DYNAMIC FORMULATION OF OPTIMAL TRANSPORT PROBLEMS.

C. JIMENEZ

Abstract. We consider the classical Monge-Kantorovich transport problem with a general cost $c(x, y) = F(y - x)$ where $F : \mathbb{R}^d \to R^+$ is a convex function and our aim is to characterize the dual optimal potential as the solution of a system of partial differential equation.
Such a characterization has been given in smooth case by L.Evans and W. Gangbo [16] for F being the Euclidian norm and by Y. Brenier [5] in the case where $F = |\cdot|^p$ with $p > 1$. We generalize these results to the case of general F and singular transported measures in the spirit of a previous work by G. Bouchitté and G. Buttazzo [7] and by adapting Y. Brenier's dynamic formulation.

Keywords: Wasserstein distance, optimal transport map, measure functionals, duality, tangential gradient, partial differential equations.

1. Introduction

In this paper, we deal with the following problem introduced by L.V. Kantorovich (see [22]):

$$(\mathcal{P}) \qquad \min_{\gamma \in \mathcal{P}(\Omega\times\Omega)} \left\{ \int_{\Omega\times\Omega} F(y-x)\ d\gamma(x,y)\ :\ \pi_1^\sharp\gamma = f_0,\ \pi_2^\sharp\gamma = f_1 \right\}$$

where f_0 and f_1 are two fixed probability measures, Ω is the closure of a bounded open subset of $\mathbb{R}^d$ and F is a convex cost (for example: $F(y - x) = |x - y|^p$, $p \geq 1$). We denote by $\pi_1^\sharp\gamma$ and $\pi_2^\sharp\gamma$ the marginals of γ.
This problem, which is central in optimal transport theory, has numerous applications in economics (see for example [12], [13]), mechanics (see for example [9]), signal theory (see for example [6] and [21]). Recently, many papers have been published around this problem, this interest is motivated by its relation with various mathematical areas like partial differential equations, geometry, probability theory (see [23])...
We focus on the link between transport theory and partial differential equations. This link has first been enlightened by L. Evans and W. Gangbo (see [16]) when f_0 and f_1 are regular measures and F is positively 1−homogeneous. G. Bouchitté and G. Buttazzo have generalized their result (see [7]) to the case where no regularity assumption is made on f_0 and f_1. The proof of G. Bouchitté and G. Buttazzo is based on the fact that, when F is 1−homogeneous, $(\mathcal{P})$ can be viewed as the dual formulation of an optimization problem on Lipschitz functions with a convex gradient constraint:

$$(\mathcal{P}^*) \qquad \sup \left\{ \int_\Omega u(x)\ d(f_1 - f_0)(x) :\ u \in \mathrm{Lip}(\Omega),\ \nabla u(x) \in C \text{ a.e. } x \right\}$$

$$C = \{x^* : < x, x^* > -F(x) \leq 0 \quad \forall x\}.$$

Again, by duality, this problem is equivalent to a third one:

$$(\tilde{\mathcal{P}}) \quad \inf\left\{\int F(\lambda):\ \lambda \in \mathcal{M}_b(\mathbb{R}^d,\mathbb{R}^d),\ \mathrm{spt}(\lambda)\subset\Omega,\ -\mathrm{div}(\lambda) = f_1 - f_0 \text{ on } \mathbb{R}^d\right\}$$

where $\lambda \mapsto \int F(\lambda)$ is defined by

$$F(\lambda) = \int_\Omega F\left(\frac{d\lambda}{d|\lambda|}(x)\right)\ d|\lambda|(x).$$

The constraint should be taken in the sense of distribution.
The crucial point of the proof is that any optimal solution u of $(\mathcal{P}^*)$ (called transport density) and the optimal vector measure $\lambda = \sigma\mu$ (μ a positive measure and $\sigma \in \mathrm{L}^1_\mu(\Omega,\mathbb{R}^d)$) of $(\tilde{\mathcal{P}})$ satisfy the following equality:

$$\int_\Omega u(x)\ d(f_1 - f_0)(x) = \int F(\sigma(x))\ d\mu(x).$$

Considering the constraint on λ , if u is regular, an integration by parts is possible and leads to:

$$\int \nabla u(x)\cdot\sigma(x)\ d\mu(x) = \int F(\sigma(x))\ d\mu(x). \tag{1.1}$$

Note that $F = \chi_C^*$ where $\chi_C(x) := 0$ if $x \in C$, $+\infty$ otherwise. Then, as $\nabla u(x) \in C$ a.e, (1.1) implies

$$\nabla u(x)\cdot\sigma(x) = F(\sigma(x)) \quad \mu-\text{a.e.} \tag{1.2}$$

If $F = |\cdot|$ and $\sigma = \frac{d\lambda}{d|\lambda|}(x)$, $\mu = |\lambda|$, (1.2) can be rewritten as

$$|\nabla u(x)| = 1, \quad \nabla u(x) = \sigma(x) \quad \mu-\text{a.e.}$$

The problem is that u is not regular in general cases. The integration by part made to reach (1.1) cannot be done in the classical sense but is still possible replacing ∇u by the tangential gradient $\nabla_\mu u$ (see [8] and section 4 of this article). The theorem that was finally proved is (in case $F = |\cdot|$):

Theorem 1.1. *(G. Bouchitté and G. Buttazzo)*
Let (u,λ) be any solutions of $(\mathcal{P}^*)$ and $(\tilde{\mathcal{P}})$. Then, (u,σ,μ) where $\mu = |\lambda|$, $\sigma = \frac{d\lambda}{d|\lambda|}$ satisfies the following system (Monge-Kantorovich equation):

$$(MK) \quad \begin{cases} \nabla_\mu u = \sigma \quad \mu-\text{a.e.} \\ -\mathrm{div}((\nabla_\mu u)\mu) = f_1 - f_0 \quad \textit{in the sense of distributions in } \mathbb{R}^d, \\ |\nabla_\mu u| = 1 \quad \mu-\text{a.e.} \\ u \in \mathrm{Lip}_1(\Omega). \end{cases}$$

Conversely, if (u,σ,μ) satisfies (MK), then u is solution of $(\mathcal{P}^*)$ and $\lambda = \sigma\mu$ is solution of $(\tilde{\mathcal{P}})$.

Remark 1.2. The equation (MK) has first been established by L. Evans and W. Gangbo ([16]) in case f_1, f_0 are absolutely continuous with respect to the Lebesgue measure and of Lipschitz density. In this case $\mu = a(x)\ dx$ with $a \in \mathrm{L}^\infty(\Omega)$ and the eikonal equation $|\nabla_\mu u(x)| = 1$ $\mu-$a.e.x becomes $|\nabla u| = 1$ almost everywhere on $\{a > 0\}$.

In the case where F is not 1−homogeneous, the dual formulation of $(\mathcal{P})$ does not involve only one application $u \in \mathrm{Lip}(\Omega)$ but two applications:

$$(\mathcal{P}^*) \quad \sup_{u,v} \left\{ \int v(x)\ df_0(x) + \int\ u(x)\ df_1(x) :\ u(x) + v(y) \leq F(y-x) \quad \forall\ x, y \in \Omega \right\}$$

(see for example [26]). This formulation cannot be linked in a direct way to a problem similar to $(\tilde{\mathcal{P}})$.
Nevertheless, assuming F is superlinear, Y. Brenier (see [5]) proved the equality between the minimum of $(\mathcal{P})$ and the infimum of the following problem:

$$\inf \left\{ \int F\left(\frac{d\lambda}{d\rho}\right) d\rho(x,t) :\ -\frac{\partial \rho}{\partial t} - \mathrm{div}(\lambda) = f_1 \otimes \delta_1 - f_0 \otimes \delta_0 \right\}. \tag{1.3}$$

Notice the introduction of the time variable t.

Y. Brenier's results where generalized to manifolds with a superlinear length cost by P. Bernard and B. Buffoni ([4]), and L. Granieri ([18]).

To understand how (1.3) was introduced, take the case $f_0 = \delta_{X_0}$, $f_1 = \delta_{X_1}$ with $X_1,\ X_2 \in \mathbb{R}^d$. Let us consider all regular curves s joining X_0 to X_1 ($s(0) = X_0$, $s(1) = X_1$) and the following associated measures:

$$\rho(x,t) = \delta_{s(t)} \otimes 1\!\!1_{[0,1]}(t)\ dt,$$
$$\lambda = \dot{s}(t)\delta_{s(t)} \otimes 1\!\!1_{[0,1]}(t)\ dt,$$

the assumption satisfied by (λ, ρ) is in this case is

$$-\frac{\partial \rho}{\partial t} - \mathrm{div}(\lambda) = f_1 \otimes \delta_1 - f_0 \otimes \delta_0.$$

Moreover it holds:

$$\begin{aligned} \inf(\mathcal{P}) &= F(X_1 - X_0) \\ &= \inf \left\{ \int_0^1 F(\dot{s}(t))\ dt :\ s(0) = X_0,\ s(1) = X_1 \right\} \\ &= \inf \left\{ \int F\left(\frac{d\lambda}{d\rho}\right) d\rho(x,t) :\ -\frac{\partial \rho}{\partial t} - \mathrm{div}(\lambda) = f_1 \otimes \delta_1 - f_0 \otimes \delta_0 \right\}. \end{aligned}$$

Generalizing this idea to any probabilities f_0 and f_1 leads to (1.3).

In this paper, we give a new approach that allows to introduce formulation like $(\mathcal{P}^*)$ and $(\tilde{\mathcal{P}})$ in both cases (F being 1−homogeneous or superlinear), more precisely, we prove:

Theorem 1.3.

$$\min(\mathcal{P}) = \max(\mathcal{Q}^*) = \min(\mathcal{Q}),$$

where $(\mathcal{Q}^*)$ *and* $(\mathcal{Q})$ *are defined by:*

$(\mathcal{Q}^*)$

$$\begin{aligned} \sup \big\{ &< f_1 \otimes \delta_1, \psi > \ - < f_0 \otimes \delta_0, \psi >:\ \psi\ \in \mathrm{Lip}(\Omega \times [0,1]), \\ &\partial_t \psi(x,t) + F^*(\nabla_x \psi(x,t)) \leq 0 \text{ a.e.}(x,t) \in \Omega \times [0,1] \big\} \end{aligned}$$

$(\mathcal{Q})$

$$\min\Big\{\int H(\chi) : \chi\in\mathcal{M}_b(\mathbb{R}^{d+1},\mathbb{R}^{d+1}),\quad \mathrm{spt}(\chi)\subset\Omega\times[0,1],\\ -\mathrm{div}_{x,t}(\chi)=f_1\otimes\delta_1-f_0\otimes\delta_0 \text{ on } \mathbb{R}^{d+1}\Big\}$$

where $H(x,t)$ is the perspective function of F.

Then, we make the link between transport theory and partial differential equations and get the following result:

Theorem 1.4. *Let $(\psi,\mu,\sigma)\in \mathrm{Lip}(\Omega\times[0,1])\times\mathcal{M}_b^+(\mathbb{R}^{d+1})\times \mathrm{L}^1_\mu(\Omega\times[0,1])^{d+1}$. If ψ and $\sigma\mu$ are solutions of $(\mathcal{Q}^*)$ and $(\mathcal{Q})$, then, the following system is satisfied:*

$$(MK_t)\quad \begin{cases} a)\ \partial_t\psi(x,t)+F^*(\nabla_x\psi(x,t))\le 0 \text{ a.e.}(x,t)\in\Omega\times[0,1],\\ b)\ -\mathrm{div}_{x,t}(\sigma\mu)=f_1\otimes\delta_1-f_0\otimes\delta_0,\quad \text{in}\quad \mathbb{R}^{d+1}\\ c)\ H(\sigma(x,t))=\sigma(x,t)\cdot D_\mu\psi(x,t)\quad \mu-\text{a.e.}(x,t)\in\Omega\times[0,1].\end{cases}$$

Conversely, if (ψ,μ,σ) satisfy (MK_t), then ψ is a solution of $(\mathcal{Q}^*)$ and $\sigma\mu$ is a solution of $(\mathcal{Q})$.

In section 2, we introduce the new cost H involving the time variable, this cost is $1-$homogeneous even if F is not. An alternative formulation of $(\mathcal{P})$ is given using the cost H.
In section 3, Theorem 1.3 is proved using classical duality and Hamilton Jacobi theory. To go further and write extremality conditions linking the solutions of $(\mathcal{Q}^*)$ and $(\tilde{\mathcal{Q}})$, it is necessary to introduce the tangent space to a measure and the tangential gradient with respect to this measure (section 4).
Finally, these last definitions make possible to prove the extremality conditions (Theorem 1.4) and to interpret them (section 5).
We illustrate our results through giving some examples (section 6).

2. Notations and preliminary results

Let Ω be a subset of $\mathbb{R}^d$, we assume Ω to be the closure of a convex open set ω. Let f_0 and f_1 be two probabilities on Ω, we denote the space of such measures $\mathcal{P}(\Omega)$. If γ is in $\mathcal{P}(\Omega\times\Omega)$, the marginals of γ will be written as $\pi_1^\sharp\gamma$ and $\pi_2^\sharp\gamma$. We also introduce the following notations:

- $\mathcal{M}(A,\mathbb{R}^d)$ with A a borelian set and $n\in N$: the space of vector borelian measures on A with values in $\mathbb{R}^d$,
- $\mathcal{M}_b^+(A)$: borelian non negative and bounded measures,
- $\mathcal{C}_o(A)$: continuous functions that vanishes at the infinity on A,
- $\mathcal{C}_b(A)$: bounded continuous functions on A,
- $\mathrm{Lip}(A)$: Lipschitz functions on A.

By abusing notations, we will denote by $\mathrm{L}^1(\Omega)$ the space $\mathrm{L}^1(\omega)$. We consider a convex continuous cost $F:\mathbb{R}^d\to\mathbb{R}^+$. We assume F is even, vanishes at 0, and satisfies:

$$\lim_{|z|\to+\infty} F(z)=+\infty, \tag{2.1}$$

where $|\cdot|$ denotes the Euclidian norm. The point is that we have not made any assumption on F about its homogeneity or superlinearity. As we have seen in the introduction, when F is positively $1-$homogeneous, it is possible to make a relation between $(\mathcal{P})$ and partial differential equations. In order to recover homogeneity, we build a new cost function depending on the time variable:

Definition 2.1.

$$\begin{array}{rcl} H: \ \mathbb{R}^d \times \mathbb{R} & \longrightarrow & [0,+\infty] \\ (z,t) & \mapsto & H(z,t) := \left\{ \begin{array}{lll} tF\left(\frac{z}{t}\right) & \text{if} & t>0, \\ F^\infty(z) & \text{if} & t=0, \\ +\infty & \text{if} & t<0; \end{array} \right. \end{array}$$

where $F^\infty(z) := \lim\limits_{t\to 0^+} tF\left(\frac{z}{t}\right) = \sup_{t>0} tF\left(\frac{z}{t}\right)$ *is the recession function of* F.

The function H is called the perspective function of F (see [24] and [20]).

Example 2.2. *We denote by* $|\cdot|$ *the Euclidian norm.*

- *Let* $F(z) := |z|$, *then:*

$$H(z,t) := \left\{ \begin{array}{lll} |z| & \text{if} & t \geq 0, \\ +\infty & \text{if} & t<0. \end{array} \right.$$

- $F(z) := |z|^p$, $p>1$*:*

$$H(z,t) := \left\{ \begin{array}{lll} \frac{|z|^p}{t^{p-1}} & \text{if} & t>0, \\ 0 & \text{if} & z=0,\ t=0, \\ +\infty & \text{if} & z\neq 0, t\leq 0. \end{array} \right.$$

Hereafter, we list some basic properties of H (Proposition 2.3, Lemma 2.4 and Proposition 2.5).

Proposition 2.3. *H is convex, lower semi-continuous and positively* $1-$*homogeneous with respect to* (z,t).
Its Fenchel transform is:

$$H^*(z^*,t^*) = \chi_K(z^*,t^*) := \left\{ \begin{array}{cc} 0 & \text{if } (z^*,t^*)\in K \\ +\infty & \text{otherwise;} \end{array} \right. \tag{2.2}$$

where K is the convex set

$$K := \{(z^*,t^*): \ F^*(z^*)+t^*\leq 0\}.$$

We have:

$$H(z,t) = H^{**}(z,t) = \sup_{(z^*,t^*)\in K} \{< z, z^* > + tt^*\}. \tag{2.3}$$

The proof of this proposition is left to the reader.

Lemma 2.4. *The interior of K is not empty.*

Proof. Note that $F^*(0)=0$ so $(0,-s)\in K$ for all $s\geq 0$. We are going to show $(0,-s)$ belongs to the interior of K for all $s>0$. It is sufficient to show 0 is in the interior of the domain of F^*, indeed, F^* will be continuous at 0 (it is a convex and l.s.c. function) which gives the desired result.
Remember we have $\lim\limits_{|z|\to+\infty} F(z)=+\infty$, so taking $A>0$, we consider t such that:

$$|z|\geq t \Rightarrow F(z)\geq A.$$

Choosing $\varepsilon > 0$ such that $t\varepsilon - A < 0$, $x^* \in B(0,\varepsilon)$, we get, for all z of norm t:

$$< x^*, \lambda z > -F(\lambda z) \leq \lambda t\varepsilon - \lambda F(z) < 0 \quad \text{for all } \lambda > 1,$$

$$< x^*, \lambda z > -F(\lambda z) \leq t\varepsilon \quad \text{for all } \lambda \leq 1.$$

The first inequality is obtained using the convexity of F combined with $F(0) = 0$ and the second is a consequence of the fact F is non-negative. Finally, we get:

$$\sup_{x\in\mathbb{R}^d} \{< x^*, x > -F(x)\} \leq t\varepsilon < +\infty \quad \text{for all } x^* \in B(0,\varepsilon). \tag{2.4}$$

We can conclude saying $B(0,\varepsilon)$ is a subset of the domain of F^*. □

Let us introduce the functional G on $\mathcal{M}(\mathbb{R}^{d+1},\mathbb{R}^{d+1})$ associated to H and the set $\Omega \times [0,1]$ which will play a rule hereafter:

$$G(\lambda) = \begin{cases} \displaystyle\int H(\lambda) := \int H\left(\frac{d\lambda}{d\mu}(x,t)\right)\, d\mu(x,t) & \text{if } \operatorname{spt}(\lambda) \subset \Omega \times [0,1], \\ +\infty & \text{otherwise.} \end{cases} \tag{2.5}$$

where $\mu \in \mathcal{M}_b^+(\mathbb{R}^{d+1})$ is any borelian measure such that $|\lambda| << \mu$. Notice that, as H is positively $1-$homogeneous, G is well defined i.e. does not depend on the choice of μ. The proof of the following result can be found in [11]:

Proposition 2.5. *G is convex, positively $1-$homogeneous, l.s.c. with respect to the weak star topology of measures.*
The following equalities hold:

$$G(\lambda) = \sup\left\{\int \psi(x,t)\, d\lambda(x,t) : \ \psi \in \mathcal{C}_o(\mathbb{R}^{d+1},\mathbb{R}^{d+1}),\ \psi(x,t) \in K,\ \forall (x,t)\right\}, \tag{2.6}$$

for any $\lambda \in \mathcal{M}(\mathbb{R}^{d+1},\mathbb{R}^{d+1})$ such that $\operatorname{spt}(\lambda) \subset \Omega \times [0,1]$,

$$G^*(\psi) = \int_{\Omega\times[0,1]} \chi_K(\psi(x,t))\, d(x,t). \tag{2.7}$$

Remark 2.6. *Let $\psi_d \in \mathcal{C}^o(\mathbb{R}^{d+1},\mathbb{R}^d)$ and $\psi_{d+1} \in \mathcal{C}^o(\mathbb{R}^{d+1},\mathbb{R})$. We have the following equivalences:*

$$\begin{aligned} G^*(\psi_d,\psi_{d+1}) = 0 \ &\Leftrightarrow\ \psi_{d+1}(x,t) + F^*(\psi_d(x,t)) \leq 0 \quad a.e.\ (x,t) \in \Omega \times [0,1] \\ &\Leftrightarrow\ (\psi_d(x,t),\psi_{d+1}(x,t)) \in K \quad a.e.\ (x,t) \in \Omega \times [0,1] \\ &\Leftrightarrow\ H^*(\psi_d(x,t),\psi_{d+1}(x,t)) = 0 \quad a.e.\ (x,t) \in \Omega \times [0,1] \\ &\Leftrightarrow\ < \omega, (\psi_d(x,t),\psi_{d+1}(x,t)) > \leq H(\omega)\ \forall \omega \in \mathbb{R}^{d+1},\ a.e.\ (x,t) \in \Omega \times [0,1]. \end{aligned}$$

Next result has a very simple proof but justifies in some sense the introduction of the function H:

Proposition 2.7.

$$\min_{\tilde\gamma\in\mathcal{P}((\Omega\times\mathbb{R}^+)^2)} \left\{\int_{(\Omega\times\mathbb{R}^+)^2} H(y-x,t-s)\, d\tilde\gamma((x,s),(y,t)) : \ \pi_1^\sharp\tilde\gamma = f_0 \otimes \delta_0,\ \pi_2^\sharp\tilde\gamma = f_1 \otimes \delta_1\right\}$$
$$= \min_{\gamma\in\mathcal{P}(\Omega^2)} \left\{\int_{\Omega^2} F(y-x)\, d\gamma(x,y) : \ \pi_1^\sharp\gamma = f_0,\ \pi_2^\sharp\gamma = f_1\right\}.$$

Proof. Let γ such that $\pi_i^\sharp\gamma = f_{i-1}$ $(i = 1, 2)$. From γ we build $\tilde{\gamma} \in \mathcal{P}((\Omega \times \mathbb{R}^+)^2)$ by setting:

$$\tilde{\gamma}((x, s), (y, t)) := \gamma(x, y) \otimes \delta_{(0,1)}(s, t).$$

Clearly $\pi_i^\sharp\tilde{\gamma} = f_{i-1} \otimes \delta_{i-1}$ $(i = 1, 2)$. Moreover we have:

$$\begin{aligned}\int_{(\Omega\times\mathbb{R}^+)^2} H(y-x, t-s)\ d\tilde{\gamma}((x, s), (y, t)) &= \int_{(\Omega\times\mathbb{R}^+)^2} H(y-x, t-s) d\gamma(x, y) \otimes \delta_{(0,1)}(s, t) \\ = \int_{\Omega^2} H(y-x, 1) d\gamma(x, y) &= \int_{\Omega^2} F(y-x) d\gamma(x, y).\end{aligned}$$

From this, we get the inequality:

$$\begin{aligned}\min_{\tilde{\gamma}\in\mathcal{P}((\Omega\times\mathbb{R}^+)^2)} &\left\{\int_{(\Omega\times\mathbb{R}^+)^2} H(y-x, t-s)\ d\tilde{\gamma}((x, s), (y, t)) :\ \pi_1^\sharp\tilde{\gamma} = f_0 \otimes \delta_0,\ \pi_2^\sharp\tilde{\gamma} = f_1 \otimes \delta_1\right\} \\ &\le \min_{\gamma\in\mathcal{P}(\Omega^2)} \left\{\int_{\Omega^2} F(y-x)\ d\gamma(x, y) :\ \pi_1^\sharp\gamma = f_0,\ \pi_2^\sharp\gamma = f_1\right\}. \qquad (2.8)\end{aligned}$$

Now, let us consider $\tilde{\gamma}$ such that $\pi_i^\sharp\tilde{\gamma} = f_{i-1} \otimes \delta_{i-1}$ $(i = 1, 2)$. For all borelian set $B \subset \Omega \times \Omega$, we set:

$$\gamma(B) = \tilde{\gamma}\left(\left\{((x, s), (y, t)) :\ (x, y) \in B,\ (s, t) \in (\mathbb{R}^+)^2\right\}\right).$$

Then $\pi_i^\sharp\gamma = f_{i-1}$ $(i = 1, 2)$. Moreover, we have:

$$\begin{aligned}\int_{\Omega\times\Omega} F(y-x) d\gamma(x, y) &= \int_{\Omega\times\{0\}\times\Omega\times\{1\}} H(y-x, t-s) d\tilde{\gamma}((x, s), (y, t)) \\ &\le \int_{(\Omega\times\mathbb{R}^+)^2} H(y-x, t-s) d\tilde{\gamma}((x, s), (y, t)).\end{aligned}$$

This inequality combined to (2.8) gives the statement. □

Remark 2.8. *In this section we have introduced costs and measure depending on the time variable $t \in \mathbb{R}^+$. The general abstract problem that one may consider is the following:*

$$\inf\left\{\int_{(\mathbb{R}^d\times\mathbb{R}^+)^2} c\big((x, s), (y, t)\big)\ d\gamma\big((x, s), (y, t)\big) :\ \pi_i^\sharp\gamma = f_i,\ i = 1, 2\right\} \qquad (2.9)$$

with $c : (\mathbb{R}^d \times \mathbb{R}^+)^2 \to \mathbb{R}^+$ and f_i $(i = 1, 2)$ two probabilities on $\mathbb{R}^d \times \mathbb{R}^+$. The formulation (2.9) may be used in different practical cases. Let us give some interpretations for time-dependent costs and measures. Let us begin with measures and consider for instance

$$f_0(x, t) = \frac{1}{2}(\delta_{(A,0)}(x, t) + \delta_{(B,0)}(x, t)), \quad \textit{and} \quad f_1(x, t) = \frac{1}{2}(\delta_{(C,1/2)}(x, t) + \delta_{(D,1)}(x, t)).$$

(A, B, C and D being fixed points in $\mathbb{R}^d$). The measures f_0 and f_1 may be viewed as follows:

- *At the beginning (at time $t = 0$), two quantities of $1/2$ each of a given material are located at A and B,*
- *Two quantities of $1/2$ each are needed at C and D at time $1/2$ and 1 respectively.*

Now, let us deal with the cost c. Many choices of costs may be relevant in applications.

Of course $c((x,s),(y,t)) = H(\frac{y-x}{t-s})$ *may be used to traduce that fast transportations are more expensive (think of* $c((x,s),(y,t)) = \frac{|y-x|^2}{|t-s|}$ *where average speed appears). Let us give another time depending cost where masses are restricted to move inside a set which may vary with time:*

$$c((x,s),(y,t)) := \inf\left\{ \int_s^t F(\dot v(t))\ dt :\ v \in W^{1,p}(]s,t[,\mathbb{R}^d), \right.$$
$$\left. v(s) = x,\ v(t) = y,\ v(\tau) \in \Gamma(\tau)\ \forall\ \tau \in [s,t] \right\},$$

$F : \mathbb{R}^d \to \mathbb{R}^+$ *is a convex regular function such that* $C_1|\cdot|^p \leq F(\cdot)$ $(p > 1)$ *for some positive constant* C_1 *and* $\Gamma(t)$ *is a closed convex subset of* $\mathbb{R}^d$ *for any positive* t.

3. Duality via Hamilton Jacobi theory

In the introduction, we have seen that, in case F is homogeneous:

$$\min(\mathcal{P}) = \sup(\mathcal{P}^*), \tag{3.1}$$

with

$$(\mathcal{P}^*) \qquad \sup\left\{ \int_\Omega u(x)\ d(f_1 - f_0)(x) :\ u \in \mathrm{Lip}(\Omega),\ \nabla u(x) \in C \text{ a.e. } x \right\}$$
$$C = \{x^* : < x, x^* > -F(x) \leq 0 \quad \forall x\}.$$

Let us now consider the optimization problem introduced in Proposition 2.7

$$\min_{\tilde\gamma \in \mathcal{P}((\Omega\times\mathbb{R}^+)^2)} \left\{ \int_{(\Omega\times\mathbb{R}^+)^2} H(y-x, t-s)\ d\tilde\gamma((x,s),(y,t)) :\ \pi_1^\sharp\tilde\gamma = f_0 \otimes \delta_0,\ \pi_2^\sharp\tilde\gamma = f_1 \otimes \delta_1 \right\}. \tag{3.2}$$

Recall H is homogeneous (Proposition 2.3), then similarly to (3.1), we should be able to show the equivalence between the problem (3.2) and the following one:

$$(\mathcal{Q}^*) \qquad \sup\big\{ < f_1 \otimes \delta_1, \psi > \ - < f_0 \otimes \delta_0, \psi >:\ \psi \in \mathrm{Lip}(\Omega \times [0,1]),$$
$$\partial_t\psi(x,t) + F^*(\nabla_x\psi(x,t)) \leq 0 \text{ a.e.}(x,t) \in \Omega \times [0,1]\big\}.$$

Using a particular viscosity solution of the Hamilton-Jacobi equation $\partial_t\psi + F^*(\nabla_x\psi) = 0$, we are going to prove that the max-value of $(\mathcal{Q}^*)$ is equal to the min-value of (3.2) or, which is equivalent, to the min-value of $(\mathcal{P})$ (cf Proposition 2.7).
To this aim, we consider the following optimization problem for which the equality (3.5) bellow holds (see for instance [26]) :

$$\max\left\{ \int_\Omega \varphi^{FF}(x)\ df_o(x) + \int_\Omega \varphi^F(y)\ df_1(y) :\ \varphi \in \mathcal{C}_b(\Omega) \right\} \tag{3.3}$$

where

$$\varphi^F(y) = \min_{x\in\Omega}\{F(y-x) - \varphi(x)\},$$
$$\varphi^{FF}(x) = \min_{y\in\Omega}\{F(y-x) - \varphi^F(y)\}.$$

It is easy to show that:

$$(\varphi^{FF})^F = \varphi^F. \tag{3.4}$$

It holds:

$$\min(\mathcal{P}) = \max(3.3). \tag{3.5}$$

Taking φ_o a solution of (3.3), we set (Lax-Oleinick formula) for all $(x,t) \in \Omega \times [0,+\infty[$:

$$\Psi_o(x,t) := \inf \left\{ -\varphi_o^{FF}(\sigma(0)) + \int_0^t F(\dot{\sigma}(\tau))\, d\tau :\ \sigma(t) = x \right\} \tag{3.6}$$

where the infimum is taken over all path $\sigma : [0,t] \to \mathbb{R}^d$ continuous and $\mathcal{C}^1$ by parts and such that $\sigma(0) \in \Omega$.
By the convexity of F, for any $t > 0$, the infimum is reached for a right line and we have an equivalent definition (Hopf-Lax formula):

$$\Psi_o(x,t) := \min_{y\in\Omega} \left\{ -\varphi_o^{FF}(y) + tF\left(\frac{y-x}{t}\right) \right\}.$$

Note that, as a consequence of the convexity of F, $\Psi_o(x,\cdot)$ is non-increasing for all $x \in \Omega$. We have (using (3.4) and formula above for $t = 1$):

$$\Psi_o(x,0) := -\varphi_o^{FF}(x), \quad \Psi_o(x,1) = \varphi_o^F(x).$$

Moreover, setting

$$S(t)u(x) = \inf \left\{ u(\sigma(0)) + \int_0^t F(\dot{\sigma}(\tau))\, d\tau :\ \sigma(t) = x \right\},$$

it is obvious that S has the semi-group property ($S(s+t)u(x) = S(t) \circ S(s)u(x)$), consequently, the following formula holds for all $y \in \Omega$, $s,t \in]0,+\infty[$, $t > s$:

$$\Psi_o(y,t) = \min_{x\in\Omega} \left\{ (t-s)F\left(\frac{x-y}{t-s}\right) + \Psi_o(x,s) \right\}. \tag{3.7}$$

We shall show Ψ_o is a Lipschitz function which satisfies the inequation of Hamilton-Jacobi $\partial_t\psi + F^*(\nabla_x\psi) \leq 0$ almost everywhere. In fact, this function is a viscosity solution of the corresponding Hamilton-Jacobi equation (we refer to [2] and [15]). The following lemmas are adapted from [15]:

Lemma 3.1. *Ψ_o is a Lipschitz function on $\Omega \times [0,+\infty[$, hence it is differentiable almost everywhere.*

Proof. • *Let us first show the existence of a constant $C_1 > 0$ such that, for every $t \leq 1/2$ and $x \in \Omega$, it holds:*

$$|\Psi_o(x,t) - \Psi_o(x,0)| \leq C_1 t.$$

In order to simplify the proof, we assume $0 \in \Omega$.
$\Psi_o(x,\cdot)$ being non-increasing, we may only show

$$\Psi_o(x,0) - \Psi_o(x,t) \leq C_1 t.$$

Using (3.7) we can easily obtain for all $y \in \Omega$:

$$\varphi_o^F(y) = \Psi_o(y,1) \leq (1-t)F\left(\frac{x-y}{1-t}\right) + \Psi_o(x,t). \tag{3.8}$$

Take $y \in \Omega$ such that

$$\varphi_o^{FF}\left(\frac{x}{1-t}\right) = F\left(\frac{y-x}{1-t}\right) - \varphi_o^F\left(\frac{y}{1-t}\right).$$

Such an y exists because we have assume $0 \in \Omega$ so $\frac{y}{1-t} \in \Omega \Rightarrow y \in \Omega$.
(If $0 \notin \Omega$, we may just replace $\frac{y}{1-t}$ by $\frac{y-tx_o}{1-t}$ with $x_o \in \Omega$. Then we have:

$$\frac{y-tx_o}{1-t} \in \Omega \Rightarrow y \in (1-t)(\Omega - x_o) + x_o \Rightarrow y \in \Omega$$

and the proof can be done exactly in the same way.) Then, by (3.8), we have:

$$\begin{aligned} \Psi_o(x,t) &\geq \varphi_o^F(y) - (1-t)F\left(\frac{x-y}{1-t}\right) \\ &= \varphi_o^F(y) - (1-t)\left(\varphi_o^F\left(\frac{y}{1-t}\right) + \varphi_o^{FF}\left(\frac{x}{1-t}\right)\right). \end{aligned}$$

Consequently (recall $\Psi_o(x,0) = -\varphi_o^{FF}(x)$), we get:

$$\Psi_o(x,0) - \Psi_o(x,t) \leq -\varphi_o^F(y) + \varphi_o^F\left(\frac{y}{1-t}\right) - t\varphi_o^F\left(\frac{y}{1-t}\right) - \varphi_o^{FF}(x) + \varphi_o^{FF}\left(\frac{x}{1-t}\right) - t\varphi_o^{FF}\left(\frac{x}{1-t}\right).$$

As F is convex l.s.c, it is Lipschitz on every compact set, hence φ_o^F and φ_o^{FF} inherit this property (in particular on the compact set $A = \cup_{t \leq 1/2} \frac{\Omega}{1-t}$) so it exists a constant $C > 0$ such that:

$$\Psi_o(x,0) - \Psi_o(x,t) \leq C\left|\frac{y}{1-t} - y\right| + C\left|\frac{x}{1-t} - x\right| - t\varphi_o^F\left(\frac{y}{1-t}\right) - t\varphi_o^{FF}\left(\frac{x}{1-t}\right).$$

Finally, using the boundedness of φ_o^F and φ_o^{FF} on A and the fact $\frac{t}{1-t} \leq t$, we get:

$$\Psi_o(x,0) - \Psi_o(x,t) \leq 2Ct \sup_{y \in \Omega} |y| + t \sup_A (|\varphi_o^F| + |\varphi_o^{FF}|).$$

•*Let $0 < s \leq t$ such that $|t-s| \leq 1/2$. Let us show that, for all $x \in \Omega$:*

$$0 \leq \Psi_o(x,s) - \Psi_o(x,t) \leq C_1|t-s|.$$

Using again (3.7), we get the existence of $u \in \Omega$ such that:

$$\begin{aligned} \Psi_o(x,s) - \Psi_o(x,t) &= \Psi_o(x,s) - sF\left(\frac{u-x}{s}\right) - \Psi_o(u,t-s) \\ &\leq -\varphi_o^{FF}(u) - \Psi_o(u,t-s) \\ &= \Psi_o(u,0) - \Psi_o(u,t-s) \\ &\leq C_1|t-s|. \end{aligned}$$

•*Let $0 < s \leq t$. Let us show that, for all $x \in \Omega$:*

$$0 \leq \Psi_o(x,s) - \Psi_o(x,t) \leq C_1|t-s|. \tag{3.9}$$

Take $m \in \mathbb{N}$ and $0 \leq \varepsilon < \frac{1}{2}$ such that $|t-s| = \frac{m}{2} + \varepsilon$. We have:

$$\begin{aligned} &\Psi_o(x,s) - \Psi_o(x,t) \\ = &\sum_{i=1}^{m}\left[\Psi_o\left(x, s + \frac{i-1}{2}\right) - \Psi_o\left(x, s + \frac{i}{2}\right)\right] + \Psi_o\left(x, s + \frac{m}{2}\right) - \Psi_o(x,t) \\ \leq &C_1\frac{m}{2} + C_1\varepsilon \\ = &C_1|t-s|. \end{aligned}$$

$\bullet$*Finally, we show the existence of $C_2 > 0$ such that for all $x, y \in \Omega$, $0 < s$:*

$$0 \leq \Psi_o(x,s) - \Psi_o(y,s) \leq C_2|y-x|. \tag{3.10}$$

Let $t = s + |x-y|$, applying again (3.7):

$$\begin{aligned} |\Psi_o(x,s) - \Psi_o(y,s)| &\leq |\Psi_o(x,s) - \Psi_o(x,t)| + |\Psi_o(x,t) - \Psi_o(y,s)| \\ &\leq C_1|t-s| + |t-s| F\left(\frac{x-y}{t-s}\right) \\ &\leq C_1|x-y| + |x-y| \sup_{B(0,1)} F. \end{aligned}$$

Consequently to (3.9) and (3.10), Ψ_o is Lipschitz, hence, by Rademacher's Theorem (see [1]), it is differentiable almost everywhere. □

Lemma 3.2. *Ψ_o satisfies the following inequation almost everywhere in $\Omega \times [0,1]$:*

$$\partial_t \Psi_o(x,t) + F^*(\nabla_x \Psi_o(x,t)) \leq 0.$$

Proof. Let x in the interior of Ω. Using (3.7), we get for all $y \in \mathbb{R}^d$:

$$\forall \varepsilon > 0 \text{ such that } x + \varepsilon y \in \Omega, \quad \Psi_o(x+\varepsilon y, t+\varepsilon) - \Psi_o(x,t) \leq \varepsilon F(y)$$

$$\forall \varepsilon > 0 \text{ such that } x - \varepsilon y \in \Omega, \quad \Psi_o(x,t) - \Psi_o(x-\varepsilon y, t-\varepsilon) \leq \varepsilon F(y).$$

Then, for all $h \in \mathbb{R}$ small enough to have $x + hy \in \Omega$:

$$\frac{\Psi_o(x+hy, t+h) - \Psi_o(x,t)}{h} \leq F(y).$$

Take $(x,t) \in \Omega \times (\mathbb{R}^+)^*$ such $\partial_t \Psi$, $\nabla_x \Psi$ exist and let h go to 0, it holds:

$$< \nabla_x \Psi(x,t), y > + \partial_t \Psi(x,t) \leq F(y),$$

then: $< \nabla_x \Psi(x,t), y > - F(y) \leq -\partial_t \Psi(x,t)$. Taking the supremum on y on the left part of the inequality gives the desired result.

□

Proposition 3.3.

$$\min(\mathcal{P}) \leq \sup(\mathcal{Q}^*).$$

Proof. The function Ψ_o defined above is admissible for $(\mathcal{Q}^*)$ (consequently to lemmas 3.1 and 3.2), moreover it satisfies $\Psi_o(x,0) := -\varphi_o^{FF}(x)$ and $\Psi_o(x,1) = \varphi_o^F(x)$ where φ_o is a solution of (3.3). These remarks imply:

$$\begin{aligned} \max(3.3) &= \int_\Omega \varphi_o^{FF}(x)\ df_o(x) - \int_\Omega \varphi_o^F(x)\ df_1(x) \\ &= \int_{\Omega \times \mathbb{R}^+} \Psi_o(x,t)\ d(f_o \otimes \delta_o)(x,t) - \int_\Omega \Psi_o(x,t)\ d(f_1 \otimes \delta_1)(x,t) \\ &\leq \sup(\mathcal{Q}^*). \end{aligned}$$

The result follows from (3.5). □

By duality, we are going to show the equivalence between $(\mathcal{Q}^*)$ and a new optimization problem:

$(\mathcal{Q})$

$$\min\left\{\int H(\chi) : \chi \in \mathcal{M}_b(\mathbb{R}^{d+1}, \mathbb{R}^{d+1}), \quad \mathrm{spt}(\chi) \subset \Omega \times [0,1], \right.$$
$$\left. -\mathrm{div}_{x,t}(\chi) = f_1 \otimes \delta_1 - f_0 \otimes \delta_0 \quad \text{on} \quad \mathbb{R}^{d+1}\right\}$$

where "$\mathrm{div}_{x,t}$" is intended in the sense of distributions, i.e. for all function $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^{d+1})$:

$$- < \mathrm{div}_{x,t}(\chi), \varphi > := < (\nabla_x \varphi, \partial_t \varphi), \chi > .$$

The quantity $\int H(\chi)$ is defined by (2.5).

Remark 3.4. When F is superlinear, the problem $(\mathcal{Q})$ can be written as:

$$\min\left\{\int F(\sigma)\, d\mu : \quad \mu \in \mathcal{M}_b^+(\mathbb{R}^{d+1}),\ \mathrm{spt}(\mu) \in \Omega \times [0,1], \right.$$
$$\left. \sigma \in \mathrm{L}^1_\mu(\Omega \times [0,1], \mathbb{R}^d),\ -\mathrm{div}_x(\sigma\mu) - \partial_t \mu = f_1 \otimes \delta_1 - f_0 \otimes \delta_0 \quad \text{sur} \quad \mathbb{R}^{d+1}\right\}.$$

Indeed, taking $\chi = (\lambda, \mu)$ with $\lambda \in \mathcal{M}_b^+(\mathbb{R}^{d+1}, \mathbb{R}^d)$ and $\mu \in \mathcal{M}_b^+(\mathbb{R}^{d+1})$, as $F^\infty(z) = +\infty$ when $z \neq 0$, we have:

$$\int H(\lambda, \mu) < +\infty \Rightarrow |\lambda| << \mu.$$

Proposition 3.5.

$$\sup(\mathcal{Q}^*) = \inf(\mathcal{Q}).$$

Before proving this result, we need to gain some regularity on the admissible functions of $(\mathcal{Q}^*)$, this will be made possible using the following:

Lemma 3.6. *Let A a closed subset of $\mathbb{R}^N$ and $\mathcal{K}$ a closed convex subset of $\mathbb{R}^N$. Let $\psi \in \mathrm{Lip}(A)$ satisfying $D\psi \in \mathcal{K}$ almost everywhere. Then, it exists $\psi_n \in \mathcal{C}^\infty(\mathbb{R}^N)$ such that:*

$$\psi_n \to \psi \quad \textit{uniformly on } A, \tag{3.11}$$
$$D\psi_n(x) \in \mathcal{K} \quad \textit{for all } x \in A. \tag{3.12}$$

Consequently to this lemma:

$$\sup(\mathcal{Q}^*) = \sup\big\{ < f_1 \otimes \delta_1 - f_0 \otimes \delta_0, \psi >: \quad \psi \in \mathcal{C}^1(\mathbb{R}^{d+1}),$$
$$\partial_t \psi(x,t) + F^*(\nabla_x \psi(x,t)) \leq 0 \quad \forall (x,t) \in \Omega \times [0,1]\big\}. \tag{3.13}$$

Proof of Lemma 3.6. By Rademacher's Theorem, ψ is differentiable almost everywhere and it exists $R > 0$ such that $|D\psi(x)| \leq R$. We introduce the following application:

$$\rho_R(z^*) := \sup\{< z, z^* >: \ z \in \mathcal{K},\ |z| \leq R\} = \chi^*_{K_R}(z^*)$$

with $K_R := \{z \in \mathcal{K},\ |z| \leq R\}$ and $\chi_{K_R}(z) = 0$ if $z \in K_R$, $+\infty$ otherwise. It is a continuous application, moreover, K_R being convex and closed, $\rho_R^*(z^*) = \chi_{K_R}$. We extend ψ outside A by setting:

$$\tilde{\psi}(x) := \inf_{y \in A} \{\psi(y) + \rho_R(x - y)\}. \tag{3.14}$$

Let us verify $\tilde{\psi}$ and ψ coincides on A. First note that the set of points $(x,y) \in A^2$ such that ψ is differentiable L^1-almost everywhere on the segment $[x,y]$ is a dense set. For all point (x,y) of this set, it exists $c \in [x,y]$ such that: $\psi(x) \leq \psi(y) + < D\psi(c), x-y >$, which implies $\psi(x) \leq \psi(y) + \rho_R(x-y)$. By continuity, this last inequality remains true for all $(x,y) \in A^2$. Consequently $\psi(x) \leq \tilde{\psi}(x)$ on A. The converse inequality is also true (take $y = x$ in (3.14)).
We can easily show that:

$$\tilde{\psi}(x) - \tilde{\psi}(y) \leq \rho_R(x-y) \quad \forall (x,y) \in \mathbb{R}^N. \tag{3.15}$$

We make a regularisation of $\tilde{\psi}$, take $(f_n)_n$ a classical sequence of regularisation kernel (f_n supported on a ball of radius $1/n$ centered at the origin), we set:

$$\psi_n(x) \ := \ \int_{B(0,1/n)} f_n(y)\tilde{\psi}(x-y)\ dy.$$

The sequence $(\psi_n)_n$ converges uniformly to ψ, moreover, by (3.15), it satisfies:

$$\psi_n(x) - \psi_n(y) \leq \rho_R(x-y).$$

This inequality implies $< D\psi_n(x), y > \leq \rho_R(y)$ for all y, i.e. $\rho_R^*(D\psi_n(x)) \leq 0$ which means $D\psi_n(x) \in \mathcal{K}$ for all x.

□

Proof of Proposition 3.5. Using (3.13), the proof reduces to show the following equality:

$$\begin{aligned} &\sup \big\{ < f_1 \otimes \delta_1 - f_0 \otimes \delta_0, \psi >: \quad \psi \in \mathcal{C}^1(\mathbb{R}^{d+1}), \\ &\qquad\qquad \partial_t \psi(x,t) + F^*(\nabla_x \psi(x,t)) \leq 0 \quad \forall (x,t) \in \Omega \times [0,1] \big\} \\ = &\inf_{\mathcal{M}_b(\mathbb{R}^{d+1})^{d+1}} \left\{ \int H(\chi) : \ \mathrm{spt}(\chi) \subset \Omega \times [0,1], \ -\mathrm{div}_{x,t}(\chi) = f_1 \otimes \delta_1 - f_0 \otimes \delta_0 \right\}. \end{aligned}$$

We introduce the operator $A : \mathcal{C}_o(\mathbb{R}^{d+1}) \to \mathcal{C}(\mathbb{R}^{d+1}, \mathbb{R}^{d+1})$ of domain $\mathcal{C}^1(\mathbb{R}^{d+1}) \cap \mathcal{C}_o(\mathbb{R}^{d+1})$ defined by:

$$A\psi = (\nabla_x \psi, \partial_t \psi) \quad \forall \psi \in \mathcal{C}^1(\mathbb{R}^{d+1}).$$

Then, as $\mathcal{C}^1(\mathbb{R}^{d+1}) \cap \mathcal{C}_o(\mathbb{R}^{d+1})$ is dense in $\mathcal{C}^1(\mathbb{R}^{d+1})$, we have (cf Remark 2.6):

$$\begin{aligned} &\sup \big\{ < f_1 \otimes \delta_1 - f_0 \otimes \delta_0, \psi >: \quad \psi \in \mathcal{C}^1(\mathbb{R}^{d+1}), \\ &\qquad\qquad \partial_t \psi(x,t) + F^*(\nabla_x \psi(x,t)) \leq 0 \quad \forall (x,t) \in \Omega \times [0,1] \big\} \\ = \ &\sup \big\{ < f_1 \otimes \delta_1 - f_0 \otimes \delta_0, \psi >: \quad \psi \in \mathcal{C}^1(\mathbb{R}^{d+1}) \cap \mathcal{C}_o(\mathbb{R}^{d+1}), \\ &\qquad\qquad \partial_t \psi(x,t) + F^*(\nabla_x \psi(x,t)) \leq 0 \quad \forall (x,t) \in \Omega \times [0,1] \big\} \\ = \ &(G^* \circ A)^*(f_1 \otimes \delta_1 - f_0 \otimes \delta_0) \end{aligned}$$

where G^* is the functional which appears in Proposition 2.5. Let us note that on the one hand the domain of A is dense in $\mathcal{C}_o(\mathbb{R}^{d+1})$. On the other hand considering v in the interior of K (which is not empty by Lemma 2.4) and the application $\psi(x,t) = v \cdot (x,t)$, G^* is continuous at $A\psi$. Then we can compute the Fenchel transform $(G^* \circ A)^*$:

$$(G^* \circ A)^*(f_1 \otimes \delta_1 - f_0 \otimes \delta_0) = \inf \left\{ G^{**}(\chi) : \ \chi \in \mathcal{M}_b(\mathbb{R}^{d+1}, \mathbb{R}^{d+1}), \ A^*(\chi) = f_1 \otimes \delta_1 - f_0 \otimes \delta_0 \right\},$$

where the infimum is in fact a minimum.
We have: $A^*(\chi) = -\mathrm{div}_{x,t}(\chi)$, and, by Proposition 2.5 :

$$G^{**}(\chi) = G(\chi) = \begin{cases} \int H(\chi) & \text{if } \mathrm{spt}(\chi) \subset \Omega \times [0,1], \\ +\infty & \text{else.} \end{cases}$$

Consequently:

$$\begin{aligned} & (G^* \circ A)^*(f_1 \otimes \delta_1 - f_0 \otimes \delta_0) \\ = & \min_{\mathcal{M}_b(\mathbb{R}^{d+1})^{d+1}} \left\{ \int H(\chi) : \ \mathrm{spt}(\chi) \subset \Omega \times [0,1], \ -\mathrm{div}_{x,t}(\chi) = f_1 \otimes \delta_1 - f_0 \otimes \delta_0 \right\}. \end{aligned}$$

□

We have already shown that $\sup(\mathcal{P}) \leq \sup(\mathcal{Q}^*) = \inf(\mathcal{Q})$. In order to get the equality between all these quantities, we only need to show:

Proposition 3.7.

$$\min(\mathcal{P}) \geq \min(\mathcal{Q}).$$

Moreover if γ is a solution of the problem $(\mathcal{P})$, we are able to construct a solution χ of $(\mathcal{Q})$, by using the formula:

$$< \chi, \Phi >:= \int_{\Omega^2} \int_0^1 \Phi\left((1-s)x_0 + sx_1, s\right) \cdot (x_1 - x_0, 1) \ ds d\gamma(x_0, x_1), \tag{3.16}$$

for all $\Phi \in \mathcal{C}_c(\mathbb{R}^{d+1}, \mathbb{R}^{d+1})$.

Proof. Let $\gamma \in \mathcal{P}(\Omega \times \Omega)$ such that $\pi_1^\sharp \gamma = f_0$ and $\pi_2^\sharp \gamma = f_1$. Let $\chi \in \mathcal{M}_b(\mathbb{R}^{d+1}, \mathbb{R}^{d+1})$ associated to γ by formula (3.16). The measure χ is admissible for $(\mathcal{Q})$. Indeed, let $\phi \in \mathcal{C}_c^\infty(\mathbb{R}^{d+1})$, applying (3.16) with $\Psi = (\nabla_x \phi, \partial_t \phi)$, we get:

$$\begin{aligned} & < \chi, (\nabla_x \phi, \partial_t \phi) > \\ = & \int_{\Omega^2} \int_0^1 \nabla_x \phi((1-s)x_0 + sx_1, s) \cdot (x_1 - x_0) + \partial_t \phi((1-s)x_0 + sx_1, s) ds d\gamma(x_0, x_1) \\ = & \int_{\Omega^2} \int_0^1 \frac{d}{dt} \left(\phi((1-s)x_0 + sx_1, s)\right) \ ds \ d\gamma(x_0, x_1) \\ = & \int_{\Omega^2} \phi(x_1, 1) - \phi(x_0, 0) \ d\gamma(x_0, x_1) \\ = & < f_1 \otimes \delta_1 - f_0 \otimes \delta_0, \phi > . \end{aligned}$$

Let $\Psi \in \mathcal{C}_c(\mathbb{R}^{d+1}, \mathbb{R}^{d+1})$ such that $\Psi(x,t) \in K$ for all (x,t). By Fenchel inequality:

$$\Psi((1-s)x_0 + sx_1, s) \cdot ((x_1 - x_0), 1) \leq H^*\left(\Psi((1-s)x_0 + sx_1, s)\right) + H(x_1 - x_0, 1).$$

Consequently (recall Remark 2.6):

$$\begin{aligned} < \chi, (\psi, \varphi) > &= \int_{\Omega^2} \int_0^1 \Psi((1-s)x_0 + sx_1, s) \cdot ((x_1 - x_0), 1) ds d\gamma(x, t) \\ &\leq \int_{\Omega^2} \left[\int_0^1 H^*\left(\Psi((1-s)x_0 + sx_1, t)\right) + H(x_1 - x_0, 1) \ ds \right] d\gamma(x, t) \\ &= \int_{\Omega^2} \int_0^1 H(x_1 - x_0, 1) \ ds d\gamma(x_0, x_1) = \int_{\Omega^2} F(x_1 - x_0) \ d\gamma(x_0, x_1). \end{aligned}$$

Then, by Proposition 2.5 :

$$\begin{aligned}\int_{\Omega\times[0,1]} H(\chi) &= \sup\left\{ <\chi,\Psi>:\ \Psi\in\mathcal{C}_c(\mathbb{R}^{d+1},\mathbb{R}^{d+1}),\ \Psi(x,t)\in K \quad \forall (x,t)\right\} \\ &\le \int_{\Omega^2} F(x_1-x_0)\ d\gamma(x_0,x_1).\end{aligned}$$

As γ is admissible for $(\mathcal{P})$, the inequality above implies:

$$\inf(\mathcal{P}) \ge \min(\mathcal{Q}).$$

□

Now, we can state the main theorem of this section which is an immediate consequence of propositions 3.3, 3.5 and 3.7:

Theorem 3.8.

$$\min(\mathcal{P}) = \max(\mathcal{Q}^*) = \min(\mathcal{Q}),$$

so is to say:

$$\begin{aligned}&\min_{\gamma\in\mathcal{P}(\Omega\times\Omega)}\left\{\int_{\Omega\times\Omega} F(y-x)\ d\gamma(x,y)\ :\ \pi_1^\sharp\gamma=f_0,\ \pi_2^\sharp\gamma=f_1\right\} \\ =\ &\max_{\partial_t\psi(x,t)+F^*(\nabla_x\psi(x,t))\le 0}\{<f_1\otimes\delta_1,\psi> - <f_0\otimes\delta_0,\psi>:\quad \psi\in \mathrm{Lip}(\Omega\times[0,1])\} \\ =\ &\inf_{\mathcal{M}_b(\mathbb{R}^{d+1})^{d+1}}\left\{\int H(\chi):\ \mathrm{spt}(\chi)\subset\Omega\times[0,1],\ -\mathrm{div}_{x,t}(\chi)=f_1\otimes\delta_1-f_0\otimes\delta_0\right\}.\end{aligned}$$

4. Tangential gradient to a measure

The notion of tangential gradient to measure has first been introduced by G. Bouchitté, G. Buttazzo and P. Seppecher (see [7], [8], [9], see also [17] and [10]). In the following subsection we try to explain their idea and why this notion is useful in our case.

4.1. **Motivation.** Take $\psi \in \mathrm{Lip}(\Omega\times[0,1])$ and $\chi\in\mathcal{M}_b(\mathbb{R}^{d+1},\mathbb{R}^{d+1})$ two solutions of $(\mathcal{Q}^*)$ and $(\mathcal{Q})$ respectively. By Theorem 3.8, we have the following equality:

$$\int H(\chi) = <\psi, f_1\otimes\delta_1> - <\psi, f_0\otimes\delta_0>.$$

Then, using the properties of χ, we may replace the measure $f_1\otimes\delta_1 - f_0\otimes\delta_0$ by the divergence of χ which leads:

$$\int H(\chi) = <\psi, -\mathrm{div}_{x,t}\chi>. \tag{4.1}$$

Let us assume for a time that ψ is regular, say $\mathcal{C}^2$, we may use an integration by parts in the right size of the equality and obtain:

$$\int H(\chi) = \int (\nabla_x\psi(x,t),\partial_t\psi)\ d\chi(x,t).$$

At this point, we may write χ as $\sigma\mu$ where $\mu\in\mathcal{M}_b^+(\mathbb{R}^{d+1})$ and $\sigma\in \mathrm{L}^1_\mu(\mathbb{R}^{d+1},\mathbb{R}^{d+1})$ (take for instance $\mu=|\chi|$ and $\sigma=\frac{d\chi}{d|\chi|}$) and reformulate the above equality as:

$$\int H(\sigma(x,t))\ d\mu(x,t) = \int (\nabla_x\psi(x,t),\partial_t\psi(x,t))\cdot\sigma(x,t)\ d\mu(x,t). \tag{4.2}$$

Then, using (2.3) we get the following optimality condition:

$$H(\sigma(x,t)) = (\nabla_x\psi(x,t), \partial_t\psi)\cdot\sigma(x,t) \quad \mu-\text{a.e.}(x,t)\in\Omega\times[0,1].$$

Unfortunately, ψ may not be regular and consequently the above argument may be false. To avoid this problem, one may use an uniform approximation of ψ by a regular sequence $(\psi_n)_n$ which we assume to be equiLipschitz, then, up to a subsequence, $(\nabla_x\psi_n, \partial_t\psi_n)_n$ has a limit ξ for the weak star topology $\sigma\left(\mathrm{L}^\infty_\mu, \mathrm{L}^1_\mu\right)$, moreover, we have:

$$<\psi, -\mathrm{div}_{x,t}\chi> = \lim_n <\psi_n, -\mathrm{div}_{x,t}\chi> = \lim_n \int (\nabla_x\psi_n, \partial_t\psi_n)\cdot\sigma\, d\mu = \int \xi\cdot\sigma\, d\mu.$$

Then (4.2) holds true with ξ instead of $(\nabla_x\psi, \partial_t\psi)$. The problem is that ξ does not make sense because it depends on the sequence $(\psi_n)_n$ we choose. To be convinced of that, let us consider an example.

Example 4.1. *Let $d=1$, $\Omega=[-\pi,\pi]$, $C := \{(x,y):\ y=x^2\}\cap(\Omega\times[0,1])$, $\mu=\mathcal{L}^1 \llcorner C$. We consider the following sequence:*

$$f_n(x,y) = \frac{\sin n(y-x^2)}{n}.$$

It converges uniformly to zero but its differential $(\nabla_x f_n, \partial_t f_n)$ converges weakly to $\xi(x,y)=(-2x,1)$ for the topology $\sigma\left(\mathrm{L}^\infty_\mu, \mathrm{L}^1_\mu\right)$. On the contrary, take $g_n=0$, the sequence $(g_n)_n$ obviously converges uniformly to zero and $(\nabla_x g_n, \partial_t g_n)$ converges weakly to $\xi(x,y)=0$.

Let us introduce the following set which plays an important rule in the following:

$$\mathcal{N} := \Big\{ \quad \xi\in\mathrm{L}^\infty_\mu(\mathbb{R}^{d+1},\mathbb{R}^{d+1}):\ \exists(u_n)_n,\ u_n\in\mathcal{C}^1(\mathbb{R}^{d+1}),$$
$$u_n\to 0 \text{ uniformly on } \mathbb{R}^{d+1},\ Du_n \overset{*}{\rightharpoonup} \xi \text{ in } \sigma\left(\mathrm{L}^\infty_\mu,\mathrm{L}^1_\mu\right) \Big\} \tag{4.3}$$

where we have denoted $D=(\nabla_x, \partial_t)$. We make tow essential remarks:

(1) it is clear that if μ is such that $\mathcal{N}$ is reduced to $\{0\}$, there is no problem,
(2) even if ψ is regular, the part of its gradient which makes sense in the above argument is $D\psi\cdot\sigma$ and not $D\psi$ in its integrality (cf (4.2)).

The idea is, for μ-almost all (x,t), to make a projection of $D\psi_n(x,t)$ (where ψ_n is an equiLipschitz and regular sequence tending uniformly to ψ) on a subspace of $\mathbb{R}^{d+1}$- called the tangent space of μ on (x,t)- in order to "kill " $\mathcal{N}$ and to conserve the part of the gradient which makes sense.

4.2. **Definition of the tangential gradient to a measure.** We consider μ a general Radon measure on $\mathbb{R}^{d+1}$. We introduce the topology τ:

$$u_n \overset{\tau}{\rightharpoonup} u \Leftrightarrow \begin{cases} u_n\to u \text{ uniformly on } \mathbb{R}^{d+1} \\ \exists C\in\mathbb{R} \text{ such that } |Du_n|_{\mathrm{L}^\infty_\mu}\le C. \end{cases} \tag{4.4}$$

An element of $\mathbb{R}^{d+1}$ will be written as "y". We denote by "$\overset{*}{\rightharpoonup}$" the convergence for the topology $\sigma(L^\infty_\mu, L^1_\mu)$. The set $\mathcal{N}$ is defined as before by (4.3). Finally, the closure of $\mathcal{N}$ for the weak star topology of $\mathrm{L}^\infty_\mu(\mathbb{R}^{d+1},\mathbb{R}^{d+1})$ will be denoted by $\overline{\mathcal{N}}$.

The aim of this subsection is to define the tangential gradient to μ of any $\psi \in \mathrm{Lip}(\Omega \times [0,1])$. The construction will be done first for $\psi \in \mathrm{Lip}(\mathbb{R}^{d+1})$, then by a localisation argument for $\psi \in \mathrm{Lip}(\Omega \times [0,1])$.

Lemma 4.2. *$\mathcal{N}$ is a vectorial subspace of $\mathrm{L}^\infty_\mu(\mathbb{R}^{d+1}, \mathbb{R}^{d+1})$ and satisfies:*

$$\forall \xi \in \mathcal{N},\ \forall \varphi \in \mathcal{C}^1(\mathbb{R}^{d+1}),\ \xi\varphi \in \mathcal{N}.$$

Using this lemma, we can define the tangent space to the measure μ:

Proposition and Definition 4.3. *It exists a multifunction T_μ from $\mathbb{R}^{d+1}$ into $\mathbb{R}^{d+1}$ such that:*

$$\eta \in \mathcal{N}^\perp \Leftrightarrow \eta(y) \in T_\mu(y) \quad \mu - \text{a.e.}y,$$
$$\xi \in \overline{\mathcal{N}} \Leftrightarrow \xi(y) \in T_\mu^\perp(y) \quad \mu - \text{a.e.}y.$$

For μ almost every y, $T_\mu(y)$ is a vectorial subspace of $\mathbb{R}^{d+1}$ called the tangential space of μ at y. We denote by $P_\mu(y,.)$ the projection on $T_\mu(y)$.

Proof. We consider the orthogonal of $\mathcal{N}$:

$$\mathcal{N}^\perp := \left\{\sigma \in \mathrm{L}^1_\mu(\mathbb{R}^{d+1}, \mathbb{R}^{d+1}) :\ \int_{\mathbb{R}^{d+1}} \xi(y)\cdot\sigma(y)\ d\mu(y) = 0,\ \forall \xi \in \mathcal{N}\right\}.$$

We first show that for all $\sigma \in \mathcal{N}^\perp$ and $A \subset \mathbb{R}^{d+1}$, we have $\mathbb{1}_A\sigma \in \mathcal{N}^\perp$. To show this, it is sufficient to consider a sequence $(\varphi_n)_n$ of $\mathrm{L}^\infty_\mu \cap \mathcal{C}^1(\mathbb{R}^{d+1})$ which converges to $\mathbb{1}_A$ for the weak star topology. Let $\xi \in \mathcal{N}$, by Lemma 4.2, $\varphi_n\xi$ is in $\mathcal{N}$ and:

$$\begin{aligned}\int_{\mathbb{R}^{d+1}} \mathbb{1}_A(y)\sigma(y)\cdot\xi(y)\ d\mu(y) &= \ <\xi\mathbb{1}_A, \sigma>_{(\mathrm{L}^\infty_\mu, \mathrm{L}^1_\mu)}\\ &= \lim_{n\to\infty} <\xi\varphi_n, \sigma>_{(\mathrm{L}^\infty_\mu, \mathrm{L}^1_\mu)} = 0.\end{aligned}$$

This shows $\mathbb{1}_A\sigma \in \mathcal{N}^\perp$.
Then, as $\mathcal{N}^\perp$ is a closed subspace, by a theorem of F. Hiai and H. Umegacki (see [19] or [25]), we get the existence of a multifunction T_μ satisfying:

$$\sigma \in \mathcal{N}^\perp \Leftrightarrow \sigma(y) \in T_\mu(y) \quad \mu - \text{a.e.}y.$$

Note that as $\mathcal{N}^\perp$ is a vectorial subspace of $\mathrm{L}^\infty_\mu(\mathbb{R}^{d+1}, \mathbb{R}^{d+1})$, the set $T_\mu(y)$ is a vectorial subspace of $\mathbb{R}^{d+1}$. Moreover, by a result of [11]:

$$\begin{aligned}\overline{\mathcal{N}} &= \mathcal{N}^{\perp\perp}\\ &= \left\{\xi \in \mathrm{L}^\infty_\mu(\mathbb{R}^{d+1}) :\ \int_{\mathbb{R}^{d+1}} \sigma(y)\cdot\xi(y)\ d\mu(y) = 0,\ \forall\sigma \in \mathcal{N}^\perp\right\}\\ &= \left\{\xi \in \mathrm{L}^\infty_\mu(\mathbb{R}^{d+1}) :\ \sup_{\sigma\in\mathcal{N}^\perp}\left(\int_{\mathbb{R}^{d+1}} \sigma(y)\cdot\xi(y)\ d\mu(y)\right) = 0\right\}\\ &= \left\{\xi \in \mathrm{L}^\infty_\mu(\mathbb{R}^{d+1}) :\ \int_{\mathbb{R}^{d+1}} \sup_{z\in T_\mu(y)} z\cdot\xi(y)\ d\mu(y) = 0\right\}\\ &= \left\{\xi \in \mathrm{L}^\infty_\mu(\mathbb{R}^{d+1}) :\ \xi(y) \in T_\mu^\perp(y) \quad \mu - \text{a.e.}y\right\}.\end{aligned}$$

□

Before going further we give some classical example of tangent spaces:

Example 4.4. • *Let* $\gamma : [0,1] \to \mathbb{R}^{d+1}$ *a regular curve and take* $\mu = \mathcal{H}^1 \llcorner \gamma([0,1])$. *Then* $T_\mu(\gamma(t)) = \{\lambda \dot{\gamma}(t) : \ \lambda \in \mathbb{R}$ *for all* $t \in [0,1]$.

- *Let* U *a bounded open subset of* $\mathbb{R}^{d+1}$ *and take* μ *the Lebesgue measure on* U, *then* $T_\mu(x,t) = \mathbb{R}^{d+1}$ *for almost every* (x,t).

One may define the tangential gradient of a regular function as the projection of its gradient at μ-almost every y on the tangent space at μ-almost every y. The following proposition allows us to extend this definition to every Lipschitzian function:

Proposition 4.5. *We consider the following operator of domain* $\mathcal{C}^1(\mathbb{R}^{d+1})$:

$$A: \begin{array}{ccc} \mathcal{C}^1(\mathbb{R}^{d+1}) & \to & \mathrm{L}^\infty_\mu(\mathbb{R}^{d+1}) \\ u & \mapsto & P_\mu(.,Du(.)). \end{array}$$

It can be extended continuously on $\mathrm{Lip}(\mathbb{R}^{d+1})$ *with respect to the topology* τ *defined by (4.4).*

Proof. We must show that if $(u_n)_n$ is a sequence of $\mathrm{L}^\infty_\mu(\mathbb{R}^{d+1})$ converging to 0 for the topology τ, then $Au_n \overset{*}{\rightharpoonup} 0$.
Let $(u_n)_n$ such a sequence. As $(Du_n)_n$ and $(Au_n)_n$ are bounded in $\mathrm{L}^\infty_\mu(\mathbb{R}^{d+1})$, up to subsequences, it exist $\eta,\ \xi \in \mathrm{L}^\infty_\mu(\mathbb{R}^{d+1})$ such that:

$$Au_n \overset{*}{\rightharpoonup} \eta, \quad Du_n \overset{*}{\rightharpoonup} \xi. \tag{4.5}$$

Let us show $\eta \in \overline{\mathcal{N}} \cap \mathcal{N}^\perp$.
a) On the one hand ξ belongs to $\mathcal{N}$. On the other hand, for $\mu-$almost every $y \in \mathbb{R}^{d+1}$, we can decompose the vector $Du_n(y)$ and get the existence of a vector $w_n(y) \in T_\mu^\perp(y)$ such that:

$$Du_n(y) = w_n(y) + Au_n(y) \quad \mu-\text{a.e.} y \in \mathbb{R}^{d+1}.$$

By definition, $w_n \in \overline{\mathcal{N}}$. Making $n \to +\infty$ we obtain that the limit $\xi - \eta$ of w_n is in $\overline{\mathcal{N}}$. Consequently, as $\xi \in \mathcal{N}$, η is in $\overline{\mathcal{N}}$.

b) By (4.5), as $Au_n \in \mathcal{N}^\perp$ for all $n \in \mathbb{N}$ and $\mathcal{N}^\perp$ is closed, η is in $\mathcal{N}^\perp$. □

Definition 4.6. *Let* $u \in \mathrm{Lip}(\mathbb{R}^{d+1})$. *We call the tangential gradient of* u *and denote by* $D_\mu u$ *the unique function* $\xi \in \mathrm{L}^\infty_\mu$ *such that:*

$$\left.\begin{array}{cc} (u_n) \in \mathrm{Lip}(\mathbb{R}^{d+1}), & \textit{equiLipschitz} \\ u_n \to u, & \textit{uniformly on } \mathbb{R}^{d+1} \end{array}\right\} \Rightarrow P_\mu(\cdot, Du_n(\cdot)) \overset{*}{\rightharpoonup} \xi.$$

Remark 4.7. If $u \in \mathcal{C}^1(\mathbb{R}^{d+1})$, then $D_\mu u = P_\mu(\cdot, Du(\cdot))$.

The next result appears while proving Proposition 4.5:

Proposition 4.8. *Let* $(u_n)_n$ *a sequence in* $\mathcal{C}^1(\mathbb{R}^{d+1})$ *such that:*

$$u_n \to u \textit{ uniformly on } \mathbb{R}^{d+1},$$

$$Du_n \overset{*}{\rightharpoonup} \xi \in \mathrm{L}^\infty_\mu(\mathbb{R}^{d+1}).$$

Then $\xi = D_\mu u + \eta$ *for some* $\eta \in \overline{\mathcal{N}}$.

4.3. **Basics properties.** The point is now to give the ad hoc integration by part formula.

Lemma 4.9. *Let* $\theta \in \mathrm{L}^1_\mu(\mathbb{R}^{d+1}, \mathbb{R}^{d+1})$ *such that* $-\mathrm{div}(\theta\mu) \in \mathcal{M}_b(\mathbb{R}^{d+1}, \mathbb{R}^{d+1})$*. It holds:*

$$\theta(y) \in T_\mu(y) \quad \mu\text{−a.e.}y.$$

Proof. By Proposition and Definition 4.3, it is sufficient to show that $\theta \in \mathcal{N}^\perp$. Let $\xi \in \mathcal{N}$, by definition it exists $(u_n)_n$ in $\mathcal{C}^1(\mathbb{R}^{d+1})$ such that $u_n \to 0$ uniformly and $Du_n \overset{*}{\rightharpoonup} \xi$. Then:

$$\begin{aligned}\int \xi(y)\cdot\theta(y)\,d\mu(y) &= \lim_{n\to+\infty}\int Du_n(y)\cdot\theta(y)\,d\mu(y)\\ &= \lim_{n\to+\infty} - < u_n, \mathrm{div}(\theta\mu) >\\ &= 0.\end{aligned}$$

□

Proposition 4.10. *(Integration by parts formula)*
Let $\psi \in \mathrm{Lip}(\mathbb{R}^{d+1})$ *and* $\theta \in \mathrm{L}^1_\mu(\mathbb{R}^{d+1}, \mathbb{R}^{d+1})$ *such that* $-\mathrm{div}(\theta\mu) \in \mathcal{M}_b(\mathbb{R}^{d+1}, \mathbb{R}^{d+1})$*. Then:*

$$< -\mathrm{div}(\theta\mu), \psi >= \int \theta(y)\cdot D_\mu\psi(y)\,d\mu(y).$$

Proof. By slightly modifying the proof of lemma 3.6, one can easily build a sequence $(\psi_n)_n$ of $\mathcal{C}^\infty_c(\mathbb{R}^{d+1})$ such that:

$$\psi_n \to \psi \text{ uniformly on } \mathbb{R}^{d+1},$$

$$D\psi_n(x,t) \leq \mathrm{Lip}(\psi) \quad \forall (x,t), \text{ where } \mathrm{Lip}(\psi) \text{ is the Lipschitz constant of } \psi.$$

By the definition of $D_\mu\psi$ we have: $D_\mu\psi_n \overset{*}{\rightharpoonup} D_\mu\psi$. On the other hand, by Lemma 4.9, $\theta(y) \in T_\mu(y)$ μ-almost everywhere, consequently

$$D\psi_n(y)\cdot\theta(y) = D_\mu\psi_n(y)\cdot\theta(y) \quad \mu - \text{almost everywhere.}$$

Thus:

$$\begin{aligned}< -\mathrm{div}(\theta\mu), \psi > &= \lim_{n\to+\infty} < -\mathrm{div}(\theta\mu), \psi_n >\\ &= \lim_{n\to+\infty}\int D\psi_n(y)\cdot\theta(y)\,d\mu(y)\\ &= \lim_{n\to+\infty}\int D_\mu\psi_n(y)\cdot\theta(y)\,d\mu(y)\\ &= \int D_\mu\psi(y)\cdot\theta(y)\,d\mu(y).\end{aligned}$$

□

As we have already said, the aim of this subsection is to build the tangential gradient $D_\mu\psi$ of any $\psi \in \mathrm{Lip}(\Omega\times[0,1])$. At this point, we have defined the tangential gradient $D_\mu\varphi$ for any $\varphi \in \mathrm{Lip}(\mathbb{R}^{d+1})$. Of course, one may consider a Lipschitz extension $\tilde{\psi}$ of $\psi \in \mathrm{Lip}(\Omega\times[0,1])$ and define

$$D_\mu\psi(x,t) := D_\mu\tilde{\psi}(x,t) \quad \mu - a.e.(x,t) \in \Omega\times[0,1].$$

The question is: does this definition depends on the Lipschitz extension $\tilde{\psi}$ we choose? By the lemma above, it does not depend on the choice of the extension and so $D_\mu \psi$ is well defined for all $\psi \in \mathrm{Lip}(\Omega \times [0,1])$. Moreover all the previous properties remain true.

Lemma 4.11. *Let B a borelian of $\mathbb{R}^{d+1}$. Then we have the implication:*

$$\left.\begin{array}{c} u \text{ Lipschitzian on } \mathbb{R}^{d+1} \\ u = 0 \ \mu - a.e. \text{ in } B \end{array}\right\} \Rightarrow D_\mu u = 0 \quad \mu - \text{almost everywhere.}$$

Proof. Without any restriction, we can assume $u \geq 0$. By the generalized coarea formula proved in [3], for any borelian function $g : \mathbb{R} \to \mathbb{R}^+$, it holds:

$$\int_{\mathbb{R}^{d+1}} g(u)|D_\mu u| \, d\mu = \int_0^{+\infty} g(t) \mathrm{per}_\mu\{u > t\} \, dt.$$

Taking $g = \mathbb{1}_{\{t=0\}}$, leads:

$$\int_B |D_\mu u| \, d\mu \leq \int_{\{u=0\}} |D_\mu u| \, d\mu = \int_0^{+\infty} g(t) \, dt = 0.$$

□

Definition 4.12. *Let $\psi \in \mathrm{Lip}(\Omega \times [0,1])$, the tangential gradient of ψ is defined by:*

$$D_\mu \psi = D_\mu \tilde{\psi}$$

where $\tilde{\psi}$ is any extension of ψ to $\mathbb{R}^{d+1}$.

The following lemma will be very useful in the next section:

Lemma 4.13. *Let $C \subset \mathbb{R}^{d+1}$ a closed convex set. Let $\psi \in \mathrm{Lip}(\Omega \times [0,1])$ such that $D\psi(y) \in C$ for all $y \in \Omega \times [0,1]$. Then it satisfies:*

$$D_\mu \psi(y) \in C + T_\mu^\perp(y) \quad \mu - a.e. \ y \in \Omega \times [0,1].$$

Proof. By Lemma 3.6, we can construct a sequence $(\psi_n)_n$ of $\mathcal{C}^\infty(\mathbb{R}^{d+1})$ with the same Lipschitz constant as ψ and such that $D\psi_n(y) \in C$ for all $y \in \Omega \times [0,1]$. For all $n \in \mathbb{N}$ and $\mu-$almost all $y \in Q$, we make the following decomposition (cf Proposition 4.8):

$$D\psi_n(y) = D_\mu \psi_n(y) + \eta_n(y)$$

where $\eta_n(y) \in T_\mu(y)^\perp$. We have $D_\mu \psi_n(y) = D\psi_n(y) - \eta(y) \in C + T_\mu(y)^\perp$. Moreover, $D_\mu \psi_n$ is bounded uniformly by the Lipschitz constant of ψ, so up to a subsequence, it admits a limit for the weak star topology of $\mathrm{L}^\infty_\mu(\mathbb{R}^{d+1})$, by definitions 4.6 and 4.12, the restriction of its limit to $\Omega \times [0,1]$ is $D_\mu \psi$.
Now, on the one hand $D_\mu \psi_n(y) \in C + T_\mu(y)^\perp$ $\mu-$almost everywhere, and, on the other hand $C + T_\mu(y)$ is a closed convex set, consequently passing to the weak limit:

$$D_\mu \psi(y) \in C + T_\mu(y)^\perp \quad \mu - \text{a.e. } y \in \Omega \times [0,1].$$

□

5. Extremality condition

We assume that the boundary of $\Omega \times [0,1]$ is Lipschitz. Returning to (4.1), the notion introduced in the above section will allow us to get a system of partial differential equations. In other words, we give the corresponding to Theorem 1.1 in our case:

Theorem 5.1.
i) Let $\psi \in \mathrm{Lip}(\Omega \times [0,1])$ *and* $\chi \in \mathcal{M}_b(\mathbb{R}^{d+1}, \mathbb{R}^{d+1})$ *solutions of* $(\mathcal{Q}^*)$ *and* $(\mathcal{Q})$ *respectively. Then, for all* $\mu \in \mathcal{M}_b^+(\mathbb{R}^{d+1})$, $\sigma \in \mathrm{L}^1_\mu(\Omega \times [0,1], \mathbb{R}^{d+1})$ *such that* $\chi = \sigma\mu$, *it holds:*

$$H(\sigma(x,t)) = \sigma(x,t) \cdot D_\mu\psi(x,t) \quad \mu - a.e.(x,t) \in \Omega \times [0,1].$$

ii) Conversely, let (σ, μ, ψ) *such that* $\psi \in \mathrm{Lip}(\Omega \times [0,1])$, $\mu \in \mathcal{M}_b(\mathbb{R}^{d+1})$ *with* $\mathrm{spt}\,\mu \subset \Omega \times [0,1]$ *and* $\sigma \in \mathrm{L}^1_\mu(\Omega \times [0,1], \mathbb{R}^{d+1})$, *we assume:*

$$(MK_t) \quad \begin{cases} a)\ D\psi(x,t) \in K \ \ a.e.(x,t), \\ b)\ -\mathrm{div}_{x,t}(\sigma\mu) = f_1 \otimes \delta_1 - f_0 \otimes \delta_0 \quad in \quad \mathbb{R}^{d+1}, \ \textit{(Diffusion equation)} \\ c)\ H(\sigma(x,t)) = \sigma(x,t) \cdot D_\mu\psi(x,t) \quad \mu - a.e.(x,t) \in \Omega \times [0,1]. \end{cases}$$

Then ψ *is a solution of* $(\mathcal{Q}^*)$ *and* $\sigma\mu$ *is a solution of* $(\mathcal{Q})$.

Proof. i) Let ψ and $\chi = \sigma\mu$ as in Theorem 5.1. Now recall the computation we have made at the beginning of section 4, the main tool of this computation was Theorem 3.8 and it leads to (4.1). Proposition 4.10 allows us to make an integration by parts and write:

$$\int H(\sigma(x,t))d\mu(x,t) = \int_{\Omega\times[0,1]} D_\mu\psi(x,t) \cdot \sigma(x,t)d\mu(x,t). \tag{5.1}$$

By Lemma 4.13:

$$D_\mu\psi(x,t) \in K + T_\mu(x,t)^\perp \quad \mu - \text{a.e.}(x,t). \tag{5.2}$$

So is to say, it exists $\eta \in \mathcal{N}^\perp$ (Proposition and Definition 4.3) such that:

$$D_\mu\psi(x,t) + \eta(x,t) \in K \quad \mu - \text{a.e.}(x,t). \tag{5.3}$$

Hence, as by Lemma 4.9 $\sigma(x,t) \in T_\mu(x,t)$, (5.1) implies:

$$\int_{\Omega\times[0,1]} H(\sigma(x,t)) - (D_\mu\psi(x,t) + \eta(x,t)) \cdot \sigma(x,t)d\mu(x,t) = 0. \tag{5.4}$$

Well then, by (2.3), we have:

$$H(\sigma(x,t)) - (D_\mu\psi(x,t) + \eta(x,t)) \cdot \sigma(x,t) \geq 0.$$

Combining this with (5.4), leads to:

$$(D_\mu\psi(x,t) + \eta(x,t)) \cdot \sigma(x,t) = H(\sigma(x,t)) \quad \mu - \text{a.e.}(x,t).$$

Finally, as $\eta(x,t) \in T_\mu(x,t)^\perp$ and $\sigma(x,t) \in T_\mu(x,t)$ $\mu-$almost everywhere:

$$D_\mu\psi(x,t) \cdot \sigma(x,t) = H(\sigma(x,t)) \quad \mu - \text{a.e.}(x,t).$$

ii) Let (σ, μ, ψ) as in Theorem 5.1. Let us show that ψ is solution of $(\mathcal{Q}^*)$ and $\sigma\mu$ a solution of $(\mathcal{Q})$. Notice first that ψ is admissible for $(\mathcal{Q}^*)$ and $\sigma\mu$ admissible for $(\mathcal{Q})$. By Theorem 3.8, it is sufficient to show:

$$< f_1 \otimes \delta_1 - f_0 \otimes \delta_0, \psi >= \int_{\Omega\times[0,1]} H(\sigma(x,t))\ d\mu(x,t). \tag{5.5}$$

The assumption $-\text{div}_{x,t}(\sigma\mu) = f_1 \otimes \delta_1 - f_0 \otimes \delta_0$ combined with Proposition 4.10 implies

$$< f_1 \otimes \delta_1 - f_0 \otimes \delta_0, \psi >= \int_{\Omega\times[0,1]} D_\mu\psi \cdot \sigma d\mu.$$

Consequently, using the assumption $H(\sigma(x,t)) = D_\mu\psi(x,t) \cdot \sigma(x,t)$ $\mu-$almost everywhere, we get (5.5). ☐

The equation $(MK_t.c)$ can be viewed as an eikonal equation, more precisely:

Proposition 5.2. *The equation $(MK_t.c)$ is equivalent to:*

$$\sigma(x,t) \in N_{K_\mu(x,t)}(D_\mu\psi(x,t))\ \mu - a.e.(x,t) \in \Omega \times [0,1] \tag{5.6}$$

where $K_\mu(x,t) = P_\mu((x,t),K)$ μ-almost every (x,t) and $N_{K_\mu(x,t)}(y,s)$ is the normal cone of $K_\mu(x,t)$ at (y,s).

When μ is the Lebesgue measure and K is the sphere of radius one, we get the classical eikonal equation: $|D\psi(x,t)| = 1$ for almost every (x,t).

Remark 5.3. *The problem $(\mathcal{Q})$ does not have a unique solution in general. Moreover, when F is a positive function (except at 0), it is always possible to find a solution $\chi = \sigma\mu$ such that $H(\sigma(x,t)) = 1$ $\mu-$almost everywhere. Indeed, take χ given by*

$$< \chi, \Phi >:= \int_{\Omega^2}\int_0^1 \Phi((1-s)x_0 + sx_1, s) \cdot (x_1 - x_0, 1)\ dsd\gamma(x_0,x_1), \tag{5.7}$$

where γ is any solution of problem $(\mathcal{P})$ and $\Phi \in \mathcal{C}_c(\mathbb{R}^{d+1},\mathbb{R}^{d+1})$. Let us denote by μ the following measure:

$$\mu = H(\chi),\ \textit{where}\ [H(\chi)](A) = \sup\left\{< \chi, \Phi >:\ \Phi \in \mathcal{C}_c(A,\mathbb{R}^{d+1}),\ \Phi(x,t) \in K\ \forall(x,t)\right\},$$

so is to say $< \mu, \varphi >:= \int_{\Omega^2}\int_0^1 \varphi((1-s)x_0 + sx_1, s)F(x_1 - x_0)\ dsd\gamma(x_0,x_1)$, for any borelian set $A \subset \mathbb{R}^{d+1}$ and $\varphi \in \mathcal{C}_c(\mathbb{R}^{d+1})$. It satisfies $|\chi| << \mu$ and $H\left(\frac{d\chi}{d\mu}(x,t)\right) = 1$ $\mu-$almost everywhere.

Remark 5.4. *The equation (5.6) implies that $\mu-$almost everywhere on the set*

$$\{(x,t) \in \Omega \times \mathbb{R}^+ :\ \sigma(x,t) \neq 0\},$$

the tangential gradient $D_\mu\psi(x,t)$ lies on the boundary of $K_\mu(x,t)$ (because the normal cone at $D_\mu\psi(x,t)$ is not reduced to 0).

6. Examples

Take $a > 0$ and consider $\Omega = [0,a] \subset \mathbb{R}$, $f_0 = \delta_0$ $f_1 = \delta_a$.

In this case we have $\inf(\mathcal{P}) = F(a)$, $\gamma = \delta_{(0,a)}$ is the only admissible measure and it is optimal.

Using formula (5.7), we can associate to this γ a solution of $(\mathcal{Q})$:

$$\chi_0 = \left(\frac{a}{\sqrt{1+a^2}}, \frac{1}{\sqrt{1+a^2}}\right) H^1 \llcorner S_0$$

where $S_0 := [(0,0),(a,1)]$. It is easy to see that if H is strictly convex, this solution is the unique solution. If it is not strictly convex, we can, for example look for solutions of type $\frac{\dot v(t)}{|\dot v(t)|}\mathcal{H}\llcorner(v([0,1]))$ with $v(0)=(0,0)$ and $v(1)=(a,1)$ and v continuous and $\mathcal{C}^1$ by parts. This reduces to find solutions of the following problem:

$$\inf\left\{\int_0^1 H(\dot v(t))\ dt:\ v(0)=(0,0),\ v(1)=(a,1),\ v\ \mathcal{C}^1 \text{ by parts}\right\}. \tag{6.1}$$

Taking v a right line and using Jensen inequality, it is easy to see χ_0 is a solution of (6.1) and:

$$\min(\mathcal{Q})=\min(6.1)=F(a).$$

Moreover, we can exhibit a solution of $(\mathcal{Q}^*)$:

$$\psi_o(x,t):=\left\{\begin{array}{cl} tF\left(\frac{x}{t}\right) & \text{if } 0\le x\le at \text{ and } t\neq 0,\\ F(a)-(1-t)F\left(\frac{a-x}{1-t}\right) & \text{if } at\le x\le a \text{ and } t\neq 1.\end{array}\right. \tag{6.2}$$

In the rest of this section, we give some solutions of $(\mathcal{Q})$ and we try to interpret the condition the optimal condition (5.6) in the two following cases:
1. $F(z)=|z|$, (linear case) 2. $F(z)=|z|^2$, (quadratic case).

6.1. **Case $F(z)=|z|$.** In this case, we have:

$$H(z,t)=\left\{\begin{array}{lll}|z| & \text{if} & t\ge 0,\\ +\infty & \text{if} & t<0,\end{array}\right. \quad F^*(z^*)=\left\{\begin{array}{ll}0 & \text{if } |z^*|\le 1,\\ +\infty & \text{otherwise,}\end{array}\right.$$

and $K=\{(z,s):\ |z|\le 1 \text{ and } s\le 0\}$.

Then, any path $v([0,1])$ ($v=(v_1,v_2)$ with $\dot v_1>0$, $\dot v_2>0$) is a solution of (6.1), so is to say, any measure of the form $\chi_v=\frac{\dot v(t)}{|\dot v(t)|}\mathcal{H}\llcorner(v([0,1]))$ is optimal for $(\mathcal{Q})$. More generally, considering a family of paths $\{S_\alpha\}_{\alpha\in A}$, where A is a set equipped with a probability P, we can give a new solution χ of $(\mathcal{Q})$ defined by

$$<\chi,\Phi>=\int_A <\chi_{v_\alpha},\Phi>\ dP(\alpha)\quad \forall\Phi\in\mathcal{C}_o(Q,\mathbb{R}^{d+1}).$$

We now give sense to the optimality condition (5.6). Note first that the particular solution ψ_o of $(\mathcal{Q}^*)$ given by (6.2) becomes $\psi_o(x,t)=x$. It is a regular solution so for all vectorial measure μ:

$$D_\mu\psi_o(x,t)=P_\mu((x,t),D\psi_o(x,t))=P_\mu((x,t),(1,0)).$$

Condition (5.6) with $\psi=\psi_o$ then reads as:

$$\sigma(x,t)\in N_{K_\mu(x,t)}(P_\mu((x,t),(1,0)))\ \mu-\text{a.e. } (x,t)\in\Omega\times[0,1]$$

where K_μ is the projection on $T_\mu(x,t)$ of the convex K. Moreover, remember that $\sigma(x,t)$ belongs to $T_\mu(x,t)$ (cf Lemma 4.9) so we deal with the condition:

$$\sigma(x,t)\in T_\mu(x,t)\cap N_{K_\mu(x,t)}(P_\mu((x,t),(1,0)))\ \mu-\text{a.e. } (x,t)\in\Omega\times[0,1]. \tag{6.3}$$

Now, as $T_\mu(x,t)$ is a vectorial subspace of $\mathbb{R}^2$, we may consider the μ-measurable sets $E_1=\{(x,t):\ T_\mu(x,t)=\mathbb{R}^2\}$ and $E_2=\{(x,t):\ T_\mu(x,t) \text{ is a line containing } 0\}$.

On E_1, we have $P_\mu((x,t)(y,s))=(y,s)$ so $K_\mu(x,t)=K$, $P_\mu((x,t),(1,0))=(1,0)$ and (see figure 1):

$$T_\mu(x,t)\cap N_{K_\mu(x,t)}(P_\mu((x,t),(1,0)))=N_K(1,0)=\{(z_1,z_2):\ z_1\ge 0,\ z_2\ge 0\}=\mathbb{R}^2_{++}.$$

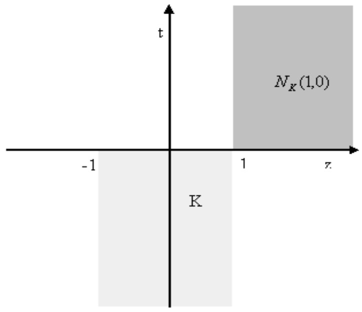

FIGURE 1. Case $T_\mu(x,t) = \mathbb{R}^2$.

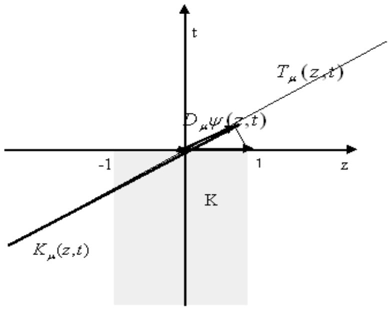

FIGURE 2. Case $T_\mu(x,t)$ is a line intersecting $\mathbb{R}^2_{++}$.

Finally, (6.3) reads as

$$\sigma(x,t) \in \mathbb{R}^2_{++}.$$

Let us study the case $(x,t) \in E_2$. Let us set $T_\mu(x,t) = \{\lambda(\cos\theta, \sin\theta) : \lambda \in \mathbb{R}\}$ with $\cos\theta > 0$. For $\mu-$almost every (x,t) and for all (y,s) we have:

$P_\mu((x,t),(y,s)) = (y\ \cos\theta + s\ \sin\theta)(\cos\theta, \sin\theta)$ and $P_\mu((x,t),(1,0)) = \cos\theta(\cos\theta, \sin\theta)$.

Then, by a simple computation, we get:

$$K_\mu(x,t) = P_\mu((x,t),K) = \{\lambda(\cos\theta, \sin\theta) : \lambda \leq \cos\theta\}, \quad \text{if } \sin\theta > 0,$$

$$K_\mu(x,t) = P_\mu((x,t),K) = \{\lambda(\cos\theta, \sin\theta) : \lambda \geq -\cos\theta\}, \quad \text{if } \sin\theta < 0,$$

$$K_\mu(x,t) = P_\mu((x,t),K) = [(-1,0),(1,0)] \quad \text{if } \sin\theta = 0.$$

In case $\sin\theta \leq 0$ (so is to say $T_\mu(x,t)$ intersects $\mathbb{R}^2_{++}$), we have (see figure 2):

$$\begin{aligned} T_\mu(x,t) \cap N_{K_\mu(x,t)}(P_\mu((x,t),(1,0))) &= \{\lambda(\cos\theta, \sin\theta) : \lambda\cos\theta = \max_{r \geq \cos\theta} \lambda r\} \\ &= \{\lambda(\cos\theta, \sin\theta) : \lambda \geq 0\} = \mathbb{R}^2_{++}. \end{aligned}$$

Finally, also in the case $T_\mu(x,t)$ is a line intersecting $\mathbb{R}^2_{++}$, (6.3) reads as:

$$\sigma(x,t) \in \mathbb{R}^2_{++}.$$

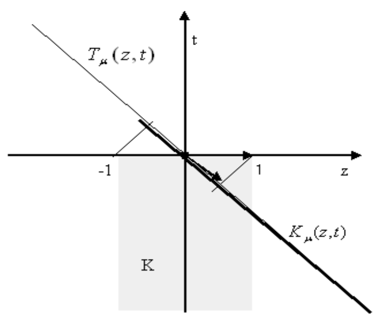

FIGURE 3. Case $T_\mu(x,t)$ is a line not intersecting $\mathbb{R}^2_{++}$.

In case $\sin\theta < 0$ (so is to say $T_\mu(x,t)$ does not intersect $\mathbb{R}^2_{++}$), we have (see figure 3):

$$T_\mu(x,t)\cap N_{K_\mu(x,t)}(P_\mu((x,t),(1,0))) = \{\lambda(\cos\theta,\sin\theta):\ \lambda\cos\theta = \max_{r\le -\cos\theta} r\lambda\} = \{(0,0)\}.$$

This leads to $\sigma(x,t) = (0,0)$. As we can always assume $\sigma \neq 0$ on a set on which μ is concentrated, this case is not relevant.

We can conclude saying that the following equivalences hold:

$$\chi = \sigma\mu \text{ is a solution of } (\mathcal{Q}) \Leftrightarrow \left\{ \begin{array}{c} -\mathrm{div}_{x,t}\chi = \delta_{(a,1)} - \delta_{(0,0)} \\ (\sigma,\mu) \text{ satisfies } (MK_t.c) \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{c} -\mathrm{div}_{x,t}\chi = \delta_{(a,1)} - \delta_{(0,0)} \\ \sigma(x,t) \in \mathbb{R}^2_{++}. \end{array} \right.$$

6.2. **Case $F(z) = |z|^2$.** (see figure 4).
In this case, we have:

$$H(x,t) := \begin{cases} \frac{|x|^2}{t} & \text{if } t > 0, \\ 0 & \text{if } x = t = 0, \\ +\infty & \text{otherwise;} \end{cases}$$

and $F^*(x^*) = \frac{|x^*|^2}{4}$, $K = \left\{(x,t):\ t \le -\frac{|x|^2}{4}\right\}$. The cost H being strictly convex, the problem $(\mathcal{Q})$ admits a unique solution which is given by:

$$\mu = \mathcal{H}^1 \llcorner [(0,0),(a,1)], \quad \sigma = \frac{(a,1)}{\sqrt{a^2+1}},$$

$$T_\mu(x,t) = \{(as,s):\ s \in \mathbb{R}\}, \quad \mu - \text{a.e. } (x,t) \in \mathrm{spt}(\mu) = [(0,0),(a,1)].$$

Let us check that it satisfies (5.6). We have

$$\psi_o(x,t) = \begin{cases} \frac{x^2}{t} & \text{if } \ 0 \le x \le at \text{ and } t \neq 0, \\ a^2 - \frac{(a-x)^2}{1-t} & \text{if } \ at \le x \le a \text{ and } t \neq 1 \end{cases}$$

$$\text{and } D\psi_o(x,t) = \begin{cases} (\frac{2x}{t}, -\frac{x^2}{t}) & \text{if } \ 0 \le x \le at \text{ and } t \neq 0, \\ (\frac{2(a-x)}{1-t}, -\frac{(a-x)^2}{(1-t)^2} & \text{if } \ at \le x \le a \text{ and } t \neq 1. \end{cases}$$

The tangent space is $T_\mu(x,t) := \{(as,s):\ s \in \mathbb{R}\}$ $\mu-$almost every (x,t), and the projection on this space is given by: $P_\mu((x,t),(y,s)) = \frac{ya+s}{a^2+1}(a,1)$ for all (y,s). Then, as $D\psi_o$ is regular, we have:

$$D_\mu\psi_o(x,t) = P_\mu((x,t),\psi_o(x,t)) = \frac{a^2}{1+a^2}(a,1) \quad \mu - \text{a.e. } (x,t) \in [(0,0),(a,1)].$$

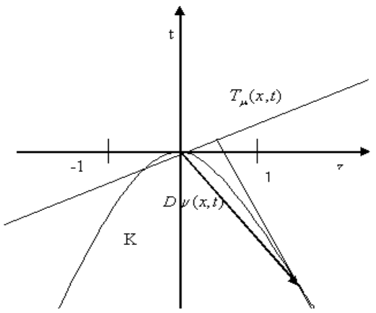

FIGURE 4. Quadratic case.

We can easily compute that:

$$\begin{aligned} K_\mu(x,t) = P_\mu((x,t),K) &= \{\frac{ya+s}{a^2+1}(a,1):\ s \leq \frac{-y^2}{4}\} \\ &= \{\lambda(a,1):\ \lambda \leq \frac{a^2}{a^2+1}\}, \end{aligned}$$

$$\begin{aligned} \text{and } N_{K_\mu(x,t)}(\frac{a^2}{1+a^2}(a,1)) &= \{(y,s):\ \frac{a}{1+a^2}(ay+s) = \max_{\lambda \leq \frac{a^2}{a^2+1}} \lambda(ay+s)\} \\ &= \{(y,s): ay+s \geq 0\}. \end{aligned}$$

Then, obviously the following condition is satisfied:

$$\sigma(x,t) \in N_{K_\mu(x,t)}(D\mu(\psi_o(x,t)).$$

Acknowledgements: The author acknowledge gratefully Guy Bouchitté for having given her the idea of this work. Sincere thanks go to him and Thierry Champion for their help and the fruitful discussions.

REFERENCES

[1] D. Azé, *Eléments d'Analyse Convexe et Variationnelle*, Ellipse, (1997).
[2] G. Barles, *Solutions de viscosité des équations de Hamilton-Jacobi*, Mathématiques et Applications (Berlin), 17. Springer-Verlag, Paris, (1994).
[3] G. Bellettini, G. Bouchitté, I. Fragalà, *BV functions with respect to a measure and relaxation of metric integral functionals*, J. Convex Anal. 6 (1999), no. 2, 349–366.
[4] P. Bernard, B. Buffoni, *Optimal mass transportation and Mather theory* to appear in Journal of the European Mathematical Society.
[5] Y. Brenier, Extended Monge-Kantorovich Theory, *Optimal transportaion and applications* (Martina Franca-2001), 91-121 Lecture Notes In Math., 1813, Springer Verlag, Berlin (2003).
[6] J. Buklew, G. Wise, *Multidimensional Asymptotic Quantization Theory with rth Power Distortion Measures*, IEEE Inform. Theory 28 (2) (1982) 239-247.
[7] G. Bouchitté, G. Buttazzo, *Characterization of optimal shapes through Monge-Kantorovich equation*, J. Eur. Math. Soc., 3 no. 2, (2001), 139-168.
[8] G. Bouchitté, G. Buttazzo, P. Seppecher, *Energies with Respect to a Measure and Applications to Low Dimensional Structures*, Calc. Var. Partial Differential Equations 5, no. 1, (1997), 35-54.
[9] G. Bouchitté, G. Buttazzo, P. Seppecher, *Shape Optimization Solutions via Monge-Kantorovich Equation*, C. R. Acad. Sci. Paris Sér. I Math. 324 (1997), no. 10, 1185–1191.

[10] G. Bouchitté, I. Fragalà, *Variational theory of weak geometric structures: the measure method and its applications.* Variational methods for discontinuous structures, 19–40, Progr. Nonlinear Differential Equations Appl., 51, Birkhäuser, Basel, 2002.
[11] G. Bouchitté, M. Valadier, *Integral Representation of Convex Functionals on a space of Measures*, Journal of Functional Analysis, 80, no.2 (1988), 398-420.
[12] G. Buttazzo, E. Oudet, E. Stepanov, *Optimal Transportation Problems with Free Dirichlet Regions*, Preprint (2002).
[13] G. Carlier, F. Santambrogio, *A Variational Model for Urban Planning with Traffic Congestion*, CVGMT preprint, Scuola Normale Superiore, Pisa (2004).
[14] B. Dacorogna, *Direct methods in the calculus of variations,* Applied Mathematical Sciences, 78. Springer-Verlag, Berlin, (1989).
[15] L. C. Evans, *Partial Differential Equation*, AMS, Graduate Studies In Mathematics, (1998).
[16] L. C. Evans, W. Gangbo, *Differential equations methods for the Monge-Kantorovich mass transfer problem*, Mem. Amer. Math. Soc. 137, no. 653, viii+66 pp., (1999).
[17] I. Fragalà, *Tangential Calculus and Variational Integrals with Respect to a Measure*, Phd Thesis, SNS of Pisa, (2000) avalaible at http://cvgmt.sns.it/people/fragala/.
[18] L. Granieri, *On action minimizing measures for the Monge-Kantorovich Problem*, to be published in Nonlinear Differential Equations Appl.
[19] F. Hiai, H. Umegaki, *Integrals, conditonal expectations and martingales functions*, J. Multivariate An. 7, 149-182, (1977).
[20] J.-B. Hiriart-Urruty, C. Lemaréchal, *Convex Analysis and Minimization Algorithms, I and II*, Springer-Verlag, Berlin, 1993.
[21] C. Jimenez, Optimisation de problèmes de transport, PhD Thesis, Univ. du Sud-Toulon Var, (2005), available at http://www.ceremade.dauphine.fr/ jimenez.
[22] L. V. Kantorovich, *On the transfer of masses*, Dokl. Akad. Nauk. SSSR 37 227-229, (1942).
[23] S. Rachev, L. Rüschendorf, *Mass Transportation Problems. Vol. I: Theory, Vol. II: Applications.* Probability and its applications. Springer-Verlag, New York, 1998.
[24] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, N. J., 1970.
[25] M. Valadier, *Quelques propriétés de l'ensemble des sections continues d'une multifonction s.c.i.,* Seminaire d'analyse convexe, Vol. 16, Exp. No. 7, 10 pp., Univ. Sci. Tech. Languedoc, Montpellier, (1986).
[26] C. Villani, *Topics in Optimal transportation*, Graduate studies in Mathematics, 58, AMS (2003).

Scuola Normale Superiore di Pisa, Piazza dei Cavalieri, 7, 56126 Pisa, Italy
E-mail address: c.jimenez@sns.it