PATH FUNCTIONALS OVER WASSERSTEIN SPACES

A. BRANCOLINI, G. BUTTAZZO AND F. SANTAMBROGIO

Abstract. Given a metric space X we consider a general class of functionals which measure the cost of a path in X joining two given points x_0 and x_1, providing abstract existence results for optimal paths. The results are then applied to the case when X is a Wasserstein space of probabilities on a given set Ω and the cost of a path depends on the value of classical functionals over measures. Conditions to link arbitrary extremal measures μ_0 and μ_1 by means of finite cost paths are given.

1. Introduction

Finding the best way to carry a given source onto a given target is a problem that received a lot of attention in recent years. The literature on Monge-Kantorovich mass transportation problems and on Wasserstein distances is very rich, as well as on the several related applications which include the shape optimization of elastic bodies (see for instance [7]), the design of public transportation networks (see for instance [9], [10], and [12]), the optimal location of production centres (one can see [8] for the asymptotics of the problem) and the study of irrigation trees (see for instance [3], and [14]).

The problem of transporting a source mass distribution onto a target mass distribution by keeping together as much mass as possible during the transport, from which tree-shaped configurations arise, has been very much studied, for instance in [3], [14], and [18]. In our new approach to this problem probability measures valued curves are considered, while the condition of keeping masses together is achieved considering only measures supported in discrete sets.

Given a source or initial probability measure μ_0 and a target or final probability measure μ_1 we look for a path γ in a Wasserstein space $\mathcal{W}_p(\Omega)$ that connects μ_0 to μ_1 and minimizes a suitable cost functional $\mathcal{J}(\gamma)$. In the present paper we consider functionals of the form

$$\mathcal{J}(\gamma) = \int_0^1 J(\gamma(t))|\gamma'|(t)\mathrm{d}t \tag{1.1}$$

where $|\gamma'|$ is the metric derivative of γ in the Wasserstein space $\mathcal{W}_p(\Omega)$ and J is a lower semicontinuous functional defined on measures. Here J may be easily seen as the coefficient of a degenerate "Riemannian distances" on the space $\mathcal{W}_p(\Omega)$.

We restrict our analysis to the case of J being a local functional over measures, an important class of functionals extensively studied by Bouchitté and Buttazzo in [4], [5], and [6]. These functionals are the key tool in our approach, and among them we can find both functionals which are finite only on concentrated measures (we can see an application of them in [11] and [16]) and functionals which are finite only on spread ones. In fact, a particular point of interest in our approach is the fact that also different kinds of "Riemannian distances" are allowed (for instance those which prefer spread measures) by a change of the functional J.

In particular, we consider here the two extreme cases, in which the functional J is chosen as one of the following:

$$G_r(\mu) = \begin{cases} \sum_{k\in\mathbb{N}}(a_k)^r & \text{if } \mu = \sum_{k\in\mathbb{N}} a_k\delta_{x_k} \\ +\infty & \text{otherwise} \end{cases} \qquad (0 \le r < 1)$$

whose domain is the space of purely atomic measures, or

$$F_q(\mu) = \begin{cases} \int_\Omega |u|^q \mathrm{d}x & \text{if } \mu = u\cdot\mathcal{L}^N \\ +\infty & \text{otherwise} \end{cases} \qquad (q > 1)$$

whose domain is the space $L^q(\Omega)$. We denote respectively by $\mathcal{G}_r$ the functional in (1.1) with J replaced by G_r and by $\mathcal{F}_q$ the same functional with J replaced by F_q.

The first case is the one in which we get a "Riemannian distance" on probabilities which makes paths passing through concentrated measures cheaper. The second case, on he contrary, allows only paths which lie on $L^q(\Omega)$.

In both cases we analyze the question of the existence of optimal paths γ_{opt} giving finite value to the functional. When the domain $\Omega \subset \mathbb{R}^N$ is compact we find for the first case:

- if μ_0 and μ_1 are atomic measures, then an optimal path γ_{opt} providing finite value to $\mathcal{G}_r$ always exists;
- if $r > 1 - 1/N$, then the same is true for any pair of measures;
- if $r \le 1 - 1/N$, then there are measures μ_0 and μ_1 such that every path connecting them has an infinite cost.

Similarly, for the second case we have:

- if μ_0 and μ_1 are in $L^q(\Omega)$, then an optimal path γ_{opt} providing finite value to $\mathcal{F}_q$ always exists;
- if $q < 1 + 1/N$, then the same is true for any pair of measures;
- if $q \ge 1 + 1/N$, then there are measures μ_0 and μ_1 such that every path connecting them has an infinite cost.

In Section 4 we also discuss the case of unbounded domains such as $\Omega = \mathbb{R}^N$.

The analysis of existence results as well as the definition of the cost functionals is made in Section 2 in an abstract metric spaces framework, which can be used for future generalizations and developments.

In relation to the other papers mentioned at the beginning of this introduction it is not difficult to see that the model we propose is different and in general provides different solutions. However, among the different features our model supplies we may cite its mathematical simplicity and the possibility of performing standard numerical computations. From the mathematical point of view, our model recalls the construction of Riemannian metrics, as we already pointed out, and in connection to this we obtain rather easy existence results for optimal paths. As far as numerics is concerned, we stress the fact that, when discretizing the metric derivative, the cost functional becomes a weighted sum of Wasserstein distances among couples of atomic probability measures which can be evaluated by well-known algorithms such as the simplex method.

Anyway, we think that taking into account the comparison with the results presented by Xia in [18] and by Maddalena, Morel and Solimini in [14] will be important for future investigations. For instance, for the model proposed in [14] conditions to link two prescribed measures by a finite cost configuration have been studied in [13] (while here and in [18] only conditions in order to link arbitrary measures are provided): we do not know if similar conditions can be achieved in our case.

2. The Metric Framework

In this section a generic metric space X with distance d is considered. Under the assumption that closed bounded subsets of X are compact, we will prove an existence result (Theorem 2.1) for variational problems with functionals of the type

$$\mathcal{J}(\gamma) = \int_0^1 J(\gamma(t))|\gamma'|(t)\ \mathrm{d}t$$

where $\gamma : [0,1] \to X$ ranges among all Lipschitz curves such that $\gamma(0) = x_0$ and $\gamma(1) = x_1$. We will refer to the value of $\mathcal{J}$ in γ as the *energy* of γ. By $|\gamma'|(t)$ we denote the *metric derivative* of γ at the point $t \in (0,1)$, i.e.

$$|\gamma'|(t) = \lim_{s\to t} \frac{d(\gamma(s),\gamma(t))}{|s-t|}.$$

As a consequence of Rademacher Theorem it can be seen (see [2]) that for any Lipschitz curve the metric derivative exists in almost every point (with respect to Lebesgue measure). Another useful result is that the variation of γ can be written in terms of the metric derivative in integral form:

$$\mathrm{Var}(\gamma) = \int_0^1 |\gamma'|(t)\ \mathrm{d}t.$$

By this formula it follows easily that $|\gamma'| \le M$ if and only if γ is M-Lipschitz, since when $s < t$

$$d(\gamma(t), \gamma(s)) \le \mathrm{Var}(\gamma, [s,t]) = \int_s^t |\gamma'|(\tau)\ \mathrm{d}\tau \le M|t-s|,$$

the converse implication being immediate.

Theorem 2.1. *Let X be a metric space such that any closed bounded subset of X is compact, $J : X \to [0, +\infty]$ be a lower semicontinuous function and x_0, x_1 be arbitrary points in X. Then the functional*

$$\mathcal{J}(\gamma) = \int_0^1 J(\gamma(t))|\gamma'|(t)\ \mathrm{d}t$$

achieves a minimum value among all Lipschitz curves $\gamma : [0,1] \to X$ such that $\gamma(0) = x_0$ and $\gamma(1) = x_1$, provided the following two assumptions are satisfied:

(H1): *there exists a curve γ_0 such that $\mathcal{J}(\gamma_0) < +\infty$;*

(H2): *it holds*

$$\int_0^\infty \inf_{B_r(x_0)} J\ \mathrm{d}r = +\infty.$$

The proof of Theorem 2.1 relies on the following reparametrization lemma whose proof can be found for example in [2].

Lemma 2.2. *Let $\gamma \in \mathrm{Lip}([0,1], X)$ and $L = \mathrm{Var}(\gamma)$ be its total variation. Then there exists a Lipschitz curve $\tilde{\gamma} \in \mathrm{Lip}([0,L], X)$ such that $|\tilde{\gamma}'| = 1$ almost everywhere in $[0,L]$ and $\tilde{\gamma}$ is a parametrization of γ.*

Proof of Theorem 2.1. Let $\{\gamma_n\}_{n\in\mathbb{N}}$ be a minimizing sequence and set $L_n = \mathrm{Var}(\gamma_n)$. Then the sequence $\{\mathcal{J}(\gamma_n)\}_{n\in\mathbb{N}}$ is bounded by a finite number M. By Lemma 2.2 there exists a sequence of curves $\tilde{\gamma}_n : [0, L_n] \to X$ parametrized with unit velocity, reparametrizing the given curves. We have:

$$M \ge \mathcal{J}(\gamma_n) = \int_0^{L_n} J(\tilde{\gamma}_n(t))\ \mathrm{d}t \ge \int_0^{L_n} \left(\inf_{B_t(x_0)} J\right)\ \mathrm{d}t.$$

Then $\{L_n\}_{n\in\mathbb{N}}$ is bounded otherwise, by assumption H2, the right hand side would be unbounded. We can reparametrize each curve γ_n at constant speed L_n, thus obtaining a new sequence $\{\hat{\gamma}_n\}_{n\in\mathbb{N}}$ in $\mathrm{Lip}([0,1], X)$, which is still a minimizing sequence, thanks to the equality $\mathcal{J}(\gamma_n) = \mathcal{J}(\hat{\gamma}_n)$. Being $\{L_n\}_{n\in\mathbb{N}}$ bounded, we get that this new minimizing sequence is uniformly bounded and uniformly Lipschitz. By Ascoli-Arzelà Theorem we can suppose up to a subsequence that $\hat{\gamma}_n \to \hat{\gamma}$ uniformly for some L-Lipschitz curve $\hat{\gamma}$ where we have taken $L = \liminf_n L_n$. By recalling the link between Lipschitz conditions and metric derivative we have

$$|\hat{\gamma}'|(t) \le L \quad \text{for a.e. } t \in [0,1].$$

Now by using the lower semicontinuity of the functional J, we obtain

$$\begin{aligned}\mathcal{J}(\hat{\gamma}) &= \int_0^1 J(\hat{\gamma}(t))|\hat{\gamma}'|(t)\ \mathrm{d}t \le L \int_0^1 \liminf_{n\to+\infty} J(\hat{\gamma}_n(t))\ \mathrm{d}t \\ &\le \liminf_{n\to+\infty} L_n \int_0^1 J(\hat{\gamma}_n(t))\ \mathrm{d}t = \liminf_{n\to+\infty} \mathcal{J}(\hat{\gamma}_n),\end{aligned}$$

that is the lower semicontinuity of $\mathcal{J}$ on the considered sequence, which achieves the proof. $\square$

Remark 2.3. Notice that the integral assumption H2 is always true if $J \ge c$ for a suitable strictly positive constant. Moreover Theorem 2.1 still holds if condition H2 is replaced by the weaker assumption that there exists a curve γ_0 such that

$$\mathcal{J}(\gamma_0) < \int_0^{+\infty} \inf_{B(x_0,r)} J\,\mathrm{d}r.$$

We give here a slightly refined version of Theorem 2.1, which will be useful in the last section. The goal here is to weaken the compactness assumption on bounded subsets of X.

Theorem 2.4. *Let (X, d, d') be a metric space endowed with two different distances, such that:*

- (K1): *$d' \le d$;*
- (K2): *all $d-$bounded sets in X are relatively compact with respect to d';*
- (K3): *the mapping $d : X \times X \to \mathbb{R}^+$ is a lower semicontinuous function with respect to the distance $d' \times d'$.*

Let $J : X \to [0, +\infty]$ be lower semicontinuous with respect to d'. Consider the functional, defined on the set of $d-$Lipschitz curves $\gamma : [0, 1] \to X$, given by

$$\mathcal{J}(\gamma) = \int_0^1 J(\gamma(t))|\gamma'|_d(t)\mathrm{d}t,$$

where $|\gamma'|_d(t)$ stands for the metric derivative of γ with respect to d. Then, with the same hypotheses H1 and H2 (where $B_r(x_0)$ are in the d-sense) of Theorem 2.1, there exists a minimum for $\mathcal{J}$.

Proof. We can take a minimizing sequence $\{\gamma_n\}_n$ and, as in Theorem 2.1, reparametrize it to obtain a sequence $\{\hat{\gamma}_n\}_n$ in which every curve has constant speed L_n. Hypothesis H2 gives us the boundedness of L_n. Hence the sequence $\{\hat{\gamma}_n\}_n$ is composed by $d-$equicontinuous functions from $[0, 1]$ to a $d-$bounded subset of X. If we endow X with the distance d' we have an equicontinuous (thanks to assumption K1) sequence of functions whose images are contained in a compact set. We can consequently use Ascoli-Arzelà Theorem to choose a subsequence (not relabeled), such that $\hat{\gamma}_n \to \gamma$, for a suitable curve γ (uniformly in the $d'-$sense). The lower semicontinuity of J with respect to d' allows us to use Fatou Lemma and shows that γ minimizes $\mathcal{J}$, as far as we can show that γ is $d-$Lipschitz with a Lipschitz constant not exceeding $\liminf_n L_n$. To do this we use assumption K3. Taken two points s, t we have in fact:

$$d(\gamma(s), \gamma(t)) \leq \liminf_n d(\hat{\gamma}_n(s), \hat{\gamma}_n(t)) \leq \liminf_n L_n|s-t|,$$

which shows the required Lipschitz property. □

3. The Case of Wasserstein Spaces

In this section we consider a compact metric space Ω equipped with a distance function c and a positive finite non-atomic Borel measure m. We consider the *p-Wasserstein metric space* $\mathcal{W}_p(\Omega)$. This is the space of Borel probability measures μ on Ω with finite momentum of order p with respect to a point x_0 (recall that the finiteness of the p-momentum with respect to a point implies the finiteness with respect to any other point)

$$\int_\Omega c(x, x_0)^p \, \mathrm{d}\mu < +\infty,$$

equipped with the *p-Wasserstein distance*

$$w_p(\mu_1, \mu_2) = \inf \left(\int_{\Omega\times\Omega} c(x,y)^p \, \lambda(\mathrm{d}x, \mathrm{d}y) \right)^{1/p}$$

where the infimum is taken on all transport plans λ between μ_1 and μ_2, that is on all probability measures λ on $\Omega \times \Omega$ whose marginals $\pi^+_\#\lambda$ and $\pi^-_\#\lambda$ coincide with μ_1 and μ_2 respectively.

Notice that, since the distance c is bounded, the space $\mathcal{W}_p(\Omega)$ consists of all probability measures. We consider functions J on $\mathcal{W}_p(\Omega)$ that can be represented in the following form:

$$J(\mu) = \int_\Omega f\left(\frac{\mathrm{d}\mu}{\mathrm{d}m}\right) \, \mathrm{d}m + \int_{\Omega\setminus A_\mu} f^\infty\left(\frac{\mathrm{d}\mu^s}{\mathrm{d}|\mu^s|}\right) \, \mathrm{d}|\mu^s| + \int_{A_\mu} g(\mu(x)) \, \mathrm{d}\#(x)$$

where

- $\mathrm{d}\mu/\mathrm{d}m$ is Radon-Nikodym derivative of μ with respect to m,
- $f : \mathbb{R} \to [0, +\infty]$ is convex, lower semicontinuous and proper (i.e. not identically $+\infty$),
- μ^s is the singular part of μ with respect to m according to the Radon-Nikodym decomposition theorem;
- f^∞ is the *recession function*

  $$f^\infty(s) := \lim_{t\to+\infty} \frac{f(s_0 + ts)}{t}$$

  (the limit is independent of the choice of s_0 in the domain of f, i.e. the set of points where f is finite),
- A_μ is the set of atoms of μ, i.e. the points such that $\mu(x) := \mu(\{x\}) > 0$,

- $g : \mathbb{R} \to [0, +\infty]$ is a lower semicontinuous subadditive function such that $g(0) = 0$
- $\#$ is the counting measure.

Note that our functional can be written in a simpler form since in our case $\mathrm{d}\mu^s/\mathrm{d}|\mu^s| = 1$ for $|\mu^s|$-a.e. of x, being μ a positive measure:

$$J(\mu) = \int_\Omega f\left(\frac{\mathrm{d}\mu}{\mathrm{d}m}\right)\ \mathrm{d}m + f^\infty(1)|\mu^s|(\Omega \setminus A_\mu) + \int_{A_\mu} g(\mu(x))\ \mathrm{d}\#(x).$$

By the results that can be found in [4] and [5], these functionals are lower semicontinuous for the weak-$*$ convergence of measures (and represent all local functionals with this semicontinuity property) whenever

$$g_0(s) := \sup_{t>0} \frac{g(st)}{t} = f^\infty(s).$$

Theorem 3.1. *Suppose that $f(s) > 0$ for $s > 0$ and $g(1) > 0$. Then we have $J \geq c > 0$. In particular, the functional $\mathcal{J}$ defined on the set of Lipschitz curves $\gamma : [0,1] \to \mathcal{W}_p(\Omega)$ with given starting and ending point achieves a minimum, provided that there exists a curve with finite cost.*

Proof. Let us fix some notation. By μ^a we mean the absolutely continuous part of μ with respect to the measure m, and by $\mu^s, \mu^\#, \mu^c$ respectively the singular part, the atomic part and the singular diffused part of μ. Then we have $\mu = \mu^a + \mu^s = \mu^a + \mu^c + \mu^\#$. Since f is convex, by Jensen inequality we have

$$\int_\Omega f\left(\frac{\mathrm{d}\mu}{\mathrm{d}m}\right)\ \mathrm{d}m \geq$$
$$m(\Omega) f\left(\frac{1}{m(\Omega)} \int_\Omega \frac{\mathrm{d}\mu}{\mathrm{d}m}\ \mathrm{d}m\right) = m(\Omega) f\left(\frac{\mu^a(\Omega)}{m(\Omega)}\right). \tag{3.2}$$

Since μ is a positive measure and f^∞ is $1-$homogeneous

$$\int_{\Omega\setminus A_\mu} f^\infty\left(\frac{\mathrm{d}\mu^s}{\mathrm{d}|\mu^s|}\right)\ \mathrm{d}|\mu^s| =$$
$$|\mu^s|(\Omega \setminus A_\mu) f^\infty(1) = m(\Omega) f^\infty\left(\frac{\mu^c(\Omega)}{m(\Omega)}\right). \tag{3.3}$$

Since g is a subadditive function

$$\int_{A_\mu} g(\mu(x))\ \mathrm{d}\#(x) =$$
$$\sum_{x\in A_\mu} g(\mu(x)) \geq g\left(\sum_{x\in A_\mu} \mu(x)\right) = g(\mu^\#(\Omega)). \tag{3.4}$$

For the recession function f^∞ it holds

$$f^\infty(x) \geq f(x+y) - f(y) \text{ for all } x, y \in \mathbb{R},$$

and so the sum of the first two terms, i.e. those given by (3.2) and (3.3), can be estimated from below by

$$m(\Omega) f\left(\frac{\mu^a(\Omega) + \mu^c(\Omega)}{m(\Omega)}\right).$$

Therefore summing up (3.2), (3.3) and (3.4) we obtain

$$J(\mu) \geq m(\Omega) f\left(\frac{\mu^a(\Omega) + \mu^c(\Omega)}{m(\Omega)}\right) + g(\mu^\#(\Omega)).$$

We set $a = \mu^\#(\Omega)$ and $1 - a = \mu^a(\Omega) + \mu^c(\Omega)$. Since the function $a \mapsto m(\Omega) f((1-a)/m(\Omega)) + g(a)$ is lower semicontinuous, it attains a minimum in the interval $[0,1]$. Thanks to our hypothesis this sum is always positive, and so we have

$$\min_{0 \leq a \leq 1} m(\Omega) f\left(\frac{1-a}{m(\Omega)}\right) + g(a) = c > 0,$$

that is, we have $J(\mu) \geq c > 0$. $\square$

We now study some special cases of the functional we defined above. In the rest of this section Ω will be a *compact convex* subset of $\mathbb{R}^N$ and the measure m will be the Lebesgue measure $\mathcal{L}^N$ on it.

3.1. **First case:** $f = +\infty$, $g(z) = |z|^r$ $(0 \leq r < 1)$**.** In this case we will denote the functional J by G_r and the corresponding functional $\mathcal{J}$ on Lipschitz paths will be called $\mathcal{G}_r$. This is the case when G_r is finite only on purely atomic measures.

We are now going to consider the question whether there exists a curve connecting two given measures keeping finite our functional. First we prove that if both the initial and the final measure are atomic the answer is positive. Then we prove that for r in a suitable subinterval of $[0,1]$ every measure can be connected to a Dirac mass, hence every measure can be connected to every other measure by a path of finite energy. Finally we show that this is not possible in general for every $r \in [0,1]$.

Theorem 3.2. *Let μ_0 and μ_1 be convex combinations of Dirac masses, i.e.,*

$$\mu_0 = \sum_{k=1}^m a_k \delta_{x_k}, \quad \mu_1 = \sum_{l=1}^n b_l \delta_{y_l}$$

with $a_k, b_l > 0$, $\sum_k a_k = \sum_l b_l = 1$. Then, there exists a Lipschitz curve $\gamma : [0,1] \to \mathcal{W}_p(\Omega)$ such that $\gamma(0) = \mu_0, \gamma(1) = \mu_1$ and

$$\mathcal{G}_r(\gamma) = \int_0^1 G_r(\gamma(t)) |\gamma'|(t) \, \mathrm{d}t < +\infty.$$

Proof. It is sufficient to prove the theorem when $a_1 = 1$, i.e. $\mu_0 = \delta_{x_1}$ since in the general case one connects the first measure μ_0 to a Dirac mass supported in an arbitrary point, then one connects that Dirac mass to the final measure μ_1. If one can keep finite the functional in both steps, then the result is proved in the general case.

We now prove that the curve $\gamma : [0,1] \to \mathcal{W}_p(\Omega)$ given by:

$$\gamma(t) = \sum_{l=1}^{n} b_l \delta_{x_1 + t(y_l - x_1)}.$$

is w_p-Lipschitz and $\mathcal{G}_r(\gamma) < +\infty$. Let t_1 and t_2 be time instants such that $t_1 < t_2$. Then the transport plan between $\sum_l b_l \delta_{x_1+t_1(y_l-x_1)}$ and $\sum_l b_l \delta_{x_1+t_2(y_l-x_1)}$ induced by the map $T(x_1 + t_1(y_l - x_1)) := x_1 + t_2(y_l - x_1)$ gives:

$$\begin{aligned} w_p(\gamma(t_2), \gamma(t_1)) &\leq \left(\int_\Omega |x - T(x)|^p \, \mathrm{d}\gamma(t_1) \right)^{1/p} \\ &= \left(\sum_{l=1}^{n} b_l |t_2 - t_1|^p |y_l - x_1|^p \right)^{1/p} \\ &= |t_2 - t_1| \left(\sum_{l=1}^{n} b_l |y_l - x_1|^p \right)^{1/p}. \end{aligned}$$

Hence the metric derivative with the respect to the Wasserstein p-distance is given by:

$$|\gamma'|(t) \leq \left(\sum_{l=1}^{n} b_l |y_l - x_1|^p \right)^{1/p} = w_p(\mu_0, \mu_1).$$

On the other hand we have:

$$G_r(\mu) = \begin{cases} \sum_{x \in A_\mu} |\mu(x)|^r & \text{if } \mu^a = \mu^c = 0 \\ +\infty & \text{otherwise.} \end{cases}$$

Then

$$G_r(\gamma(t)) = \begin{cases} 1 & \text{if } t = 0 \\ \sum_{l=1}^{n} b_l^r & \text{if } t > 0. \end{cases}$$

Hence

$$\mathcal{G}_r(\gamma) = \int_0^1 G_r(\gamma(t)) |\gamma'|(t) \, \mathrm{d}t \leq \sum_{l=1}^{n} |b_l|^r \left(\sum_{l=1}^{n} b_l |y_l - x_1|^p \right)^{1/p} < +\infty. \qquad \square$$

Remark 3.3. By repeating the proof of Theorem 3.2 one obtains that the statement still holds for infinite sums of Dirac masses (i.e. $m = n = +\infty$) provided $G_r(\mu_0)$ and $G_r(\mu_1)$ are finite, that is $\sum_k a_k^r < +\infty$ and $\sum_l b_l^r < +\infty$.

The proof of the next theorem is related to the one of Proposition 3.1 of [18].

Theorem 3.4. *Let $1-1/N < r \le 1$. Then given two arbitrary μ_0 and μ_1 in $\mathcal{W}_p(\Omega)$, there exists a curve joining them such that the functional $\mathcal{G}_r$ is finite.*

Proof. It is sufficient to prove that every measure can be joined to a Dirac mass in an arbitrary point. We prove first the statement for $\Omega = [0,1]^N$. The *dyadic subdivision of order k* of $Q = [0,1]^N$ is given by the family of closed N-dimensional cubes $\{Q_h^k\}_{h\in I_k}$ where $I_k = \{1,2,3,\ldots,2^k\}^N$ obtained by Q dividing each edge into 2^k pieces of equal length. We will refer to the elements of $\{Q_h^k\}_{h\in I_k}$ as k-cubes. To every Borel regular finite measure μ we associate the following sequence of measures:

$$\mu_k = \sum_{h\in I_k} b_h^k \delta_{y_h}$$

where $b_h = \mu(Q_h^k)$ and y_h is the center of Q_h^k. It is straightforward to see that $\mu_k \rightharpoonup^* \mu$ as $k \to +\infty$.

The idea is now simple (see Figure 1): first join μ_k to μ_{k+1} with an arc length parametrization γ_k, second put together all these curves to obtain a path from a Dirac mass to the measure μ. At every step a k-cube is divided in 2^N parts which are $(k+1)$-cubes. To bring the Dirac mass in the centre of the k-cube to the 2^N centres of the $(k+1)$-cubes with the right weights at each centre one splits the centre of the k-cube into 2^N parts moving towards the centres of the adjacent $(k+1)$-cubes in such a way that each point moves with unitary speed. At each step (see Figure 1 where the first three steps are represented) we obtain a curve γ_k defined on an interval of length $(1/2)^k d/2$ (d is the diagonal of Q) such that $|\gamma_k'|(t) = 1$ for all t.

Let us now compute the value of the functional on the curve γ made by joining all curves γ_k above. Since the function $f(x_1,\ldots,x_n) = \sum_i^n x_i^r$ with the constraint $\sum_i^n x_i = 1$ reaches its maximum at point $(1/n,\ldots,1/n)$ we have:

$$\mathcal{G}_r(\gamma) = \sum_{k=1}^{+\infty}\left(\frac{1}{2}\frac{1}{2^k}d\sum_{h\in I_k}(b_h^k)^r\right) \le \sum_{k=1}^{+\infty}\left(\frac{1}{2}\frac{1}{2^k}d2^{Nk}\left(\frac{1}{2^{Nk}}\right)^r\right).$$

Since $1-1/N < r \le 1$ the sum considered above is convergent.

In the case of a general Ω it is sufficient to consider a large cube containing the support of the measure μ such that the centre is contained in Ω. □

The bound given by $r > 1-1/N$ is sharp. We have in fact the following result.

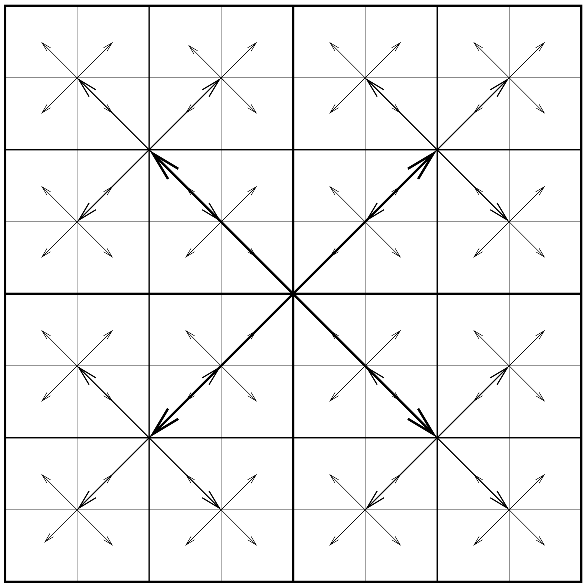

FIGURE 1. Approximation at step $k = 3$.

Theorem 3.5. *Suppose $r \leq 1 - 1/N$. Then there exists a probability measure μ on Ω such that every nonconstant w_p-Lipschitz path γ such that $\gamma(0) = \mu$ is such that $\mathcal{G}_r(\gamma) = +\infty$.*

Proof. Let Ω be the cube $[0,1]^N$ and μ the Lebesgue measure on it. We want to estimate from below

$$\inf \{G_r(\nu) \mid w_p(\mu, \nu) \leq t\}$$

and we will show it to be larger than $ct^{-N(1-r)}$. Therefore, if γ is a w_p-Lipschitz path with constant speed which originates from μ, the integral defining $\mathcal{G}_r$ diverges. We can simply consider $t = 2^{-k}$. To estimate $G_r(\nu)$, when ν is such that $w_p(\mu, \nu) \leq t$, consider a partition of Ω by small cubes of side ε. Let k be the number of those cubes Q_i such that $\nu(Q_i) \leq \mu(Q_i)/2 = \varepsilon^N/2$. In all these cubes we have a zone in which the optimal transport map s between μ and ν must take values outside the cube; this zone, given by $Q_i \setminus s^{-1}(Q_i)$, has a measure of at least $\varepsilon^N/2$. We want to estimate from below the contribute of this zone to the total transport cost between μ and ν. For this contribute we may write

$$\begin{aligned}\int_{Q_i\setminus s^{-1}(Q_i)} d(x,\partial Q_i)^p \mathrm{d}x =& \int_0^{(\varepsilon/2)^p} |\left(Q_i\setminus s^{-1}(Q_i)\right)\cap\{d(x,\partial Q_i)^p>\tau\}|\mathrm{d}\tau\\ \geq& \int_0^{(\varepsilon/2)^p}\left(\frac{\varepsilon^N}{2}-|\{d(x,\partial Q_i)^p\leq\tau\}|\right)\mathrm{d}\tau\\ \geq& \int_0^{B^p\varepsilon^p}\left(\frac{\varepsilon^N}{2}-|\{d(x,\partial Q_i)\leq B\varepsilon\}|\right)\mathrm{d}\tau\\ \geq& c_1\varepsilon^p\varepsilon^N,\end{aligned}$$

where B is sufficiently small and c_1 is a positive constant. By recalling that the total transport cost (i.e. the $p-$th power of the distance w_p) is less than t^p, we have

$$kc_1\varepsilon^{N+p}\leq t^p. \tag{3.5}$$

On the other hand, the value of G_r can be estimated from below by means of the other cubes and we have

$$G_r(\nu)\geq(\varepsilon^{-N}-k)c_2\varepsilon^{Nr}.$$

Let us now choose $\varepsilon=mt$ with m an integer such that $c_1m^p>1$ and, by using (3.5), we have

$$G_r(\nu)\geq t^{-N}(m^{-N}-m^{-N-p}/c_1)c_2m^{Nr}t^{Nr}=c_3t^{-N(1-r)},$$

where the constant c_3 is positive.

For general Ω we can simply use a cube contained in Ω and show that the Lebesgue measure on it, rescaled to a probability measure, cannot be reached keeping finite the value of the integral. $\square$

Example 3.6 (Y-shaped paths versus V-shaped paths)**.** Consider the example in Figure 2, where we suppose that l and h are fixed. We define for $0\leq t\leq l_0$

$$x(t)=(t,0)$$

and for $l_0\leq t\leq l_0+\sqrt{l_1^2+h^2}$

$$\begin{aligned}x_1(t)=&(l_0+l_1\frac{t-l_0}{\sqrt{l_1^2+h^2}},h\frac{t-l_0}{\sqrt{l_1^2+h^2}})\\ x_2(t)=&(l_0+l_1\frac{t-l_0}{\sqrt{l_1^2+h^2}},-h\frac{t-l_0}{\sqrt{l_1^2+h^2}}).\end{aligned}$$

Let us consider the curve $\gamma:[0,l_0+\sqrt{l_1^2+h^2}]\to\mathcal{W}_p(\Omega)$ defined by

$$\gamma(t)=\begin{cases}\delta_{x(t)} & \text{if } 0\leq t<l_0\\ \frac{1}{2}\delta_{x_1(t)}+\frac{1}{2}\delta_{x_2(t)} & \text{if } l_0\leq t\leq l_0+\sqrt{l_1^2+h^2}.\end{cases}$$

It easy to see that $|\gamma'|(t)=1$ and that

$$\mathcal{G}_r(\gamma)=l_0+2^{1-r}\sqrt{(1-l_0)^2+h^2}.$$

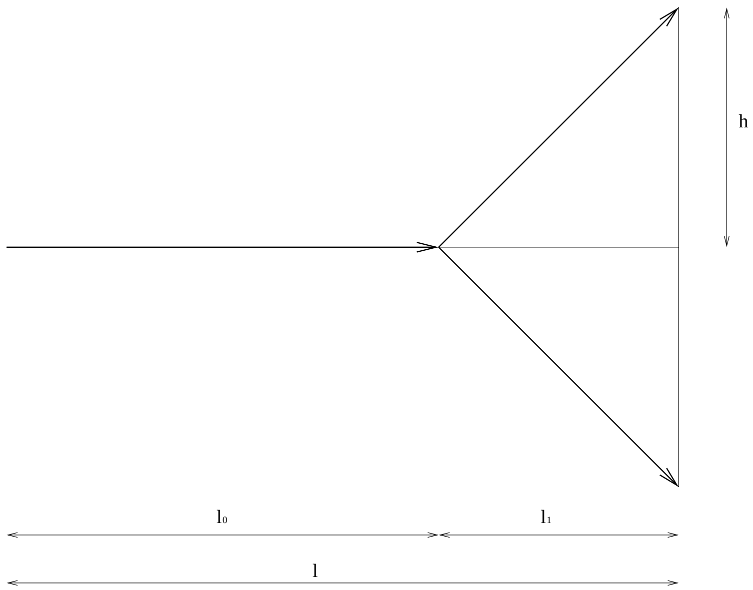

FIGURE 2. A Y-shaped path for $r = 1/2$.

Then the minimum is achieved for

$$l_0 = l - \frac{h}{\sqrt{4^{1-r} - 1}}.$$

In particular, when $r = 1/2$ we have a Y-shaped path (similar to the one of Figure 2) when $l > h$, while the path is V-shaped when $l \leq h$.

Remark 3.7. The result given by Theorem 3.4 can clearly be improved for particular choices of μ_0 and μ_1. For instance, we can connect a Dirac mass to the k-dimensional Hausdorff measure on a smooth k-surface for all $r \in [1 - 1/k, 1]$ (see also [14]).

3.2. **Second case: $f(z) = |z|^q$ $(q > 1), g = +\infty$.** We follow the same structure of the previous section. In this case we will denote the functional J by F_q and $\mathcal{J}$ by $\mathcal{F}_q$.

We start by proving that when $F_q(\mu_0)$ and $F_q(\mu_1)$ are finite, that is μ_0 and μ_1 are measures with $L^q(\Omega)$ densities, the optimal path problem admits a solution with finite energy.

Theorem 3.8. *Assume that $\mu_0 = u_0 \cdot \mathcal{L}^N, \mu_1 = u_1 \cdot \mathcal{L}^N$ with u_0, $u_1 \in L^q(\Omega)$. Then μ_0 and μ_1 can be joined by a finite energy path.*

The proof of this result relies on the notion of *displacement convexity*. This notion has been developed by McCann in [15] and is described also in [1] and [17].

Definition 3.9 (Displacement interpolation)**.** Let μ_0 and μ_1 be two absolutely continuous probability measures on Ω and let $T : \Omega \to \Omega$ be an optimal transport map (unique if $p > 1$) between μ_0 and μ_1 with respect to the cost function $|x-y|^p$. The map $\gamma^T : [0,1] \to \mathcal{W}_p(\Omega)$ given by

$$t \mapsto \gamma^T(t) := [(1-t)\mathrm{Id} + tT]_{\#}\mu_0 \tag{3.6}$$

is called a *displacement interpolation.*

Remark 3.10. It is well-known (see [1]) that the curve defined in (3.6) is a geodesic in $\mathcal{W}_p(\Omega)$, parametrized in such a way that

$$|(\gamma^T)'|(t) = w_p(\mu_0, \mu_1) \text{ for a.e. } t.$$

Definition 3.11 (Displacement convexity)**.** A functional F defined on all absolutely continuous measures (with respect to the Lebesgue measure) of $\mathcal{W}_p(\Omega)$ is said to be *displacement convex* if all the maps

$$t \mapsto F(\gamma^T(t))$$

are convex on $[0,1]$ for every choice of μ_0, μ_1 absolutely continuous measures and T optimal transport map as in Definition 3.9.

The proof of the following criterion for displacement convexity can be found in [15] and [17] (for the case $p=2$) but one of the most complete, since it deals also with the case $p \neq 2$, consists in Proposition 9.3.9 of [1].

Theorem 3.12. *Consider the following functional on the space $\mathcal{W}_p(\Omega)$, where Ω is any convex subset of $\mathbb{R}^N$:*

$$F(\mu) = \begin{cases} \int_\Omega U\left(\frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{L}^N}\right)\,\mathrm{d}x & \textit{if } \mu \textit{ is absolutely continuous} \\ +\infty & \textit{otherwise} \end{cases}$$

where $U : [0,+\infty] \to [0,+\infty]$ is a lower semicontinuous function with $U(0) = 0$. If the map

$$r \to r^N U(r^{-N})$$

is convex non-increasing on $]0,+\infty[$, then the functional F is displacement convex.

Remark 3.13. The assumption $r \to r^N U(r^{-N})$ being convex and non-increasing easily implies that U itself is convex.

Proof of Theorem 3.8. By Theorem 3.12 the functional F_q is displacement convex, so that

$$F_q(\gamma^T(t)) \leq (1-t)F_q(\mu_0) + tF_q(\mu_1).$$

Then

$$\int_0^1 F_q(\gamma^T(t))|(\gamma^T)'|(t)\,\mathrm{d}t \leq$$

$$w_p(\mu_0,\mu_1)\int_0^1 [(1-t)F_q(\mu_0) + tF_q(\mu_1)]\,\mathrm{d}t =$$

$$\frac{1}{2}(F_q(\mu_0) + F_q(\mu_1))w_p(\mu_0,\mu_1).$$

Since $F_q(\mu_0)$ and $F_q(u_1)$ are finite, we have that the path $t \mapsto \gamma^T(t)$ provides a finite value for the energy functional $\mathcal{F}_q$. $\square$

Next step will be the existence of an admissible path for arbitrary extremal measures, if q satisfies some additional constraints.

Recall that if μ_0 and μ_1 are probability measures given by L^1 densities (u_0 and u_1 respectively) and T is a transport map between them with sufficient regularity we have:

$$u_1(y) = u_0(T^{-1}(y))|\det \mathrm{D}T^{-1}(y)|.$$

Lemma 3.14. *Let $q < 1 + 1/N$. Let also $\mu = u \cdot \mathcal{L}^N$ with $u \in L^q(\Omega)$ and $\nu = \sum_{j=1}^k b_j \delta_{y_j}$ with $\sum_{j=1}^k b_j = 1$. Then there exists a path between μ and ν with finite energy.*

Proof. Let T be an optimal transport map between μ and ν. Let $B_j := T^{-1}(y_j)$. We now show that the path γ^T has a finite energy. Let us set $T_t = (1-t)\mathrm{Id} + tT$. If $x \in B_j$, then $T_t(x) = (1-t)x + ty_j$ and $\det \mathrm{D}T_t(x) = (1-t)^N$. Let u_t be the density of the measure $(T_t)_{\#}\mu$, that is to say:

$$u_t(y) = u(T_t^{-1}(y))|\det \mathrm{D}T_t^{-1}(y)|.$$

We then have:

$$\begin{aligned}
\int |u_t(y)|^q \,\mathrm{d}y &= \sum_{j=1}^k \int \left| u\left(\frac{y - ty_j}{1-t}\right)\right|^q \frac{1}{(1-t)^{Nq}} \,\mathrm{d}y \\
&= \sum_{j=1}^k \int |u(z)|^q (1-t)^{N(1-q)} \,\mathrm{d}z \\
&= (1-t)^{N(1-q)} \int |u(z)|^q \,\mathrm{d}z.
\end{aligned}$$

Moreover, thanks to Remark 3.10, the metric derivative $|\gamma'|(t)$ is constantly equal to the Wasserstein distance $w_p(\mu,\nu)$. Then,

$$\mathcal{F}_q(\gamma) = w_p(\mu,\nu) \int_0^1 \int |u_t(y)|^q \,\mathrm{d}y \,\mathrm{d}t = \frac{w_p(\mu,\nu)}{N + 1 - Nq} \int_\Omega |u|^q \,\mathrm{d}x$$

which is finite since $q < 1 + 1/N$. $\square$

Theorem 3.15. *Let $q < 1 + 1/N$. Then every couple of measures can be joined by a path with finite energy.*

Proof. It is enough to link any measure ν to a fixed L^q measure μ (for instance, the normalized Lebesgue measure) with a finite energy path. Let $\{\nu_k\}_{k\in\mathbb{N}}$ be a sequence of atomic measures approximating ν in the Wasserstein distance w_p. By Lemma 3.14, for every k there is a path γ_k with energy

$$\mathcal{F}_q(\gamma_k) = C w_p(\mu, \nu_k)$$

where C is a constant which only depends on N, q, Ω (and of course μ). Extracting a convergent subsequence of $\{\gamma_k\}_{k\in\mathbb{N}}$ provides a path γ

such that, by repeating the lower semicontinuity argument of Theorem 2.1, gives

$$\mathcal{F}_q(\gamma) \le \liminf_{k\to+\infty} \mathcal{F}_q(\gamma_k) = \lim_{k\to+\infty} C w_p(\mu, \nu_k) = C w_p(\mu, \nu).$$

Since γ_k connects μ to ν_k, then γ connects μ to ν and the result is established. □

As in the previous section, we show that the previous result is sharp, as it can be seen from the following statement which is valid in a more general setting. In fact, we prove an estimate which holds for every w_p-Lipschitz curve not only valued in $\mathcal{P}(\Omega)$, but also in $\mathcal{P}(\mathbb{R}^N)$.

Theorem 3.16. *Suppose $q \ge 1 + 1/N$. Then there exists $\mu \in \mathcal{W}_p(\Omega)$ such that every non-constant w_p-Lipschitz path γ with $\gamma(0) = \mu$ gives $\mathcal{F}_q(\gamma) = +\infty$.*

Proof. Let us choose $\mu = \delta_0$ (supposing, up to a translation, that $0 \in \Omega$). It is sufficient to prove that

$$\inf \{F_q(\nu) \,|\, \nu \in \mathcal{P}(\Omega),\, w_p(\mu, \nu) \le t\} \ge C t^{-N(q-1)}, \qquad (3.7)$$

with $C > 0$. In fact, by reparametrization, it is sufficient to prove that the functional is infinite on constant speed paths. Taken such a path γ, with constant speed $L > 0$, we then have

$$\mathcal{F}_q(\gamma) = L \int_0^1 F_q(\gamma(t))\mathrm{d}t \ge L \int_0^1 C(Lt)^{-N(q-1)}\mathrm{d}t = +\infty,$$

where the integral diverges thanks to the assumption on q. To prove (3.7) we can suppose that $\Omega = \mathbb{R}^N$, which is the worst case. This shows that the result does depend neither on the compactness nor on the convexity of Ω. By considering the map that associates to every probability measure ρ the measure $\nu = (m_t)_{\#}\rho$, where $m_t(x) = tx$, one has a one-to-one correspondence between the probabilities whose Wasserstein distance from δ_0 is less than 1 and those whose distance is less than t. It is easy to see that ν is L^q if and only if the same happens for ρ and that the density of ν is the function $x \mapsto t^{-N}u(x/t)$, where u is the density of ρ. Therefore

$$F_q(\nu) = \int \frac{u^q(x/t)}{t^{Nq}}\mathrm{d}x = \int u^q(y) t^{-Nq} t^N \mathrm{d}y = F_q(\rho) t^{-N(q-1)}.$$

Consequently, it is now sufficient to evaluate the infimum in (3.7) when $t = 1$, and this number will be the constant C we are looking for. We will show that this infimum is in fact a minimum, thus obtaining that it is strictly positive. This problem is quite similar to those studied in [16]. To get the existence of a minimum we recall that the functional F_q is sequentially lower semicontinuous with respect to weak-$*$ topology on probability measures, while the set $\{\nu \in \mathcal{P}(\mathbb{R}^n) \,|\, w_p(\delta_0, \nu) \le 1\}$ is sequentially compact with respect to the same topology (in fact every sequence in it turns out to be tight). □

Remark 3.17. As in the previous case, it is possible that two measures could be connected by a finite energy path even when q is greater than $1 + 1/N$. For instance, with $N = 2$, the path given by

$$\gamma(t) = \frac{1}{4t} 1_{[-1,1]\times[-t,t]} \cdot \mathcal{L}^2$$

is a Lipschitz path in $\mathcal{W}_p([-1,1] \times [-1,1])$ joining $\gamma_0 = 1/2\mathcal{H}^1 \llcorner [-1,1]$ to $\gamma_1 = 1/4\mathcal{L}^2$ (it is in fact a Wasserstein geodesic between them). The energy is finite as far as

$$\int_0^1 \frac{4t}{(4t)^q} \mathrm{d}t < +\infty.$$

This condition is fullfilled when $1 - q > -1$, i.e. when $q < 2$, instead of the condition $q < 1 + 1/2$ found in Theorem 3.15.

4. The Case of Unbounded Domains

The existence results of the previous section were based on two important facts: the compactness of Wasserstein spaces $\mathcal{W}_p(\Omega)$ when Ω itself is compact and $1 \le p < +\infty$, and the estimate like $F_q \ge c > 0$, proven in Theorem 3.1, that can be obtained when $|\Omega| < +\infty$. Both the facts do not hold when $\Omega = \mathbb{R}^N$, for instance. This is the reason why we developed in Section 1 some tools giving the existence of optimal paths under weaker assumption, even in the abstract metric setting. To replace the compactness of Ω we need to use Theorem 2.4, while to deal with the fact that we do not have $F_q(\nu) \ge c > 0$ in the case where ν runs over all $\mathcal{W}_p(\mathbb{R}^N)$ we can use the weaker assumption given by hypothesis H2. In this Section we only deal with the case of $\mathcal{F}_q$-like functionals studied in the compact case in Section 3.2; the case of atomic measures and $\mathcal{G}_r$-like functionals of Section 3.1 still present some extra difficulties when Ω is unbounded. We stress the fact that most of the techniques we use can be adapted to deal with several different cases, i.e. Ω unbounded but not necessarily the whole space, or the space $\mathcal{W}_\infty(\Omega)$ (where the distance is given by transport costs computed in a supremal way instead of an integral one). Notice that the use of Theorem 2.4 is necessary because in general, if Ω is not compact, the corresponding Wasserstein spaces are not even locally compact (and the same happens when we take Ω compact but we choose to consider the space $\mathcal{W}_\infty(\Omega)$), thus we cannot have the compactness of closed balls.

First, we show some lemmas in order to use Theorem 2.4.

Lemma 4.1. *The weak topology (i.e. the one induced by the duality with the space $C_b(\Omega)$ of bounded continuous functions on Ω) on the space $\mathcal{W}_p(\Omega)$ can be metrized by a distance d' such that $d' \le w_1 \le w_p$.*

Proof. The usual distance metrizing the weak topology is given by

$$d(\mu,\nu)=\sum_{k=1}^{\infty}2^{-k}\left|\int \phi_k \mathrm{d}(\mu-\nu)\right|,$$

where $(\phi_k)_k$ is a dense sequence in the unit ball of $C_b(\Omega)$. We can choose these functions to be Lipschitz continuous and let, for every index k, c_k be the Lipschitz constant of ϕ_k. Then

$$d'(\mu,\nu)=\sum_{k=1}^{\infty}\frac{2^{-k}}{1+c_k}\left|\int \phi_k \mathrm{d}(\mu-\nu)\right|$$

is a distance which metrizes the same topology. Being $\phi_k/(1+c_k)$ a $1-$Lipschitz function, thanks to the dual formulation of Monge's problem, we have

$$\left|\int \frac{\phi_k}{1+c_k}\mathrm{d}(\mu-\nu)\right|\leq w_1(\mu,\nu),$$

and so, by summing up on k, we get $d'\leq w_1$ as required. □

The following two lemmas are well known.

Lemma 4.2. *The distance w_p is lower semicontinuous on $\mathcal{W}_p(\Omega)\times\mathcal{W}_p(\Omega)$ endowed with the weak $\times$ weak convergence.*

Proof. Take $\mu_n\rightharpoonup\mu$ and $\nu_n\rightharpoonup\nu$. Let γ_n be an optimal transport plan for the cost $|x-y|^p$ between μ_n and ν_n: the sequence of this plans turns out to be tight thanks to tightness of the sequence of the marginal measures, and so we may suppose $\gamma_n\rightharpoonup\gamma$. We can now see that γ is a transport plan between μ and ν and so it holds

$$\begin{aligned} w_p(\mu,\nu) &\leq \left(\int |x-y|^p \mathrm{d}\gamma\right)^{1/p}\\ &\leq \liminf_{n\to+\infty}\left(\int |x-y|^p \mathrm{d}\gamma_n\right)^{1/p}=\liminf_{n\to+\infty} w_p(\mu_n,\nu_n). \quad \square \end{aligned}$$

Lemma 4.3. *All bounded sets in $\mathcal{W}_p(\mathbb{R}^N)$ are relatively compact with respect to weak topology.*

Proof. Just notice that, in a bounded set, every sequence of probability measures turns out to be tight. The limits up to subsequences (that exist in the weak sense) still belong to the space $\mathcal{W}_p(\mathbb{R}^N)$ as a consequence of the lower semicontinuity of the functional $\mu\mapsto w_p(\mu,\delta_0)$ (which is nothing but the $p-$th momentum of the measure). □

We can give now our result.

Theorem 4.4. *Let F_q and $\mathcal{F}_q$ be defined as in Section 3.2 respectively on $\mathcal{W}_p(\mathbb{R}^N)$ and on the set of Lipschitz path in $\mathcal{W}_p(\mathbb{R}^N)$ joining two measures μ_0 and μ_1. Then*

- *if $q < 1 + 1/N$ for every μ_0 and μ_1 there exists a path giving finite and minimal value to $\mathcal{F}_q$;*
- *if $q \geq 1 + 1/N$ there exist measures μ_0 such that $\mathcal{F}_q = +\infty$ on every nonconstant path starting from μ_0.*

Proof. Let us start by the case $q < 1 + 1/N$: thanks to Lemma 4.2 and 4.3 we can use Theorem 2.4 and so we just need to verify the two assumptions H1 and H2. The existence of a finite-energy path can be achieved in the same way as in Theorem 3.15, by passing through a fixed L^q probability measure. Notice that, in order to have the convergence of a subsequence and the lower semicontinuity in the approximation by atomic measures, we will argue as in the proof of Theorem 2.4 instead of Theorem 2.1. In order to estimate the integral in H2 we will use the same estimate given in Theorem 3.16, to achieve

$$\inf \{F_q(\nu) \mid \nu \in \mathcal{P}(\Omega),\, w_p(\mu, \nu) \leq t\} \geq Ct^{-N(q-1)},$$

so that the integral diverges as far as $q < 1 + 1/N$.

By repeating the arguments of Theorem 3.16, we can then prove also the second part of our result, because $\mu = \delta_0$ cannot be joined to any other probability measure by a finite energy path. □

Remark 4.5. In the previous theorem we did not mention the possibility to link, for arbitrary $q > 1$, two measures $\mu_0, \mu_1 \in L^q(\mathbb{R}^N)$. It is easy to check that the same construction used in Theorem 3.8 can be used in this setting too. We get in such a way the existence of a path providing a finite value to $\mathcal{F}_q$, but some problems arise when we look for a minimal one. In fact, for arbitrary q, condition H2 is no longer fulfilled and this prevents us from applying the general existence results.

To conclude this section, we highlight the difference between the case we dealt with (about the functional $\mathcal{F}_q$) and the other important case, represented by the functional $\mathcal{G}_r$. In this latter case it is not necessary to pass through the divergence of the integral in assumption H2, because we actually have $G_r \geq 1$, as already shown. On the other hand, some difficulties arise in verifying assumption H1. In fact the construction we made to build a finite energy path linking δ_0 to a probability measure μ strongly uses the compactness of the support of μ. In order to get a similar construction for the case $\Omega = \mathbb{R}^N$ we would need an estimate like

$$\inf\{w_p(\mu, \nu) \mid \#\,\mathrm{spt}(\nu) \leq k\} \leq C(\mu)k^{-1/N},$$

where $C(\mu)$ is a finite constant depending on the measure μ. It is easy to get a similar estimate when μ has compact support, but the constant may depend on the diameter of its support. The existence of a similar estimate for arbitrary measures μ is linked to the asymptotics of the rescaled location problem in $\mathbb{R}^N$. A theory on this asymptotic problem has been explicitly developed (for instance in [8]) only in the case of compact support. However, it leads to a condition like $\mu^{N/(N+p)} \in L^1$, which is always fulfilled for μ compactly supported, while it may fail for general probability measures in $\mathcal{W}_p(\mathbb{R}^N)$. A positive answer to this estimate would easily imply a theorem similar to Theorem 4.4 for the case of $\mathcal{G}_r$, but we cannot say this to be necessary.

Acknowledgements. This work is part of the European Research Training Network *"Homogenization and Multiple Scales"(HMS2000)* under contract HPRN-2000-00109. The authors also acknowledge the support of the project *"Problemi di Ottimizzazione in Teoria del Trasporto ed Applicazioni a Problemi di Pianificazione Urbana"* of the Italian GNAMPA and of the project *"Calcolo delle Variazioni"* of the Italian Ministry of Education.

References

1. L. Ambrosio, N. Gigli, and G. Savarè, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Lectures in Mathematics, ETH Zürich, vol. 7, Birkhäuser, Basel, 2005.
2. L. Ambrosio and P. Tilli, *Selected Topics on "Analysis on Metric Spaces"*, Appunti dei Corsi Tenuti da Docenti della Scuola, Scuola Normale Superiore, Pisa, 2000.
3. M. Bernot, V. Caselles, and J. M. Morel, *Are there infinite irrigation trees?*, Preprint available at `http://www.cmla.ens-cachan.fr/Cmla/`, 2003.
4. G. Bouchitté and G. Buttazzo, *New Lower Semicontinuous Results for Nonconvex Functionals defined on Measures*, Nonlinear Analysis, Theory, Methods & Applications **15** (1990), no. 7, 679–692.
5. ______, *Integral Representation of Nonconvex Functionals defined on Measures*, Ann. Inst. H. Poincaré Anal. Non Linéaire. **9** (1992), no. 1, 101–117.
6. ______, *Relaxation for a Class of Nonconvex Functionals defined on Measures*, Ann. Inst. H. Poincaré Anal. Non Linéaire **10** (1993), no. 3, 345–361.
7. ______, *Characterization of Optimal Shapes and Masses through Monge-Kantorovich Equation*, J. Eur. Math. Soc. (JEMS) **3** (2001), no. 2, 139–168.
8. G. Bouchitté, C. Jimenez, and M. Rajesh, *Asymptotique d'un Problème de Positionnement Optimal*, C. R. Acad. Sci. Paris **335** (2002), 853–858.
9. A. Brancolini and G. Buttazzo, *Optimal networks for mass transportation problems*, ESAIM Control Optim. Calc. Var. **11** (2005), no. 1, 88–101.
10. G. Buttazzo, E. Oudet, and E. Stepanov, *Optimal Transportation Problems with Free Dirichlet Regions*, Progress in Nonlinear Differential Equations and their Applications **51** (2002), 41–65, Proceedings in *"Variational Methods for Discontinuous Structures"*, Cernobbio 2001.
11. G. Buttazzo and F. Santambrogio, *A model for the Optimal Planning of an Urban Area*, Preprint available at `http://cvgmt.sns.it/`, to appear on SIAM J. Math. Anal., 2003.
12. G. Buttazzo and E. Stepanov, *Optimal Transportation Networks as Free Dirichlet Regions for the Monge-Kantorovich Problem*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **2** (2003), 631–678.
13. G. Devillanova and S. Solimini, *On the dimension of an irrigable measure*, Preprint in preparation.
14. F. Maddalena, J.M. Morel, and S. Solimini, *A Variational Model of Irrigation Patterns*, Interfaces and Free Boundaries **5** (2003), 391–415.

15. R. J. McCann, *A Convexity Principle for Interacting Gases*, Adv. Math. **128** (1997), no. 1, 153–159.
16. F. Santambrogio, *Misure Ottime per Costi di Trasporto e Funzionali Locali*, 2003, Tesi di Laurea. Available at `http://cvgmt.sns.it/`.
17. C. Villani, *Topics in Optimal Transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003.
18. Q. Xia, *Optimal Paths Related to Transport Problems*, Comm. Cont. Math. **5** (2003), no. 2, 251–279.

A. Brancolini, Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy, `a.brancolini@sns.it`

G. Buttazzo, Dipartimento di Matematica, Università di Pisa, Largo Bruno Pontecorvo 5, 56127 Pisa, Italy, `buttazzo@dm.unipi.it`

F. Santambrogio, Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy, `f.santambrogio@sns.it`