omega)$. On voit sur cet exemple qu'une fonction $u \in L_\mu^p(\Omega)$ admet de nombreuses dérivées : l'ensemble $\pi_0^{-1}(\{u\}) \cap \bar{G}$ n'est pas réduit à un singleton.

Montrons maintenant le théorème suivant, qui garantit que $W_\mu^{k,p}(\Omega)$ est un bon espace fonctionnel.

Théorème 2.5. *L'espace $(W_\mu^{k,p}(\Omega), \|\cdot\|_{k,p,\mu})$ est un espace de Banach réflexif et un espace de Hilbert dans le cas $p = 2$.*

Démonstration. Le premier constat est que \bar{G} est un sous-espace fermé de $L_\mu^p(\Omega, S^{\leq k} \mathbf{R}^d)$, c'est donc un espace de Banach réflexif et un espace de Hilbert dans le cas $p = 2$: en effet, un sous-espace fermé d'un espace de Banach réflexif est un espace de Banach réflexif, on pourra se référer à [1, page 50 corollaire 3]. Et comme on l'a fait remarquer, en notant $\pi_0^{\bar{G}}$ la restriction de π_0 à \bar{G} , $(W_\mu^{k,p}(\Omega), \|\cdot\|_{k,p,\mu})$ s'identifie isométriquement à l'espace $\bar{G}/\ker(\pi_0^{\bar{G}})$ muni de la norme quotient.

Or le quotient par rapport à un sous-espace fermé d'un espace de Banach réflexif est un espace de Banach réflexif (on pourra se référer par exemple à [1, page 47 exercice 5 et page 62 exercice 1]), et le quotient par rapport à un sous-espace fermé d'un espace de Hilbert est un espace de Hilbert [1, page 71 proposition 6]. \square

Terminons par un résultat qui explique en quoi une connaissance précise de la structure de \bar{G} est nécessaire pour des résultats de relaxation, et donc que la donnée de $W_\mu^{k,p}(\Omega)$ comme espace de Banach ne suffit pas. Le théorème présenté ci-dessous s'inspire de théorèmes comme [2, Théorème 3.1] pour le cas $k = 1$ et [3, Théorème 4.3] dans le cas $k = 2$. Dans ces deux cas, le résultat est plus fort car la structure de \bar{G} pour les cas $k = 1$ et $k = 2$ est exploitée ; et les hypothèses moins restrictives que celles que l'on va faire ci-dessous, le but du théorème étant de présenter la philosophie générale, pas de donner les hypothèses optimales.

On se donne $h : \Omega \times S^{\leq k} \mathbf{R}^d \rightarrow \mathbf{R}_+$ une fonction continue positive telle que :

- pour μ -presque tout $x \in \Omega$, la fonction $z \in S^{\leq k} \mathbf{R}^d \mapsto h(x, z)$ soit convexe,
- il existe deux constantes $C_2 \geq C_1 > 0$ telles que pour μ -presque tout $x \in \Omega$ et pour tout $z = (z_0, z_1, \dots, z_k) \in S^{\leq k} \mathbf{R}^d$, on ait

$$C_1(|z_1|^p + |z_2|^p + \dots + |z_k|^p) \leq h(x, z) \leq C_2(1 + |z_0|^p + |z_1|^p + |z_2|^p + \dots + |z_k|^p). \quad (2.19)$$

On note ensuite H la fonctionnelle définie sur $L_\mu^p(\Omega)$ par

$$H(\varphi) := \begin{cases} \int_{\Omega} h(x, \varphi(x), D\varphi(x), \dots, D^k\varphi(x)) \, d\mu(x) & \text{si } \varphi \in \mathcal{D}(\Omega) \\ +\infty & \text{sinon} \end{cases}. \quad (2.20)$$

On note \bar{H} la relaxée de H , et l'on remarquera que si l'on choisit correctement la fonction h , la fonctionnelle \bar{H} correspond à la version « rigoureuse » de l'énergie (1.2) discutée dans l'introduction. On a alors la formule suivante :

Théorème 2.6. *La relaxée \bar{H} de H n'est finie que sur $W_\mu^{k,p}(\Omega)$, et si $u \in W_\mu^{k,p}(\Omega)$ alors*

$$\bar{H}(u) = \inf \left\{ \int_{\Omega} h(x, v(x)) \, d\mu(x); v \in \bar{G} \text{ et } \pi_0(v) = u \right\}. \quad (2.21)$$

On remarque la grande similitude avec la proposition 2.4 (elle correspond au cas particulier $h(x, z) = |z|^p$), la preuve du théorème est d'ailleurs proche de celle de cette proposition. On pourrait montrer que l'infimum dans la formule (2.21) donnant $\bar{H}(u)$ est un minimum, mais le $v \in \bar{G}$ optimal dépend de la fonction h : c'est pour cette raison que l'on a besoin de la connaissance de tout le graphe \bar{G} lorsque l'on veut calculer la relaxée d'une fonctionnelle comme H .

Démonstration. Notons $I(u)$ le membre de droite de l'équation (2.21), de sorte que $I : W_\mu^{k,p}(\Omega) \rightarrow \mathbf{R}$.

Soit $u \in L_\mu^p(\Omega)$ tel que $\bar{H}(u) < +\infty$. On peut trouver une suite $(\varphi_n)_{n \in \mathbf{N}}$ d'éléments de $\mathcal{D}(\Omega)$ telle que $\varphi_n \rightarrow u$ et $H(\varphi_n) \rightarrow \bar{H}(u)$ lorsque $n \rightarrow +\infty$. Grâce à l'hypothèse (2.19), on voit que $(D\varphi_n)_{n \in \mathbf{N}}, (D^2\varphi_n)_{n \in \mathbf{N}}, \dots, (D^k\varphi_n)_{n \in \mathbf{N}}$ sont bornées. Quitte à extraire une sous-suite, cela signifie qu'elles convergent faiblement vers respectivement v_1, v_2, \dots, v_k et que $(u, v_1, v_2, \dots, v_k) \in \bar{G}$, en particulier $u \in W_\mu^{k,p}(\Omega)$. On a alors besoin du lemme technique suivant dont on pourra trouver la démonstration dans [8, Théorème 4.5].

Lemme 2.7. *L'application $\hat{H} : L_\mu^p(\Omega, S^{\leq k}\mathbf{R}^d) \rightarrow \mathbf{R}$, définie par*

$$\hat{H}(w) := \int_{\Omega} h(x, w(x)) \, d\mu(x) \quad (2.22)$$

pour tout $w \in L_\mu^p(\Omega, S^{\leq k}\mathbf{R}^d)$, est semi-continue inférieurement pour la convergence faible sur $L_\mu^p(\Omega, S^{\leq k}\mathbf{R}^d)$.

À l'aide de ce lemme, peut écrire que

$$\bar{H}(u) = \lim_{n \rightarrow +\infty} H(\varphi_n) \quad (2.23)$$

$$= \lim_{n \rightarrow +\infty} \hat{H}(\varphi_n, D\varphi_n, \dots, D^k \varphi_n) \quad (2.24)$$

$$\geq \hat{H}(u, v_1, \dots, v_k) \quad (2.25)$$

$$\geq I(u). \quad (2.26)$$

Pour prouver l'inégalité dans l'autre sens, on se donne $u \in W_\mu^{k,p}(\Omega)$ et une suite $(v_n)_{n \in \mathbf{N}}$ d'éléments de \bar{G} telle que $\pi_0(v_n) = u$ et $\hat{H}(v_n) \rightarrow I(u)$ lorsque $n \rightarrow +\infty$. On remarque que la fonctionnelle \hat{H} est semi-continue supérieurement pour la convergence forte sur $L_\mu^p(\Omega, S^{\leq k} \mathbf{R}^d)$: il suffit d'utiliser la majoration (2.19) que satisfait h par hypothèse et le lemme de Fatou. En conséquence, par densité de G dans \bar{G} , pour tout $n \in \mathbf{N}^*$ on peut trouver $\varphi_n \in \mathcal{D}(\Omega)$ tel que $\|\varphi_n - \pi_0(v_n)\|_{p,\mu} \leq \frac{1}{n}$ et $H(\varphi_n) = \hat{H}(\varphi_n, D\varphi_n, \dots, D^k \varphi_n) \leq \hat{H}(v_n) + \frac{1}{n}$. On voit alors que $\varphi_n \rightarrow u$ lorsque $n \rightarrow +\infty$ et que $\liminf_n H(\varphi_n) \leq \liminf_n \hat{H}(v_n) = I(u)$ de sorte que $\bar{H}(u) \leq I(u)$.

□

Si jamais il existe des injections compactes, c'est-à-dire si un contrôle des dérivées d'ordre k permet de déduire une convergence forte des dérivées d'ordre inférieur, alors on peut affaiblir les hypothèses sur h , notamment imposer une minoration plus faible, du type $h(x, z) \geq C_1 |z_k|^p$.

3 Le cas des dérivées premières : l'espace tangent

Dans cette section $k = 1$, c'est-à-dire que l'on s'intéresse à l'espace $W_\mu^{1,p}(\Omega)$. Comme on va le voir, la description du graphe \bar{G} est plutôt explicite : elle fait intervenir une notion d'espace tangent à la mesure μ qui est un espace vectoriel qui dépend du point de Ω considéré et qui donne la « direction » de μ en ce point. On aura besoin dans un premier temps de présenter quelques outils permettant de parler de la mesurabilité de cet espace tangent, puis on pourra décrire le graphe \bar{G} . On montrera ensuite pourquoi la définition de notre espace tangent, tirée de [17], coïncide avec celle donnée dans [2]. Enfin, à l'aide de [12], on donnera une description complète de $W_\mu^{1,p}(I)$ dans le cas où I est un intervalle de \mathbb{R} (le cas $d = 1$).

3.1 Une digression sur les multifonctions mesurables

Comme expliqué ci-dessus, on va avoir besoin de parler de la mesurabilité des applications à valeurs dans les sous-espaces vectoriels de \mathbb{R}^d . Le livre de Castaing et Valadier [5] développe une théorie des applications mesurables à valeurs dans les parties fermées d'un espace métrique complet séparable. Ce qui est présenté dans la suite s'en inspire fortement mais est restreint à ce dont on aura besoin dans le reste de ce mémoire. On démontrera notamment le théorème 3.3 qui n'est pas d'une grande profondeur conceptuelle, mais qui est nécessaire dans ce qui suit et qui n'a pas été prouvé (voire explicitement énoncé) dans [17] ou [2].

Dans la suite, l'ensemble des sous-espaces vectoriels de \mathbb{R}^d est noté $\text{sev}(\mathbb{R}^d)$, et Ω est naturellement muni de sa tribu borélienne.

Définition 3.1. *Une application $T : \Omega \rightarrow \text{sev}(\mathbb{R}^d)$ est dite mesurable s'il existe une famille dénombrable $(u_n)_{n \in \mathbb{N}}$ de fonctions $\Omega \rightarrow \mathbb{R}^d$ mesurables telle que pour tout $x \in \Omega$, $T(x)$ soit égal au sous-espace vectoriel engendré par l'ensemble $\{u_n(x); n \in \mathbb{N}\}$.*

Remarque. Le cadre donné dans [5] est le suivant : si (E, \mathcal{E}) est un espace mesurable, (X, d) un espace métrique complet séparable et Γ une application de E à valeurs dans les parties fermées non vides de X , les trois assertions suivantes sont équivalentes :

- (i) Il existe une famille dénombrable $(u_n)_{n \in \mathbb{N}}$ de fonctions $E \rightarrow X$ mesurables telle que pour tout $e \in E$, $\Gamma(e)$ soit égal à l'adhérence de l'ensemble $\{u_n(e); n \in \mathbb{N}\}$.
- (ii) Pour tout ouvert O de X , l'ensemble $\{e \in E \text{ tel que } O \cap \Gamma(e) \neq \emptyset\}$ est une partie mesurable de E .
- (iii) Pour tout $x \in X$, l'application $e \mapsto d(x, \Gamma(e))$ est mesurable.

Si elles sont satisfaites, on dit que Γ est une application mesurable. Si l'on se restreint aux applications à valeurs dans $\text{sev}(\mathbf{R}^d)$, on pourrait montrer que cette définition coïncide avec la définition 3.1.

Le lemme suivant permet d'obtenir une paramétrisation maniable d'une application mesurable à valeurs dans $\text{sev}(\mathbf{R}^d)$.

Lemme 3.2. *Soit $T : \Omega \rightarrow \text{sev}(\mathbf{R}^d)$ une application mesurable. Notons $d_T : \Omega \rightarrow \{0, 1, \dots, d\}$ l'application qui à un point $x \in \Omega$ associe la dimension de $T(x)$. Alors d_T est mesurable. De plus, il existe des fonctions $v_1, v_2, \dots, v_d : \Omega \rightarrow \mathbf{R}^d$ mesurables telles que pour tout $x \in \Omega$, $(v_1(x), v_2(x), \dots, v_{d_T(x)}(x))$ soit une base orthonormée de $T(x)$ et $(v_{d_T(x)+1}(x), \dots, v_d(x)) = (0, \dots, 0)$.*

Le d -uplet (v_1, v_2, \dots, v_d) est dit adapté à T .

Démonstration. Soit $(u_n)_{n \in \mathbf{N}}$ une famille dénombrable de fonctions mesurables telle que pour tout $x \in \Omega$, $T(x)$ soit égal au sous-espace vectoriel engendré par l'ensemble $\{u_n(x); n \in \mathbf{N}\}$. Si on se fixe $x \in \Omega$, il n'est pas dur d'extraire de $\{u_n(x); n \in \mathbf{N}\}$ une base de $T(x)$, puis d'appliquer le processus d'orthonormalisation de Gram-Schmidt pour la rendre orthonormée. La difficulté est de réaliser cette opération pour des x différents de façon à ce que la base obtenue dépende de façon mesurable de x .

Pour $a \in \{1, 2, \dots, d\}$, notons $(I_m^a)_{m \in \mathbf{N}}$ une énumération de \mathbf{N}^a (c'est-à-dire des a -uplets d'entiers) : si $m \in \mathbf{N}$, les éléments de I_m^a sont notés $I_m^a(1), I_m^a(2), \dots, I_m^a(a)$ et $u_{I_m^a}$ désigne le a -uplet de fonctions $(u_{I_m^a(1)}, u_{I_m^a(2)}, \dots, u_{I_m^a(a)})$.

Mesurabilité de d_T . Remarquons tout d'abord que si $a \in \{1, 2, \dots, d\}$ et $m \in \mathbf{N}$, l'ensemble $\{u_{I_m^a} \text{ est liée}\}$ est un borélien de Ω . En effet, si $M(x)$ désigne la matrice de $(u_{I_m^a(1)}(x), u_{I_m^a(2)}(x), \dots, u_{I_m^a(a)}(x))$ dans la base canonique de \mathbf{R}^d , dire que la famille $u_{I_m^a}$ est liée revient à dire que le déterminant de toute sous-matrice de taille $a \times a$ de M est nul.

Or il est clair que

$$\{d_T < a\} = \bigcap_{m \in \mathbf{N}} \{u_{I_m^a} \text{ est liée}\}. \quad (3.1)$$

Une intersection dénombrable de boréliens étant un borélien, cela suffit pour conclure que $\{d_T < a\}$ en est un et donc que d_T est mesurable.

Construction d'une base orthonormée. Fixons $a \in \{0, 1, \dots, d\}$. On va construire des fonctions mesurables w_1, w_2, \dots, w_a , définies sur $\{d_T = a\}$ et à valeurs dans \mathbf{R}^d , telles que $(w_1(x), w_2(x), \dots, w_a(x))$ soit une base de $T(x)$ pour tout $x \in \{d_T = a\}$. Pour $x \in \{d_T = a\}$ fixé, on pose $(w_1(x), w_2(x), \dots, w_a(x)) := u_{I_m^a}(x)$ où m est le plus petit entier tel que $u_{I_m^a}(x)$

soit libre (il existe au moins un m tel que $u_{I_m^a}(x)$ soit libre car $d_T(x) = a$). Construites ainsi, les fonctions w_1, w_2, \dots, w_a sont mesurables car on peut écrire

$$(w_1, w_2, \dots, w_a) := \sum_{m=0}^{+\infty} u_{I_m^a} \mathbb{1}_{\left\{u_{I_{m'}^a} \text{ liée pour } m' < m \text{ et } u_{I_m^a} \text{ libre}\right\}}. \quad (3.2)$$

Remarquons que la somme est bien définie, car à x fixé il n'y a qu'un seul terme non nul dans celle-ci. On applique ensuite le procédé d'orthonormalisation de Gram-Schmidt à la famille (w_1, w_2, \dots, w_a) , c'est-à-dire que l'on pose, pour $b \in \{1, 2, \dots, a\}$,

$$\begin{cases} \hat{v}_b := w_b - \sum_{c=0}^{b-1} w_b \cdot v_c \\ v_b := \frac{\hat{v}_b}{|\hat{v}_b|} \end{cases} \quad (3.3)$$

où \cdot et $|\cdot|$ désignent respectivement le produit scalaire et la norme euclidienne sur \mathbf{R}^d . Les fonctions v_1, v_2, \dots, v_a sont alors clairement mesurables. Toujours sur $\{d_T = a\}$, on définit aussi $(v_{a+1}, v_{a+2}, \dots, v_d) := (0, 0, \dots, 0)$.

Les ensembles $\{d_T = a\}$ forment une partition mesurable de Ω lorsque a parcourt $\{0, 1, \dots, d\}$, on peut donc effectuer cette construction sur chacun des éléments de la partition pour obtenir des fonctions v_1, v_2, \dots, v_d mesurables, définies sur Ω et à valeurs dans \mathbf{R}^d , telles que (v_1, v_2, \dots, v_d) soit adapté à T . □

Si $F \in \text{sev}(\mathbf{R}^d)$, on note $P_F : \mathbf{R}^d \rightarrow \mathbf{R}^d$ la projection orthogonale sur F . Si $u : \Omega \rightarrow \mathbf{R}^d$ et $T : \Omega \rightarrow \text{sev}(\mathbf{R}^d)$, la fonction $P_T u : \Omega \rightarrow \mathbf{R}^d$ est naturellement définie par $(P_T u)(x) := P_{T(x)}(u(x))$ pour tout $x \in \Omega$.

Théorème 3.3. *Soient $u : \Omega \rightarrow \mathbf{R}^d$ et $T : \Omega \rightarrow \text{sev}(\mathbf{R}^d)$ des applications mesurables. Alors $P_T u$ est mesurable.*

Démonstration. Soit (v_1, v_2, \dots, v_d) un d -uplet adapté à T . Alors

$$P_T u = \sum_{a=1}^d (u \cdot v_a) v_a, \quad (3.4)$$

ce qui est suffisant pour conclure à la mesurabilité de $P_T u$. □

À partir de ce théorème découlent d'autres résultats techniques.

Proposition 3.4. Soit $T : \Omega \rightarrow \text{sev}(\mathbf{R}^d)$ une application mesurable. Alors l'application $T^\perp : \Omega \rightarrow \text{sev}(\mathbf{R}^d)$, définie par $T^\perp(x) := (T(x))^\perp$ pour tout $x \in \Omega$, est mesurable.

Démonstration. Notons (e_1, e_2, \dots, e_d) une base de \mathbf{R}^d . Pour $a \in \{1, 2, \dots, d\}$, on note $u_a : \Omega \rightarrow \mathbf{R}^d$ la fonction définie par $u_a(x) := P_{T^\perp(x)}e_a = e_a - P_{T(x)}e_a$ pour tout $x \in \Omega$, elle est mesurable grâce au théorème 3.3. Il est alors facile de vérifier que pour tout $x \in \Omega$, $T^\perp(x)$ coïncide avec le sous-espace vectoriel engendré par $\{u_a(x); a = 1, 2, \dots, d\}$. \square

On en vient à définir les espaces fonctionnels suivants :

Définition 3.5. Soit $T : \Omega \rightarrow \text{sev}(\mathbf{R}^d)$ une application mesurable. L'espace $L_\mu^p(\Omega, T)$ est défini comme l'ensemble des fonctions u appartenant à $L_\mu^p(\Omega, \mathbf{R}^d)$ et telles que $u(x) \in T(x)$ pour μ -presque tout $x \in \Omega$.

Au sens de la dualité $L_\mu^p - L_\mu^{p'}$, le résultat suivant découle du théorème précédent :

Proposition 3.6. Soit $T : \Omega \rightarrow \text{sev}(\mathbf{R}^d)$ une application mesurable. On a l'identité $(L_\mu^p(\Omega, T))^\perp = L_\mu^{p'}(\Omega, T^\perp)$.

Démonstration. L'inclusion $L_\mu^{p'}(\Omega, T^\perp) \subset (L_\mu^p(\Omega, T))^\perp$ est trivialement vraie puisque si $w \in L_\mu^{p'}(\Omega, T^\perp)$ et $v \in L_\mu^p(\Omega, T)$, alors $v \cdot w = 0$ μ -presque partout, donc $\int_\Omega v \cdot w \, d\mu = 0$.

Pour obtenir l'autre inclusion, on se donne $w \in (L_\mu^p(\Omega, T))^\perp \subset L_\mu^{p'}(\Omega, \mathbf{R}^d)$. La fonction $P_T w$ est mesurable d'après le théorème 3.3, on pose donc

$$v := \frac{P_T w}{|P_T w|} \mathbb{1}_{\{P_T w \neq 0\}}, \quad (3.5)$$

de sorte que v est une fonction mesurable et bornée, en particulier $v \in L_\mu^p(\Omega, T)$. Or il se trouve que

$$\int_\Omega v \cdot w \, d\mu = \int_\Omega |P_T w| \, d\mu, \quad (3.6)$$

ce qui implique que $P_T w = 0$ μ -presque partout, donc que $w \in T^\perp$ μ -presque partout, c'est-à-dire $w \in L_\mu^{p'}(\Omega, T^\perp)$. \square

Terminons avec la définition de la μ -union essentielle d'une famille de fonctions mesurables, elle sera l'outil permettant de construire des applications mesurables à valeurs dans les sous-espaces vectoriels de \mathbf{R}^d .

Théorème 3.7. Soit U un sous-espace vectoriel de $L_\mu^p(\Omega, \mathbf{R}^d)$. Il existe une application mesurable $T : \Omega \rightarrow \text{sev}(\mathbf{R}^d)$, unique à un ensemble de μ -mesure nulle près, telle que pour

tout $u \in U$, pour μ -presque tout $x \in \Omega$, $u(x) \in T(x)$, et qui soit minimale au sens de l'inclusion : si $Q : \Omega \rightarrow \text{sev}(\mathbf{R}^d)$ est mesurable et vérifie pour tout $u \in U$, pour μ -presque tout $x \in \Omega$, $u(x) \in Q(x)$, alors on a l'inclusion $T(x) \subset Q(x)$ pour μ -presque tout $x \in \Omega$.

L'application T est appelée la μ -union essentielle de la famille U .

Démonstration. Comme l'espace $L_\mu^p(\Omega, \mathbf{R}^d)$ est métrique séparable, on peut trouver une famille dénombrable $(u_n)_{n \in \mathbf{N}}$ d'éléments de U dense dans celui-ci. On pose alors, pour $x \in \Omega$,

$$T(x) := \text{Vect}\{u_n(x); n \in \mathbf{N}\}. \quad (3.7)$$

L'application $T : \Omega \rightarrow \text{sev}(\mathbf{R}^d)$ est alors mesurable par construction.

Si $Q : \Omega \rightarrow \mathbf{R}^d$ est telle que pour tout $u \in U$, pour μ -presque tout $x \in \Omega$, $u(x) \in Q(x)$, cela signifie que pour μ -presque tout $x \in \Omega$, pour tout $n \in \mathbf{N}$, $u_n(x) \in Q(x)$. Puisque $Q(x) \in \text{sev}(\mathbf{R}^d)$ pour tout $x \in \Omega$, l'inclusion $T \subset Q$ a bien lieu μ -presque partout.

Soit $u \in U$. Par densité, on peut trouver une suite $(n_m)_{m \in \mathbf{N}}$ telle que $u_{n_m} \rightarrow u$ dans $L_\mu^p(\Omega, \mathbf{R}^d)$ lorsque $m \rightarrow +\infty$. Quitte à extraire une sous-suite (on ne la renomme pas), on peut supposer que la convergence a lieu μ -presque partout. Soit A un borélien tel que $\mu(A) = 0$ et $u_{n_m}(x) \rightarrow u(x)$ pour tout $x \notin A$ lorsque $m \rightarrow +\infty$. Comme $u_{n_m}(x) \in T(x)$ pour tout $m \in \mathbf{N}$ et que $T(x)$ est fermé, $u(x) \in T(x)$ pour tout $x \notin A$.

L'unicité de T à un ensemble de μ -mesure nulle près est une conséquence immédiate de sa minimalité au sens de l'inclusion. \square

Remarque. Si U est un sous-espace vectoriel de $L_\mu^p(\Omega, \mathbf{R}^d)$, la μ -union essentielle de l'adhérence de U coïncide μ -presque partout avec celle de U . Pour le voir, il suffit de remarquer que si une famille $(u_n)_{n \in \mathbf{N}}$ est dense dans U , elle l'est aussi dans \bar{U} .

Remarque. Dans [15], un résultat beaucoup plus général est énoncé et démontré : on se donne (E, \mathcal{E}, ν) un espace mesuré et (X, d) un espace métrique complet séparable. Alors toute famille quelconque (pas nécessairement dénombrable) $(\Gamma_\alpha)_{\alpha \in A}$ d'applications mesurables de E à valeurs dans les parties fermées non vides de X possède une borne supérieure, c'est-à-dire qu'il existe une application Γ mesurable à valeurs dans les parties fermées non vides de X telle que pour tout $\alpha \in A$, pour ν -presque tout $e \in E$, $\Gamma_\alpha(e) \subset \Gamma(e)$, et telle que Γ soit minimale au sens de l'inclusion ν -presque partout.

Le théorème 3.7 est le cas particulier $A = U$ et $\Gamma_u : x \mapsto \{u(x)\}$ pour $u \in U$. Sa démonstration est bien plus simple que dans le cas général car la séparabilité de $L_\mu^p(\Omega, \mathbf{R}^d)$ permet de se ramener au cas où A est dénombrable.

La proposition suivante permet de mieux comprendre ce que représente la μ -union essentielle et elle sera utile par la suite.

Proposition 3.8. *Si $T : \Omega \rightarrow \text{sev}(\mathbf{R}^d)$ est mesurable, alors la μ -union essentielle de $L_\mu^p(\Omega, T)$ est T .*

Démonstration. Notons Q la μ -union essentielle de la famille $L_\mu^p(\Omega, T)$. Comme tout élément v de $L_\mu^p(\Omega, T)$ vérifie $v(x) \in T(x)$ pour μ -presque tout $x \in \Omega$, cela signifie, par minimalité de Q , que $Q \subset T$ μ -presque partout. Définissons le « reste » par $R := (Q + T^\perp)^\perp$ de telle sorte que $\mathbf{R}^d = Q \oplus R \oplus T^\perp$ et $T = Q \oplus R$, les sommes étant directes et orthogonales, on cherche alors à montrer que $R = \{0\}$ μ -presque partout.

L'application $R : \Omega \rightarrow \text{sev}(\mathbf{R}^d)$ est mesurable (c'est la proposition 3.4 qui permet de l'affirmer), on peut donc considérer un d -uplet (v_1, v_2, \dots, v_d) de fonctions adapté à R . En particulier, les v_a appartiennent à $L_\mu^p(\Omega, R)$ donc à $L_\mu^p(\Omega, T)$ pour $a \in \{1, 2, \dots, d\}$. Si $x \in \Omega$ est tel que $v_1(x) \neq 0$, cela signifie que $v_1(x) \notin Q(x)$, donc $\mu(\{v_1 \neq 0\}) = 0$ par définition de Q . Mais $\{v_1 \neq 0\}$ coïncide avec $\{R \neq \{0\}\}$, c'est donc suffisant pour conclure $Q = T$ μ -presque partout. \square

3.2 L'espace tangent de Zhikov et la structure du graphe

On présente ici la description de l'espace tangent à μ utilisée par Zhikov [17] et reprise par Louet [12], elle permet d'avoir une description précise du graphe \bar{G} .

Rappelons que $\bar{G} \subset L_\mu^p(\Omega, S^{\leq 1}\mathbf{R}^d) = L_\mu^p(\Omega, \mathbf{R}) \times L_\mu^p(\Omega, \mathbf{R}^d)$. Si $u \in W_\mu^{1,p}(\Omega)$, on note $\Gamma(u) \subset L_\mu^p(\Omega, \mathbf{R}^d)$ l'ensemble de ses gradients défini par

$$\Gamma(u) := \pi_1(\pi_0^{-1}(\{u\}) \cap \bar{G}) \quad (3.8)$$

$$= \{v \in L_\mu^p(\Omega, \mathbf{R}^d) \text{ tel que } (u, v) \in \bar{G}\} \quad (3.9)$$

$$= \{v \in L_\mu^p(\Omega, \mathbf{R}^d) \text{ tel que } \exists(\varphi_n)_{n \in \mathbf{N}} \subset \mathcal{D}(\Omega)^{\mathbf{N}} \text{ avec } \varphi_n \rightarrow u \text{ et } D\varphi_n \rightarrow v\}. \quad (3.10)$$

La deuxième caractérisation permet en particulier de voir que $\Gamma(u)$ est fermé. De plus, comme \bar{G} est un sous-espace vectoriel, on voit que $\Gamma(u)$ est un espace affine dirigé par l'ensemble des gradients de 0, c'est-à-dire $\Gamma(0)$. Or il se trouve que $\Gamma(0)$ vérifie une propriété de « localisation » utile pour étudier sa structure :

Lemme 3.9. *Soient w une fonction mesurable bornée à valeurs dans \mathbf{R} et $v \in \Gamma(0)$. Alors $wv \in \Gamma(0)$.*

Démonstration. Démontrons tout d'abord que si ψ est un élément de $\mathcal{D}(\Omega)$ et $v \in \Gamma(0)$ alors $\psi v \in \Gamma(0)$. En effet, dans ce cas, grâce à la caractérisation (3.10) de $\Gamma(0)$, on peut trouver une suite $(\varphi_n)_{n \in \mathbf{N}}$ d'éléments de $\mathcal{D}(\Omega)$ qui converge vers 0 dans $L_\mu^p(\Omega)$ et telle

que $D\varphi_n \rightarrow v$ dans $L_\mu^p(\Omega, \mathbf{R}^d)$ lorsque $n \rightarrow +\infty$. La suite $(\psi\varphi_n)_{n \in \mathbf{N}}$ tend aussi vers 0, et $D(\psi\varphi_n) = \varphi_n D\psi + \psi D\varphi_n$ tend vers ψv lorsque $n \rightarrow +\infty$.

Dans le cas général, si w est une fonction mesurable bornée, on peut trouver une suite $(\psi_n)_{n \in \mathbf{N}}$ d'éléments de $\mathcal{D}(\Omega)$ telle que $\psi_n \rightarrow w$ μ -presque partout lorsque $n \rightarrow +\infty$ et telle que $|\psi_n(x)|$ soit uniformément borné en $x \in \Omega$ et $n \in \mathbf{N}$ (il suffit d'utiliser le théorème de Lusin pour voir w comme une limite μ -presque partout de fonctions continues uniformément bornées, puis de les régulariser pour obtenir des fonctions de $\mathcal{D}(\Omega)$). À l'aide du théorème de convergence dominée, on voit que $\psi_n v \rightarrow wv$ dans $L_\mu^p(\Omega, \mathbf{R}^d)$ lorsque $n \rightarrow +\infty$. Mais $\psi_n v \in \Gamma(0)$ et $\Gamma(0)$ est fermé, ce qui suffit pour conclure. \square

On en est maintenant en mesure d'énoncer le théorème décrivant la structure de $\Gamma(0)$: grâce au lemme précédent, on va voir que $\Gamma(0)$ admet une description « point par point ».

Théorème 3.10. *Il existe une application $T_\mu^p : \Omega \rightarrow \text{sev}(\mathbf{R}^d)$ mesurable telle que $\Gamma(0) = L_\mu^p(\Omega, (T_\mu^p)^\perp)$.*

L'application T_μ^p est appelée l'espace tangent à μ .

L'espace tangent donne la « direction » locale de μ . Comme on le voit, il n'y a besoin d'aucune hypothèse sur μ pour le définir (seulement que c'est une mesure de Radon positive), une contrepartie étant, comme on le verra plus tard (cf. la remarque page 26), qu'il peut dépendre de l'exposant p .

Démonstration. Construction de l'espace tangent. Soit $R_\mu^p : \Omega \rightarrow \text{sev}(\mathbf{R}^d)$ la μ -union essentielle de la famille $\Gamma(0)$ (cf. théorème 3.7). Pour la définir, il suffit de prendre une famille $(v_n)_{n \in \mathbf{N}}$ de $\Gamma(0)$ dense dans celui-ci et de poser, pour $x \in \Omega$,

$$R_\mu^p(x) := \text{Vect}\{v_n(x); n \in \mathbf{N}\}. \quad (3.11)$$

On définit ensuite $T_\mu^p := (R_\mu^p)^\perp$, c'est une application mesurable d'après la proposition 3.4.

Première inclusion. Par construction il est clair que si $v \in \Gamma(0)$, alors $v(x) \in (T_\mu^p)^\perp(x) = R_\mu^p(x)$ pour μ -presque tout x , ce qui se lit $\Gamma(0) \subset L_\mu^p(\Omega, (T_\mu^p)^\perp)$.

Inclusion réciproque. Plutôt que de montrer directement $\Gamma(0) \supset L_\mu^p(\Omega, (T_\mu^p)^\perp)$, puisque les ensembles $\Gamma(0)$ et $L_\mu^p(\Omega, (T_\mu^p)^\perp)$ sont fermés, on préfère « passer » à l'orthogonal et montrer que $\Gamma(0)^\perp \subset (L_\mu^p(\Omega, (T_\mu^p)^\perp))^\perp$ au sens de la dualité $L_\mu^p - L_\mu^{p'}$. Grâce à la proposition 3.6, cela revient à montrer $\Gamma(0)^\perp \subset L_\mu^{p'}(\Omega, T_\mu^p)$, on se donne donc un élément w de $\Gamma(0)^\perp$. Si $v \in \Gamma(0)$ et B est un borélien, la fonction $\mathbb{1}_B v$ appartient à $\Gamma(0)$ d'après

le lemme 3.9, en conséquence $\int_B v \cdot w = 0$. Comme B est arbitraire, on en déduit que $v \cdot w = 0$ μ -presque partout. Les v_n (qui interviennent dans la définition de R_μ^p) sont en nombre dénombrable, donc il existe un borélien A tel que $\mu(A) = 0$ et tel que pour tout $x \notin A$ et tout $n \in \mathbb{N}$, $w(x) \cdot v_n(x) = 0$. On en déduit que pour tout $x \notin A$, $w(x) \perp R_\mu^p(x)$, c'est-à-dire $w(x) \in (T_\mu^p)(x)$. Cela suffit pour dire que $w \in L_\mu^p(\Omega, T_\mu^p)$ et conclut la preuve. \square

À partir de là, si $u \in W_\mu^{1,p}(\Omega)$ et $v \in \Gamma(u)$, la composante de v sur $(T_\mu^p)^\perp$ peut prendre n'importe quelle valeur tandis que celle sur T_μ^p ne dépend que de u , cette dernière sera appelée gradient intrinsèque à u et notée $D_\mu u$. Pour l'écrire formellement :

Théorème 3.11. *Il existe une application linéaire $D_\mu : W_\mu^{1,p}(\Omega) \rightarrow L_\mu^p(\Omega, T_\mu^p)$ telle que*

$$\bar{G} = \{(u, D_\mu u + v); u \in W_\mu^{1,p}(\Omega) \text{ et } v \in L_\mu^p(\Omega, (T_\mu^p)^\perp)\}. \quad (3.12)$$

Si de plus $\varphi \in \mathcal{D}(\Omega)$, alors $D_\mu \varphi = P_{T_\mu^p}(D\varphi)$. Quant à la norme sur $W_\mu^{1,p}(\Omega)$, elle est donnée par

$$\|u\|_{1,p,\mu} = \left(\|u\|_{p,\mu}^p + \|D_\mu u\|_{p,\mu}^p \right)^{\frac{1}{p}}. \quad (3.13)$$

On peut remarquer que l'identité $D_\mu \varphi = P_{T_\mu^p}(D\varphi)$ si $\varphi \in \mathcal{D}(\Omega)$ garantit que l'opérateur D_μ est local, puisque les valeurs de $D\varphi$ et T_μ^p en un point $x \in \Omega$ dépendent uniquement du comportement de respectivement φ et μ au voisinage de x .

Démonstration. Soit $u \in W_\mu^{1,p}(\Omega)$ et w un élément quelconque de $\Gamma(u)$. On définit $D_\mu u$ par $D_\mu u := P_{T_\mu^p} w$. Cette définition ne dépend pas du choix de w car deux éléments de $\Gamma(u)$ diffèrent par un élément de $\Gamma(0)$, et $P_{T_\mu^p}$ s'annule sur $\Gamma(0)$ d'après le théorème 3.10. De plus, $D_\mu u \in \Gamma(u)$ car $D_\mu u = w - P_{(T_\mu^p)^\perp} w$ et $P_{(T_\mu^p)^\perp} w$ est un élément de $L_\mu^p(\Omega, (T_\mu^p)^\perp)$ donc de $\Gamma(0)$ d'après le théorème 3.10. En conséquence, $\Gamma(u) = D_\mu u + \Gamma(0) = D_\mu u + L_\mu^p(\Omega, (T_\mu^p)^\perp)$ et c'est précisément ce que dit l'équation (3.12).

Si de plus $\varphi \in \mathcal{D}(\Omega)$, alors $D\varphi \in \Gamma(\varphi)$ donc $D_\mu \varphi = P_{T_\mu^p}(D\varphi)$.

Et pour obtenir la formule donnant la norme, il suffit de remarquer que si $(u, D_\mu u + v) \in \bar{G}$ avec $v \in L_\mu^p(\Omega, (T_\mu^p)^\perp)$ alors

$$\|(u, D_\mu u + v)\|_{p,\mu} = \left(\|u\|_{p,\mu}^p + \|D_\mu u\|_{p,\mu}^p + \|v\|_{p,\mu}^p \right)^{\frac{1}{p}}, \quad (3.14)$$

et cette quantité est minimale lorsque $v = 0$ et vaut dans ce cas le membre de droite de (3.13). \square

Exemple. On reprend l'exemple de la masse de Dirac donné page 10. Dans ce cas, puisque $\tilde{G} = L_\mu^p(\Omega, S^{\leq 1}\mathbf{R}^d)$, on voit que $\Gamma(0) = L_\mu^p(\Omega, \mathbf{R}^d)$, de sorte que $T_\mu^p = \{0\}$ μ -presque partout.

De manière générale, plus T_μ^p est petit et plus une fonction possède de dérivées, tandis que si $T_\mu^p = \mathbf{R}^d$ μ -presque partout, un élément de $W_\mu^{1,p}(\Omega)$ possède une unique dérivée.

Exemple. Montrons que si S est une sous-variété de \mathbf{R}^d de dimension m , et si $\mu := \mathcal{H}^m \llcorner S$ est la mesure de Hausdorff m -dimensionnelle restreinte à S , alors T_μ^p coïncide μ -presque partout avec l'espace tangent (de la géométrie différentielle) à S .

En effet, *via* un changement de coordonnées, on est ramené à traiter le cas $S = \Omega \cap (\mathbf{R}^m \times \{0\}^{d-m})$ et $\mu = \theta \mathcal{H}^m \llcorner S$, où $\theta : S \rightarrow \mathbf{R}_+$ est une fonction régulière strictement positive.

Soient $v \in \Gamma(0)$ et $(\varphi_n)_{n \in \mathbf{N}}$ une suite d'éléments de $\mathcal{D}(\Omega)$ qui converge vers 0 dans $L_\mu^p(\Omega)$ et telle que $D\varphi_n \rightarrow v$ dans $L_\mu^p(\Omega, \mathbf{R}^d)$ lorsque $n \rightarrow +\infty$. Si $a \in \{1, 2, \dots, m\}$ et si $\psi \in \mathcal{D}(\Omega) \subset L_\mu^p(\Omega)$, alors

$$\int_\Omega \frac{\partial \varphi_n}{\partial x_a} \psi \, d\mu = \int_{\mathbf{R}^m} \frac{\partial \varphi_n}{\partial x_a}(x_1, \dots, x_m, 0, \dots, 0) \psi(x_1, \dots, x_m, 0, \dots, 0) \theta(x_1, \dots, x_m) \, dx_1 \, dx_2 \dots \, dx_m \quad (3.15)$$

$$= - \int_{\mathbf{R}^m} \varphi_n(x_1, \dots, x_m, 0, \dots, 0) \frac{\partial}{\partial x_a} [\theta \psi](x_1, \dots, x_m, 0, \dots, 0) \, dx_1 \, dx_2 \dots \, dx_m \quad (3.16)$$

$$\rightarrow 0 \quad (3.17)$$

lorsque $n \rightarrow +\infty$ puisque $\varphi_n \rightarrow 0$ dans $L_\mu^p(\Omega)$. Mais comme on sait aussi que cette quantité converge vers $\int_\Omega v_a \psi \, d\mu$ (où v_a est la a -ième composante de v) et comme ψ est arbitraire, on en conclut que v_a est nulle. Or $a \in \{1, 2, \dots, m\}$ est arbitraire, c'est-à-dire qu'en notant Q l'espace $\{0\}^m \times \mathbf{R}^{m-d}$, on obtient $v \in L_\mu^p(\Omega, Q)$.

Réciproquement, si $v \in L_\mu^p(\Omega, Q)$, on peut approcher v dans $L_\mu^p(\Omega, \mathbf{R}^d)$ par une suite $(\psi_n)_{n \in \mathbf{N}}$ de fonctions régulières et à valeurs dans Q . On se donne alors $\chi \in \mathcal{D}(\Omega)$ qui vaut 1 au voisinage de S . En posant

$$\varphi_n(x_1, x_2, \dots, x_d) := \chi(x_1, x_2, \dots, x_d) \psi_n(x_1, x_2, \dots, x_m, 0, \dots, 0) \cdot (x_{m+1}, \dots, x_d), \quad (3.18)$$

on voit que φ_n s'annule sur S et que $D\varphi_n$ coïncide sur S avec ψ_n , ce qui montre que $\psi_n \in \Gamma(0)$. En faisant tendre n vers $+\infty$, on obtient $v \in \Gamma(0)$, ce qui permet d'obtenir

$$\Gamma(0) = L_\mu^p(\Omega, Q).$$

La proposition 3.8 permet de dire que Q est la μ -union essentielle de $\Gamma(0)$, or cette dernière est égale à $(T_\mu^p)^\perp$ par définition, de sorte que l'espace tangent T_μ^p coïncide μ -presque partout avec Q^\perp qui est précisément l'espace tangent à S au sens de la géométrie différentielle.

Pour terminer, montrons comment, à l'aide la description donnée par le théorème 3.11, on peut améliorer le résultat de relaxation donné par le théorème 2.6 et retrouver le théorème de Bouchitté [2, Théorème 3.1].

On se donne $h : \Omega \times \mathbf{R}^d \rightarrow \mathbf{R}_+$ une fonction continue positive et convexe en sa deuxième variable. On suppose qu'il existe deux constantes $C_2 \geq C_1 > 0$ telles que pour tout $(x, z) \in \Omega \times \mathbf{R}^d$,

$$C_1|z|^p \leq h(x, z) \leq C_2(1 + |z|^p). \quad (3.19)$$

On aura besoin de la fonction $h_\mu : \Omega \times \mathbf{R}^d \rightarrow \mathbf{R}_+$ définie par, pour $(x, z) \in \Omega \times \mathbf{R}^d$,

$$h_\mu(x, z) := \inf\{h(x, z + v); v \in (T_\mu^p)^\perp(x)\}. \quad (3.20)$$

Pour vérifier que cette fonction est mesurable, on se donne une famille dénombrable $(v_n)_{n \in \mathbf{N}}$ de fonctions $\Omega \rightarrow \mathbf{R}^d$ mesurables telle que pour μ -presque tout $x \in \Omega$, $(T_\mu^p)^\perp(x)$ soit le sous-espace vectoriel engendré par $\{v_n(x); n \in \mathbf{N}\}$. On note alors $(w_n)_{n \in \mathbf{N}}$ le \mathbf{Q} -espace vectoriel engendré par $(v_n)_{n \in \mathbf{N}}$ de sorte que pour μ -presque tout $x \in \Omega$, $(T_\mu^p)^\perp(x)$ soit égal à l'adhérence de la famille $\{w_n(x); n \in \mathbf{N}\}$. On peut alors écrire, par continuité de h , que pour μ -presque tout $x \in \Omega$ et pour tout $z \in \mathbf{R}^d$,

$$h_\mu(x, z) = \inf\{h(x, z + w_n(x)); n \in \mathbf{N}\}. \quad (3.21)$$

De même que dans le théorème 2.6, on note H la fonctionnelle définie sur $L_\mu^p(\Omega)$ par

$$H(\varphi) := \begin{cases} \int_\Omega h(x, D\varphi(x)) \, d\mu(x) & \text{si } \varphi \in \mathcal{D}(\Omega) \\ +\infty & \text{sinon} \end{cases}. \quad (3.22)$$

En notant \bar{H} la relaxée de H , on a le résultat suivant :

Théorème 3.12. *La fonctionnelle \bar{H} n'est finie que sur $W_\mu^{1,p}(\Omega)$, et si $u \in W_\mu^{1,p}(\Omega)$ alors*

$$\bar{H}(u) = \int_\Omega h_\mu(x, D_\mu u(x)) \, d\mu(x). \quad (3.23)$$

Démonstration. On note $I(u)$ le membre de droite de l'équation (3.23). En combinant le théorème 2.6 donnant l'expression générale de \bar{H} et le théorème 3.11 donnant la structure de \bar{G} dans le cas qui nous intéresse, on obtient

$$\bar{H}(u) = \inf \left\{ \int_{\Omega} h(x, D_{\mu}(x) + v(x)) \, d\mu(x); v \in L_{\mu}^p(\Omega, (T_{\mu}^p)^{\perp}) \right\}. \quad (3.24)$$

Au vu de la définition de h_{μ} , on en déduit directement que $\bar{H} \geq I$. Afin de montrer l'inégalité dans l'autre sens, il suffit d'échanger l'infimum et le symbole \int . C'est un résultat technique d'analyse convexe que l'on ne démontrera pas ici, on peut le trouver dans [4, Théorème 1] par exemple. □

3.3 Un autre point de vue : la définition de Bouchitté

Il se trouve que la définition de l'espace tangent à une mesure présentée ci-dessus n'est pas la seule présente dans la littérature, et il n'est pas inintéressant de voir comment les différentes définitions sont reliées entre elles, c'est d'ailleurs l'objet de l'article de Fragalà [6]. On ne le discutera cependant pas ici, la question relevant plus du domaine de la théorie géométrique de la mesure et ne concernant pas vraiment les espaces de Sobolev. On se contentera de montrer pourquoi notre définition coïncide avec celle de Bouchitté [2]. La preuve présentée s'inspire d'une remarque faite dans [3], on a tenté d'explicitier complètement l'argument.

L'idée est d'utiliser l'identité $\bar{G} = (G^{\perp})^{\perp}$ au sens de la dualité $L_{\mu}^p - L_{\mu}^{p'}$. Or, dire que $(u, v) \in L_{\mu}^{p'}(\Omega) \times L_{\mu}^{p'}(\Omega, \mathbf{R}^d)$ est dans G^{\perp} revient à dire que pour toute fonction $\varphi \in \mathcal{D}(\Omega)$,

$$\int_{\Omega} D\varphi \cdot v \, d\mu = - \int_{\Omega} \varphi u \, d\mu. \quad (3.25)$$

Cette observation motive la définition suivante :

Définition 3.13. On note $\operatorname{div}_{\mu} : L_{\mu}^{p'}(\Omega, \mathbf{R}^d) \rightarrow L_{\mu}^{p'}(\Omega)$ l'opérateur défini sur

$$D(\operatorname{div}_{\mu}) := \left\{ v \in L_{\mu}^{p'}(\Omega, \mathbf{R}^d) \text{ tel que } \exists C, \forall \varphi \in \mathcal{D}(\Omega), \left| \int_{\Omega} D\varphi \cdot v \, d\mu \right| \leq C \|\varphi\|_{p, \mu} \right\} \quad (3.26)$$

et tel que pour tout $v \in D(\operatorname{div}_{\mu})$,

$$\forall \varphi \in \mathcal{D}(\Omega), \int_{\Omega} \varphi \operatorname{div}_{\mu}(v) \, d\mu = - \int_{\Omega} D\varphi \cdot v \, d\mu. \quad (3.27)$$

On pourra d'ailleurs remarquer que si $v \in D(\operatorname{div}_\mu)$, alors $\operatorname{div}(v\mu) = \operatorname{div}_\mu(v) \cdot \mu$ au sens des distributions. Il est alors clair que $(u, v) \in G^\perp$ si et seulement si $v \in D(\operatorname{div}_\mu)$ et $u = \operatorname{div}_\mu(v)$. On en déduit sans trop de peine l'identité suivante :

Proposition 3.14. *On a l'égalité $\Gamma(0) = D(\operatorname{div}_\mu)^\perp$.*

Démonstration. Il suffit d'écrire

$$w \in \Gamma(0) \Leftrightarrow (0, w) \in \bar{G} \quad (3.28)$$

$$\Leftrightarrow (0, w) \perp G^\perp \quad (3.29)$$

$$\Leftrightarrow \forall (u, v) \in G^\perp, \int_\Omega v \cdot w \, d\mu = 0 \quad (3.30)$$

$$\Leftrightarrow \forall v \in D(\operatorname{div}_\mu), \int_\Omega v \cdot w \, d\mu = 0, \quad (3.31)$$

ce qui est exactement le résultat annoncé. \square

On en déduit la définition de l'espace tangent T_μ^p utilisée dans l'article [2].

Théorème 3.15. *L'espace tangent T_μ^p est la μ -union essentielle de la famille $D(\operatorname{div}_\mu)$.*

Démonstration. Notons Q la μ -union essentielle de la famille $D(\operatorname{div}_\mu)$. Comme on l'a fait remarquer à la suite de la preuve du théorème 3.7, Q est aussi la μ -union essentielle de la famille $\overline{D(\operatorname{div}_\mu)}$.

Mais grâce à la proposition 3.14, en prenant l'orthogonal, on obtient la relation $\Gamma(0)^\perp = \overline{D(\operatorname{div}_\mu)}$. Or il se trouve que $\Gamma(0)^\perp = L_\mu^{p'}(\Omega, T_\mu^p)$ (on utilise le théorème 3.10 et la proposition 3.6), donc $L_\mu^{p'}(\Omega, T_\mu^p) = \overline{D(\operatorname{div}_\mu)}$. Or la μ -union essentielle de $L_\mu^{p'}(\Omega, T_\mu^p)$ est T_μ^p (cf. proposition 3.8) ce qui permet d'écrire $Q = T_\mu^p$ et conclut la preuve. \square

3.4 Une description explicite en dimension un

Dans le cas où $d = 1$, Louet [12] a donné une description complètement explicite de l'espace tangent et de l'espace de Sobolev $W_\mu^{1,p}$. Puisqu'elle sera particulièrement utile pour l'étude du transport optimal avec pénalisation en gradient, on la reproduit ici. Dans la suite, I désigne un intervalle ouvert et borné de \mathbb{R} et μ est une mesure de Radon positive et finie sur I , tandis que la mesure de Lebesgue sur I est notée \mathcal{L} .

On note $\mu =: \mu_a + \mu_s$ la décomposition de μ en une mesure μ_a absolument continue par rapport à \mathcal{L} et une mesure μ_s étrangère à \mathcal{L} . On désigne par A_μ un borélien tel que $\mathcal{L}(A_\mu) = 0$ et $\mu_s(I \setminus A_\mu) = 0$, tandis que $f \in L^1(I)$ est la fonction positive telle que $\mu_a = f \cdot \mathcal{L}$.

Le « mauvais » ensemble M_μ^p est défini comme le complémentaire du plus grand ouvert sur lequel $f^{-\frac{1}{p-1}}$ est localement intégrable, à savoir

$$M_\mu^p := \left\{ x \in I \text{ tel que } \forall \varepsilon > 0, \int_{]x-\varepsilon, x+\varepsilon[} f^{-\frac{1}{p-1}} d\mathcal{L} = +\infty \right\}. \quad (3.32)$$

L'ensemble M_μ^p est un fermé et il est construit de telle sorte que l'injection de $L_{\mu_a}^p(I \setminus M_\mu^p)$ dans $L_{\text{loc}}^1(I \setminus M_\mu^p)$ soit continue (cf. proposition 2.1). L'espace tangent ne peut prendre que deux valeurs possibles puisque \mathbf{R} possède seulement deux sous-espaces vectoriels, et l'on a la description suivante :

Théorème 3.16. *Pour μ -presque tout x de I ,*

$$T_\mu^p(x) = \begin{cases} \mathbf{R} & \text{si } x \notin (M_\mu^p \cup A_\mu) \\ \{0\} & \text{si } x \in (M_\mu^p \cup A_\mu) \end{cases}. \quad (3.33)$$

La démonstration de ce théorème (surtout pour ce qui est de montrer que $T_\mu^p = \{0\}$ sur M_μ^p) est assez technique, on ne la reproduira pas ici et on renvoie à [12, Théorème 1.3.1].

Remarque. L'ensemble M_μ^p dépend de l'exposant p , en conséquence l'espace tangent T_μ^p peut dépendre de p . Il suffit par exemple de prendre $I :=]-1, 1[$ et $f(x) := |x|_{]-1, 1[}$ puis de remarquer que $M_{f, \mathcal{L}}^p = \{0\}$ pour $p \leq 2$ et $M_{f, \mathcal{L}}^p = \emptyset$ sinon.

Pour montrer que la dépendance de T_μ^p en x peut être assez compliquée, on va étendre légèrement un résultat prouvé par Louet [12, proposition 1.4.1] en montrant que M_μ^p peut être *a priori* n'importe quel fermé, même si la densité f est strictement positive \mathcal{L} -presque partout.

Proposition 3.17. *Soit F un fermé de I . Alors il existe une fonction $f \in L^1(I)$ strictement positive \mathcal{L} -presque partout telle que l'ensemble des points autour desquels $f^{-\frac{1}{p-1}}$ n'est pas intégrable soit exactement F (en notation abrégée $M_{f, \mathcal{L}}^p = F$).*

Démonstration. Remarquons tout d'abord que quitte à considérer f^{p-1} , on peut se ramener au cas $p = 2$, c'est-à-dire que l'on cherche à imposer que f^{-1} ne soit pas intégrable uniquement au voisinage des points de F .

Définition de f . Soit $(x_n)_{n \in \mathbf{N}}$ une famille dénombrable d'éléments (pas nécessairement deux à deux distincts) de I dense dans F . Pour $\eta > 0$ fixé, on note U_η l'ouvert

$$U_\eta := \left(\bigcup_{n \in \mathbf{N}} \left] x_n - \frac{\eta}{2^n}, x_n + \frac{\eta}{2^n} \right[\right) \cap I, \quad (3.34)$$

de telle sorte que $(U_\eta)_{\eta>0}$ forme une famille croissante d'ensembles mesurables. Une majoration immédiate montre que $\mathcal{L}(U_\eta) \leq 4\eta$.

On définit ensuite la fonction f mesurable positive par

$$f := \left(1 + \sum_{m=1}^{+\infty} \mathbb{1}_{U_{\frac{1}{m}}} \right)^{-1}. \quad (3.35)$$

Elle est clairement bornée par 1 donc $f \in L^1(I)$. De plus $f(x) = 0$ si et seulement si x appartient à $U_{\frac{1}{m}}$ pour une infinité de m ce qui revient à dire, par décroissance de la suite $(U_{\frac{1}{m}})_{m \in \mathbb{N}^*}$, que x appartient à $U_{\frac{1}{m}}$ pour tout $m \in \mathbb{N}^*$. Or puisque

$$\mathcal{L} \left(\bigcap_{m \in \mathbb{N}^*} U_{\frac{1}{m}} \right) = \lim_{m \rightarrow +\infty} \mathcal{L} \left(U_{\frac{1}{m}} \right) = 0, \quad (3.36)$$

on peut conclure que $f > 0$ \mathcal{L} -presque partout.

Intégrabilité de f^{-1} en dehors de F . Soit x un point de l'ouvert $I \setminus F$. Il existe $\varepsilon > 0$ tel que $]x - 2\varepsilon, x + 2\varepsilon[\cap F = \emptyset$. Si $y \in]x - \varepsilon, x + \varepsilon[$, alors y est à distance au moins ε de F donc $y \notin U_{\frac{1}{m}}$ pour $m > \varepsilon^{-1}$. Cela montre que $f(y) > (1 + \varepsilon^{-1})^{-1}$ pour tout $y \in]x - \varepsilon, x + \varepsilon[$, et donc que f^{-1} est intégrable sur $]x - \varepsilon, x + \varepsilon[$.

Non intégrabilité de f^{-1} au voisinage des points de F . Soient $x \in F$ et $\varepsilon > 0$ tels que $]x - 2\varepsilon, x + 2\varepsilon[\subset I$. Par densité il existe $n_0 \in \mathbb{N}$ tel que $x_{n_0} \in]x - \varepsilon, x + \varepsilon[$. En particulier, pour tout $m > \varepsilon^{-1}$, on a l'inclusion

$$\left] x_{n_0} - \frac{1}{2^{n_0}m}, x_{n_0} + \frac{1}{2^{n_0}m} \right[\subset]x - 2\varepsilon, x + 2\varepsilon[, \quad (3.37)$$

ce qui permet d'écrire que

$$\mathcal{L} \left(U_{\frac{1}{m}} \cap]x - 2\varepsilon, x + 2\varepsilon[\right) \geq \mathcal{L} \left(\left] x_{n_0} - \frac{1}{2^{n_0}m}, x_{n_0} + \frac{1}{2^{n_0}m} \right[\cap]x - 2\varepsilon, x + 2\varepsilon[\right) \quad (3.38)$$

$$= \mathcal{L} \left(\left] x_{n_0} - \frac{1}{2^{n_0}m}, x_{n_0} + \frac{1}{2^{n_0}m} \right[\right) \quad (3.39)$$

$$= \frac{1}{2^{n_0-1}m}. \quad (3.40)$$

Cette minoration permet d'en déduire

$$\int_{x-2\varepsilon}^{x+2\varepsilon} f^{-1} d\mathcal{L} = 4\varepsilon + \sum_{m=1}^{+\infty} \int_{x-2\varepsilon}^{x+2\varepsilon} \mathbb{1}_{U_{\frac{1}{m}}} d\mathcal{L} \quad (3.41)$$

$$= 4\varepsilon + \sum_{m=1}^{+\infty} \mathcal{L} \left(U_{\frac{1}{m}} \cap]x-2\varepsilon, x+2\varepsilon[\right) \quad (3.42)$$

$$\geq 4\varepsilon + \sum_{m > \varepsilon^{-1}} \frac{1}{2^{n_0-1} m} \quad (3.43)$$

$$= +\infty, \quad (3.44)$$

ce qui est exactement le résultat voulu. \square

À partir de la structure de l'espace tangent T_{μ}^p on aimerait pouvoir en déduire celle de $W_{\mu}^{1,p}(I)$. Ce que l'on sait grâce au théorème 3.11 est que si $\varphi \in \mathcal{D}(I)$, alors $D_{\mu}\varphi = \varphi' \mathbb{1}_{I \setminus (M_{\mu}^p \cup A_{\mu})}$. On pourrait donc croire que $W_{\mu}^{1,p}(I)$ est l'ensemble des fonctions $u \in L_{\mu}^p(I)$ telles que la dérivée u' de u au sens des distributions, bien définie sur $I \setminus M_{\mu}^p$, soit telle que $u' \mathbb{1}_{I \setminus (M_{\mu}^p \cup A_{\mu})} \in L_{\mu}^p(I \setminus M_{\mu}^p)$, c'est notamment ce qu'a affirmé Louet [12, Proposition 1.3.1]. Mais on doit prendre garde au fait que la dérivée au sens des distributions ne voit pas les ensembles \mathcal{L} -négligeables alors que la mesure μ_s est précisément concentrée sur un tel ensemble. On va présenter deux exemples qui invalident le résultat de Louet tel qu'il l'a formulé.

Premier exemple. On prend $p := 2$, $I :=]-1, 1[$, μ_a est la mesure de Lebesgue sur I (c'est-à-dire $f := 1$) et μ_s est une masse de Dirac en 0 : en notation condensée, $\mu := \mathcal{L} + \delta_0$. Ici $M_{\mu}^p = \emptyset$ et $A_{\mu} = \{0\}$ convient. Prenons pour u la fonction $\mathbb{1}_{\{0\}}$, c'est-à-dire qui vaut 1 uniquement en 0 et 0 partout ailleurs : elle appartient clairement à $L_{\mu}^p(I)$. Comme $\{0\}$ est \mathcal{L} -négligeable, u est nulle au sens des distributions donc sa dérivée au sens des distributions est un élément de $L_{\mu}^p(I)$. Cependant, u n'est pas un élément de $W_{\mu}^{1,p}(I)$: si on se donne une suite $(\varphi_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{D}(I)$ qui converge vers u dans $L_{\mu}^p(I)$, la présence de la masse de Dirac force $(\varphi_n(0))_{n \in \mathbb{N}}$ à converger vers $u(0) = 1$, tandis qu'à extraction près $(\varphi_n)_{n \in \mathbb{N}}$ converge \mathcal{L} -presque partout vers 0. Une minoration élémentaire montre qu'alors la norme de φ_n' dans $L_{\mu_a}^p$ ne peut que tendre vers $+\infty$ lorsque $n \rightarrow +\infty$ et donc que $(\varphi_n')_{n \in \mathbb{N}}$ ne peut être une suite convergente de $L_{\mu}^p(I)$.

On peut cependant remarquer que dans ce cas toute fonction de $W_{\mu}^{1,p}(I)$ possède un représentant continu (il s'agit des injections de Sobolev « classiques » puisque $\mu_a = \mathcal{L}$), et il suffirait d'imposer que la valeur en 0 d'une fonction de $W_{\mu}^{1,p}(I)$ soit celle de ce représentant continu. Le deuxième exemple montre une situation légèrement plus délicate.

Deuxième exemple. On prend $p := 2$, $I :=]-1, 1[$ et pour f la fonction définie sur $]-1, 0[$ par $f(x) := |x|$ et sur $[0, 1[$ par $f(x) := 1$. On note $\mu_a = f \cdot \mathcal{L}$ tandis que μ_s est une masse de Dirac en 0 : en notation condensée, $d\mu(x) := \mathbb{1}_{\{-1 < x < 0\}}|x| \cdot d\mathcal{L}(x) + \mathbb{1}_{\{0 \leq x < 1\}} \cdot d\mathcal{L}(x) + d\delta_0(x)$. On a alors $M_\mu^p = \{0\}$ et $A_\mu = \{0\}$ convient. Considérons la fonction u définie par -1 sur $]-1, 0[$ et 1 sur $]0, 1[$.

La question est de savoir quelle valeur attribuer à u en 0 : comme on le verra par la suite, puisque $f^{-\frac{1}{p-1}} = f^{-1}$ est intégrable sur $[0, 1[$ mais pas $]-1, 0]$, la fonction u doit être continue sur $[0, 1[$ mais pas nécessairement sur $]-1, 0]$ et il faut donc imposer $u(0) = 1$. C'est alors une condition suffisante et la fonction u ainsi définie appartient à $W_\mu^{1,p}(I)$.

Ces exemples montrent que des choses délicates peuvent se passer lorsque d'éventuels atomes de μ se situent sur M_μ^p . On arrive cependant à formuler le résultat suivant :

Théorème 3.18. *Un élément u de $L_\mu^p(I)$ appartient à $W_\mu^{1,p}(I)$ si et seulement si u possède un représentant qui soit continu sur tout segment S tel que $f^{-\frac{1}{p-1}} \in L^1(S)$ et si $(u|_{I \setminus M_\mu^p})' \in L_{\mu_a}^p(I \setminus M_\mu^p)$ (où la dérivée est prise au sens des distributions). Dans ce cas $D_\mu u = (u|_{I \setminus M_\mu^p})' \mathbb{1}_{I \setminus (M_\mu^p \cup A_\mu)}$ et*

$$\|u\|_{1,p,\mu} = \left(\|u\|_{p,\mu}^p + \left\| (u|_{I \setminus M_\mu^p})' \right\|_{p,\mu_a}^p \right)^{\frac{1}{p}}. \quad (3.45)$$

Sur la partie $I \setminus (M_\mu^p \cup A_\mu)$, tout se passe comme si on avait un espace de Sobolev « classique » puisque l'on peut définir les dérivées au sens des distributions. Au contraire, sur $M_\mu^p \cup A_\mu$, l'espace de Sobolev coïncide avec $L_\mu^p(M_\mu^p \cup A_\mu)$.

La démonstration est assez longue et essentiellement technique : la difficulté (qui apparaît sur les exemples) est la présence éventuelle d'atomes dans la mesure μ .

Démonstration. Sens direct. Commençons par énoncer une version adaptée à notre cas des injections de Sobolev :

Lemme 3.19. *Soit $S \subset I$ un segment tel que $f^{-\frac{1}{p-1}} \in L^1(S)$. Si $v \in W_\mu^{1,p}(I)$, alors v possède un représentant continu sur S dont le module de continuité est majoré par une quantité ne dépendant que de f et de $\|v\|_{1,p,\mu}$.*

Démonstration du lemme. Grâce à l'inégalité de Hölder, la même que celle utilisée dans

la preuve de la proposition 2.1, si φ est une fonction régulière, $x, y \in S$ et $x \leq y$,

$$|\varphi(y) - \varphi(x)| = \left| \int_x^y \varphi'(t) dt \right| \quad (3.46)$$

$$\leq \|\varphi' \mathbb{1}_{[x,y]}\|_1 \quad (3.47)$$

$$\leq \|\varphi'\|_{p,\mu_a} \left\| f^{-\frac{1}{p-1}} \mathbb{1}_{[x,y]} \right\|_1^{\frac{1}{p'}}. \quad (3.48)$$

On se donne $v \in W_\mu^{1,p}(I)$ et une suite $(\varphi_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{D}(I)$ qui converge vers v et telle que $(\varphi'_n)_{n \in \mathbb{N}}$ soit bornée dans $L_\mu^p(I)$. À extraction près, on peut supposer que la convergence a lieu μ -presque partout. Mais grâce à la majoration (3.48), puisque $\left\| f^{-\frac{1}{p-1}} \mathbb{1}_{[x,y]} \right\|_1^{\frac{1}{p'}}$ tend vers 0 lorsque $x \rightarrow y$ d'après le théorème de convergence monotone, on voit que le module de continuité des φ_n sur S est uniformément borné en $n \in \mathbb{N}$. Puisque de plus la convergence a lieu μ -presque partout, on peut conclure que la convergence de la suite $(\varphi_n)_{n \in \mathbb{N}}$ a lieu uniformément sur S . Cela permet de déduire que v possède un représentant continu sur S tel que pour tous $x, y \in S$ avec $x \leq y$,

$$|v(y) - v(x)| \leq \|v\|_{1,p,\mu} \left\| f^{-\frac{1}{p-1}} \mathbb{1}_{[x,y]} \right\|_1^{\frac{1}{p'}}. \quad (3.49)$$

□

Soit $u \in W_\mu^{1,p}(I)$. Grâce au lemme 3.19 on sait que u possède un représentant continu sur tout segment S tel que $f^{-\frac{1}{p-1}} \in L^1(S)$.

On se donne ensuite une suite $(\varphi_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{D}(I)$ telle que $\varphi_n \rightarrow u$ et $D\varphi_n \rightarrow D_\mu u$ dans $L_\mu^p(I)$ lorsque $n \rightarrow +\infty$. Comme $\mu_a \leq \mu$ et que $L_{\mu_a}^p(I \setminus M_\mu^p)$ s'injecte continûment dans $L_{\text{loc}}^1(I \setminus M_\mu^p)$ donc dans les distributions, la suite $(\varphi_n|_{I \setminus M_\mu^p})_{n \in \mathbb{N}}$ converge au sens des distributions vers $u|_{I \setminus M_\mu^p}$. En conséquence, la dérivée au sens des distributions de $u|_{I \setminus M_\mu^p}$ coïncide \mathcal{L} -presque partout avec la limite de la suite $(D\varphi_n|_{I \setminus M_\mu^p})_{n \in \mathbb{N}}$ à savoir $D_\mu u|_{I \setminus M_\mu^p}$. Puisque $D_\mu u$ appartient à $L_\mu^p(I)$ qui est inclus dans $L_{\mu_a}^p(I)$, on a $(u|_{I \setminus M_\mu^p})' \in L_{\mu_a}^p(I \setminus M_\mu^p)$. Comme de plus l'espace tangent est nul sur $M_\mu^p \cup A_\mu$, on peut écrire $D_\mu u = (u|_{I \setminus M_\mu^p})' \mathbb{1}_{I \setminus (M_\mu^p \cup A_\mu)}$. L'équation (3.45) est alors une conséquence directe de l'équation (3.13).

Sens réciproque. On se donne une fonction u continue sur tout segment S tel que $f^{-\frac{1}{p-1}} \in L^1(S)$ et telle que $(u|_{I \setminus M_\mu^p})' \in L_{\mu_a}^p(I \setminus M_\mu^p)$ (où la dérivée est prise au sens des distributions). On suppose dans un premier temps que cette fonction est bornée par un réel C .

On se fixe maintenant $\varepsilon > 0$. Pour prouver que $u \in W_\mu^{1,p}(I)$, on va approximer u dans

$L^p_\mu(I)$ à ε près par une fonction φ de classe C^1 dont la dérivée sera bornée dans $L^p_{\mu_a}(I)$ indépendamment de ε , ce qui est en fait suffisant pour assurer que $D_\mu\varphi$ est bornée dans $L^p_\mu(I)$ indépendamment de ε puisque l'espace tangent est réduit au singleton $\{0\}$ sur A_μ .

Sens réciproque, définition du « bon » ensemble B_μ^p . On note \mathcal{S} l'ensemble des segments non réduits à des singletons sur lesquels $f^{-\frac{1}{p-1}}$ est intégrable et $B_\mu^p := \bigcup_{S \in \mathcal{S}} S$. On vérifie facilement l'inclusion $I \setminus M_\mu^p \subset B_\mu^p$ mais elle peut être stricte : on pourra s'en convaincre en regardant le deuxième exemple (page 29) où $B_\mu^p = I$ tandis que $M_\mu^p = \{0\} \neq \emptyset$. En fait, un point $x \in I$ appartient à $I \setminus M_\mu^p$ s'il existe un voisinage de x sur lequel $f^{-\frac{1}{p-1}}$ est intégrable, et à B_μ^p si $f^{-\frac{1}{p-1}}$ est intégrable sur un voisinage à gauche ou un voisinage à droite de x . Mais quoiqu'il en soit on peut écrire B_μ^p comme une union dénombrable d'éléments de \mathcal{S} . On peut donc trouver un nombre fini de segments $S_1 = [x_1, y_1], S_2 = [x_2, y_2], \dots, S_N = [x_N, y_N]$ tels que :

- les segments ne soient pas réduits à des singletons c'est-à-dire $x_i < y_i$ pour $i \in \{1, 2, \dots, N\}$,
- les segments soient deux à deux disjoints, on impose donc $y_i < x_{i+1}$ pour $i \in \{1, 2, \dots, N-1\}$,
- la fonction $f^{-\frac{1}{p-1}}$ soit intégrable sur chacun des S_i pour $i \in \{1, 2, \dots, N\}$ mais ne le soit sur aucune des composantes connexes de $I \setminus (S_1 \cup S_2 \cup \dots \cup S_N)$,
- ces segments recouvrent presque tout B_μ^p , c'est-à-dire $\mu(B_\mu^p \setminus (S_1 \cup S_2 \cup \dots \cup S_N)) \leq \varepsilon^p$.

Pour $i \in \{1, 2, \dots, N\}$, on peut approcher dans $L^p_{\mu_a}(S_i)$ la fonction u' par une suite $(\psi_n)_{n \in \mathbb{N}}$ de fonctions C^∞ qui s'annulent aux bords de S_i . Pour $n \in \mathbb{N}$, on définit φ_n comme la primitive de ψ_n qui vaut $u(x_i)$ en x_i . Grâce à la majoration (3.49), puisque u et φ_n coïncident en x_i , si $x \in S_i$ alors

$$|u(x) - \varphi_n(x)| \leq \|u' - \psi_n\|_{p, \mu_a} \left\| f^{-\frac{1}{p-1}} \mathbb{1}_{S_i} \right\|_1^{\frac{1}{p}}, \quad (3.50)$$

ce qui montre que $(\varphi_n)_{n \in \mathbb{N}}$ converge uniformément vers u sur S_i , en particulier la convergence a aussi lieu dans $L^p_\mu(S_i)$. En prenant $n \in \mathbb{N}$ assez grand et en répétant cette opération N fois, on peut trouver une fonction φ définie sur $S_1 \cup S_2 \cup \dots \cup S_N$ de classe C^∞ , telle que

$$\|(\varphi - u)|_{S_1 \cup S_2 \cup \dots \cup S_N}\|_{p, \mu} \leq \varepsilon \quad (3.51)$$

et dont la dérivée est nulle aux bords de chacun des S_i tout en étant proche de u' dans $L^p_{\mu_a}$ donc bornée :

$$\|\varphi'|_{S_1 \cup S_2 \cup \dots \cup S_N}\|_{p, \mu_a} \leq \|u' \mathbb{1}_{I \setminus M_\mu^p}\|_{p, \mu_a} + 1. \quad (3.52)$$

Sens réciproque, traitement de M_μ^p : approximation par une fonction constante par morceaux. Notons $y_0 < x_{N+1}$ les bornes de I de telle sorte que le complémentaire de $S_1 \cup S_2 \cup \dots \cup S_N$ soit l'union des $]y_i, x_{i+1}[$ pour i parcourant $\{0, 1, \dots, N\}$.

On se fixe alors $i \in \{0, 1, \dots, N\}$. Les fonctions constantes par morceaux (parfois appelées fonctions en escalier) sont denses dans les fonctions continues au sens de la convergence uniforme et ces dernières sont denses dans $L_\mu^p(]y_i, x_{i+1}[)$. En conséquence, on peut trouver une fonction v constante par morceaux telle que $\|(u - v)|_{]y_i, x_{i+1}[}\|_{p, \mu} \leq \frac{\varepsilon}{N+1}$. Quitte à créer une discontinuité supplémentaire au voisinage des bords on peut d'ailleurs la choisir telle que $v(y_i) = \varphi(y_i)$ et $v(x_{i+1}) = \varphi(x_{i+1})$, où φ est la fonction définie ci-dessus sur $S_1 \cup S_2 \cup \dots \cup S_N$. Notons alors $y_i < z_1 < z_2 < \dots < z_L < x_{i+1}$ les L points de discontinuité de v .

Il s'agit ensuite de déplacer les points de discontinuité de v de telle sorte à ce qu'ils se trouvent sur M_μ^p . Plus précisément, on remarque que $M_\mu^p \cap [y_i, x_{i+1}]$ est un fermé non vide (puisque $f^{-\frac{1}{p-1}}$ n'est pas intégrable sur $]y_i, x_{i+1}[$), donc pour $j \in \{1, 2, \dots, L\}$ on peut noter \hat{z}_j le point de $M_\mu^p \cap [y_i, x_{i+1}]$ qui minimise la distance de z_j à $M_\mu^p \cap [y_i, x_{i+1}]$ (s'il y en a deux on choisit arbitrairement le plus petit). On a alors les inégalités $y_i \leq \hat{z}_1 \leq \hat{z}_2 \leq \dots \leq \hat{z}_L \leq x_{i+1}$ et on définit la fonction \hat{v} en « déplaçant » les discontinuités de v sur les \hat{z}_j . Concrètement, on pose par convention $z_0 = \hat{z}_0 := y_i$ et $z_{L+1} = \hat{z}_{L+1} := x_{i+1}$, et si $\hat{z}_j < \hat{z}_{j+1}$ alors \hat{v} prend sur $] \hat{z}_j, \hat{z}_{j+1}[$ la valeur que v prend sur $]z_j, z_{j+1}[$ (cf. la figure 1 pour un exemple). Si un des \hat{z}_j est un atome pour la mesure μ alors \hat{v} prend la même valeur que v en \hat{z}_j , et sur les autres points de discontinuité de \hat{v} on impose de manière arbitraire que cette dernière soit continue à droite. Pour estimer de combien \hat{v} s'écarte de v , remarquons que l'ensemble des points où \hat{v} diffère de v n'intersecte pas M_μ^p donc est contenu dans B_μ^p , et sur celui-ci la différence est au plus $2C$ (où C est un majorant de u , donc de v). En conséquence $\|\hat{v} - v\|_{p, \mu} \leq 2C\mu(]y_i, x_{i+1}[\cap B_\mu^p)^{\frac{1}{p}}$.

Sens réciproque, traitement de M_μ^p : régularisation autour des points de discontinuité. Pour transformer \hat{v} en une fonction continue on utilise le lemme suivant :

Lemme 3.20. *Soit $S = [w_0, w_1]$ un segment sur lequel $f^{-\frac{1}{p-1}}$ n'est pas intégrable. Pour tout $\eta > 0$ et tout couple (v_0, v_1) , il existe une fonction φ de classe C^∞ telle que :*

- $\varphi(w_0) = v_0$ et $\varphi(w_1) = v_1$,
- la fonction φ soit monotone donc bornée par $\max(|v_0|, |v_1|)$,
- la dérivée de φ soit nulle aux bords : $\varphi'(w_0) = 0$ et $\varphi'(w_1) = 0$,
- la dérivée de φ soit petite dans $L_{\mu_a}^p(S)$: $\|\varphi'\|_{p, \mu_a} \leq \eta$.

Preuve du lemme. On peut supposer que $v_1 \neq v_0$: dans le cas contraire la fonction constante et égale à v_0 convient. Grâce à la proposition 2.1 (qui est l'ingrédient clé de la

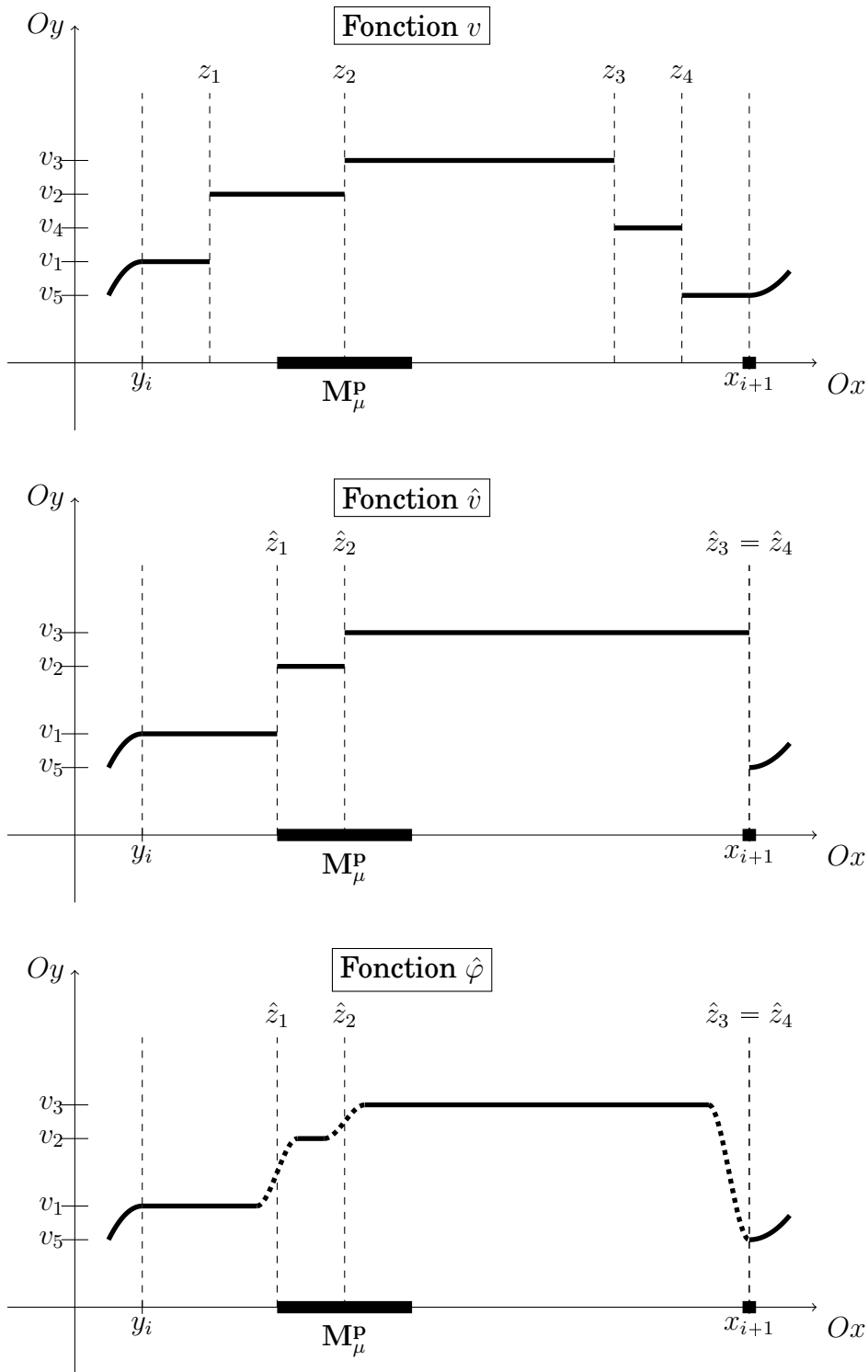


FIGURE 1 – Exemple de fonction v constante par morceaux dont on déplace les points de discontinuité sur M_μ^p pour obtenir \hat{v} , puis que l'on régularise autour de ses points de discontinuité (en pointillés) pour obtenir la fonction $\hat{\varphi}$ de classe C^1 . Ici, $M_\mu^p \cap [y_i, x_{i+1}]$ est constitué de l'union d'un segment (en gras) et du singleton $\{x_{i+1}\}$.

démonstration), une fois η fixé, on peut trouver une fonction g mesurable positive bornée telle que $\|g\|_1 = 1$ mais $\|g\|_{p,\mu_a} < \frac{\eta}{v_1 - v_0}$. En utilisant le théorème de Lusin on peut trouver une suite $(\psi_n)_{n \in \mathbb{N}}$ de fonctions positives de classe C^∞ s'annulant en w_0 et w_1 qui converge \mathcal{L} -presque partout vers g et de manière uniformément bornée en $x \in S$ et $n \in \mathbb{N}$. Quitte à les normaliser, on peut supposer de plus que $\|\psi_n\|_1 = 1$ pour tout $n \in \mathbb{N}$. On trouve $n_0 \in \mathbb{N}$ assez grand tel que $\|\psi_{n_0}\|_{p,\mu_a} \leq \frac{\eta}{|v_1 - v_0|}$ et on définit, pour $x \in S$,

$$\varphi(x) := v_0 + (v_1 - v_0) \int_{w_0}^x \psi_{n_0}(y) dy. \quad (3.53)$$

En effet, φ est une fonction de classe C^∞ , monotone (car ψ_{n_0} est positive) et telle que $\varphi(w_0) = v_0$ et $\varphi(w_1) = v_1$. De plus, comme $\varphi' = (v_1 - v_0)\psi'_{n_0}$, la dérivée est bien nulle aux bords et $\|\varphi'\|_{p,\mu_a} \leq \eta$. \square

On considère maintenant \hat{z}_j un point de discontinuité de \hat{v} , il appartient donc à M_μ^p . Si \hat{z}_j n'est pas un atome de μ ou appartient à B_μ^p , on peut trouver un segment S , inclus dans $[y_i, x_{i+1}]$, contenant \hat{z}_j mais pas d'autres points de discontinuité de \hat{v} , tel que $\mu(S) \leq \frac{\varepsilon^p}{L(N+1)^p} + \mu(B_\mu^p \cap \{\hat{z}_j\})$ (où L , rappelons-le, majore le nombre de points de discontinuité de \hat{v}), et tel que $f^{-\frac{1}{p-1}}$ ne soit pas intégrable sur S . Le lemme 3.20 garantit que sur S on peut remplacer \hat{v} par une fonction $\hat{\varphi}$ de classe C^∞ , telle que la jonction aux extrémités de S soit C^1 (notamment car \hat{v} est constante donc de dérivée nulle sur $S \setminus \{\hat{z}_j\}$), et telle que $\|\hat{\varphi}'|_S\|_{p,\mu_a} \leq \frac{\varepsilon}{L(N+1)}$.

Si \hat{z}_j est un atome de μ et n'appartient pas à B_μ^p cela signifie que $f^{-\frac{1}{p-1}}$ n'est intégrable ni au voisinage à gauche, ni au voisinage à droite de \hat{z}_j . On peut donc trouver deux segments $S_g = [z_g, \hat{z}_j]$ et $S_d = [\hat{z}_j, z_d]$ ne contenant pas de points de discontinuité de \hat{v} à part \hat{z}_j , tels que $y_i < z_g < \hat{z}_j < z_d < x_{i+1}$, $\mu(]z_g, \hat{z}_j[) \leq \frac{\varepsilon^p}{2L(N+1)^p}$, $\mu(]\hat{z}_j, z_d[) \leq \frac{\varepsilon^p}{2L(N+1)^p}$ et $f^{-\frac{1}{p-1}}$ ne soit intégrable ni sur S_g , ni sur S_d . On utilise ensuite le lemme 3.20 sur S_g et S_d pour obtenir une fonction $\hat{\varphi}$ de classe C^1 qui coïncide avec \hat{v} en z_g, z_d et \hat{z}_j , dont la dérivée s'annule en z_g, z_d et \hat{z}_j , et telle que $\|\hat{\varphi}'|_{S_g \cup S_d}\|_{p,\mu_a} \leq \frac{\varepsilon}{L(N+1)}$.

En effectuant ce traitement pour chacun des points de discontinuité de \hat{v} et en définissant $\hat{\varphi} := \hat{v}$ sinon, on construit une fonction $\hat{\varphi}$ (cf. la figure 1 pour un exemple), définie sur $[y_i, x_{i+1}]$, de classe C^1 et qui se raccorde de façon C^1 avec la fonction φ construite sur $S_1 \cup S_2 \cup \dots \cup S_N$ (car on a imposé que les dérivées de la fonction φ soient nulles aux bords de $S_1 \cup S_2 \cup \dots \cup S_N$). De plus, $\hat{\varphi}$ diffère de \hat{v} seulement au voisinage des points de discontinuités de cette dernière, plus précisément sur un ensemble dont la μ -mesure

est majorée par $\frac{\varepsilon^p}{(N+1)^p} + \mu(B_\mu^p \cap]y_i, x_{i+1}[)$. Comme chacune des deux fonctions est majorée par C , on a l'estimation suivante :

$$\left\| (\hat{\varphi} - \hat{v})|_{]y_i, x_{i+1}[} \right\|_{p, \mu} \leq 2C \left(\frac{\varepsilon^p}{(N+1)^p} + \mu(B_\mu^p \cap]y_i, x_{i+1}[) \right)^{\frac{1}{p}} \quad (3.54)$$

$$\leq 4C \left(\frac{\varepsilon}{N+1} + \mu(B_\mu^p \cap]y_i, x_{i+1}[)^{\frac{1}{p}} \right). \quad (3.55)$$

Combinée au fait que $\| (u - v)|_{]y_i, x_{i+1}[} \|_{p, \mu} \leq \frac{\varepsilon}{N+1}$ et $\|\hat{v} - v\|_{p, \mu} \leq 2C\mu(]y_i, x_{i+1}[\cap B_\mu^p)^{\frac{1}{p}}$, cette majoration conduit à

$$\left\| (\hat{\varphi} - u)|_{]y_i, x_{i+1}[} \right\|_{p, \mu} \leq (4C + 1) \frac{\varepsilon}{N+1} + 6C\mu(]y_i, x_{i+1}[\cap B_\mu^p)^{\frac{1}{p}}. \quad (3.56)$$

Quant à la dérivée, puisque cette dernière n'est non nulle qu'au voisinage des points de discontinuité de \hat{v} , et au vu des majorations précédentes,

$$\left\| \hat{\varphi}'|_{]y_i, x_{i+1}[} \right\|_{p, \mu_a} \leq \frac{\varepsilon}{N+1}. \quad (3.57)$$

Sens réciproque, l'approximation de u est bien globale. On effectue la construction ci-dessus pour chacun des intervalles $]y_i, x_{i+1}[$ pour $i \in \{0, 1, \dots, N\}$. En raccordant avec la fonction déjà définie sur $S_1 \cup S_2 \cup \dots \cup S_N$, on obtient une fonction φ , définie sur I et de classe C^1 . En combinant l'estimation (3.51) et les estimations (3.56) pour $i \in \{0, 1, \dots, N\}$, et puisque ces dernières portent sur des intervalles deux à deux disjoints, on obtient

$$\|u - \varphi\|_{p, \mu} \leq \varepsilon + (4C + 1)\varepsilon + 6C\mu(B_\mu^p \setminus (S_1 \cup S_2 \cup \dots \cup S_N))^{\frac{1}{p}} \quad (3.58)$$

$$\leq (10C + 2)\varepsilon. \quad (3.59)$$

Cela montre que φ approche u aussi près que voulu dans $L_\mu^p(I)$. Quant à la dérivée, les majorations (3.52) et (3.57) montrent que

$$\|\varphi'\|_{p, \mu_a} \leq \|u' \mathbb{1}_{I \setminus M_\mu^p}\|_{p, \mu_a} + 1 + \varepsilon, \quad (3.60)$$

ce qui est une majoration uniforme lorsque $\varepsilon \rightarrow 0$ et suffit pour conclure que $u \in W_\mu^{1,p}(I)$.

Sens réciproque, le cas où u est n'est pas bornée. Dans le cas général, on se ramène

au cas d'une fonction bornée en posant, pour $n \in \mathbb{N}^*$,

$$u_n := n \arctan \left(\frac{u}{n} \right), \quad (3.61)$$

qui est une fonction bornée et dont la dérivée sur $I \setminus M_\mu^p$ est bornée par celle de u , donc $u_n \in W_\mu^{1,p}(I)$. De plus, la suite $(u_n)_{n \in \mathbb{N}^*}$ converge simplement vers u en étant bornée par u et la suite $\left(\left(u_n|_{I \setminus M_\mu^p} \right)' \right)_{n \in \mathbb{N}^*}$ est bornée par $\left(u|_{I \setminus M_\mu^p} \right)'$, le théorème de convergence dominée assure alors que $(u_n)_{n \in \mathbb{N}^*}$ converge vers u dans $L_\mu^p(I)$ tout en conservant une dérivée bornée dans $L_{\mu_a}^p(I \setminus M_\mu^p)$, ce qui suffit pour montrer que $u \in W_\mu^{1,p}(I)$. \square

À l'aide de ce théorème, on peut décrire l'espace $W_\mu^{1,p}(I)$ pour les deux exemples donnés précédemment.

Premier exemple (page 28). Une fonction $u \in L_\mu^2(]-1, 1[)$ appartient à $W_\mu^{1,2}(]-1, 1[)$ si et seulement si elle possède un représentant continu sur $]-1, 1[$ et si sa dérivée, bien définie sur $]-1, 1[$ au sens des distributions, appartient à $L_{\mu_a}^2(]-1, 1[)$. Dans ce cas, sa norme est donnée par

$$\|u\|_{1,2,\mu}^2 = \int_{-1}^1 |u(x)|^2 dx + |u(0)|^2 + \int_{-1}^1 |u'(x)|^2 dx. \quad (3.62)$$

Deuxième exemple (page 29). Une fonction $u \in L_\mu^2(]-1, 1[)$ appartient à $W_\mu^{1,2}(]-1, 1[)$ si et seulement si elle possède un représentant continu sur $]-1, 0[$ et sur $[0, 1[$ et si sa dérivée, bien définie sur $]-1, 0[\cup]0, 1[$ au sens des distributions, appartient à $L_{\mu_a}^2(]-1, 0[\cup]0, 1[)$. Dans ce cas, sa norme est donnée par

$$\|u\|_{1,2,\mu}^2 = \int_{-1}^1 |u(x)|^2 f(x) dx + |u(0)|^2 + \int_{-1}^1 |u'(x)|^2 f(x) dx. \quad (3.63)$$

4 Application au transport optimal

4.1 Le transport optimal avec pénalisation en gradient

Explicitons les hypothèses techniques que l'on fait, la motivation du problème de transport optimal avec pénalisation en gradient ayant déjà été faite dans l'introduction.

- Comme précédemment, μ désigne une mesure de Radon positive finie sur un domaine *borné* Ω de \mathbf{R}^d . On se donne une autre mesure de Radon ν positive sur un autre domaine *borné* Ω' de \mathbf{R}^d telle que $\mu(\Omega) = \nu(\Omega')$.
- On se donne une fonction de coût $c : \Omega \times \Omega' \rightarrow \mathbf{R}_+$ positive, *continue et bornée*.

On note $W_\mu^{1,p}(\Omega, \mathbf{R}^d)$ l'espace $(W_\mu^{1,p}(\Omega))^d$ et l'on peut alors donner un sens à l'équation (1.5) discutée dans l'introduction, à savoir

$$\inf \left\{ \int_{\Omega} c(x, T(x)) d\mu(x) + \|D_\mu T\|_{p,\mu}^p; T \in W_\mu^{1,p}(\Omega, \mathbf{R}^d) \text{ et } T\#\mu = \nu \right\}. \quad (4.1)$$

On dit que le problème de Monge avec pénalisation en gradient consistant à envoyer μ sur ν pour le coût c possède une solution si l'infimum dans l'équation (4.1) est un minimum.

Fixons quelques notations : on note $\mathcal{J}_{\mu,G}$ la fonctionnelle définie sur $W_\mu^{1,p}(\Omega, \mathbf{R}^d)$ par

$$\mathcal{J}_{\mu,G}(T) := \int_{\Omega} c(x, T(x)) d\mu(x) + \|D_\mu T\|_{p,\mu}^p, \quad (4.2)$$

tandis que \mathcal{J}_μ , sans l'indice G , est définie pour toute fonction mesurable et correspond au coût de transport sans le terme en gradient :

$$\mathcal{J}_\mu(T) := \int_{\Omega} c(x, T(x)) d\mu(x). \quad (4.3)$$

Quant à la notation $J_G(\mu, \nu)$ elle désigne le coût optimal, c'est-à-dire

$$J_G(\mu, \nu) := \inf \{ \mathcal{J}_{\mu,G}(T); T \in W_\mu^{1,p}(\Omega, \mathbf{R}^d) \text{ et } T\#\mu = \nu \}, \quad (4.4)$$

et de même

$$J(\mu, \nu) := \inf \{ \mathcal{J}_\mu(T); T \text{ mesurable et } T\#\mu = \nu \}. \quad (4.5)$$

Dire que le problème de Monge consistant à envoyer μ sur ν pour le coût c possède une solution signifie que l'infimum dans (4.5) est un minimum.

Plusieurs questions, toutes abordées dans la thèse de Louet [12], surgissent :

- Savoir si le problème de Monge avec pénalisation en gradient possède une solution.
- Si la réponse à la première question est positive, on peut chercher à caractériser les applications de transport optimales. Par exemple lorsque $d \geq 2$, on peut perturber légèrement une application $T \in W_{\mu}^{1,p}(\Omega, \mathbf{R}^d)$ tout en conservant la contrainte $T\#\mu = \nu$, ce qui permet d'obtenir une équation de type Euler-Lagrange.
- Si la réponse à la première question est positive, on peut considérer une pénalisation en gradient de plus en plus petite c'est-à-dire que l'on définit

$$\mathcal{J}_{\mu,G,\varepsilon}(T) := \int_{\Omega} c(x, T(x)) d\mu(x) + \varepsilon \|D_{\mu}T\|_{p,\mu}^p, \quad (4.6)$$

et l'on étudie le comportement de cette fonctionnelle (au sens de la Γ -convergence) lorsque $\varepsilon \rightarrow 0$. Dans certains cas on peut prouver qu'il y a convergence vers \mathcal{J}_{μ} , la question est alors de savoir à quelle vitesse elle a lieu.

On se contentera dans ce mémoire d'aborder la première question en montrant deux résultats d'existence différents : dans le cas où les mesures sont « sympathiques », c'est-à-dire à densité par rapport à la mesure de Lebesgue et avec des densités régulières, et dans le cas de la dimension un mais pour des mesures quelconques. Comme on le verra, la différence avec le transport optimal « classique » est que la méthode directe du calcul des variations fonctionne mieux, puisque l'on peut contrôler les dérivées d'une suite minimisante de transports, ce qui permet d'espérer une convergence presque partout et de passer à la limite la contrainte sur la mesure image.

Commençons par le cas des mesures « sympathiques » en définissant tout d'abord ce qu'est un ouvert « sympathique »

Définition 4.1. *Un domaine $\Omega \subset \mathbf{R}^d$ est dit C^1 -polaire s'il existe un point $x_0 \in \mathbf{R}^d$ et une fonction $\gamma : \mathbf{S}^{d-1} \rightarrow]0, +\infty[$ de classe C^1 tels que*

$$\Omega = \left\{ x \in \mathbf{R}^d; |x - x_0| < \gamma \left(\frac{x - x_0}{|x - x_0|} \right) \right\}. \quad (4.7)$$

En particulier un tel domaine est étoilé et possède une frontière de classe C^1 . À partir de cette définition, on peut énoncer le résultat suivant :

Théorème 4.2. *Soient Ω et Ω' deux domaines C^1 -polaires, f et g deux fonctions définies sur respectivement Ω et Ω' de classe $C^{0,\alpha}$ pour un $\alpha \in]0, 1[$ et uniformément minorées par une constante strictement positive. On définit $\mu := f \cdot \mathcal{L}$ et $\nu := g \cdot \mathcal{L}$ et l'on suppose*

$\mu(\Omega) = \nu(\Omega')$. Alors le problème de Monge avec pénalisation en gradient consistant à envoyer μ sur ν pour le coût c possède une solution.

Démonstration. Les hypothèses techniques faites sur les domaines et sur la régularité de f et g permettent d'assurer qu'il existe au moins un transport T lipschitzien tel que $T\#\mu = \nu$: on pourra trouver une démonstration de ce résultat dans [12, Théorème 3.2.1]. Or, comme f est minorée par une constante strictement positive, $f^{-\frac{1}{p-1}} \in L^1_{\text{loc}}(\Omega)$ et l'on peut donc définir les dérivées au sens des distributions, le caractère lipschitzien de T assure que sa différentielle est bornée \mathcal{L} -presque partout donc μ -presque partout, ce qui montre que $T \in W^1_{\mu}(\Omega, \mathbf{R}^d)$.

On peut alors utiliser la méthode directe du calcul des variations : comme $\mathcal{J}_{\mu,G} \geq 0$, on peut considérer une suite $(T_n)_{n \in \mathbf{N}}$ d'éléments de $W^1_{\mu}(\Omega, \mathbf{R}^d)$ telle que $T_n\#\mu = \nu$ pour tout $n \in \mathbf{N}$ et $\mathcal{J}_{\mu,G}(T_n) \rightarrow J_G(\mu, \nu)$ lorsque $n \rightarrow +\infty$. Le fait que la suite $(\mathcal{J}_{\mu,G}(T_n))_{n \in \mathbf{N}}$ soit bornée implique que la suite $(\|DT_n\|_{p,\mu})_{n \in \mathbf{N}}$ est bornée donc $(\|DT_n\|_p)_{n \in \mathbf{N}}$ aussi puisque f est minorée par une constante strictement positive. Comme de plus pour tout $n \in \mathbf{N}$, T_n est à valeurs dans Ω' qui est borné, on voit que $(T_n)_{n \in \mathbf{N}}$ est bornée dans $W^{1,p}(\Omega, \mathbf{R}^d)$.

Les injections de Sobolev garantissent qu'à extraction près (on ne renomme pas les suites extraites), on peut supposer que $(T_n)_{n \in \mathbf{N}}$ converge fortement dans $L^p(\Omega, \mathbf{R}^d)$, faiblement dans $W^{1,p}(\Omega, \mathbf{R}^d)$ et \mathcal{L} -presque partout vers $T \in W^{1,p}(\Omega, \mathbf{R}^d)$. Mais comme f est de classe $C^{0,\alpha}$ elle est bornée, donc $T \in W^1_{\mu}(\Omega, \mathbf{R}^d)$ et la convergence a lieu faiblement dans $W^1_{\mu}(\Omega, \mathbf{R}^d)$.

La convergence \mathcal{L} -presque partout permet de passer à la limite dans la relation $T_n\#\mu = \nu$ lorsque $n \rightarrow +\infty$, de sorte que $T\#\mu = \nu$, et la convergence faible ainsi que \mathcal{L} -presque partout permettent d'écrire que

$$\mathcal{J}_{\mu,G}(T) \leq \liminf_{n \rightarrow +\infty} \mathcal{J}_{\mu,G}(T_n) \tag{4.8}$$

$$= J_G(\mu, \nu) \tag{4.9}$$

ce qui conclut la preuve. □

On voit que ce résultat est assez grossier : on se donne des hypothèses suffisamment restrictives pour pouvoir se ramener au cas de $W^{1,p}(\Omega, \mathbf{R}^d)$ et utiliser les injections de Sobolev. En effet, le point clé est que ces dernières garantissent une convergence \mathcal{L} -presque partout qui permet de passer à la limite la contrainte sur la mesure image.

4.2 Un résultat d'existence en dimension un

Dans le cas de la dimension un, grâce à la connaissance précise de l'espace $W_\mu^{1,p}$, on arrive à obtenir un résultat assez fin pour l'existence d'une solution au problème de Monge avec pénalisation en gradient. On se donne donc I et I' deux intervalles ouverts bornés de \mathbf{R} , μ et ν deux mesures de Radon positives de même masse sur respectivement I et I' . Louet a déjà énoncé un résultat d'existence dans le cas d'un coût quadratique [12, Théorème 1.3.2], on va montrer comment on peut l'étendre à n'importe quel type de coût.

On utilise les mêmes notations que dans la section 3.4 : μ_a désigne la partie absolument continue par rapport à \mathcal{L} de μ , et f est la densité de μ_a par rapport à \mathcal{L} . On note M_μ^p le complémentaire du plus grand ouvert sur lequel $f^{-\frac{1}{p-1}}$ est intégrable. La lettre \mathcal{S} désigne l'ensemble des segments non réduits à des singletons sur lesquels $f^{-\frac{1}{p-1}}$ est intégrable et $B_\mu^p := \bigcup_{S \in \mathcal{S}} S$ (cf. page 31). On aura besoin l'ensemble (au plus dénombrable) des atomes de μ sur la partie $I \setminus B_\mu^p$, on le not \mathcal{A} .

On définit alors la partie « régulière » de μ par $\mu_r := \mu \llcorner B_\mu^p$, la partie « atomique » par

$$\mu_{\text{atm}} := \sum_{x \in \mathcal{A}} \mu(\{x\}) \delta_x, \quad (4.10)$$

et la partie « irrégulière » par

$$\mu_i := \mu \llcorner (I \setminus B_\mu^p) - \mu_{\text{atm}}. \quad (4.11)$$

Cette dernière est par construction une mesure sans atomes, et l'on a la décomposition $\mu = \mu_r + \mu_{\text{atm}} + \mu_i$. Enfin, rappelons qu'au vu de l'expression de D_μ (théorème 3.18), la fonction de coût prend la forme

$$\mathcal{J}_{\mu,G}(T) = \int_I c(x, T(x)) d\mu(x) + \int_{I \setminus M_\mu^p} |T'(x)|^p d\mu_a(x) \quad (4.12)$$

pour $T \in W_\mu^{1,p}(I)$. On a maintenant tous les éléments nécessaires pour énoncer et démontrer le résultat suivant :

Théorème 4.3. *On suppose qu'il existe au moins un transport $T \in W_\mu^{1,p}(I)$ tel que $T\#\mu = \nu$. On suppose aussi que pour toute mesure de Radon ν_i positive vérifiant $\nu_i \leq \nu$ et $\nu_i(I') = \mu_i(I)$, le problème de Monge consistant à envoyer μ_i sur ν_i de façon optimale pour le coût c possède une solution. Alors le problème de Monge avec pénalisation en gradient consistant à envoyer μ sur ν pour le coût c possède une solution.*

Remarque. L'existence d'au moins une application $T \in W_\mu^{1,p}(I)$ transportant μ sur ν n'est pas forcément facile à vérifier avec la simple donnée des mesures μ et ν , on ne discutera cependant pas cette hypothèse ici et on la supposera vérifiée.

Remarque. On se ramène à une question de transport optimal *sans* pénalisation en gradient, à savoir l'hypothèse que le problème de Monge entre μ_i et ν_i possède une solution. La mesure μ_i ne possède d'ailleurs pas d'atomes, c'est une propriété souvent utile pour prouver l'existence d'une solution au problème de Monge.

Démonstration. On applique la méthode directe du calcul des variations : comme la fonctionnelle $\mathcal{J}_{\mu,G}$ est positive, $J_G(\mu, \nu) \geq 0$ et donc on peut trouver une suite $(T_n)_{n \in \mathbb{N}}$ d'éléments de $W_\mu^{1,p}(I)$ telle que $T_n \# \mu = \nu$ pour tout $n \in \mathbb{N}$ et $\mathcal{J}_{\mu,G}(T_n) \rightarrow J_G(\mu, \nu)$ lorsque $n \rightarrow +\infty$. L'idée est la suivante : sur la partie B_μ^p , on va pouvoir extraire de $(T_n)_{n \in \mathbb{N}}$ une suite qui converge simplement ce qui permet de passer à la limite la contrainte sur la mesure image, tandis que sur $I \setminus B_\mu^p$, le terme en gradient n'intervient pas donc le transport optimal pour le problème de Monge *sans* pénalisation en gradient convient.

Dans la suite de la preuve on ne renomme pas les sous-suites lorsque l'on effectue des extractions.

Construction de l'application de transport optimale. On note, pour $n \in \mathbb{N}$,

$$\begin{cases} \nu_{r,n} & := T_n \# \mu_r \\ \nu_{\text{atm},n} & := T_n \# \mu_{\text{atm}} \\ \nu_{i,n} & := T_n \# \mu_i \end{cases} \quad (4.13)$$

Les suites $(\nu_{r,n})_{n \in \mathbb{N}}$, $(\nu_{\text{atm},n})_{n \in \mathbb{N}}$ et $(\nu_{i,n})_{n \in \mathbb{N}}$ sont bornées et à support inclus dans \bar{I}' qui est compact, quitte à extraire une sous-suite on peut supposer qu'elles convergent faiblement au sens des mesures (c'est-à-dire faiblement au sens de la dualité avec les fonctions continues et bornées sur \mathbb{R}) vers des mesures positives ν_r , ν_{atm} et ν_i . Les relations $\mu_r(I) = \nu_r(I')$, $\mu_{\text{atm}}(I) = \nu_{\text{atm}}(I')$, $\mu_i(I) = \nu_i(I')$ et $\nu_r + \nu_{\text{atm}} + \nu_i = \nu$ découlent directement de la convergence faible.

On choisit une application de transport $T^i : I \rightarrow I'$, optimale pour le problème de Monge (sans pénalisation en gradient) entre μ_i et ν_i pour le coût c , une telle application existe par hypothèse.

Pour $x \in \mathcal{A}$, la suite $(T_n(x))_{n \in \mathbb{N}}$ est à valeurs dans I' donc bornée, à extraction près on peut supposer qu'elle converge. Via un procédé d'extraction diagonal on peut supposer que c'est le cas pour chaque $x \in \mathcal{A}$. On note alors T^{atm} la fonction définie sur \mathcal{A} , qui à un élément x de \mathcal{A} associe la limite de la suite $(T_n(x))_{n \in \mathbb{N}}$.

Pour la partie sur B_μ^p , on remarque que pour tout $n \in \mathbb{N}$, la fonction T_n est à valeurs dans I' donc elle est bornée, de plus puisque la suite $(\mathcal{J}_{\mu,G}(T_n))_{n \in \mathbb{N}}$ est bornée, $(T_n)_{n \in \mathbb{N}}$ est bornée dans $W_\mu^{1,p}(I)$. On trouve alors un ensemble dénombrable \mathcal{S}_d inclus dans \mathcal{S} tel que tout segment $S \in \mathcal{S}$ soit inclus dans un segment de \mathcal{S}_d . Si on fixe $S \in \mathcal{S}_d$, grâce au lemme 3.19 (ou plus précisément la majoration (3.49)), on voit que la suite $(T_n|_S)_{n \in \mathbb{N}}$ est bornée et équicontinue, le théorème d'Ascoli garantit qu'à extraction près elle converge uniformément sur S vers une fonction continue. Via un procédé d'extraction diagonal, on peut répéter cette opération pour chaque segment de \mathcal{S}_d , on obtient alors une fonction T^r , définie sur B_μ^p , et telle que $(T_n)_{n \in \mathbb{N}}$ converge uniformément vers T^r sur chaque segment de \mathcal{S}_d . En particulier l'application T^r est continue sur chaque segment de \mathcal{S} et la convergence a lieu simplement sur B_μ^p .

Si $x \in I$, on définit alors

$$T(x) := \begin{cases} T^r(x) & \text{si } x \in B_\mu^p \\ T^{\text{atm}}(x) & \text{si } x \in \mathcal{A} \\ T^i(x) & \text{sinon} \end{cases} . \quad (4.14)$$

L'application T appartient à $W_\mu^{1,p}(I)$. Par construction T est bien continue sur tous les segments de \mathcal{S} . De plus, comme la suite $(\mathbb{1}_{B_\mu^p} T_n)_{n \in \mathbb{N}}$ est bornée dans $W_{\mu_a}^{1,p}(I)$ qui est un espace de Banach réflexif (c'est le théorème 2.5), on peut supposer, quitte à extraire une sous-suite, qu'elle converge faiblement. Mais sur B_μ^p , à cause de l'injection continue de $L_{\mu_a}^p$ dans L_{loc}^1 , la convergence faible implique la convergence au sens des distributions donc T^r coïncide sur B_μ^p avec la limite faible de la suite $(\mathbb{1}_{B_\mu^p} T_n)_{n \in \mathbb{N}}$. De plus l'inégalité

$$\int_{B_\mu^p} |T|^p d\mu_a + \int_{I \setminus M_\mu^p} |T'|^p d\mu_a = \|\mathbb{1}_{B_\mu^p} T\|_{1,p,\mu_a}^p \quad (4.15)$$

$$\leq \liminf_{n \rightarrow +\infty} \|\mathbb{1}_{B_\mu^p} T_n\|_{1,p,\mu_a}^p \quad (4.16)$$

$$= \liminf_{n \rightarrow +\infty} \left(\int_{B_\mu^p} |T_n|^p d\mu_a + \int_{I \setminus M_\mu^p} |T'_n|^p d\mu_a \right), \quad (4.17)$$

combinée au fait que

$$\lim_{n \rightarrow +\infty} \int_{B_\mu^p} |T_n|^p d\mu_a = \int_{B_\mu^p} |T|^p d\mu_a \quad (4.18)$$

puisque $(T_n)_{n \in \mathbb{N}}$ est bornée et converge simplement vers T sur B_μ^p , permet d'écrire que

$$\int_{I \setminus M_\mu^p} |T'|^p d\mu_a \leq \liminf_{n \rightarrow +\infty} \int_{I \setminus M_\mu^p} |T'_n|^p d\mu_a. \quad (4.19)$$

Cela montre que $T' \in L^p_{\mu_a}(I \setminus M^p_\mu)$. Puisque de plus T est mesurable sur $I \setminus B^p_\mu$ et à valeurs dans I' donc est borné, on peut conclure que $T \in W^{1,p}_\mu(I)$.

L'application T envoie μ sur ν . On sait déjà que l'application T envoie μ_i sur ν_i , il reste à vérifier qu'elle envoie $\mu_r + \mu_{\text{atm}}$ sur $\nu_r + \nu_{\text{atm}}$. Pour cela, on se donne une fonction φ , définie sur \mathbf{R} , continue et bornée et l'on remarque que

$$\int_I (\varphi \circ T) d(\mu_r + \mu_{\text{atm}}) = \int_I (\varphi \circ T - \varphi \circ T_n) d(\mu_r + \mu_{\text{atm}}) + \int_I (\varphi \circ T_n) d(\mu_r + \mu_{\text{atm}}) \quad (4.20)$$

$$= \int_I (\varphi \circ T - \varphi \circ T_n) d(\mu_r + \mu_{\text{atm}}) + \int_{I'} \varphi d(\nu_{r,n} + \nu_{\text{atm},n}), \quad (4.21)$$

où l'on a utilisé le fait que $T_n \# (\mu_r + \mu_{\text{atm}}) = \nu_{r,n} + \nu_{\text{atm},n}$. Par convergence faible, des mesures le second terme tend vers $\int \varphi d(\nu_r + \nu_{\text{atm}})$ lorsque $n \rightarrow +\infty$. Quant au premier, puisque $(T_n)_{n \in \mathbf{N}}$ converge simplement vers T sur $(B^p_\mu \cup \mathcal{A})$, le théorème de convergence dominée assure qu'il tend vers 0 lorsque $n \rightarrow +\infty$.

L'application de transport T est optimale. Pour le prouver, pour tout $n \in \mathbf{N}$, on définit l'application auxiliaire \hat{T}_n par

$$\hat{T}_n(x) := \begin{cases} T_n(x) & \text{si } x \in (B^p_\mu \cup \mathcal{A}) \\ T^i(x) & \text{sinon} \end{cases}. \quad (4.22)$$

Elle ne transporte plus μ sur ν , en revanche $(\hat{T}_n)_{n \in \mathbf{N}}$ converge simplement vers T . En conséquence, on peut passer à la limite à l'aide du théorème de convergence dominée dans le terme $\int_I c(x, \hat{T}_n(x)) d\mu(x)$, ce qui, combiné à l'inégalité (4.19), permet d'écrire

$$\mathcal{J}_{\mu,G}(T) \leq \liminf_{n \rightarrow +\infty} \mathcal{J}_{\mu,G}(\hat{T}_n). \quad (4.23)$$

On remarque ensuite que la différence entre \hat{T}_n et T_n ne porte que sur $I \setminus (B^p_\mu \cup \mathcal{A})$, donc

$$\mathcal{J}_{\mu,G}(T_n) - \mathcal{J}_{\mu,G}(\hat{T}_n) = \int_{I \setminus B^p_\mu} c(x, T_n(x)) d\mu_i(x) - \int_{I \setminus B^p_\mu} c(x, \hat{T}_n(x)) d\mu_i(x) \quad (4.24)$$

$$= \int_{I \setminus B^p_\mu} c(x, T_n(x)) d\mu_i(x) - J(\mu_i, \nu_i) \quad (4.25)$$

$$\geq J(\mu_i, \nu_{i,n}) - J(\mu_i, \nu_i), \quad (4.26)$$

où l'on a d'abord utilisé le fait que sur $I \setminus (B^p_\mu \cup \mathcal{A})$ l'application \hat{T}_n coïncide avec T^i et transporte donc μ_i sur ν_i de façon optimale, c'est-à-dire avec le coût $J(\mu_i, \nu_i)$, puis que T_n

transporte μ_i sur $\nu_{i,n}$ et ne peut le faire qu'avec un coût plus grand que $J(\mu_i, \nu_{i,n})$. Or il se trouve que c'est un résultat classique de transport optimal que J est une fonction continue de ses arguments pour la convergence faible des mesures (on pourra consulter par exemple [14, Théorème 1.6.11]), ce qui montre que le membre de droite de l'inéquation (4.26) tend vers 0 lorsque $n \rightarrow +\infty$. Cela peut se réécrire

$$\liminf_{n \rightarrow +\infty} \mathcal{J}_{\mu,G}(\hat{T}_n) \leq \liminf_{n \rightarrow +\infty} \mathcal{J}_{\mu,G}(T_n) \quad (4.27)$$

$$= J_G(\mu, \nu). \quad (4.28)$$

En utilisant aussi l'inégalité (4.23), on en déduit que $\mathcal{J}_{\mu,G}(T) \leq J_G(\mu, \nu)$ ce qui prouve l'optimalité de T . □

Exemple. On reprend le deuxième exemple de la section 3.4, il se trouve page 29. Dans ce cas $B_\mu^p = I$ donc $\mu = \mu_r$, de sorte que s'il existe au moins un transport $T \in W_\mu^{1,p}(I)$ tel que $T\#\mu = \nu$, alors le problème de Monge avec pénalisation en gradient consistant à envoyer μ sur ν pour le coût c possède une solution.

Dans le cas où $\mu_i \neq 0$, on peut se restreindre à certaines fonctions de coût pour lesquelles on sait sous quelles conditions il existe une solution au problème de Monge. Commençons par retrouver le résultat donné par Louet [12, Théorème 1.3.2] :

Théorème 4.4. *On se place dans le cas où il existe une fonction h strictement convexe et positive telle que, pour $x \in I$ et $y \in I'$, on ait $c(x, y) = h(x - y)$. On suppose qu'il existe au moins une application $T \in W_\mu^{1,p}(I)$ telle que $T\#\mu = \nu$. Alors le problème de Monge avec pénalisation en gradient consistant à envoyer μ sur ν pour le coût c possède une solution.*

Démonstration. Pour appliquer le théorème 4.3, il suffit de vérifier que pour toute mesure positive ν_i vérifiant $\nu_i \leq \nu$ et $\mu_i(I) = \nu_i(I')$, le problème de Monge consistant à envoyer μ_i sur ν_i pour le coût c possède une solution. Mais par construction μ_i ne possède pas d'atomes (ils se trouvent tous dans μ_{atm}), il suffit donc d'utiliser un résultat classique de transport optimal (voir par exemple [14, Théorème 2.2.1]). □

Traitons aussi le cas où le coût est « concave », de la forme $c(x, y) = \sqrt{|x - y|}$ par exemple.

Théorème 4.5. *On se place dans le cas où il existe une fonction h strictement concave et positive telle que, pour $x \in I$ et $y \in I'$, on ait $c(x, y) = h(|x - y|)$. On suppose que μ_i s'annule sur $\text{spt}(\nu)$ et qu'il existe au moins une application $T \in W_\mu^{1,p}(I)$ telle que $T\#\mu = \nu$. Alors le problème de Monge avec pénalisation en gradient consistant à envoyer μ sur ν pour le coût c possède une solution.*

Démonstration. Si ν_i est une mesure positive satisfaisant $\nu_i \leq \nu$ et $\mu_i(I) = \nu_i(I')$, alors μ_i s'annule sur $\text{spt}(\nu_i)$, on peut donc utiliser [7, Théorème 1.4] pour conclure à l'existence d'une solution au problème de Monge consistant à envoyer μ_i sur ν_i pour le coût c , ce qui permet d'appliquer le théorème 4.3. \square

Références

- [1] B. Beauzamy. *Introduction to Banach spaces and their geometry*. Number 86 in Notas de matemática. North-Holland, 1985.
- [2] Guy Bouchitté, Giuseppe Buttazzo, and Pierre Seppecher. Energies with respect to a measure and applications to low dimensional structures. *Calculus of Variations and Partial Differential Equations*, 5(1) :37–54, 1996.
- [3] Guy Bouchitté and Ilaria Fragalà. Second-order energies on thin structures : variational theory and non-local effects. *Journal of Functional Analysis*, 204(1) :228 – 267, 2003.
- [4] Guy Bouchitté and Michel Valadier. Integral representation of convex functionals on a space of measures. *Journal of Functional Analysis*, 80(2) :398 – 420, 1988.
- [5] G. Castaing and M. Valadier. *Convex analysis and measurable multifunctions*. Lecture notes in mathematics. 1977.
- [6] Ilaria Fragalà and Carlo Mantegazza. On some notions of tangent space to a measure. *Proceedings of the Royal Society of Edinburgh : Section A Mathematics*, 129 :331–342, 1 1999.
- [7] Wilfrid Gangbo and Robert J. McCann. The geometry of optimal transportation. *Acta Mathematica*, 177(2) :113–161, 1996.
- [8] E. Giusti. *Direct Methods in the Calculus of Variations*. World Scientific, 2003.
- [9] Piotr Hajlasz. Sobolev spaces on metric-measure spaces. *Contemporary mathematics*, 338 :173–218, 2003.
- [10] Opic Bohumír Kufner, Alois. How to define reasonably weighted Sobolev spaces. *Commentationes Mathematicae Universitatis Carolinae*, 025(3) :537–554, 1984.
- [11] L.D. Landau, E.M. Lifshitz, A.M. Kosevich, and Pitaevskii. *Theory of Elasticity*.
- [12] Jean Louet. *Problèmes de transport optimal avec pénalisation en gradient*. PhD thesis, 2014. Thèse de doctorat dirigée par Santambrogio, Filippo.
- [13] Paolo Tilli Luigi Ambrosio. *Topics on analysis in metric Spaces*. Oxford Lecture Series in Mathematics and Its Applications. 2003.

- [14] Filippo Santambrogio. *Optimal Transport for Applied Mathematicians*. 2015.
- [15] M. Valadier. Multi-applications mesurables à valeurs convexes compactes. *J. Math. Pures Appl.*, 50 :265–297, 1971.
- [16] C. Villani. *Topics in Optimal Transportation*. Graduate studies in mathematics. American Mathematical Society, 2003.
- [17] V V Zhikov. Homogenization of elasticity problems on singular structures. *Izvestiya : Mathematics*, 66(2) :299, 2002.

A Relaxée d'une fonction

Soient (X, d) un espace métrique et $F : X \rightarrow \mathbf{R}_+ \cup \{+\infty\}$ une fonction positive.

Définition A.1. La relaxée de F , notée \bar{F} , est la plus grande fonction semi-continue inférieurement, définie sur X et à valeurs dans $\mathbf{R}_+ \cup \{+\infty\}$, qui soit majorée par F .

Puisqu'un supremum de fonctions semi-continues inférieurement est semi-continu inférieurement et qu'il existe au moins une fonction semi-continue inférieurement majorée par F (à savoir la fonction nulle), \bar{F} est bien définie.

La propriété suivante justifie l'intérêt de la relaxée en calcul des variations :

Proposition A.2. On a l'égalité $\inf_X \bar{F} = \inf_X F$.

Démonstration. Par définition $\bar{F} \leq F$ donc $\inf_X \bar{F} \leq \inf_X F$. Pour prouver l'inégalité dans l'autre sens, on remarque que la fonction constante et égale à $\inf_X F$ est semi-continue inférieurement et majorée par F , donc $\bar{F} \geq \inf_X F$ d'où $\inf_X \bar{F} \geq \inf_X F$. \square

Supposons que l'on se donne une partie Y dense et une fonction F que l'on sait définir sur Y . On peut alors l'étendre à X tout entier en posant $F(x) = +\infty$ si $x \notin Y$. La relaxée \bar{F} est la seule fonction « raisonnable » qui s'approche de F et que l'on cherchera à minimiser : elle est semi-continue inférieurement et son infimum coïncide avec celui de F sur Y .

Et l'on aura aussi besoin d'une représentation maniable de la relaxée :

Proposition A.3. Si $x \in X$ on a l'égalité suivante :

$$\bar{F}(x) = \inf \left\{ \liminf_{n \rightarrow +\infty} F(x_n); x_n \rightarrow x \right\}. \quad (\text{A.1})$$

Démonstration. Appelons $G(x)$ le membre de droite de l'équation (A.1). En prenant pour $(x_n)_{n \in \mathbf{N}}$ la suite constante égale à x , on voit que $G(x) \leq F(x)$. De plus, fixons $x \in X$ et $(x_n)_{n \in \mathbf{N}}$ une suite qui converge vers x . Pour tout $n \in \mathbf{N}^*$, on peut trouver un point $y_n \in X$ tel que $d(x_n, y_n) \leq \frac{1}{n}$ et $F(y_n) \leq G(x_n) + \frac{1}{n}$ par définition de G . On vérifie alors que $y_n \rightarrow x$ lorsque $n \rightarrow +\infty$, donc $\liminf_n F(y_n) \geq G(x)$ et on en conclut que $\liminf_n G(x_n) \geq G(x)$, c'est-à-dire que G est une fonction semi-continue inférieurement. Comme on sait qu'elle est majorée par F , on peut en déduire que $G \leq \bar{F}$.

Pour prouver l'inégalité dans l'autre sens, on se fixe $x \in X$ et une suite $(x_n)_{n \in \mathbf{N}}$ qui converge vers x . Puisque $\bar{F} \leq F$, on a l'inégalité $\liminf_n \bar{F}(x_n) \leq \liminf_n F(x_n)$, donc par semi-continuité inférieure $\bar{F}(x) \leq \liminf_n F(x_n)$. En passant à l'infimum sur le membre de droite, on obtient bien $\bar{F}(x) \leq G(x)$. \square