Annali di Matematica 184, 215–238 (2005)

Digital Object Identifier (DOI) 10.1007/s10231-004-0109-5

Aldo Pratelli

Equivalence between some definitions for the optimal mass transport problem and for the transport density on manifolds

Received: March 18, 2003; in final form: July 22, 2003
Published online: May 25, 2004 –

Abstract. In this paper we consider three problems, which are related to the classical Monge's optimal mass transport problem and which are known to be equivalent when the ambient space is an open, convex and bounded subset of $\mathbb{R}^n$; to these problems there correspond different definitions of particular measures (often called *transport densities*), which are also known to be equivalent. Here we will generalize the setting of these problems and the resulting definitions of transport densities to the case of a Riemannian manifold endowed with a finslerian semidistance, and we will see that the equivalences still hold.

Mathematics Subject Classification (2000). 46G10, 49Q20, 58C35

Key words. mass transportation – transport density – shape optimization

1. Introduction

Here we briefly introduce the original context of the problems we will study.

The celebrated mass transport problem, first formulated by Monge over two centuries ago, can be today expressed as the search of γ which minimizes the cost

$$C(\gamma) = \iint_{\Omega\times\Omega} c(x, y)\, d\gamma(x, y) \tag{1.1}$$

among the *transports*, i.e. the measures $\gamma \in \mathcal{M}^+(\Omega \times \Omega)$ such that $\gamma_1 = f^+$ and $\gamma_2 = f^-$, where f^+ and f^- are given positive measures of equal (and finite) total mass on Ω and we denote by γ_1 and γ_2 the projections of γ on Ω. The "cost" c is a function from $\Omega \times \Omega$ to $\mathbb{R}^+$ such that $c(x, y) = c(y, x)$: the original problem was with $\Omega \subseteq \mathbb{R}^n$ and $c(x, y) = |x - y|$ (the Euclidean norm of $\mathbb{R}^n$), but also other situations (for example $c(x, y) = |x - y|^p$) have been considered. This way to express the problem was first suggested by Kantorovich ([15], [16]), and then it has been deeply studied (here there is a vast literature, to give an example only [11] or [14] are cited). Two of the main and best known results are collected in the following theorem, which holds for a generic topological space Ω:

A. Pratelli: Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy, e-mail: a.pratelli@sns.it

Theorem 1.1. *If c is l.s.c. and bounded there exists an optimal transport γ. If, moreover, c is a semidistance, for each optimal transport γ there exists a function u (often called the Kantorovich potential) such that $u(x) - u(y) \leq c(x, y)$ for any couple (x, y) and $u(x) - u(y) = c(x, y)$ for γ−a.e. $(x, y) \in \Omega \times \Omega$.*

Given a transport γ, the measure $\mu \in \mathcal{M}^+(\Omega)$ defined by

$$\langle \mu, \varphi \rangle := \iint_{\Omega\times\Omega} \left(\int_\Omega \varphi(z)\, d\mathcal{H}^1_{xy}(z) \right) d\gamma(x, y) \qquad \forall\, \varphi \in \mathrm{C}_b(\Omega), \tag{1.2}$$

is usually called the *transport density*; with $\mathcal{H}^1_{xy}$ we denote the restriction of the one-dimensional Hausdorff measure $\mathcal{H}^1$ to the segment xy. We interpret this formulation in terms of the original transport problem by saying that γ is the transport which takes a mass $\gamma(A \times B)$ from A and transports it onto B, and then the transport density results in nothing else than the distribution in Ω of the work done by the transport γ. We will call the *optimal transport density* each transport density corresponding to a transport γ of minimal cost.

The second problem we consider was proposed by Brenier (see [8] and [5]), and consists of finding the best *configuration* which transports f^+ onto f^-. A configuration is a set of scalar measures f_t and vectorial measures $\overline{E}_t$ on Ω, with $t \in [0, 1]$, such that the compatibility condition $f_t' + \operatorname{div} \overline{E}_t = 0$ holds in the distribution sense; by saying that $\{f_t, \overline{E}_t\}$ "transports f^+ onto f^-" we intend that $f_0 = f^+$ and $f_1 = f^-$. The "goodness" of a configuration is given by its cost

$$C\left(\{f_t, \overline{E}_t\}\right) := \int_0^1 \|\overline{E}_t\|\, dt, \tag{1.3}$$

where $\|\overline{E}_t\| = |\overline{E}_t|(\Omega)$ is the norm of $\overline{E}_t$ in the space of the vectorial measures on Ω (see the Appendix for the definition of the measure $|\overline{E}_t|$); of course a configuration is better if its cost is lower. The interpretation of these configurations in terms of the mass transport problem is clear: we imagine that a transport from f_0 to f_1 is made in a unitary time, and then f_t and $\overline{E}_t$, respectively, say which is the distribution of mass and how it has to be moved at time t. For each configuration $\{f_t, \overline{E}_t\}$ we consider the *work density*, i.e. the measure μ given by

$$\mu := \int_0^1 |\overline{E}_t|\, dt. \tag{1.4}$$

We will call the *optimal work density* each work density corresponding to a configuration $\{f_t, \overline{E}_t\}$ of minimal cost.

About the interest and the different physical interpretations of this problem see the introduction of [5].

The last problem to be considered is the *shape optimization*, which can be reformulated (see [7]) as the search of the extremals of the following dual problems:

$$\sup \left\{ \int_\Omega u\, d(f^+ - f^-) :\ u \text{ is 1-Lipschitz} \right\} \tag{1.5}$$

$$\inf \left\{ \int_\Omega d|\nu| :\ -\operatorname{div} \nu = f^+ - f^- \right\}. \tag{1.6}$$

The *optimal shapes* are the measures $\mu = |\nu|$ where ν is a solution of (1.6). In [7] it was also shown the connection between the shape optimization problem and the following PDE equation, whose formulation in the context of the Monge–Kantorovich problem is due to Evans and Gangbo (see [12]):

$$\begin{cases} -\operatorname{div}(\mu Du) = f^+ - f^- & \text{on } \Omega \\ |Du| = 1 & \mu - \text{a.e.}, \end{cases} \tag{1.7}$$

where $\mu \in \mathcal{M}^+(\Omega)$ and u is a 1-Lipschitz function on Ω; we will discuss the connections which hold on the manifolds in Section 3.5.

The equivalence between these problems (with, as usual, $c(x, y) = |x - y|$ for the Monge problem) can be collected in the following:

Theorem 1.2. $\inf(1.1) = \inf(1.3) = \inf(1.6)$. *Moreover, the set of the optimal transport densities, the set of the optimal work densities and the set of the optimal shapes are the same.*

Let us discuss for a moment the second part of the last assertion: in general, there are several optimal transport densities, corresponding to different transports of minimal cost, as there are several optimal work densities and optimal shapes; however, a measure μ on Ω is an optimal transport density if and only if it is an optimal work density and if and only if it is an optimal shape.

Some discussions about the equivalences between the problems and the proofs of the assertions of the preceding theorem can be found in [1], [6] and [10], under the hypothesis that Ω is an open, bounded and convex subset of $\mathbb{R}^n$; a generalization that has been already considered (see for example [6]) is to imagine a closed Dirichlet region $\Sigma \in \Omega$ on which the transport is free, while the case of a Riemannian manifold was considered also in [9].

In this work we will consider the more general situation in which the ambient space is a Riemannian manifold M; moreover, the Riemannian distance will be replaced by a more general finslerian semidistance, i.e. the infinitesimal length of a path σ is given by $j(\sigma, \sigma')$, where we are given a positively 1-homogeneous function $j : TM \longrightarrow \mathbb{R}^+$: throughout the paper, when $v \in T_x M$ we will always write $j(x, v)$ instead of simply $j(v)$. Note that d reduces to the standard Riemannian distance if $j(x, v) = |v|$, while the case of the Dirichlet region Σ discussed above is the situation in which $j(x, v) = 0$ if $x \in \Sigma$ and $j(x, v) = |v|$ otherwise.

The plan of the paper is the following: in Section 2 we list and discuss the different hypotheses we will use through the paper; in Section 3 we will set the problems and their relaxations in our context; in Section 4 we will give our results, namely that the extremals of the problems are equal, that the problems – or their relaxed version – have always a solution and that the optimal measures of the three problems are the same; moreover, we will discuss the non-intersection of the transport rays in the strictly convex case.

We state here a result of measure theory we will often use through the paper (by $\mathcal{P}(X)$ we denote the set of the probability measures on X):

Theorem 1.3 (disintegration theorem). *Let $f : X \longrightarrow Y$ and $\mu \in \mathcal{M}(X)$. Then there exist $\mu_y \in \mathcal{P}(X)$ such that μ_y is concentrated on $\{x : f(x) = y\}$ for any $y \in Y$ and*

$$\int_X \varphi(x)\, d\mu(x) = \int_Y \left(\int_X \varphi(x)\, d\mu_y(x) \right) d f_\# \mu(y) \qquad \forall \varphi \in C_b(X).$$

Moreover the measures μ_y are univocally determined up to a negligible set with respect to $f_\# \mu$.

The proof of this assertion can be found, for instance, in [2].

2. Ambient space and hypotheses

In this paper M is an $n-$dimensional connected Riemannian manifold, not necessarily complete and possibly with a boundary, whose Riemannian metric is denoted by d_M. We want to define a finslerian semidistance on M (to find a general discussion on finslerian metrics, refer to [4]); to do this, we are given a function $j : TM \longrightarrow \mathbb{R}^+$ 1-homogeneous, i.e. $j(x, \lambda v) = |\lambda| j(x, v)$. Then we denote by S the set of the paths in M which are Lipschitz with respect to d_M, and we regard S as a metric space by means of the distance

$$d_S(\sigma, \tau) := \max_{t \in [0,1]} d_M\left(\sigma(t), \tau(t)\right), \tag{2.1}$$

here we think, as we will always do, $\sigma : [0, 1] \longrightarrow M$ as the parameterization at constant speed of any element $\sigma \in S$. Now we define the length of each path $\sigma \in S$ as

$$l(\sigma) := \int_{t=0}^{1} j\big(\sigma(t), \sigma'(t)\big)\, dt,$$

finally the finslerian semidistance d between two points is simply the infimum of the lengths of the paths that connect them, i.e.

$$d(x, y) := \inf \left\{ l(\sigma) : \sigma \in S,\ \sigma(0) = x,\ \sigma(1) = y \right\}.$$

We will call M a *manifold of type 0* if the following holds:

0-a) The metric space (M, d) is bounded;
0-b) j is locally bounded, i.e. for each $x \in M$ there exists a neighborhood U of x and a constant K such that $j(z, v) \leq K|v|$ for each $z \in U$ and $v \in T_z M$;
0-c) for each $x \in M$ the restriction $j(x, \cdot)$ is a convex function.

If M is of type 0, we will call it a *manifold of type 1* if the following holds:

1-a) M is a complete manifold, possibly with boundary;
1-b) j is l.s.c.;

1-c) there exists an idempotent map $ret : S \longrightarrow S$ (i.e. $ret^2 = ret$) such that $ret(\sigma)$ has the same extremal points of σ, $\mathcal{H}^1_j \llcorner ret(\sigma) \leq \mathcal{H}^1_j \llcorner \sigma$ – see (3.1) – and, denoted by l_M the Riemannian length of the paths, it holds that

$$l_M(ret(\sigma)) \leq C + \alpha\, l(\sigma), \tag{2.2}$$

we will call "retraction" the function ret and its push-forward $ret_\# : \mathcal{M}^+(M) \to \mathcal{M}^+(M)$ and "retracted" a path σ or a transport η if $ret(\sigma) = \sigma$ or $ret_\#(\eta) = \eta$.

Finally, if M is of type 0, we will call it a *manifold of type 2* if, denoted by $(\hat{M}, d_M)$ and $(\hat{S}, d_S)$ the completions of (M, d_M) and (S, d_S), respectively, the following holds:

2-a) $\hat{M}$ is a manifold of type 1, where we extend j from TM to the whole $T\hat{M}$ by lower semicontinuity;

2-b) for each $\sigma \in \hat{S}$, there exists a sequence $\sigma_i \to \sigma$ in S such that $l(\sigma_i) \to l(\sigma)$. It is easily noted that $\hat{S}$ is the metric space of the paths on $\hat{M}$ which are Lipschitz with respect to d_M, where the distance is given as in (2.1).

2.1. About the hypotheses. Here we briefly discuss the hypotheses we made: first of all, let us note that our hypotheses are very weak; in fact, they cover many more situations than the ones which have been already considered for the Monge–Kantorovich problem, and in particular all the compact manifolds with a boundary, equipped with any Riemannian metric and possibly with a Dirichlet region. In the manifolds of type 0, the more general we can work with, the problems presented in Section 1 will be suitably restated, and we will prove that the extremals are still equal. Note that asking M to be a bounded metric space allows us to avoid situations in which all the extremals are $+\infty$, which has nothing of particular interesting and forces to make a lot of boring distinctions; the fact that j is locally bounded ensures that the distance d is continuous with respect to the Riemannian topology; finally, the convexity assumption is standard when dealing with a finslerian semidistance.

It is easy to imagine that we can not hope the extremals to be reached without some completeness hypotheses: in fact, we will prove the existence of solutions and the fact that the optimal measures of the problems are the same for the manifold of type 1, which are complete, and for the manifold of type 2, in which we will consider a relaxed setting of the problems passing to the completion. To convince ourselves that these hypotheses are nearly essential to get the existence of solutions, we point out that:

1) If j were not l.s.c., we could not hope to have paths of minimal length, which is clearly essential to reach the extremals of the problems; moreover, with standard convexity arguments the lower semicontinuity of j implies the one of $l : S \to \mathbb{R}^+$.

2) The hypothesis *1-c)* seems at first glance a bit strange, but it is quite natural: the idea is that if for example σ contains some closed subpath, then $ret(\sigma)$ is without them; or if in σ there is a long – in the Riemannian sense – part in a Dirichlet region (it is possible since it has null finslerian length), then in $ret(\sigma)$ it will be replaced by a short one with the same extremal points. This hypothesis is useful because if there are short paths with great Riemannian length, we need some corresponding paths which are short also in the Riemannian sense; if not, there could be a lack of paths of minimal length, and this would of course prevent the minimal being achieved. However, if for example on M there is a Dirichlet region Σ and $j(x, v) \geq \varepsilon|v|$ for $x \notin \Sigma$, (2.2) holds when $\alpha = \varepsilon^{-1}$ and C is the sum of the diameters of the arcwise connected components of Σ (if finite). Vice versa, if there is no ε such that $j(x, v) \geq \varepsilon v$ there are counter-examples to the existence of minimizers.

3) The hypothesis *2-b)* is fundamental to made a good relaxation of the problems in the non-complete case since, if it does not hold, the extremals could change passing to the relaxed formulations: in fact, there could be paths in $\hat{M}$ connecting two points with a length strictly smaller than the distance between them. However, note that this is not a strong hypotheses: it holds, for example, if j is continuous on $\hat{M} \setminus M$.

3. Setting of the problems and relaxation

In this section we will set the problems in the more general context of the manifolds of type 0; then, we will relax them in the non-complete case of the manifold of type 2; then, we discuss the possibility of generalizing problem (1.7); finally, we present some operations which can be done with the transports. In the following, f^+ and f^- will be positive measures on M of equal total mass, and we will denote $f = f^+ - f^-$.

3.1. The Monge–Kantorovich problem. Let us recall that, in the classical case of a convex subset Ω of $\mathbb{R}^n$, a transport is simply a measure γ on $\Omega \times \Omega$, whose meaning is that a mass $\gamma(x, y)$ has to be moved from x to y; γ does not say *how* to move this mass, but this does not give difficulties since clearly the best way is to follow the segment from x to y. In our new setting the situation is no longer so easy, because, given x and y in M, there could be more than one shortest path from x to y, or there could be no shortest path at all: to say that a certain mass must travel from x to y is then not sufficient, we must specify which path has to be used. The best definition for a transport is then the following:

Definition 3.1. *Given f^+, $f^- \in \mathcal{M}^+(M)$ of equal total mass, a transport from f^+ to f^- is a measure $\eta \in \mathcal{M}^+(S)$ such that $\pi_{0\#}\eta = f^+$ and $\pi_{1\#}\eta = f^-$.*

Here, and also in the rest of the paper, $\pi_t : S \longrightarrow M$ is the function defined by $\pi_t(\sigma) = \sigma(t)$, recalling that we always think of σ as parameterized at constant speed. It is easy to understand how we think of such a measure η as a transport: $\eta(\sigma)$ indicates the quantity of mass which must be moved following the path σ, and then $\pi_{0\#}\eta$ and $\pi_{1\#}\eta$ are the distributions of mass that start and arrive, respectively. We can note that in this new setting, where giving a transport means saying precisely how to move the mass, the transports are much more similar to the configurations (in the sense of Brenier). In fact, the difference between transports and configurations will be better discussed in Remark 4.3.

Having defined the transports, we have now to specify their cost; since $\eta(\sigma)$ is the quantity of mass which is moved through the path σ, it is not difficult to generalize (1.1) defining the cost of η, which can also be $+\infty$, as

$$C(\eta) := \int_S l(\sigma)\, d\eta(\sigma). \tag{MK}$$

Note that what we are presenting is a generalization of the classical transport problem in the topological space M with, as cost, the semidistance d, and we still have inf(1.1)=inf(MK): in fact, we can associate to the "new" transport η the "old" transport $\gamma = (\pi_0, \pi_1)_{\#}\eta$, and it is clearly $C(\gamma) \leq C(\eta)$; on the other hand, if $\beta : M \times M \longrightarrow S$ is a measurable selection of paths which are optimal up to ε and we associate to the "old" γ the "new" $\eta = \beta_{\#}\gamma$, it is $C(\eta) \leq C(\gamma) + \varepsilon\|\gamma\|$ (as usual, throughout the paper we will denote the norm of a positive measure $\alpha \in \mathcal{M}(X)$ with $\|\alpha\| = \alpha(X)$). This argument also shows that a minimizing transport (if there exists one) is concentrated on geodesics for the semidistance d.

In order to introduce the transport densities, as it can be understood recalling (1.2), we have first to define $\mathcal{H}^1_j \llcorner \sigma$ for $\sigma \in S$, which is the 1-dimensional Hausdorff measure weighed with j on σ: formally, if $\varphi \in C_b(M)$, it can be computed by

$$\left\langle \mathcal{H}^1_j \llcorner \sigma, \varphi \right\rangle := \int_0^1 \varphi\left(\sigma(t)\right) j\big(\sigma(t), \sigma'(t)\big)\, dt. \tag{3.1}$$

We can now define the transport density associated to a transport η: the direct generalization of (1.2) turns out to be

$$\mu := \int_S \mathcal{H}^1_j \llcorner \sigma \, d\eta(\sigma); \tag{3.2}$$

note also that $l(\sigma) = \langle \mathcal{H}^1_j \llcorner \sigma, 1\rangle$ and then $C(\eta) = \langle \mu, 1\rangle = \mu(M)$.

3.2. The Brenier problem. The new formulation of the Brenier problem will be the same as the old one, and we will need only to extend the definitions of the objects we work with to our new setting. Recalling the definitions given in the Appendix, we start defining the *configuration* of any set $\{f_t \in \mathcal{M}^+(M),\ \overline{E}_t \in \mathcal{M}^n(M), t \in [0,1]\}$ such that the compatibility condition

$$f_t' + \operatorname{div} \overline{E}_t = 0 \tag{3.3}$$

holds. The meaning of (3.3) is that for each $\varphi \in C_b^\infty(M)$ the equality $d/dt\langle f_t, \varphi\rangle = \langle \overline{E}_t, D\varphi\rangle$ holds in the weak sense: then, if $\{f_t, \overline{E}_t\}$ is a configuration, for any $\varphi \in C_b^\infty(M)$ the function $t \to \langle \overline{E}_t, D\varphi\rangle$ is in $L^1_{\text{loc}}([0,1])$ and $t \to \langle f_t, \varphi\rangle$ is in $W^{1,1}_{\text{loc}}([0,1])$, so in particular it is continuous. This implies that the measure f_t is well defined for any time t, and then it makes sense to say that $\{f_t, \overline{E}_t\}$ is a *configuration from* f^+ *to* f^-; in particular the measures f_t are somehow an interpolation between f_0 and f_1. With an easy generalization of (1.3), we define the cost of a configuration as

$$C\left(\{f_t, \overline{E}_t\}\right) := \int_0^1 \int_M j\big(x, d\overline{E}_t(x)\big)\, dt, \tag{Br}$$

where, writting any $\nu \in \mathcal{M}^n(M)$ as $\nu = \bar{\omega}|\nu|$ as discussed in the Appendix, we will throughout the paper, intend

$$j(x, \nu) := j(x, \bar{\omega})\, |\nu| \quad \in \mathcal{M}^+(M);$$

this means that, for any continuous and bounded function φ,

$$\langle j(x, \nu), \varphi\rangle = \int_M \varphi(x)\, j(x, \nu) = \int_M \varphi(x)\, j(x, \bar{\omega})\, d|\nu|.$$

Finally, the definition of the work density which generalizes (1.4) is

$$\mu := \int_0^1 j(x, d\overline{E}_t(x))\, dt. \tag{3.4}$$

As was made for the Monge–Kantorovich problem, we can note that $C\left(\{f_t, \overline{E}_t\}\right) = \langle \mu, 1\rangle = \mu(M)$.

3.3. The shape optimization problem. As we said in the introduction, the shape optimization problem consists of finding two dual extremals; to generalize them, we consider first of all the set $\mathrm{Lip}_j(M)$ of the functions on M which are 1-Lipschitz with respect to d (recall that d depends on j). So, instead of (1.5), we will try now to find

$$\sup \int_M u\, df \tag{3.5}$$

among all $u \in \mathrm{Lip}_j(M)$; note that, thanks to the local boundedness of j, the elements of $\mathrm{Lip}_j(M)$ are locally Lipschitz, and so continuous, with respect to d_M; moreover, since (M, d) is bounded by hypothesis *0-a)*, then the elements of $\mathrm{Lip}_j(M)$ are also bounded: thus the integral in (3.5) is well defined.

On the other hand, the search of the minimizers of (1.6) becomes the search of

$$\inf \int_M j(x, \nu) \tag{SO}$$

among all $\nu \in \mathcal{M}^n(M)$ such that $-\operatorname{div} \nu = f$, i.e. $\langle f, \varphi\rangle = \langle \nu, D\varphi\rangle$ for each $\varphi \in C_b^\infty(M)$. Since we refer to (3.5) and (SO) as "dual problems", it can be imagined that we will prove these extremals are equal; we prove now the first inequality, the second one will follow from the proof of Theorem A:

Proposition 3.2. $\sup(3.5) \leq \inf(SO)$.

Proof. Let us take a smooth 1-Lipschitz function u and a vectorial measure $\nu \in \mathcal{M}^n(M)$ such that $-\operatorname{div}\nu = f$; recalling the definition of the semidistance d and the fact that $u \in \mathrm{Lip}_j(M)$ it is easy to deduce that

$$Du(x) \cdot \bar{\omega}(x) \leq j\left(x, \bar{\omega}(x)\right) \qquad |\nu|\text{−a.e.}, \tag{3.6}$$

where we write, as in the Appendix, $\nu = \bar{\omega}|\nu|$. We can then write

$$\int_M u\,df = \int_M Du \cdot d\nu = \int_M Du \cdot \bar{\omega}\,d|\nu| \leq \int_M j\left(x, \bar{\omega}\right) d|\nu| = \int_M j(x, \nu). \tag{3.7}$$

The extension of the result for any $u \in \mathrm{Lip}_j(M)$ can be obtained using the definition and the properties of the tangential derivatives given in [7] as generalized in [13]: in fact, the elements of $\mathrm{Lip}_j(M)$ are easily seen to be in $\mathrm{W}^{1,2}_{|\nu|,\mathrm{loc}}$, thus the tangential derivative $D_{|\nu|}u$ is defined $|\nu|$−a.e.; moreover, since $u \in \mathrm{Lip}_j(M)$ then (3.6) generalizes as

$$D_{|\nu|}u(x) \cdot \bar{\omega}(x) \leq j\left(x, \bar{\omega}(x)\right) \qquad |\nu|\text{−a.e.}$$

Finally, since for any $v \in \mathrm{W}^{1,2}_{\mu,\mathrm{loc}}$ and any φ such that $\operatorname{div}(\varphi\mu) \in \mathcal{M}$

$$\int D_\mu v \cdot \varphi\,d\mu = -\langle \operatorname{div}(\varphi\mu), v\rangle$$

holds, setting $\varphi = \bar{\omega}$, $\mu = |\nu|$ and $v = u$ we find

$$\int_M D_{|\nu|}u \cdot \bar{\omega}\,d|\nu| = \int_M u\,df,$$

and then, as in (3.7), we conclude that

$$\int u\,df \leq \int j\left(x, \bar{\omega}\right) d|\nu| = \int j(x, \nu).$$

Since the last inequality holds for a generic $u \in \mathrm{Lip}_j(M)$ and ν such that $-\operatorname{div}\nu = f$, taking the sup in the left side and the inf in the right side we conclude the proof. □

3.4. Relaxation of the problems. As we noted in Section 2, the completeness of the space M is essential to hope for the existence of solutions for our problems. For example, if M is a ball in $\mathbb{R}^n$ without the center and with the Riemannian distance, x and y are two opposite points and $f^+ = \delta_x$ and $f^- = \delta_y$, then clearly $\inf(MK) = d(x, y)$, but any path between x and y is strictly longer than $d(x, y)$ and then the inf can not be reached; the same reason shows that no configuration is optimal and that also (SO) is not a minimum. Only in (3.5) is the maximum always attained, and in fact we will not need to relax that problem.

Let us then consider a manifold M of type 2: we will relax the problems simply by considering them with $\hat{M}$ and $\hat{S}$ in place of M and S. Then a *relaxed transport* from f^+ to f^- is nothing else than a measure $\eta \in \mathcal{M}^+(\hat{S})$ such that $\pi_{0\#}\eta = f^+$ and $\pi_{1\#}\eta = f^-$; in the same way, a *relaxed configuration* is $\{f_t, \overline{E}_t\}$, where the f_t and the $\overline{E}_t$ are now scalar and vectorial measures on $\hat{M}$, if the compatibility condition (3.3) holds and $f_0 = f^+$, $f_1 = f^-$; finally, a *relaxed shape* is now the measure $j(x, \nu)$ if $-\operatorname{div}\nu = f$ and ν is a vectorial measure on $\hat{M}$. The relaxed problems will be then the search of the relaxed transports, or configurations, or shapes, which are extremals for (MK), (Br) or (SO), where the integrals must be intended over $\hat{M}$ and $\hat{S}$ instead of M and S. Finally, note that the problem to maximize (3.5) need not be extended: in fact, the 1-Lipschitz functions on M or $\hat{M}$ are exactly the same by construction, and the integral $\int u\,df$ over M or $\hat{M}$ does not change, since f is a measure on M. Moreover, it is easily seen that (3.5) admits a maximum, recalling that $\int df = 0$ and applying the Ascoli–Arzelà theorem to a maximizing sequence of functions with 0 mean.

3.5. The extension of problem (1.7). As we said in the introduction, in the classical setting there is a connection between the shape optimization problem and the PDE equation (1.7), namely that (u, μ) solves (1.7) if and only if u and $\nu = \mu Du$ separately solve (1.5) and (1.6). Since our finslerian context is too general to give a notion of gradient, in this paper we are not interested in a generalization of (1.7); however, in this section we will briefly describe how this can be done in a particular case, which contains most of the most interesting situations. To give a meaningful definition of "gradient", the strict convexity of j is needed: let us then assume, only for this section, that there is a closed Dirichlet region Σ such that j is continueous on $M \setminus \Sigma$ and, if $x \in \Sigma$ then $j(x, \cdot) \equiv 0$, while otherwise $j(x, \cdot)$ is strictly convex on T_xM, clearly except for the multiple vectors (recall the 1-homogeneity). Then, if $u \in W^{1,\infty}_{\mathrm{loc}}(M)$ and $x \notin \Sigma$ is a point where $Du(x)$ is defined, it is $u(x + \varepsilon v) \approx u(x) + \varepsilon\, Du(x) \cdot v$ and $d(x, x + \varepsilon v) \approx \varepsilon\, j(x, v)$, and then the slope is

$$v \longrightarrow \frac{Du(x) \cdot v}{j(x, v)}. \tag{3.8}$$

On the unit sphere of T_xM this is a continuous function, and then by compactness it admits a maximum $\hat{v}$, which is unique thanks to the strict convexity of j on T_xM. For each $u \in W^{1,\infty}_{\mathrm{loc}}(M)$ and $x \notin \Sigma$ where $Du(x)$ is defined, then, we define the *gradient of u at x* as

$$\nabla_j u(x) := \frac{Du(x) \cdot \hat{v}}{j(x, \hat{v})^2}\hat{v},$$

which is the vector on the direction of the maximal slope such that $j(x, \nabla_j u(x))$ equals that slope; then it easily follows that $\mathrm{Lip}_j(M)$ are exactly the continuous elements of $W^{1,\infty}_{\mathrm{loc}}(M)$ such that $j\big(x, \nabla_j u(x)\big) \leq 1$ for a.e. $x \in M \setminus \Sigma$. Now we need to generalize this concept with respect to a measure: to do this, suppose we are given a measure μ and let $x \notin \Sigma$ such that $D_\mu u(x)$, the μ-tangential gradient (see [7] and [13]), is defined. Let then $\hat{v}$, as before, be the unique maximum of (3.8) on the unit sphere of T_xM: the *μ-tangential gradient of u at x* is

$$\nabla_{j,\mu} u(x) := \frac{D_\mu u(x) \cdot \hat{v}}{j(x, \hat{v})^2}\hat{v},$$

which is clearly defined μ−a.e.; now, we can finally generalize problem (1.7) to the search of a measure $\mu \in \mathcal{M}^+(M)$ and a function $u \in \mathrm{Lip}_j(M)$ such that

$$\begin{cases} -\operatorname{div}(\mu\nabla_{j,\mu} u) = f & \text{on } M \setminus \Sigma \\ j(x, \nabla_{j,\mu} u) = 1 & \mu - \text{a.e. on } M \setminus \Sigma. \end{cases} \tag{3.9}$$

This is a good extension of the classical setting, since we can prove the:

Proposition 3.3. *There are solutions for problem (3.9); in particular, if u and ν are optimal for (3.5) and (SO), then* $(u, j(x, \nu))$ *solves (3.9).*

Proof. Let us take u and ν solving (3.5) and (SO) – Theorem B will prove that such extremals exist – and, defining $\mu = j(x, \nu)$, write $\nu = \theta\mu + \tilde{\nu}$ where $\tilde{\nu} = \nu \llcorner \Sigma$: then $j(x, \theta) \equiv 1$ μ−a.e.; recall now that $D_\mu u(x) \cdot v \leq j(x, v)$ for μ−a.e. x and each v thanks to the 1-Lipschitz property of u and that $\langle \tilde{\nu}, Du\rangle = 0$, as one can infer from the continuity of u and the fact that $-\operatorname{div}\nu = f \in \mathcal{M}(M)$. Since, as we will note with Corollary 4.1, inf (SO) = sup (3.5), we have

$$\begin{aligned} \int_M d\mu &= \int_M j(x, \nu) = \inf\,(SO) = \sup\,(3.5) = \int_M u\, df = \langle u, -\operatorname{div}(\theta\mu + \tilde{\nu})\rangle \\ &= \int_M D_\mu u \cdot \theta\, d\mu \leq \int_M j(x, \theta)\, d\mu = \int_M d\mu; \end{aligned}$$

it follows that $D_\mu u \cdot \theta = j(x, \theta)$, and this implies that $\nabla_{j,\mu} u = \theta$; since problem (3.9) requires the validity of $-\operatorname{div}(\mu\nabla_{j,\mu} u) = f$ only on $M \setminus \Sigma$, (u, μ) solves (3.9). □

The converse of the theorem is only partially true: in fact, if (u, μ) solves (3.9) then u is extremal for (3.5) and $\nu = \mu\nabla_{j,\mu} u$ is such that $\int j(x, \nu) = (SO)$; but $-\operatorname{div}\nu = f$ holds only out of Σ and not necessarily on the whole M (but we can change ν in Σ such that the equation holds on M). However, since $j(x, \cdot) \equiv 0$ in Σ, it follows that the measures μ in the solutions of (3.9) are exactly the optimal shapes.

3.6. Useful operations in the theory of mass transport. Here we define four useful operations we can have to do with the transports and prove their main properties.

3.6.1. The restriction. Let us begin with the restriction: if $\eta \in \mathcal{M}^+(S)$ is a transport from f^+ to f^- and $h \leq f^+$, we define the *restriction* of η to h (or, more precisely, the restriction on the first component) as the transport

$$\eta_{|h} := \left(h_{f^+} \circ \pi_0\right) \eta,$$

where $h_{f^+} \in L^1_{f^+}(M)$ is the density of h with respect to f^+, i.e. $h = h_{f^+} f^+$. It is immediatly noted that $\pi_0 \eta_{|h} = h$, which means that $\eta_{|h}$ is a transport moving h somewhere, and that if $f^+ = g + h$ then $\eta = \eta_{|g} + \eta_{|h}$ and $C(\eta) = C(\eta_{|g}) + C(\eta_{|h})$. In the same way one can make the restriction on the second component to $h \leq f^-$, that we will denote $\eta^{|h}$.

3.6.2. The composition. The second operation we consider is the *composition* of transports: if η transports f^+ onto g and ν transports g onto f^-, it is clear that we would gladly consider some transport $\nu \circ \eta$ that moves f^+ onto f^-; it is not difficult to write down the definition which corresponds to what we have in mind, that turns out out be

$$\langle \nu \circ \eta, \varphi \rangle := \int_{x \in M} \left(\iint_{(\sigma,\tau) \in S \times S} \varphi(\sigma \circ \tau) d\nu_x(\sigma)\, d\eta_x(\tau) \right) dg(x),$$

where ν_x and η_x denote the disintegrations of ν and η with respect to π_0 and π_1 in the sense of Theorem 1.3; note that $\sigma \circ \tau$ is well defined, since for ν_x−a.e. σ and η_x−a.e. τ we have $\sigma(0) = \tau(1) = x$. Some straightforward calculations prove the following:

Proposition 3.4. *The composition $\nu \circ \eta$ is a transport that moves f^+ onto f^-; moreover, $\zeta \circ (\nu \circ \eta) = (\zeta \circ \nu) \circ \eta$; finally the transport density is additive, in the sense that $\mu_{\nu\circ\eta} = \mu_\nu + \mu_\eta$, and then $C(\nu \circ \eta) = C(\nu) + C(\eta)$.*

3.6.3. The simplification. We want to define now the *simplification* of transports: given η_0, which transports $f^+ + \varrho_0$ onto $f^- + \varrho_0$, we would like to build another transport moving f^+ onto f^- in a reasonable way (we mean doing somehow what η_0 does; anyway, this will be more clear with Proposition 3.5): to make this, let us consider $\eta_{0|f^+}$, and write it as $\eta_{0,a} + \eta_{0,b}$, where the first part transports mass on f^- and the second on ϱ_0; formally,

$$\eta_{0,a} := \left(\eta_{0|f^+}\right)^{|\pi_{1\#}(\eta_{0|f^+}) \wedge f^-} \qquad\qquad \eta_{0,b} = \eta_{0|f^+} - \eta_{0,a};$$

on the other hand, we divide $\eta_{0|\varrho_0}$ as $\eta_{0,c} + \eta_{0,d}$, where

$$\eta_{0,c} := \left(\eta_{0|\varrho_0}\right)_{|\pi_{1\#}\eta_{0,b}} \qquad\qquad \eta_{0,d} = \eta_{0|\varrho_0} - \eta_{0,c};$$

so we write $\eta_0 = \eta_{0,a} + \eta_{0,b} + \eta_{0,c} + \eta_{0,d}$, and $\eta_{0,a}$ transports part of f^+ onto a part of f^-, $\eta_{0,b}$ transports the rest of f^+ onto a part of ϱ_0, $\eta_{0,c}$ transports this part of ϱ_0 somewhere and finally $\eta_{0,d}$ concludes everything; note that one can make the composition $\eta_{0,c} \circ \eta_{0,b}$. Let us then define $\tau_0 = 0$, $\tau_1 = \eta_{0,a}$, $\eta_1 = \eta_{0,c} \circ \eta_{0,b} + \eta_{0,d}$, $\varrho_1 = \pi_{0\#}\eta_{0,d}$ and, finally, $f_0^+ = f^+$, $f_0^- = f^-$, $f_1^+ = \pi_{0\#}\eta_1 - \varrho_1$ and $f_1^- = \pi_{1\#}\eta_1 - \varrho_1$. So, since $\|f^+\| \le \|f_0^+ - f_1^+\| + \|f_1^+\|$, recalling that $\tau_0 + \tau_1$ and $\eta_{0,b}$ transport $f_0^+ - f_1^+$ onto $f_0^- - f_1^-$ and f_1^+ onto $\varrho_0 - \varrho_1$, respectively, we obtain

$$\|f^+\|_{\mathcal{M}^+(M)} \le \|\tau_0 + \tau_1\|_{\mathcal{M}^+(S)} + \|\varrho_0 - \varrho_1\|_{\mathcal{M}^+(M)}.$$

But now we have η_1 that transports $f_1^+ + \varrho_1$ onto $f_1^- + \varrho_1$, and then we can iterate our construction finding η_i, τ_i, ϱ_i, f_i^+ and f_i^-; so $\tau_0 + \cdots + \tau_i$ transports $f^+ - f_i^+$ on $f^- - f_i^-$, and the last inequality generalizes as

$$\|f^+\|_{\mathcal{M}^+(M)} \le \|\tau_0 + \cdots + \tau_i\|_{\mathcal{M}^+(S)} + \|\varrho_{i-1} - \varrho_i\|_{\mathcal{M}^+(M)}; \tag{3.10}$$

now, since $\varrho_0 \ge \varrho_1 \ge \cdots \ge \varrho_i$ and $\|\tau_0 + \cdots + \tau_i\| \le \|f^+\|$, we infer that $\|\varrho_{i-1} - \varrho_i\| \to 0$ and $\tau_0 + \cdots + \tau_i \to \tau$; then, thanks to (3.10), it follows that τ transports exactly f^+ onto f^-. This measure τ is what we call the *simplification* of η with respect to ϱ_0: we have the:

Proposition 3.5. *If τ is the simplification of η, $\mu_\tau \leq \mu_\eta$. Recalling that, as noted in Section 3.1, $C(\tau) = \|\mu_\tau\|$ and $C(\eta) = \|\mu_\eta\|$, it follows that $C(\tau) \leq C(\eta)$.*

Proof. The construction made before, together with Proposition 3.4, assures us that $\mu_\eta = \mu_{\tau_0+\cdots+\tau_i} + \mu_{\eta_i}$; then $\mu_{\tau_0+\cdots+\tau_i}$ is an increasing sequence of measures bounded by μ_η, and so $\mu_{\tau_0+\cdots+\tau_i} \to \mu_\infty$ with $\mu_\infty \leq \mu_\eta$. We will conclude by proving that $\mu_\tau \leq \mu_\infty$: in fact, given $0 \leq \varphi \in C(M)$, $\langle \mathcal{H}^1_j \llcorner \sigma, \varphi\rangle$ is a l.s.c. function of σ (as we will better discuss in Lemma 4.5), and then the convergence of $\tau_0 + \cdots + \tau_i$ to τ and the definition (3.2) of μ imply $\langle \mu_\tau, \varphi\rangle \leq \liminf \langle \mu_{\tau_0+\cdots+\tau_i}, \varphi\rangle = \langle \mu_\infty, \varphi\rangle$. □

3.6.4. The fattening. The last operation we consider is, in a certain sense, the inverse of the simplification: given η which transports f^+ onto f^- and $\varrho \in \mathcal{M}^+(M)$, we could need another transport moving $f^+ + \varrho$ onto $f^- + \varrho$; it will be simply $\eta + Id_\varrho$, where Id_ϱ is the trivial transport that moves ϱ onto itself without moving anything: we will call it the *fattening* of η with respect to ϱ; note that the fattening leaves the transport density unchanged.

4. Results

Here are the results of this paper: we begin by proving that on a manifold of type 0 the extremals of the three problems are equal, then we study how to make a passage between the items of the different problems; then we will state and prove our main results, namely that the problems on the manifolds of type 1 and the relaxed problems on the manifolds of type 2 have solutions, and the corresponding optimal measures are the same; finally we will prove the non-intersection of the transport rays in the strictly convex case.

4.1. Equality of the extremals. Here we generalize the first part of Theorem 1.2:

Theorem A. *If M is a manifold of type 0,* $\inf(MK) = \inf(Br) = \inf(SO)$.

Proof. We will prove this theorem in three steps.

<u>*Step 1:*</u> $\inf(MK) \geq \inf(Br)$.

Given the transport η, we define $f_t := \pi_{t\#}\eta$ and $\langle \overline{E}_t, \varphi\rangle := \int_S \varphi(\sigma(t)) \cdot \sigma'(t) d\eta$: it is easy to check that it is an admissible configuration, and its cost is less than $C(\eta)$ by the Jensen inequality (we will make the explicit computation in Proposition 4.2).

<u>*Step 2:*</u> $\inf(Br) \geq \inf(SO)$.

Given the configuration $\{f_t, \overline{E}_t\}$, let us define $\nu = -\int_0^1 \overline{E}_t$: from the compatibility condition (3.3) we infer that $-\operatorname{div}\nu = f$; moreover, the Jensen inequality immediately implies $\int j(x, \nu) \leq C(\{f_t, \overline{E}_t\})$.

<u>*Step 3:*</u> $\inf(SO) \geq \inf(MK)$.

We have already noted that inf(1.1) = inf(MK); let us then for a moment consider the “old” transport problem: thanks to Theorem 1.1 and the continuity of d, we know that there exists an optimal “old” transport γ (while, as we discussed in the previous section, there could be no optimal "new" transport) and there exists a Kantorovich potential u; so

$$\sup(3.5) \geq \int u\, df = \iint u(x) - u(y)\, d\gamma = \iint d(x,y)\, d\gamma = \inf(1.1) \\ = \inf(MK),$$

recalling that $\gamma_1 = f^+$ and $\gamma_2 = f^-$: then this step follows from Proposition 3.2. □

Corollary 4.1. *The statement of Proposition 3.2 can be strengthened by saying that* $\sup(3.5) = \inf(SO)$*: in fact, to prove the previous theorem we showed that* $\sup(3.5) \geq \inf(MK) \geq \inf(Br) \geq \inf(SO)$.

4.2. Passing from one problem to another. Here we will describe how, starting from an object of one of the problems, it is possible to build a "better" object for one of the other problems: we will prove three propositions, which will be the tools with which in Section 4.3 we will prove our main theorems. Through this section, we assume we are dealing with a manifold of type 1 (so M, and consequently S, are complete). Let us begin with the first result:

Proposition 4.2. *For each transport η there exists a configuration $\{f_t, \overline{E}_t\}$ such that $\tilde{\mu} \leq \mu$ and then $C(\{f_t, \overline{E}_t\}) \leq C(\eta)$, where μ and $\tilde{\mu}$ are the transport density and the work density associated to η and $\{f_t, \overline{E}_t\}$, respectively.*

Proof. First of all, given η we define $\{f_t, \overline{E}_t\}$ as

$$f_t := \pi_{t\#}\eta \qquad\qquad \langle \overline{E}_t, \varphi\rangle := \int_S \varphi(\sigma(t)) \cdot \sigma'(t)\, d\eta(\sigma);$$

$\{f_t, \overline{E}_t\}$ is admissible, since $f_0 = f^+$ and $f_1 = f^-$ is clear, and

$$\langle f_t', u\rangle = \frac{d}{dt}\int_M u(x)\, df_t(x) = \frac{d}{dt}\int_S u(\sigma(t))\, d\eta(\sigma) = \int_S Du(\sigma(t)) \cdot \sigma'(t)\, d\eta(\sigma) \\ = \langle \overline{E}_t, Du\rangle = -\langle \operatorname{div} \overline{E}_t, u\rangle.$$

Moreover, let us denote by $\{\eta_x\}$ the disintegration of η with respect to $\pi_t : S \longrightarrow M$, in the sense of Theorem 1.3, and let $h(x) = \int_S \sigma'(t)\, d\eta_x(\sigma)$; then $\overline{E}_t = h f_t$, since

$$\langle \overline{E}_t, \varphi\rangle = \int_S \varphi(\sigma(t)) \cdot \sigma'(t)\, d\eta(\sigma) = \int_M \left(\int_S \varphi(\sigma(t)) \cdot \sigma'(t)\, d\eta_x(\sigma)\right) d\pi_{t\#}\eta(x) \\ = \int_M \varphi(x) \cdot \left(\int_S \sigma'(t)\, d\eta_x(\sigma)\right) df_t(x) = \int_M \varphi(x) \cdot h(x)\, df_t(x) = \langle h f_t, \varphi\rangle$$

(recall that $\sigma(t) = \pi_t(\sigma)$ is equal to x for η_x−a.e. σ). Finally, we can prove $\tilde{\mu} \leq \mu$: in fact, if $0 \leq \varphi \in C_b(M)$ we have

$$\begin{aligned}
\langle \tilde{\mu}, \varphi \rangle &= \int_0^1 \int_M \varphi(x) j(x, d\overline{E}_t)\, dt = \int_0^1 \int_M \varphi(x) j(x, h(x))\, df_t(x)\, dt \\
&= \int_0^1 \int_M \varphi(x) j\left(x, \int_S \sigma'(t)\, d\eta_x(\sigma)\right) df_t(x)\, dt \\
&\leq \int_0^1 \int_M \varphi(x) \left(\int_S j(x, \sigma'(t)) d\eta_x(\sigma)\right) df_t(x)\, dt \\
&= \int_0^1 \int_S \varphi(\sigma(t)) j\big(\sigma(t), \sigma'(t)\big)\, d\eta(\sigma)\, dt = \int_S \left\langle \mathcal{H}^1_j \llcorner \sigma, \varphi \right\rangle d\eta(\sigma) = \langle \mu, \varphi \rangle
\end{aligned}$$

recalling (3.4), the Jensen inequality, the properties of disintegration described in Theorem 1.3, (3.1) and (3.2). This gives $\tilde{\mu} \leq \mu$ and, since $C(\{f_t, \overline{E}_t\}) = \langle \tilde{\mu}, 1 \rangle$ and $C(\eta) = \langle \mu, 1 \rangle$, the proof is completed. □

Remark 4.3. In the previous proposition, $\{f_t, \overline{E}_t\}$ has been built exactly following the paths used by the measure η, so one could think that $C(\eta) = C(\{f_t, \overline{E}_t\})$, while the inequality we proved can also hold strictly. The reason for this fact relies essentially on the main difference between the configuration and the transports: the first are made "effectively", on M, while the second only "in theory", on S. Then the different paths are covered independently by a transport, while they interact when covered by a configuration. In fact, the preceding proposition shows that this interaction lowers the cost thanks to the convexity of j; and if the convexity is strict, also the inequality in the preceding proposition can hold strictly.

Proposition 4.4. *For each configuration $\{f_t, \overline{E}_t\}$ there exists ν such that $-\operatorname{div} \nu = f$, the shape $j(x, \nu)$ is less than the work density μ and then $\int j(x, \nu) \leq C(\{f_t, \overline{E}_t\})$; on the other hand, to each ν such that $-\operatorname{div} \nu = f$ we can associate a configuration $\{f_t, \overline{E}_t\}$ such that $\mu = j(x, \nu)$ and then $C(\{f_t, \overline{E}_t\}) = \int j(x, \nu)$.*

Proof. Given $\{f_t, \overline{E}_t\}$, we define $\nu = -\int_0^1 \overline{E}_t\, dt$; the fact that $-\operatorname{div} \nu = f$ follows from the compatibility condition (3.3), while $j(x, \nu) \leq \mu$ comes directly from the Jensen inequality; then $\int j(x, \nu) \leq \int \mu = C(\{f_t, \overline{E}_t\})$.

On the other hand, given ν let us define

$$f_t := (1-t) f^+ + t f^- \qquad\qquad \overline{E}_t := -\nu;$$

the compatibility condition holds since $\operatorname{div} \overline{E}_t = f = -df_t/dt$, and moreover, $\mu = \int_0^1 j(x, \overline{E}_t) = j(x, \nu)$; then the thesis follows. □

To finish our comparison, we should pass from a shape to a transport and this will require a bit of work from us: first of all, we need the following:

Lemma 4.5. *If $\{\eta_i\}$ is a sequence of retracted transports with $C(\eta_i) \leq H < +\infty$ and $\|\eta_i\| \leq H$, there exists η such that, up to a subsequence, $\eta_i \xrightarrow[\mathcal{M}^+(S)]{*} \eta$ and $\mu \leq \liminf \mu_i$ (i.e. $\langle \mu, \varphi \rangle \leq \liminf \langle \mu_i, \varphi \rangle$ for any $\varphi \geq 0$), where μ and μ_i are the transport densities relative to η and η_i.*

Proof. Let us define

$$\eta_{i,k} := \begin{cases} \eta_i \llcorner \{\sigma \in S : \ l_M(\sigma) < 1\} & \text{if } k = 0 \\ \eta_i \llcorner \{\sigma \in S : \ 2^{k-1} \leq l_M(\sigma) < 2^k\} & \text{otherwise,} \end{cases}$$

and note that for each $k \in \mathbb{N}$ $\{\eta_{i,k}\}$ is, varying i, a sequence of measures on $S_k = \{\sigma \in S : \ l_M(\sigma) \leq 2^k\}$, which is easily seen to be complete and totally bounded and then a compact metric space: then, since $\|\eta_{i,k}\| \leq \|\eta_i\| \leq H$, we find $\eta_k \in \mathcal{M}^+(S)$ such that, up to a subsequence, $\eta_{i,k} \xrightarrow[\mathcal{M}^+(S_k)]{*} \eta_k$. Moreover, for each $k \geq 1$ we have, recalling (2.2), $C(\eta_{i,k}) \geq (2^{k-1} - C)\alpha^{-1}\|\eta_{i,k}\|$, and then

$$\|\eta_{i,k}\|_{\mathcal{M}^+(S)} \leq \frac{H\alpha}{2^{k-1} - C} \qquad \qquad \|\eta_k\|_{\mathcal{M}^+(S)} \leq \frac{H\alpha}{2^{k-1} - C},$$

where the second inequality follows by the first one with a limit for $i \to +\infty$: it is then well defined $\eta = \sum \eta_k$ in $\mathcal{M}(S)$. Moreover we have $\eta_i \xrightarrow{*} \eta$: in fact, given $\varphi \in C(S)$, we have that $\langle \eta_{i,0} + \cdots + \eta_{i,k}, \varphi\rangle \to \langle \eta_0 + \cdots + \eta_k, \varphi\rangle$ and that the two terms differ from $\langle \eta_i, \varphi\rangle$ and $\langle \eta, \varphi\rangle$ less then $H\alpha\|\varphi\| \sum_{l=k}^{+\infty}(2^{l-1} - C)^{-1}$.

To finish the proof, we have to check that $\mu \leq \liminf \mu_i$; first we note that, given $0 \leq \varphi \in C(M)$, $\langle \mathcal{H}^1_j \llcorner \sigma, \varphi\rangle$ is a l.s.c. function of σ; in fact, we know that $\int_0^1 j(\sigma, \sigma')\, dt = \liminf_{\tau \to \sigma} \int_0^1 j(\tau, \tau')\, dt$ for each $\sigma \in S$ and then, since φ is continuous and the image of σ compact, we have also

$$\int_0^1 \varphi(\sigma) j(\sigma, \sigma')\, dt = \liminf_{\tau \to \sigma} \int_0^1 \varphi(\tau) j(\tau, \tau')\, dt,$$

that, recalling the definition (3.1) of $\mathcal{H}^1_j \llcorner \sigma$, gives the searched lower semicontinuity of $\langle \mathcal{H}^1_j \llcorner \sigma, \varphi\rangle$; since $\eta_i \xrightarrow{*} \eta$, (3.2) assures that $\langle \mu, \varphi\rangle \leq \liminf\langle \mu_i, \varphi\rangle$. $\square$

Then let us prove the:

Proposition 4.6. *For each ν such that $-\operatorname{div} \nu = f$ there exists a transport η such that $\mu_\eta \leq j(x, \nu)$ and then $C(\eta) \leq \int j(x, \nu)$, where μ_η is the transport density associated to η.*

Proof. We will make the proof in three steps:

<u>*Step 1:*</u> The smooth case.
Let us first suppose that f and ν are smooth and supported on $M \setminus \partial M$; then, following the construction made in Theorem 4.2 of [1], we prove the thesis: first of all, choose $\varepsilon > 0$ and writing $\omega_t(x) = tf^-(x) + (1-t)f^+(x) + \varepsilon$, we define the smooth vector field

$$\overline{v}_t(x) := \frac{\nu(x)}{\omega_t(x)},$$

which suggests how to move the mass; in other words, we define the smooth flow $\Psi : [0, 1] \times M \longrightarrow M$ as

$$\begin{cases} \Psi(0, x) = x \\ \dfrac{d}{dt}\Psi(t, x) = \overline{v}_t(\Psi(t, x)), \end{cases}$$

and consider the transport $\tilde{\eta} = \xi_{\#}\omega_0$, where $\xi(x)\,(s) = \Psi(s, x)$. We note then that

$$\Psi(t, \cdot)_{\#}\omega_0 = \omega_t\ ; \tag{4.1}$$

in fact, both the measures are easily seen to solve the ODE

$$\begin{cases} \zeta(0, \cdot) = \omega_0 \\ \zeta' = -\operatorname{div}(\overline{v}_t\zeta), \end{cases}$$

and then, all being smooth, equation (4.1) follows from the uniqueness of the solution. Moreover we have that, also thanks to (4.1),

$$\begin{aligned} \langle \mu_{\tilde{\eta}}, \varphi\rangle &= \int_S \big\langle \mathcal{H}^1_j \llcorner \sigma, \varphi\big\rangle\, d\tilde{\eta}(\sigma) = \int_M \big\langle \mathcal{H}^1_j \llcorner \xi(x), \varphi\big\rangle\, d\omega_0(x) \\ &= \int_M \int_0^1 \varphi\big(\Psi(t, x)\big)\, j\big(\Psi(t, x), \overline{v}_t(\Psi(t, x))\big)\, dt\, d\omega_0(x) \\ &= \int_0^1 \int_M \varphi(x)\, j(x, \overline{v}_t(x))\, d\omega_t(x)\, dt = \int_0^1 \langle j(x, \nu), \varphi\rangle\, dt = \langle j(x, \nu), \varphi\rangle. \end{aligned}$$

Note now that $\tilde{\eta}$ is a transport that moves $\omega_0 = f^+ + \varepsilon$ onto $\omega_1 = f^- + \varepsilon$, and then the thesis follows from Proposition 3.5, calling η the simplification of $\tilde{\eta}$.

Step 2: The case when j is Lipschitz.
Let us first suppose that $j(\cdot, \omega)$ is Lipschitz for $|\omega| \leq 1$ uniformly on ω: this means that there exists a constant L such that whenever $|\omega| \leq 1$ the function $j(\cdot, \omega)$ is Lipschitz with constant L (since $j(x, \cdot)$ is 1-homogeneous, it follows that in general $j(\cdot, \omega)$ is Lipschitz with constant $L|\omega|$). Define now ν_i, $\nu_{i,\varepsilon}$ and ν_ε as in (A.1). Take now $0 \leq \varphi \in C_b(M)$; fix i and making, for simplicity, the calculation in $\mathbb{R}^n$ instead of in U_i, we have

$$\nu_{\varepsilon,i}(x) = \int_y \rho_{\varepsilon/2}(x - y)\, d\nu_i\big((1 - \varepsilon)y\big) = \int_y \rho_{\varepsilon/2}(x - \tilde{y})\, d\nu_i(y),$$

writing $\tilde{y} = y/(1 - \varepsilon)$; set also $\bar{\omega}$ such that $\nu_i = \bar{\omega}|\nu_i|$: since $|\bar{\omega}(y)| = 1$ and thanks to the Lipschitz assumption on j, for each δ there exists $\bar{\varepsilon}$ such that $j(x, \omega(y)) \leq j(y, \omega(y)) + \delta$ if $\varepsilon \leq \bar{\varepsilon}$ and $\rho_{\varepsilon/2}(x - \tilde{y}) > 0$. Then, using the fact that $j(x, \cdot)$ is

1-homogeneous and the Jensen inequality, we have

$$\begin{aligned}\langle j(x,\nu_{\varepsilon,i}),\varphi\rangle &= \int_x \varphi(x)\, j(x,\nu_{\varepsilon,i}(x))\,dx = \int_x \varphi(x)\, j\left(x, \int_y \rho_{\varepsilon/2}(x-\tilde y)\,d\nu_i(y)\right) dx\\ &\le \int_x \varphi(x)\int_y j\big(x,\rho_{\varepsilon/2}(x-\tilde y)\,d\nu_i(y)\big)\,dx\\ &= \int_x\int_y \varphi(x)\rho_{\varepsilon/2}(x-\tilde y)\, j\big(x,\bar\omega(y)\big)\,d|\nu_i|(y)\,dx\\ &\le \int_x\int_y \varphi(x)\rho_{\varepsilon/2}(x-\tilde y)\, j\left(y,\bar\omega(y)\right)\,d|\nu_i|(y)\,dx + \delta\,\|\varphi\|\,\|\nu_i\|\\ &= \int_y \varphi_\varepsilon(y) j(y,\nu_i(y)) + \delta\,\|\varphi\|\,\|\nu_i\| = \langle j(x,\nu_i),\varphi_\varepsilon\rangle + \delta\,\|\varphi\|\,\|\nu_i\|,\end{aligned}$$

writing $\varphi_\varepsilon(y) = \varphi * \rho_{\varepsilon/2}(\tilde y)$. Since clearly $\langle j(x,\nu_i),\varphi_\varepsilon\rangle \to \langle j(x,\nu_i),\varphi\rangle$, recalling that δ was arbitrary and arguing as in Theorem A.2. to pass from an estimate in a single chart to a global estimate, we deduce $\liminf j(x,\nu_\varepsilon) \le j(x,\nu)$. Now, using the first step, to each ν_ε we can associate a transport $\tilde\eta_\varepsilon$ which moves f_ε^+ onto f_ε^- with $\mu_{\tilde\eta_\varepsilon} \le j(x,\nu_\varepsilon)$ and moreover, written $\eta_\varepsilon = ret_\#\tilde\eta_\varepsilon$, thanks to hypothesis *1-c)* of Section 2 we know that $\mu_{\eta_\varepsilon} \le \mu_{\tilde\eta_\varepsilon}$. Since $\liminf j(x,\nu_\varepsilon) \le j(x,\nu)$ we know that, up to a subsequence, $C(\eta_\varepsilon) \le H$; moreover, $f_\varepsilon \xrightarrow{*} f$ thanks to Theorem A.2. and then $\|\eta_\varepsilon\| = \|\tilde\eta_\varepsilon\| = \|f_\varepsilon^+\| \le \|f_\varepsilon\| \le H$: so we can apply Lemma 4.5 to the sequence η_ε obtaining η. Now, $\mu_\eta \le \liminf \mu_{\eta_\varepsilon} \le \liminf j(x,\nu_\varepsilon) \le j(x,\nu)$ and, again using $f_\varepsilon \xrightarrow{*} f$, η is a transport from $f^+ + \tau$ to $f^- + \tau$ with a suitable measure $\tau \in \mathcal{M}(M)$. So the thesis follows, possibly applying to η a fattening and a simplification with respect to τ^- and τ^+, respectively.

Step 3: The general case.
We have now only to skip the Lipschitz assumption on j: first of all note that the construction of the preceding step allows us, given ν, to build a transport η which does not depend on j; then, since j is l.s.c., we can choose $j_\varepsilon \nearrow j$ with j_ε Lipschitz and we already know that

$$\mu_\varepsilon := \int_S \mathcal{H}^1_{j_\varepsilon} \llcorner \sigma\, d\eta(\sigma) \le j_\varepsilon(x,\nu) \le j(x,\nu);$$

then the thesis will follow by the fact that $\mu_\varepsilon \xrightarrow{*} \mu$; finally, this fact is easily proved by twice applying the monotone convergence theorem (once to an integral in $[0,1]$ and once to an integral in S). □

4.3. Existence of extremals and equivalence between the optimal measures. We prove here the main results of this paper, which generalize Theorem 1.2 for the manifolds of type 1 and 2, in the second case in the relaxed sense. First of all we have the

Theorem B. *If M is a manifold of type 1, there exist extremals for problems (MK), (Br) and (SO). Moreover, the associated optimal transport densities, work densities and shapes are exactly the same.*

Remark 4.7. As discussed after the statement of Theorem 1.2, in general there is not a uniqueness of the optimal transport densities, work densities and shape: however, as the theorem above claims, each measure $\mu \in \mathcal{M}^+(M)$ is an optimal transport density if and only if it is an optimal work density, and if and only if it is an optimal shape.

Proof of Theorem B. Let us begin by taking an optimal sequence of transports $\{\eta_i\}$ from f^+ to f^-, which we can assume to be retracted since the retraction lowers the costs: thanks to Lemma 4.5, which we can apply since for each i we have $\|\eta_i\| = \|f^+\|$, we get a transport η, and this is also a transport from f^+ to f^- since $\eta_i \xrightarrow{*} \eta$. But η is an optimal transport because $\{\eta_i\}$ was an optimal sequence and $C(\eta) = \langle \mu_\eta, 1\rangle \leq \liminf \langle \mu_{\eta_i}, 1\rangle = \liminf C(\eta_i)$: so we proved that the Monge–Kantorovich problem has a solution. Then, all we stated immediately follows by Propositions 4.2, 4.4 and 4.6. □

Then we prove that the extension of the problems we defined for the manifolds of type 2 is a "good" relaxation:

Theorem C. *The problems defined in Section 3.4 are a standard relaxation in the case when M is a manifold of type 2. In particular, the relaxed problems have solutions and the optimal measures associated to the different problems are the same. Moreover, the extremals* $\inf(MK)$*,* $\inf(Br)$ *and* $\inf(SO)$ *of the original problems do not change with the relaxation. Finally, each relaxed solution is a weak* limit of non-relaxed items such that also the associated measures weak* converges.*

Proof. The first fact comes directly from Theorem B, used on the manifold $\hat{M}$ which, by hypothesis *2-a*, is a manifold of type 1. To prove the other assertions, let us first define the functions $\tau_i : \hat{S} \longrightarrow S$ such that $|l(\sigma) - l(\tau_i(\sigma))| \leq 2^{-i}$, $d_S(\sigma, \tau_i(\sigma)) \leq 2^{-i}$ and, if both the extremals of $\sigma \in \hat{S}$ are in M, then they are the same as $\tau_i(\sigma)$: this can be done thanks to hypothesis *2-b*. If we define now $\eta_i := \tau_{i\#}\eta$, it follows that the η_i are non-relaxed transports; moreover, by the definition of τ_i, it holds that the inequality $C(\eta_i) \leq C(\eta) + 2^{-i}\|\eta\|$: this assures that the extremals have not decreased by relaxing and, since they clearly cannot have been increased (recall, that the relaxed items are more than the non-relaxed ones), also the second assertion follows. In our construction, we have also that $\eta_i \xrightarrow{*} \eta$ and $\mu_{\eta_i} \xrightarrow{*} \mu_\eta$, where the last convergence holds since $C(\eta) = \lim C(\eta_i)$; so the last assertion is now proved for the Monge–Kantorovich problem. To prove it for the other two problems it suffices to use again Propositions 4.2, 4.4 and 4.6. □

Note that proving these theorems we have given some constructions that allow us, starting from a solution of a problem, to build a solution of one of the other problems preserving the optimal measure.

Remark 4.8. We can point out that also sup(3.5) is attained, both in the manifolds of type 1 and 2 (since it was the only problem we did not need to relax in Section 3.4): in fact, using Theorem 1.1 we can take a Kantorovich potential u associated to an optimal transport and, recalling Step 3 in Theorem A, it follows that u is optimal for (3.5).

Remark 4.9. In [9] the question was left open as to whether or not each optimal shape on a Riemannian manifold can be obtained integrating the $\mathcal{H}^1$ measures on some geodesics. Thanks to the theorems we proved, we can give an affirmative answer to this question, also in our more general contexts.

4.4. Non-intersection of transport rays. In this section, we will discuss the generalization of the non-intersection property of the transport rays, which holds in the Euclidean case; through the section, M will be a manifold of type 1: then the same results hold on $\hat{M}$ if M is a manifold of type 2.

It is a well-known fact that, in the Euclidean case, the transport rays – i.e. the paths of a given optimal transport – do not intersect, in the sense that all the different transport rays that pass for a fixed point are parallel: this can be expressed with the equation

$$\int_0^1 \left|\overline{E}_t\right| \, dt = \left| \int_0^1 \overline{E}_t \, dt \right| , \tag{4.2}$$

that holds for each optimal Brenier configuration and can be found for example in [1]. It is easy to understand that the non-intersection can not be generalized in our setting without an hypothesis of strict convexity for j: for example, if in $\mathbb{R}^2$ we set $j(x, v) = v_1$, the horizontal component of a vector, and consider the problem to translate horizontally an homogeneous vertical block, it is clear that it can be done in many optimal ways also with a great intersection of rays. However, let us note the following:

Lemma 4.10. *If $\{f_t, \overline{E}_t\}$ is optimal, then*

$$\int_0^1 j(x, \overline{E}_t) \, dt = j\left(x, \int_0^1 \overline{E}_t \, dt \right) . \tag{4.3}$$

Proof. This is a very easy fact; define $\nu = -\int_0^1 \overline{E}_t$ and calling μ the work density relative to $\{f_t, \overline{E}_t\}$, we have $j(x, \nu) = \mu$: in fact, $j(x, \nu) \leq \mu$ as in Proposition 4.4, and on the other hand the inequality cannot hold strictly since $\{f_t, \overline{E}_t\}$ is optimal. Recalling (3.4), this equality is exactly (4.3) and then the thesis is proved. □

Clearly (4.3) is the perfect parallel to (4.2); let us see how this can be read as a non-intersection result in the strictly convex case: first we recall that, being j 1-homogeneous, we call it strictly convex on T_xM if $j(tv + (1-t)w) \leq tj(x, v) + (1-t)j(x, w)$ for each $v,\ w \in T_xM$ and $t \in (0, 1)$, and the equality holds only if v and w are parallel. Then the hypothesis of strict convexity allows us to prove the:

Theorem D. (Non-intersection of transport rays) *If $\{f_t, \overline{E}_t\}$ is optimal and j is strictly convex on each T_xM, there are vectors $h(x) \in T_xM$ such that $\overline{E}_t(x)$ is parallel to $h(x)$ for a.e. (x, t) (with respect to the measure $|\overline{E}_t| \otimes dt$).*

Proof. First of all, let us write as usual $\overline{E}_t = \omega_t\, d|\overline{E}_t|$, and define $\alpha := |\overline{E}_t| \otimes dt$, a positive measure on $M \times [0,1]$; moreover, we call π the projection of $M \times [0,1]$ on M and operate a disintegration of α with respect to π, finding the measures α_x on $[0,1]$ – to be precise, the α_x are measures on $M \times [0,1]$ concentrated on the couples (y,t) with $y = x$, and then we consider them as measures on $[0,1]$. Setting now

$$h(x) := \int_0^1 \omega_t(x)\, d\alpha_x(t) \qquad\qquad \xi := \pi_\# \alpha, \tag{4.4}$$

we have $\int_0^1 \overline{E}_t = h\, d\xi$, since for each smooth vector field φ it is

$$\begin{aligned}
\Big\langle \int_0^1 \overline{E}_t, \varphi \Big\rangle &= \int_0^1 \int_M \varphi(x) \cdot d\overline{E}_t(x)\, dt = \iint_{M\times[0,1]} \varphi(x)\cdot \omega_t(x)\, d\big(|\overline{E}_t| \otimes dt\big)(x,t) \\
&= \int_M \left(\int_0^1 \varphi(x) \cdot \omega_t(x)\, d\alpha_x(t) \right) d\pi_\#\alpha(x) \\
&= \int_M \varphi(x)\cdot h(x)\, d\xi = \langle h\, d\xi, \varphi\rangle,
\end{aligned}$$

recalling the properties of disintegration; then we can evaluate

$$\begin{aligned}
\int_M j\left(x, \int_0^1 \overline{E}_t\, dt\right) &= \int_M j(x, h(x))\, d\xi(x) = \int_M j\left(x, \int_0^1 \omega_t(x)\, d\alpha_x(t)\right) d\xi(x) \\
&\le \int_M \int_0^1 j(x, \omega_t(x))\, d\alpha_x(t)\, d\xi(x) \\
&= \int_M \int_0^1 j(x, \omega_t(x))\, d\alpha(x,t) \\
&= \int_0^1 \int_M j(x, \omega_t(x))\, d|\overline{E}_t|(x)\, dt = \int_0^1 \int_M j(x, \overline{E}_t)\, dt,
\end{aligned}$$

using the Jensen inequality. But thanks to (4.3), the inequality must be an equality, and using the strict convexity of j it follows that the unit vectors $\omega_t(x)$ are equal, varying t; to be more precise, for ξ–a.e. x the vectors $\omega_t(x)$ are equal for α_x–a.e. t. Recalling that $\overline{E}_t = \omega_t |\overline{E}_t|$ and (4.4); the thesis follows. □

Remark 4.11. Note that, in the presence of a Dirichlet region Σ, the thesis of the theorem cannot be true: since moving in Σ does not cost, there can be intersecting paths also in optimal transports; in fact, if $x \in \Sigma$ then $j(x,\cdot) \equiv 0$ and then the strict convexity fails. However, the assertion of the previous theorem can clearly be strengthened by saying that the different transport rays cannot cross in the points x such that j is strictly convex on $T_x M$ or, equivalently, that for each x such that $j(x,\cdot)$ is strictly convex there exists a vector $h(x)$ on $T_x M$ such that $\overline{E}_t(x)$ is parallel to $h(x)$ for a.e. t.

Appendix A. Vectorial measures on manifolds

Let M be an n-dimensional connected manifold, possibly with a boundary; the space $\mathcal{M}^n(M)$ of the vectorial measures on the manifold M can be defined as the dual space of the continuous sections of the tangent bundle TM. Given a measure $\nu \in \mathcal{M}^n(M)$, one would clearly find it useful (as in the Euclidean case) to write $\nu = \omega|\nu|$ with an unitary vector field ω and a suitable definition of the scalar measure $|\nu|$: it turns out that (exactly as in the Euclidean case) the correct definition for $|\nu| \in \mathcal{M}^+(M)$ has to be

$$\langle |\nu|, h\rangle := \sup\{\langle \nu, \omega h\rangle,\ |\omega| \leq 1\} \qquad \forall\, 0 \leq h \in C_b(M).$$

We define then L^1_ν as the completion of $C_b(M)$ embedded with the norm $\|\varphi\|_{L^1_\nu} := \langle |\nu|, \varphi\rangle$; finally, one can easily prove the following generalization of Riesz's theorem, either using it on the charts of an atlas or directly by intrinsic duality arguments:

Theorem A.1. *There exists a $|\nu|$-measurable unit vector field $\bar\omega$ such that $\nu = \bar\omega|\nu|$: this means that $\langle \nu, \varphi\rangle = \langle |\nu|, \varphi\cdot\bar\omega\rangle$ whenever φ is a continuous section of the tangent bundle TM.*

We could want to approximate a measure $\nu \in \mathcal{M}^n(M)$ with regular measures ν_ε; to do this, let us first fix an atlas for M: so we have a locally finite open covering $\{U_i\}$ of M and the corresponding functions $\psi_i : U_i \to \mathbb{R}^n$, where each ψ_i is a diffeomorphism between U_i and the image $\psi_i(U_i)$ and that image is the open unitary hypercube $(0,1)^n$ if $U_i \cap \partial M = \emptyset$, while it is $(0,1)^{n-1} \times [0,1)$ if $U_i \cap \partial M \neq \emptyset$: moreover, in the last case it is $\psi_i^{-1}\big((0,1)^{n-1} \times \{0\}\big) = U_i \cap \partial M$. Let us define now $\tau_\varepsilon : \mathbb{R}^n \to \mathbb{R}^n$ as $\tau_\varepsilon(x) = (1-\varepsilon)x$ and denote by $\{\varphi_i\}$ a partition of the unity corresponding to the covering $\{U_i\}$ and by ρ_ε the standard convolution kernel. Finally, we set

$$\begin{cases} \nu_i := \varphi_i \nu; \\ \nu_{\varepsilon,i} := {\psi_i^{-1}}_\#\big((\tau_\varepsilon \circ \psi_i)_\# \nu_i * \rho_{\varepsilon/2}\big); \\ \nu_\varepsilon = \sum_i \nu_{\varepsilon,i}. \end{cases} \tag{A.1}$$

This definition is what we needed, since we have the:

Theorem A.2. *$\nu_\varepsilon \in C^\infty(M \setminus \partial M)$ and $\nu_\varepsilon \xrightarrow[\mathcal{M}^n(M)]{*} \nu$. Moreover, if $-\operatorname{div}\nu = f$ then $f_\varepsilon := -\operatorname{div}\nu_\varepsilon \xrightarrow[\mathcal{M}(M)]{*} f$.*

Proof. The smoothness assertion is obvious, since ν_ε is a locally finite sum of smooth functions supported in the interior of M. To prove the weak* convergence of ν_ε to ν, let ω be a continuous field of vectors of norm less than 1: a standard computation, starting from the definition of the convolution and using Fubini's theorem, shows that $\langle \nu_{\varepsilon,i}, \omega\rangle \leq \langle |\nu_i|, 1\rangle$, so the generality of ω assures that $\langle |\nu_{\varepsilon,i}|, 1\rangle \leq \langle |\nu_i|, 1\rangle$. Now, from the definition it follows that $\nu_i = |\nu_i|\,\bar\omega$ (i.e. the ν_i are all parallel to ν, since $\bar\omega$ works for each i), and then

$$\sum_i \langle |\nu_i|, 1\rangle = \sum_i \langle \nu_i, \bar\omega\rangle = \langle \nu, \bar\omega\rangle < +\infty;$$

the existence of a finite set I such that

$$\sum_{i \notin I} \langle |\nu_i|, 1 \rangle \leq \delta \tag{A.2}$$

follows. Then given a continuous vector field h, since $\langle \nu, h \rangle = \sum_i \langle \nu_i, h \rangle$ we can assume, possibly passing to a greater finite set I, that

$$\left| \langle \nu, h \rangle - \sum_{i \in I} \langle \nu_i, h \rangle \right| \leq \delta. \tag{A.3}$$

On the other hand, thanks to (A.2) we have

$$\sum_{i \notin I} \langle \nu_{\varepsilon,i}, h \rangle \leq \|h\|_{C_b(M)} \sum_{i \notin I} \langle |\nu_{\varepsilon,i}|, 1 \rangle \leq \|h\| \sum_{i \notin I} \langle |\nu_i|, 1 \rangle \leq \delta \|h\|. \tag{A.4}$$

Since clearly for each i it is $\nu_{\varepsilon,i} \xrightarrow[\varepsilon \to 0]{*} \nu_i$ and I is finite, if $\varepsilon \ll 1$ we have

$$\left| \sum_{i \in I} \langle \nu_{\varepsilon,i} - \nu_i, h \rangle \right| \leq \delta. \tag{A.5}$$

Putting together (A.3), (A.4) and (A.5), we infer

$$\left| \langle \nu, h \rangle - \langle \nu_\varepsilon, h \rangle \right| \leq \delta \left(2 + \|h\|_{C_b(M)} \right),$$

and from the generality of h it follows that $\nu_\varepsilon \xrightarrow{*} \nu$. The same argument, made in the scalar case, also proves the last assertion, just applying the definition in order to write f_ε as a convolution also. □

Acknowledgements. I wish to thank L. Ambrosio and G. Buttazzo for their help and for many hints during the preparation of this work, and L. De Pascale for the useful discussions we had on this subject.

I am also indebted to the referee of this paper for his really careful reading and his precious comments, which helped make this paper more accurate and readable.

References

1. Ambrosio, L.: Lecture notes on optimal transport problems. In: Mathematical aspects of evolving interfaces. Lecture Notes in Mathematics, Vol. 1812, pp. 1–52. Springer 2003
2. Ambrosio, L., Fusco, N., Pallara, D.: Functions of Bounded Variation and Free Discontinuity Problems. Oxford: Oxford University Press 2000
3. Bangert, V.: Minimal measures and minimizing closed normal one-currents. Geom. Funct. Anal. **9**, 413–427 (1999)
4. Bao, D., Chern, S.S., Shen, Z.: An introduction to Riemann-Finsler geometry. Graduate Texts in Mathematics, Vol. 200. New York: Springer 2000
5. Benamou, J.-D., Brenier, Y.: A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. Numer. Math. **84**, 375–393 (2000)
6. Bouchitté, G., Buttazzo, G.: Characterization of optimal shapes and masses through Monge–Kantorovich equation. J. Eur. Math. Soc. (JEMS) **3**, 139–168 (2001)

7. Bouchitté, G., Buttazzo, G., Seppecher, P.: Shape Optimization Solutions via Monge–Kantorovich Equation. C. R. Acad. Sci. Paris, Sér. I, Math. **324**, 1185–1191 (1997)
8. Brenier, Y.: A Homogenized Model for Vortex Sheets. Arch. Ration. Mech. Anal. **138**, 319–353 (1997)
9. De Pascale, L.: On a constrained shape optimization problem. Commun. Appl. Anal. **7** (2003)
10. De Pascale, L., Pratelli, A.: Regularity properties for Monge transport density and for solutions of some shape optimization problem. Calc. Var. Partial Differ. Equ. **14**, 249–274 (2002)
11. Evans, L.C.: Partial Differential Equations and Monge–Kantorovich Mass Transfer. Current Developments in Mathematics, pp. 65–126, 1997
12. Evans, L.C., Gangbo, W.: Differential Equations Methods for the Monge–Kantorovich Mass Transfer Problem. Mem. Am. Math. Soc. **137** (1999)
13. Fragalà, I., Mantegazza, C.: On some notions of tangent space to a measure. Proc. R. Soc. Edinb., Sect. A, Math. **129**, 331–342 (1999)
14. Gangbo, W., McCann, R.J.: The geometry of optimal transportation. Acta Math. **177**, 113–161 (1996)
15. Kantorovich, L.V.: On the transfer of masses. Dokl. Akad. Nauk, Ross. Akad. Nauk **37**, 227–229 (1942)
16. Kantorovich, L.V.: On a problem of Monge. Usp. Mat. Nauk **3**, 225–226 (1948)