DECOMPOSITION OF GEODESICS IN THE WASSERSTEIN SPACE AND THE GLOBALIZATION PROBLEM

FABIO CAVALLETTI

ABSTRACT. We will prove a decomposition for Wasserstein geodesics in the following sense: let (X, d, m) be a non-branching metric measure space verifying $\mathsf{CD}_{loc}(K, N)$ or equivalently $\mathsf{CD}^*(K, N)$. Then every geodesic μ_t in the L^2-Wasserstein space, with $\mu_t \ll m$, is decomposable as the product of two densities, one corresponding to a geodesic with support of codimension one verifying $\mathsf{CD}^*(K, N - 1)$, and the other associated with a precise one dimensional measure, provided the length map enjoys local Lipschitz regularity. The motivation for our decomposition is in the use of the component evolving like CD^* in the globalization problem.

For a particular class of optimal transportation we prove the linearity in time of the other component, obtaining therefore the global $\mathsf{CD}(K, N)$ for μ_t. The result can be therefore interpret as a globalization theorem for $\mathsf{CD}(K, N)$ for this class of optimal transportation, or as a "self-improving property" for $\mathsf{CD}^*(K, N)$.

Assuming more regularity, namely in the setting of infinitesimally strictly convex metric measure space, the one dimensional density is the product of two differentials giving more insight on the density decomposition.

CONTENTS

1. Introduction

The class of metric measure spaces with generalized lower bounds on the Ricci curvature formulated in terms of optimal transportation, has been introduced by Sturm in [20, 21] and independently by Lott and Villani in [18]. The spaces belonging to this class are called $\mathsf{CD}(K, N)$-spaces and the condition characterizing them is denoted with $\mathsf{CD}(K, N)$.

In the *curvature-dimension condition* $\mathsf{CD}(K, N)$ the two parameters K and N play the role of a curvature lower bound and a dimension upper bound, respectively. Among the many relevant properties enjoyed by $\mathsf{CD}(K, N)$, the following one also serves as a motivation: a complete Riemannian manifold satisfies $\mathsf{CD}(K, N)$ if and only if its Ricci curvature is bounded from below by K and its dimension from above by N.

Roughly speaking curvature-dimension condition $\mathsf{CD}(K, N)$ prescribes how the volume of a given set is affected by curvature when it is moved via optimal transportation. It imposes that the distortion is ruled by a coefficient denoted by $\tau_{K,N}^{(t)}(\theta)$ depending on the curvature K, on the dimension N, on the time of the evolution t and on the point length θ. The main feature of $\tau_{K,N}^{(t)}(\theta)$ is that it is obtained mixing two different volume distortions: an $(N-1)$-dimensional distortion depending on the curvature K and a one dimensional evolution that doesn't contain any curvature information. Namely

$$\tau_{K,N}^{(t)}(\theta) = t^{1/N} \sigma_{K,N-1}^{(t)}(\theta)^{(N-1)/N},$$

where $\sigma_{K,N-1}^{(t)}(\theta)^{(N-1)/N}$ contains the information on the $(N-1)$-dimensional volume distortion and the evolution in the remaining direction is ruled just by $t^{1/N}$. The coefficient $\sigma_{K,N}^{(t)}(\theta)$ is the solution (in time) of the second order differential equation

$$y'' + \theta^2 \frac{K}{N} y = 0, \qquad y(0) = 0, \quad y'(0) = 1.$$

The previous equation appears naturally in the study of the Jacobian of the differential of the exponential map in the context of differential geometry, and indeed it rules the part of the Jacobian associated to the restriction to an hyperplane of the differential of the exponential map, see [21] for more details.

A broad variety of geometric and functional analytic properties can be deduced from the curvature-dimension condition $\mathsf{CD}(K, N)$: the Brunn-Minkowski inequality, the Bishop-Gromov volume comparison theorem, the Bonnet-Myers theorem, the doubling property and local Poincaré inequalities on balls. All these listed results are in a quantitative form (volume of intermediate points, volume growth, upper bound on the diameter and so on) depending on K, N.

One of the most important questions on $\mathsf{CD}(K, N)$ that are still open, and we will try to understand in this note, is whether this notion enjoys a globalization property: can we say that a metric measure space (X, d, m) satisfies $\mathsf{CD}(K, N)$ provided $\mathsf{CD}(K, N)$ holds true locally on a family of sets X_i covering X?

A first tentative of answer this problem was given by Bacher and Sturm in [7]: they proved that a non-branching metric measure space (X, d, m) verifies the local curvature-dimension condition $\mathsf{CD}_{loc}(K, N)$ if and only if it verifies the global reduced curvature-dimension condition $\mathsf{CD}^*(K, N)$. The latter is obtained from $\mathsf{CD}(K, N)$ imposing that the volume distortion, during the evolution through an optimal transportation, is ruled by $\sigma_{K,N}^{(t)}(\theta)$ instead of $\tau_{K,N}^{(t)}(\theta)$. The $\mathsf{CD}^*(K, N)$ is a priori weaker than $\mathsf{CD}(K, N)$ and the converse comparison can be obtained only changing the value of the lower bound on the curvature: condition $\mathsf{CD}^*(K, N)$ implies $\mathsf{CD}(K^*, N)$ where $K^* = K(N-1)/N$ (for $K \geq 0$ and for $K < 0$ a suitable formula holds). Therefore the curvature condition contained in $\mathsf{CD}^*(K, N)$ is a priori weaker than the one contained $\mathsf{CD}(K, N)$.

Roughly speaking, the main reason why the globalization property holds for the reduced curvature-dimension condition stays in the good behavior (in time) of $\sigma_{K,N}^{(t)}(\theta)$, which in turn can be led back to the previous second order differential equation. The same approach applied to $\mathsf{CD}(K, N)$, that is try to prove the globalization property for $\mathsf{CD}(K, N)$ directly from the properties of $\tau_{K,N}^{(t)}(\theta)$, seems to not work.

A different approach to the problem has been presented by the author together with Sturm in [11]. The approach in [11] was, in the case of an optimal transportation between a diffuse measure and a Dirac delta, to isolate a local $(N-1)$-dimensional evolution ruled by $\sigma_{K,N-1}^{(t)}(\theta)$ and then using the nice properties

of $\sigma_{K,N-1}^{(t)}(\theta)$, obtain a global $(N-1)$-dimensional evolution ruled by the coefficient $\sigma_{K,N-1}^{(t)}(\theta)$. Then using Hölder inequality and the linear behavior of the other direction, pass from the $(N-1)$-dimensional version to the full-dimensional version with coefficient $\tau_{K,N}^{(t)}(\theta)$.

So the strategy was to reproduce in the setting of metric measure spaces the calculations done in the Riemannian framework where, taking advantage of parallel transport, from $Ric \geq K$ it is possible to split the Jacobian determinant of the differential of the exponential map into two components: one of codimension 1 evolving accordingly to $\sigma_{K,N-1}$ and one representing the distortion in the direction of motion that is concave.

To be more precise in [11] it is proved that if (X, d, m) is a non-branching metric measure space that verifies $\mathsf{CD}_{loc}(K, N)$ then it verifies the weaker $\mathsf{MCP}(K, N)$. While $\mathsf{CD}(K, N)$ is a condition on the optimal transport between any pair of absolutely continuous (w.r.t. m) probability measure on X, $\mathsf{MCP}(K, N)$ is a condition on the optimal transport between a Dirac delta and the uniform distribution m on X. Indeed to detect the $(N-1)$-dimensional evolution it is necessary to decompose the whole evolution. Considering the optimal transport between a Dirac mass in o and the uniform distribution m permits to immediately understand that the family of spheres around o provides the correct $(N-1)$-dimensional support of the evolving measures. So the choice of a Dirac delta as second marginal was really crucial and strongly influenced the geometry of the optimal transportation.

The aim of this paper is to identify, in the general case of optimal transportation between any measures, the $(N-1)$-dimensional evolution verifying $\mathsf{CD}^*(K, N-1)$ and, starting from that, provide a decomposition for densities of geodesics that can be interpret as a parallel transport. The N-dimensional density will be written as the product between the $(N-1)$-dimensional density verifying $\mathsf{CD}^*(K, N-1)$ and of a 1-dimensional density not necessarily associated to a 1-dimensional geodesic. In the framework of infinitesimally strictly convex spaces, the 1-dimensional density will be obtained as the product of two differential, producing then a more direct decomposition.

We will construct a full decomposition for any optimal transportation verifying a local Lipschitz regularity, see Assumption 1 and 2 for the precise hypothesis. We apply this decomposition to the globalization problem for $\mathsf{CD}(K, N)$. With this approach we are able to reduce the problem to prove the concavity in time of the 1-dimensional density, provided Assumption 1 and Assumption 2 are verified. It is important to underline here that in the framework of Riemannian manifolds endowed with volume measures both Assumption 1 and 2 are proved to hold.

Moreover in the particular case of optimal transport plans giving the same speed to geodesics leaving from the same level set of the associated Kantorovich potential, we prove indeed both regularity and linearity of the 1-dimensional factor and we get the full $\mathsf{CD}(K, N)$ inequality. So we prove the global estimate of $\mathsf{CD}(K, N)$ for a certain class of optimal transportation, clearly including all the cases treated in [11].

We now present the paper in more details.
Let (X, d, m) be a non-branching metric measure space verifying the local curvature dimension condition and $\mu_t = \varrho_t m$ be a geodesic (in the L^2-Wasserstein space) that we want to decompose as stated before. The first difficulty we have to handle with is to find a suitable partition of the space. Unlikely optimal transportations connecting measures to deltas, there is not just a universal family of sets but one for each $t \in [0, 1]$: if φ is a Kantorovich potential associated to (μ_0, μ_1), then

$$\{\gamma_t : \varphi(\gamma_0) = a, \gamma \in \mathrm{supp}(\boldsymbol{\gamma})\}_{a \in \mathbb{R}}$$

is the family of partitions, one for each $t \in [0, 1]$, that will be considered. Here $\boldsymbol{\gamma} \in \mathcal{P}(\mathcal{G}(X))$ is a dynamical optimal transference plan of μ_t and $\mathcal{P}(\mathcal{G}(X))$ denotes the space of probability measures over $\mathcal{G}(X)$, the space of geodesic in X endowed with the uniform topology inherited as a subset of $C([0, 1], X)$.

The intuitive reason suggesting that the previous family is the right one, stays in the Brenier-McCann Theorem for optimal transportation on manifold that gives a precise formula for the optimal maps:

$$T_t(x) = \exp_x(-t\nabla\varphi(x)), \qquad (T_t)_\sharp \mu_0 = \mu_t.$$

By definition, geodesics on manifold verify $\nabla_{\dot\gamma}\dot\gamma = 0$, where ∇ only here denotes the Levi-Civita connection, meaning that there is no curvature in the direction of γ. Hence the direction orthogonal to the motion should be the one carrying all the curvature information. Since $\dot\gamma_0 = -\nabla\varphi$, (here $\nabla\varphi$ is the gradient of φ) the family of sets orthogonal to the motion are the level sets of φ.

On the rigorous mathematical side, the reason why that family is the right one stays in the following property: the set
$$\{(\gamma_0, \gamma_1) \in X \times X : \varphi(\gamma_0) = a\}$$
is d-cyclically monotone (Proposition 4.1). Hence for $\gamma \neq \hat{\gamma} \in \mathrm{supp}(\boldsymbol{\gamma})$ with $\varphi(\gamma_0) = \varphi(\gamma_1)$ it holds
$$\gamma_s \neq \gamma_t, \quad \forall s, t \in (0, 1).$$
Therefore for $s \neq t$, $\{\gamma_s : \varphi(\gamma_0) = a\}$ and $\{\gamma_t : \varphi(\gamma_0) = a\}$ are disjoint. This key property permits to consider the evolution of each "slice" of the geodesic μ_t, where with "slice" we mean its conditional measure with respect to the level sets of the chosen Kantorovich potential.

Here the structure is very rich. Using this new property of d-cyclical monotonicity, it is possible to construct L^2-Wasserstein geodesics with also d-monotone support. The whole construction does not rely on any curvature bound of the space and its interest goes beyond the scope of this paper. For this reason we commit Section 4 to the presentation of these results in their fully generality.

As it is well known, any d-monotone set is formed by family of geodesics that do not intersect at any time. For this reason a translation along this geodesics is well defined. Denote by ϕ_a a Kantorovich potential associated to the d-monotone set $\{(\gamma_s, \gamma_t) : \gamma \in G_a, s \leq t \in [0, 1]\}$. The crucial idea to construct L^2-geodesics is to move via "translation" level sets of ϕ_a to level sets of ϕ_a. As proved in Lemma 4.6 and Proposition 4.7 this will produce a geodesic in the L^2-Wasserstein space, showing a new connection between L^1 and L^2 optimal transportation problems.

The relevance of this construction for the globalization problem stays in the following property: the family of geodesics obtained in Section 4 have a linear structure on each geodesic forming the d-monotone set. Therefore there is one degree of freedom to play with. This property, that was already present in [11] but somehow hidden, will be fundamental here to improve the curvature estimates for the element of codimension one passing from N to $N - 1$.

Coming back the the decomposition, if we want to perform a dimensional reduction argument on measures the right tool is Disintegration Theorem (Theorem 2.18): (Proposition 5.2)
$$\boldsymbol{\gamma} = \int_{\varphi(\mu_0)} \boldsymbol{\gamma}_a \mathcal{L}^1(da), \qquad \boldsymbol{\gamma}_a \in \mathcal{P}(G), \quad \boldsymbol{\gamma}_a(\{\gamma \in G : \varphi(\gamma_0) = a\}) = \|\boldsymbol{\gamma}_a\|,$$
where $\varphi(\mu_0) = \varphi(\mathrm{supp}(\mu_0))$ and G is the support of $\boldsymbol{\gamma}$. Since
$$\mu_t = (e_t)_\sharp \boldsymbol{\gamma} = \int_{\varphi(\mu_0)} (e_t)_\sharp \boldsymbol{\gamma}_a \mathcal{L}^1(da),$$
the geodesics of codimension one that should verify curvature estimates like $\mathsf{CD}^*(K, N - 1)$ is $t \mapsto (e_t)_\sharp(\boldsymbol{\gamma}_a)$, for all $a \in \varphi(\mu_0)$. Since curvature properties in metric measure spaces are formulated in terms of a reference measure and $(e_t)_\sharp(\boldsymbol{\gamma}_a)$ is singular with respect to m, it is not obvious which reference measures of codimension one we have to choose. One option could be to consider for each $t \in [0, 1]$, the family
$$\{\gamma_t : \varphi(\gamma_0) = a, \gamma \in G\}_{a \in \varphi(\mu_0)}.$$
Then for each $t \in [0, 1]$, by d^2-cyclical monotonicity, the family is a partition of $e_t(G)$ and hence we have (Proposition 5.2 and Lemma 5.4)
$$m \llcorner_{e_t(G)} = \int_{\varphi(\mu_0)} \hat{m}_{a,t} \mathcal{L}^1(da), \qquad \hat{m}_{a,t}(\{\gamma_t : \varphi(\gamma_0) = a\}) = \|\hat{m}_{a,t}\|.$$
But the $(N - 1)$-dimensional measures $\hat{m}_{a,t}$ are not the right reference measures to prove $\mathsf{CD}^*(K, N - 1)$ estimate for the densities of $(e_t)_\sharp \boldsymbol{\gamma}_a$. Indeed if $(e_t)_\sharp \boldsymbol{\gamma} = \mu_t = \varrho_t m$, then,
$$\int \varrho_t \hat{m}_{a,t} \mathcal{L}^1(da) = \varrho_t m \llcorner_{e_t(G)} = \mu_t = \int (e_t)_\sharp \boldsymbol{\gamma}_a \mathcal{L}^1(da)$$
and by uniqueness of disintegration $(e_t)_\sharp \boldsymbol{\gamma}_a = \varrho_t \hat{m}_{a,t}$ and therefore the density is ϱ_t and no gain in dimension is possible.

The correct reference measures are built as follows. For each $a \in \varphi(\mu_0)$, consider the following family of sets
$$\{\gamma_t : \varphi(\gamma_0) = a, \gamma \in G\}_{t \in [0,1]},$$

that is for a fixed a we take all the evolutions for $t \in [0,1]$ of the level set a of φ.

By d-cyclical monotonicity, they are disjoint (Lemma 4.2). If $\bar{\Gamma}_a(1) := \cup_{t \in [0,1]} \{\gamma_t : \varphi(\gamma_0) = a, \gamma \in G\}$, then (Proposition 5.6)

$$m \llcorner_{\bar{\Gamma}_a(1)} = \int_{[0,1]} m_{a,t} \mathcal{L}^1(dt), \qquad m_{a,t}(\{\gamma_t : \varphi(\gamma_0) = a\}) = \|m_{a,t}\|.$$

Since in the disintegration above the quotient measure is supported on $[0,1]$, that is the range of the time variable, $m_{a,t}$ should be interpret as the conditional measure moving (with t) in the same direction of the optimal transportation.

In order to apply the results of Section 4 to get an improvement of curvature estimates, we have to show that $(e_t)_\sharp \boldsymbol{\gamma}_a \ll m_{a,t}$. After having that, to get the improvement one could use the "linear" structure of geodesics of Section 4 together with the curvature bound estimate they have to satisfy because of $(e_t)_\sharp \boldsymbol{\gamma}_a \ll m_{a,t}$.

So suppose that we have already proved $(e_t)_\sharp (\boldsymbol{\gamma}_a) = h_{a,t} m_{a,t}$ and $t \mapsto h_{a,t}(\gamma_t)$ satisfies the local (and hence the global) reduced curvature-dimension condition $\mathsf{CD}^*_{loc}(K, N-1)$. Then the situation would be

$$h_{a,t} m_{a,t} = e_{t\,\sharp} \boldsymbol{\gamma}_a = \varrho_t \hat{m}_{a,t}.$$

Our final scope is to prove properties on ϱ_t, and to translate information on $h_{a,t}$ into information on ϱ_t is necessary to put in relation the two different reference measures of codimension one $m_{a,t}$ and $\hat{m}_{a,t}$.

Actually the path we will adopt in the note will be the other way round. First we will show that $\lambda_t m_{a,t} = \hat{m}_{a,t}$ for some function λ_t defined on $e_t(G)$ and then from that we deduce that $(e_t)_\sharp (\boldsymbol{\gamma}_a)$ can be written as $h_{a,t} m_{a,t}$. After that we will prove $\mathsf{CD}^*(K, N-1)$ for $h_{a,t}$. We will obtain a decomposition of the following type

$$\varrho_t = \frac{1}{\lambda_t} h_{a,t}$$

and therefore to prove curvature estimate for ϱ_t also information on λ_t are needed.

We have additional properties of λ_t, that will permit to prove the full $\mathsf{CD}(K, N)$ estimate for ϱ_t, in the particular case of optimal transportation giving constant speed to geodesics leaving from the same level sets and not inverting the level sets of φ during the evolution, that is

$$L(\gamma) = f(\varphi(\gamma_0)), \quad \boldsymbol{\gamma} - a.e. \ \gamma \in \mathcal{G}(X),$$

with $f : \varphi(\text{supp}[\mu_0]) \to \mathbb{R}$ such that $a \mapsto a - f^2/2$ is a non increasing function of a. This condition permits to say, see Lemma 5.1, that a level set of φ after time t is moved to a level set of φ_t and this produce a simplification on the geometry of the optimal transportation. Indeed under this assumption, the map $t \mapsto \lambda_t(\gamma_t)$ is linear.

Due to the relevance of this family of optimal transportations and to better explain why λ_t is linear, we will first present part of the decomposition procedure in Section 5 under this additional assumption on the length of geodesics. In particular in Section 5 we will show that (Proposition 5.2, Lemma 5.4 and Proposition 5.6)

$$(1.1) \qquad m \llcorner_{e_t(G)} = \int_{\varphi(\mu_0)} \hat{m}_{a,t} \mathcal{L}^1(da), \qquad m \llcorner_{\bar{\Gamma}_a(1)} = \int_{[0,1]} m_{a,t} \mathcal{L}^1(dt),$$

and (Proposition 5.2 and Lemma 5.4)

$$(1.2) \qquad \hat{m}_{a,t} \ll \mathcal{S}^h \llcorner_{e_t(G_a)}, \qquad m_{a,t} \ll \mathcal{S}^h \llcorner_{e_t(G_a)}.$$

The latter will be fundamental in order to compare $m_{a,t}$ to $\hat{m}_{a,t}$. Here $\mathcal{S}^h$ denotes the spherical Hausdorff measure of codimension one, see Section 2.2. The proofs of these results will be easier and shorter compared to the one in the general case.

In Section 6 we prove (1.1) and (1.2) without the extra assumption on the shape of the Wasserstein geodesic. Anyway while (1.1) can be proven with no difficulties, the proof of (1.2) necessary relies on some regularity property of two important function and it is here that we have to introduce Assumption 1 and Assumption 2. The functions are the length map at time t for $t \in (0,1)$, that is $L_t : e_t(G) \to (0, \infty)$ defined by

$$L_t(\gamma_t) = L(\gamma).$$

And the map $\Phi_t : e_t(G) \to \mathbb{R}$ defined by $\Phi_t(\gamma_t) = \varphi(\gamma_0)$. Thanks to the non branching assumption on the space, both functions are well defined. Note that here we also observe that in the hypothesis of Section 5, both Assumption 1 and Assumption 2 are verified by L_t and Φ_t. Moreover we prove that Assumption 1 and Assumption 2 hold if (X, d, m) is a Riemannian manifold with Riemannian volume.

In Section 7 through a careful blow-up analysis (Proposition 7.3, Proposition 7.5 and Lemma 7.6), we prove that

$$\hat{m}_{a,t} \ll m_{a,t}.$$

If $\hat{m}_{a,t} = \lambda_t m_{a,t}$, we also prove (Theorem 7.8) that

$$\frac{1}{\lambda_t(\gamma_t)} = \lim_{s \to 0} \frac{\Phi_t(\gamma_t) - \Phi_t(\gamma_{t+s})}{s}.$$

This result is a key step in the proof of the aforementioned decomposition of ϱ_t. It clarifies the expression of one of the two function decomposing ϱ_t. Moreover as a consequence (Corollary 7.7) for every $t \in [0,1]$, we have $(e_t)_\sharp(\boldsymbol{\gamma}_a) \ll m_{a,t}$.

In Section 8 we show that if $h_{a,t}$ is the density introduced before, then $t \mapsto h_{a,t}(\gamma_t)$ satisfies the local reduced curvature-dimension condition $\mathsf{CD}^*_{loc}(K, N-1)$ (Theorem 8.2) and therefore $\mathsf{CD}^*(K, N-1)$. Here the main point, as already said before, is to use the results of Section 4 and consider a geodesic in the Wasserstein space, absolute continuous with respect to m, moving in the same direction of $t \mapsto (e_t)_\sharp \boldsymbol{\gamma}_a$ Taking inspiration from the Riemannian framework, the volume distortion affects only $(N-1)$ dimensions.

So up to normalization constant

$$h_{a,t} m_{a,t} = (e_t)_\sharp(\boldsymbol{\gamma}_a) = \varrho_t \hat{m}_{a,t} = \varrho_t \lambda_t m_{a,t},$$

with $h_{a,t}$ verifying $\mathsf{CD}^*(K, N-1)$. We have therefore proved the following result (Theorem 8.3)

Theorem 1.1. *Let (X, d, m) be a non-branching metric measure space verifying $\mathsf{CD}_{loc}(K, N)$ or $\mathsf{CD}^*(K, N)$ and let $\{\mu_t\}_{t \in [0,1]} \subset \mathcal{P}_2(X, d, m)$ be a geodesic with $\mu_t = \varrho_t m$. Assume moreover Assumption 1 and Assumption 2. Then*

$$\varrho_t(\gamma_t) = C(a) \frac{1}{\lambda_t(\gamma_t)} h_{a,t}(\gamma), \qquad \boldsymbol{\gamma} - a.e. \ \gamma \in G,$$

where $a = \varphi(\gamma_0)$ and $C(a) = \|\boldsymbol{\gamma}_a\|$ is a constant depending only on a. The map $[0,1] \ni t \mapsto h_{a,t}(\gamma)$ verifies $\mathsf{CD}^*(K, N-1)$ for $\boldsymbol{\gamma}$-a.e. $\gamma \in G$ and

$$\frac{1}{\lambda_t(\gamma_t)} = \lim_{s \to 0} \frac{\Phi_t(\gamma_t) - \Phi_t(\gamma_{t+s})}{s}.$$

The constant $C(a)$ of Theorem 1.1 has the following explicit formula

$$C(a) = \left(\int \varrho_t(z) \hat{m}_{a,t}(dz) \right)$$

where $a = \varphi(\gamma_0)$. Note again that the value of the integral does not depend on time, but just on a and therefore in order to prove $\mathsf{CD}(K, N)$-like estimates, the integral can be dropped out.

In the second part of Section 8 we prove that under the same assumptions of Section 5 the function $\lambda_t(\gamma_t)$ is linear in t (Proposition 8.4). Hence we have obtained the other main result of this note (Theorem 8.5).

Theorem 1.2. *Let (X, d, m) be a non-branching metric measure space verifying $\mathsf{CD}_{loc}(K, N)$ or $\mathsf{CD}^*(K, N)$ and let $\{\mu_t\}_{t \in [0,1]} \subset \mathcal{P}_2(X, d, m)$ be a geodesic with $\mu_t = \varrho_t m$. Assume moreover that*

$$L(\gamma) = f(\varphi(\gamma_0)),$$

for some $f : \varphi(\mu_0) \to (0, \infty)$ such that $\varphi(\mu_0) \ni a \mapsto a - f^2/a$ is non increasing. Then

$$\varrho_t(\gamma_t)^{-1/N} \geq \varrho_0(\gamma_0)^{-1/N} \tau_{K,N}^{(1-t)}(d(\gamma_0, \gamma_1)) + \varrho_1(\gamma_1)^{-1/N} \tau_{K,N}^{(s)}(d(\gamma_0, \gamma_1)),$$

for every $t \in [0, 1]$ and for $\boldsymbol{\gamma}$-a.e. $\gamma \in G$.

The family of geodesics verifying the hypothesis of Theorem 1.2 includes for instance all of those optimal transportation having as Kantorovich potential

$$\varphi(x) = \frac{1}{2} d^2(x, A)$$

for any $A \subset X$. Indeed such φ is d^2-concave and its weak upper gradient is always one. No assumption on A is needed and therefore no assumption on the shape of φ^c.

We conclude the note with Section 9 where assuming the space to be infinitesimally strictly convex (see (2.9)), we prove that (Proposition 9.2)

$$\frac{1}{\lambda_t(\gamma_t)} = D\Phi_t(\nabla \varphi_t)(\gamma_t), \qquad \boldsymbol{\gamma} - a.e.\gamma,$$

and hence the general decomposition: up to a constant (in time) factor become

$$\varrho_t = D\Phi_t(\nabla \varphi_t)h_t.$$

We conclude the note with a formal calculation in the Euclidean space putting in relation $D\Phi_t(\nabla \varphi_t)$ with the Hessian of φ_t.

Our starting hypothesis can be chosen to be equivalently $\mathsf{CD}_{loc}(K, N)$ or $\mathsf{CD}^*(K, N)$. Hence the results proved can be read from two different perspective, accordingly to $\mathsf{CD}_{loc}(K, N)$ or $\mathsf{CD}^*(K, N)$. From the point of view of $\mathsf{CD}^*(K, N)$, where the globalization property is already known, the main result is that for nice optimal transportations the entropy inequality can be improved to the curvature-dimension condition, giving a "self-improving" type of result. From the point of view of $\mathsf{CD}_{loc}(K, N)$ clearly the main issue is the globalization problem. Here the main statement is that the local-to-global property is true for nice optimal transportations and in the general case under the aforementioned regularity properties, is almost equivalent to the concavity of the 1-dimensional density λ_t. The latter it is in turn strongly linked to the composition property of the differential operator D.

The last comment is for the assumption of non branching property for (X, d, m). As shown by Rajala and Sturm in [19], strong $\mathsf{CD}(K, \infty)$-spaces and Riemannian $\mathsf{CD}(K, N)$ for $N \in \mathbb{R} \cup \{\infty\}$ have the property that for any couple of probability measures μ_0, μ_1 with $\mu_0, \mu_1 \ll m$ all the L^2-optimal transportations are concentrated on a set of non branching geodesics. That is all $\boldsymbol{\gamma} \in \mathcal{P}(\mathcal{G}(X))$, dynamical optimal plans with starting point μ_0 and ending point μ_1 are such that the evaluation map for each $t \in [0, 1)$

$$e_t : G \to X$$

is injective, even if the space is not assumed to be non branching, where G is the support of $\boldsymbol{\gamma}$.

Since our construction relies not only on the L^2-optimal dynamical plan but on the strong interplay between d^2-cyclically monotone sets and d-cyclically monotone sets, the substitution of the non branching property of the space with RCD-condition or with the strong $\mathsf{CD}(K, \infty)$ is a delicate task that would go beyond the scope of this note. For instance RCD-condition will not prevent the following "bad" situation: $\gamma, \hat{\gamma} \in G_a$ so that they have a common point $z = \gamma_s = \hat{\gamma}_t$ for $t \neq s$. In particular the proof of Lemma 4.2, that is one the building block of our analysis, does not work only assuming non branching support of $\boldsymbol{\gamma}$.

2. PRELIMINARIES

Let (X, d) be a metric space. The length $L(\gamma)$ of a continuous curve $\gamma : [0, 1] \to X$ is defined as

$$L(\gamma) := \sup \sum_{k=1}^{n} d(\gamma(t_{k-1}), \gamma(t_k))$$

where the supremum runs over $n \in \mathbb{N}$ and over all partitions $0 = t_0 < t_1 < \cdots < t_n = 1$. Note that $L(\gamma) \geq d(\gamma(0), \gamma(1))$. A curve is called *geodesic* if and only if $L(\gamma) = d(\gamma(0), \gamma(1))$. If this is the case, we can assume γ to have constant speed, i.e. $L(\gamma_{\llcorner[s,t]}) = |s - t|L(\gamma) = |s - t|d(\gamma(0), \gamma(1))$ for every $0 \leq s \leq t \leq 1$.

Denote by $\mathcal{G}(X)$ the space of geodesic $\gamma : [0, 1] \to X$ in X, regarded as subset of $C([0, 1], M)$ of continuous functions equipped with the topology of uniform convergence.

(X, d) is said to be a *length space* if and only if for every $x, y \in X$,

$$d(x, y) = \inf L(\gamma)$$

where the infimum runs over all continuous curves joining x and y. It is said to be a *geodesic space* if all x and y are connected by a geodesic. A point z will be called t-intermediate point of points x and y if $d(x,z) = td(x,y)$ and $d(z,y) = (1-t)d(x,y)$.

Definition 2.1. A geodesic space (X,d) is *non-branching* if and only if for every $r \geq 0$ and $x, y \in X$ such that $d(x,y) = r/2$, the set

$$\{z \in X : d(x,z) = r\} \cap \{z \in X : d(y,z) = r/2\}$$

consists of a single point.

Throughout the following we will denote by $B_r(z)$ the open ball of radius r centered in z. A standard map in optimal transportation is the evaluation map: for a fixed $t \in [0,1]$, $e_t : \mathcal{G}(X) \to X$ is defined by $e_t(\gamma) := \gamma_t$. The push-forward of a given measure, say η, via a map f will be denoted by $f_\sharp \eta$ and is defined by $f_\sharp \eta(A) := \eta(f^{-1}(A))$, for any measurable A.

2.1. **Geometry of metric measure spaces.** What follows is contained [21].

A *metric measure space* is a triple (X,d,m) where (X,d) is a complete separable metric space and m is a locally finite measure (i.e. $m(B_r(x)) < \infty$ for all $x \in X$ and all sufficiently small $r > 0$) on X equipped with its Borel σ-algebra. We exclude the case $m(X) = 0$. A *non-branching* metric measure space will be a metric measure space (X,d,m) such that (X,d) is a non-branching geodesic space.

$\mathcal{P}_2(X,d)$ denotes the L^2-Wasserstein space of Borel probability measures on X and W_2 the corresponding L^2-Wasserstein distance. The subspace of m-absolutely continuous measures is denoted by $\mathcal{P}_2(X,d,m)$.

The following are well-known results in optimal transportation theory and are valid for general metric measure spaces.

Lemma 2.2. *Let (X,d,m) be a metric measure space. For each geodesic $\mu : [0,1] \to \mathcal{P}_2(X,d)$ there exists a probability measure $\boldsymbol{\gamma}$ on $\mathcal{G}(X)$ such that*

- *$e_{t\sharp}\boldsymbol{\gamma} = \mu_t$ for all $t \in [0,1]$;*
- *for each pair (s,t) the transference plan $(e_s, e_t)_\sharp \boldsymbol{\gamma}$ is an optimal coupling for W_2.*

Consider the Rényi entropy functional

$$\mathcal{S}_N(\cdot|m) : \mathcal{P}_2(X,d) \to \mathbb{R}$$

with respect to m, defined by

$$(2.1) \qquad \mathcal{S}_N(\mu|m) := -\int_X \varrho^{-1/N}(x)\mu(dx)$$

for $\mu \in \mathcal{P}_2(X)$, where ϱ is the density of the absolutely continuous part μ^c in the Lebesgue decomposition $\mu = \mu^c + \mu^s = \varrho m + \mu^s$.

Given two numbers $K, N \in \mathbb{R}$ with $N \geq 1$, we put for $(t,\theta) \in [0,1] \times \mathbb{R}_+$,

$$(2.2) \qquad \tau_{K,N}^{(t)}(\theta) := \begin{cases} \infty, & \text{if } K\theta^2 \geq (N-1)\pi^2, \\ t^{1/N}\left(\dfrac{\sin(t\theta\sqrt{K/(N-1)})}{\sin(\theta\sqrt{K/(N-1)})}\right)^{1-1/N} & \text{if } 0 < K\theta^2 \leq (N-1)\pi^2, \\ t & \text{if } K\theta^2 < 0 \text{ or} \\ & \text{if } K\theta^2 = 0 \text{ and } N = 1, \\ t^{1/N}\left(\dfrac{\sinh(t\theta\sqrt{-K/(N-1)})}{\sinh(\theta\sqrt{-K/(N-1)})}\right)^{1-1/N} & \text{if } K\theta^2 \leq 0 \text{ and } N > 1. \end{cases}$$

That is, $\tau_{K,N}^{(t)}(\theta) := t^{1/N}\sigma_{K,N-1}^{(t)}(\theta)^{(N-1)/N}$ where

$$\sigma_{K,N}^{(t)}(\theta) = \frac{\sin(t\theta\sqrt{K/N})}{\sin(\theta\sqrt{K/N})},$$

if $0 < K\theta^2 < N\pi^2$ and with appropriate interpretation otherwise. Moreover we put

$$\varsigma_{K,N}^{(t)}(\theta) := \tau_{K,N}^{(t)}(\theta)^N.$$

The coefficients $\tau_{K,N}^{(t)}(\theta), \sigma_{K,N}^{(t)}(\theta)$ and $\varsigma_{K,N}^{(t)}(\theta)$ are the volume distortion coefficients with K playing the role of curvature and N the one of dimension.

The curvature-dimension condition $\mathsf{CD}(K, N)$ is defined in terms of convexity properties of the entropy functional. In the following definitions K and N will be real numbers with $N \geq 1$.

Definition 2.3 (Curvature-Dimension condition)**.** We say that (X, d, m) satisfies $\mathsf{CD}(K, N)$ if and only if for each pair $\mu_0, \mu_1 \in \mathcal{P}_2(X, d, m)$ there exists an optimal coupling π of $\mu_0 = \varrho_0 m$ and $\mu_1 = \varrho_1 m$, and a geodesic $\mu : [0, 1] \to \mathcal{P}_2(X, d, m)$ connecting μ_0 and μ_1 such that

$$
(2.3) \quad \begin{aligned}
\mathcal{S}_{N'}(\mu_t | m) \leq -\int_{X \times X} \Big[& \tau_{K,N'}^{(1-t)}(d(x_0, x_1)) \varrho_0^{-1/N'}(x_0) \\
& + \tau_{K,N'}^{(t)}(d(x_0, x_1)) \varrho_1^{-1/N'}(x_1) \Big] \pi(dx_0 dx_1),
\end{aligned}
$$

for all $t \in [0, 1]$ and all $N' \geq N$.

The following is a variant of $\mathsf{CD}(K, N)$ and it has been introduced in [7].

Definition 2.4 (Reduced Curvature-Dimension condition)**.** We say that (X, d, m) satisfies $\mathsf{CD}^*(K, N)$ if and only if for each pair $\mu_0, \mu_1 \in \mathcal{P}_2(X, d, m)$ there exists an optimal coupling π of $\mu_0 = \varrho_0 m$ and $\mu_1 = \varrho_1 m$, and a geodesic $\mu : [0, 1] \to \mathcal{P}_2(X, d, m)$ connecting μ_0 and μ_1 such that (2.3) holds true for all $t \in [0, 1]$ and all $N' \geq N$ with the coefficients $\tau_{K,N}^{(t)}(d(x_0, x_1))$ and $\tau_{K,N}^{(1-t)}(d(x_0, x_1))$ replaced by $\sigma_{K,N}^{(t)}(d(x_0, x_1))$ and $\sigma_{K,N}^{(1-t)}(d(x_0, x_1))$, respectively.

For both definitions there is a local version. Here we state only the local counterpart of $\mathsf{CD}(K, N)$, being clear what would be the one for $\mathsf{CD}^*(K, N)$.

Definition 2.5 (Local Curvature-Dimension condition)**.** We say that (X, d, m) satisfies $\mathsf{CD}_{loc}(K, N)$ if and only if each point $x \in X$ has a neighborhood $X(x)$ such that for each pair $\mu_0, \mu_1 \in \mathcal{P}_2(X, d, m)$ supported in $X(x)$ there exists an optimal coupling π of $\mu_0 = \varrho_0 m$ and $\mu_1 = \varrho_1 m$, and a geodesic $\mu : [0, 1] \to \mathcal{P}_2(X, d, m)$ connecting μ_0 and μ_1 such that (2.3) holds true for all $t \in [0, 1]$ and all $N' \geq N$.

It is worth noticing that in the previous definition the geodesic μ can exit from the neighborhood $X(x)$.

One of the main property of the reduced curvature dimension condition is the globalization one: under the non-branching assumption conditions $\mathsf{CD}_{loc}^*(K, N)$ and $\mathsf{CD}^*(K, N)$ are equivalent. Moreover it holds:

- $\mathsf{CD}_{loc}^*(K, N)$ is equivalent to $\mathsf{CD}_{loc}(K, N)$;
- $\mathsf{CD}(K, N)$ implies $\mathsf{CD}^*(K, N)$;
- $\mathsf{CD}^*(K, N)$ implies $\mathsf{CD}(K^*, N)$ where $K^* = K(N-1)/N$.

Hence it is possible to pass from CD_{loc} to CD at the price of passing through CD^* and therefore worsening the lower bound on the curvature. For all of these properties, see [7].

If a non-branching (X, d, m) satisfies $\mathsf{CD}(K, N)$ then geodesics are unique $m \otimes m$-a.e..

Lemma 2.6. *Assume that (X, d, m) is non-branching and satisfies $\mathsf{CD}(K, N)$ for some pair (K, N). Then for every $x \in \mathrm{supp}[m]$ and m-a.e. $y \in X$ (with the exceptional set depending on x) there exists a unique geodesic between x and y.*

Moreover there exists a measurable map $\gamma : X^2 \to \mathcal{G}(X)$ such that for $m \otimes m$-a.e. $(x, y) \in X^2$ the curve $t \mapsto \gamma_t(x, y)$ is the unique geodesic connecting x and y.

Under non-branching assumption is possible to formulate $\mathsf{CD}(K, N)$ in an equivalent point-wise version: (X, d, m) satisfies $\mathsf{CD}(K, N)$ if and only if for each pair $\mu_0, \mu_1 \in \mathcal{P}_2(X, d, m)$ and each dynamical optimal plan γ,

$$
(2.4) \quad \varrho_t(\gamma_t(x_0, x_1)) \leq \Big[\tau_{K,N'}^{(1-t)}(d(x_0, x_1)) \varrho_0^{-1/N'}(x_0) + \tau_{K,N'}^{(t)}(d(x_0, x_1)) \varrho_1^{-1/N'}(x_1) \Big]^{-N},
$$

for all $t \in [0, 1]$, and $(e_0, e_1)_\sharp \gamma$-a.e. $(x_0, x_1) \in X \times X$. Here ϱ_t is the density of the geodesic $(e_t)_\sharp \gamma$. Recall that $\gamma \in \mathcal{P}(\mathcal{G}(X))$ is a dynamical optimal plan if $\pi = (e_0, e_1)_\sharp \gamma \in \Pi(\mu_0, \mu_1)$ is optimal and the map $t \mapsto \mu_t := e_{t\sharp} \gamma$ is a geodesic in the 2-Wasserstein space.

We conclude with a partial list of properties enjoyed by metric measure spaces satisfying $\mathsf{CD}^*(K, N)$ (or $\mathsf{CD}_{loc}(K, N)$). If (X, d, m) verifies $\mathsf{CD}^*(K, N)$ then:

- m is a doubling measure;
- m verifies Bishop-Gromov volume growth inequality;
- m verifies Brunn-Minkowski inequality;

with all of these properties stated in a quantitative form.

2.2. **Spherical Hausdorff measure of codimension 1 and Coarea formula.** What follows is contained in [3] and is valid under milder assumption than $\mathsf{CD}^*(K,N)$ (or $\mathsf{CD}_{loc}(K,N)$) but for an easier exposition we assume (X,d,m) to satisfy $\mathsf{CD}^*(K,N)$.

Recall that for $K \geq 0$ the measure m is doubling that is $m(B_{2r}(x)) \leq (C_D/2)m(B_r(x))$ where C_D is the doubling constant of m. If $K < 0$ the measure m is locally uniformly doubling, i.e. $m(B_{2r}(x)) \leq (C_R/2)m(B_r(x))$ for any $r \leq R$ and some constant C_R depending on R but not on x.

If $B(X)$ is the set of balls, define the function $h : B(X) \to [0,\infty]$ as

$$h(\bar{B}_r(x)) := \frac{m(\bar{B}_r(x))}{r}.$$

Due to the (locally uniformly) doubling properties of m, the function h turns out to be a (locally uniformly) doubling function. Then, using the Carathéodory construction, we may define the generalized Hausdorff spherical measure $\mathcal{S}^h$ as

$$(2.5) \qquad \mathcal{S}^h(A) := \lim_{r\downarrow 0} \inf \left\{ \sum_{i\in\mathbb{N}} h(B_i) : B_i \in B(X), A \subset \bigcup_{i\in\mathbb{N}} B_i, \operatorname{diam}(B_i) \leq r \right\}.$$

The space of functions of bounded variation $BV(X)$ and the perimeter measure have been studied in [2], [3], [6], [17]. If $u \in BV(X)$, its total variation measure will be denoted with $|Du|$. We will use the following Coarea formula.

Theorem 2.7 ([6], Theorem 4.3, Theorem 4.4)**.** *For every $u \in BV(X)$ and every Borel set $A \subset X$ it holds*

$$|Du|(A) = \int_{-\infty}^{\infty} P(\{u > t\}, A)dt.$$

Moreover for any set $E \subset X$ of finite perimeter, the measure $P(E,\cdot)$ is concentrated on a subset of the essential boundary $\partial^* E$ and for any Borel set $B \subset X$

$$\frac{1}{c}\mathcal{S}^h(B \cap \partial^* E) \leq P(E,B) \leq c\,\mathcal{S}^h(B \cap \partial^* E)$$

with $c > 0$ depending only on K and N.

If u is a Lipschitz function, its total variation is equivalent as measure to $\|\nabla u\|m$, where

$$(2.6) \qquad \|\nabla u\|(x) := \liminf_{r\to 0} \frac{1}{r} \sup_{y \in \bar{B}_r(x)} |u(y) - u(x)|.$$

The following comparison is taken from [17]: for any Borel set $A \subset X$

$$(2.7) \qquad c_0 \int_A \|\nabla u\|(x)m(dx) \leq |Du|(A) \leq \int_A \|\nabla u\|(x)m(dx),$$

for some constant $c_0 > 0$ depending again only on K, N. The last result we would like to recall is a particular form of Coarea formula for Lipschitz functions.

Proposition 2.8 ([3], Proposition 5.1)**.** *For any u Lipschitz function defined on X and any B Borel set we have*

$$\int_{\mathbb{R}} \mathcal{S}^h(B \cap u^{-1}(t))dt \leq Lip(u)m(B).$$

2.3. **Gradients and differentials.** This part is taken from [15]. A curve $\gamma \in C([0,1], X)$ is said to be absolutely continuous provided there exists $f \in L^1([0,1])$ such that

$$d(\gamma_s, \gamma_t) \leq \int_s^t f(\tau) d\tau, \qquad \forall s, t \in [0,1], s \leq t.$$

Let $AC([0,1], X)$ denote the set of absolutely continuous curves. If $\gamma \in AC([0,1], X)$ then the limit

$$\lim_{\tau \to 0} \frac{d(\gamma_{t+\tau}, \gamma_t)}{\tau}$$

exists for a.e. $t \in [0,1]$, is called metric derivative and denoted by $|\dot{\gamma}_t|$.

Given Borel functions $f : X \to \mathbb{R}, G : X \to [0,\infty]$ we say that G is an upper gradient of f provided

$$|f(\gamma_0) - f(\gamma_1)| \leq \int_0^1 G(\gamma_t)|\dot{\gamma}_t| dt, \quad \forall \gamma \in AC([0,1], M),$$

where $|\dot{\gamma}_t|$ is the metric derivative of γ in t. For $f : X \to \mathbb{R}$ the local Lipschitz constant $|Df| : X \to [0,\infty]$ is defined by

$$|Df|(x) := \limsup_{y \to x} \frac{|f(y) - f(x)|}{d(y,x)}$$

if x is not isolated, and 0 otherwise. Define

$$|D^+ f|(x) := \limsup_{y \to x} \frac{(f(y) - f(x))^+}{d(y,x)}, \quad |D^- f|(x) := \limsup_{y \to x} \frac{(f(y) - f(x))^-}{d(y,x)},$$

the ascending and descending slope respectively. If f is locally Lipschitz, then $|D^{\pm} f|, |Df|$ are all upper gradients of f. In order to give a weaker notion of slope, consider the following family: $\boldsymbol{\gamma} \in \mathcal{P}(C([0,1], X))$ is a *test plan* if

$$e_{t\,\sharp}\boldsymbol{\gamma} \leq Cm, \quad \forall t \in [0,1], \qquad \text{and} \qquad \int \int_0^1 |\dot{\gamma}_t| dt \boldsymbol{\gamma}(d\gamma) < \infty,$$

where C is a positive constant. Therefore we have the following.

Definition 2.9. A Borel map $f : X \to \mathbb{R}$ belongs to the Sobolev class $S^2(X, d, m)$ (resp. $S^2_{loc}(X, d, m)$) if there exists a non-negative function $G \in L^2(X, m)$ (resp. $L^2_{loc}(X, m)$) such that

$$(2.8) \qquad \int |f(\gamma_0) - f(\gamma_1)| \boldsymbol{\gamma}(d\gamma) \leq \int \int_0^1 G(\gamma_s)|\dot{\gamma}_s| ds \boldsymbol{\gamma}(d\gamma), \qquad \forall \boldsymbol{\gamma} \text{ test plan.}$$

If this is the case, G is called *weak upper gradient*.

For $f \in S^2(X, d, m)$ there exists a minimal function G, in the m-a.e. sense, in $L^2(X, m)$ such that (2.8) holds. Denote such minimal function with $|Df|_w$. Accordingly define the semi-norm $\|f\|_{S^2(X,d,m)} := \||Df|_w\|_{L^2(X,m)}$.

We now state a result on the weak upper gradient of Kantorovich potentials also known as metric Brenier's Theorem.

Proposition 2.10 ([5], Theorem 10.3). *Let (X, d, m) verify $\mathsf{CD}(K, N)$ for $K \in \mathbb{R}$ and $N \geq 1$ and be non-branching. Let $\mu_0, \mu_1 \in \mathcal{P}_2(X, d, m)$, φ be a Kantorovich potential. Then for every $\boldsymbol{\gamma}$ optimal dynamical transference plan it holds*

$$d(\gamma_0, \gamma_1) = |D\varphi|_w(\gamma_0) = |D^+\varphi|(\gamma_0), \qquad for \ \boldsymbol{\gamma} - a.e. \gamma.$$

If moreover the densities of μ_0 and of μ_1 are both in $L^\infty(X, m)$, then

$$\lim_{t \downarrow 0} \frac{\varphi(\gamma_0) - \varphi(\gamma_t)}{d(\gamma_0, \gamma_t)} = d(\gamma_0, \gamma_1), \qquad in \ L^2(\mathcal{G}(X), \boldsymbol{\gamma}).$$

In order to compute higher order derivatives, we introduce the following.

Definition 2.11. Let $f, g \in S^2(X, d, m)$. The functions

$$D^+ f(\nabla g) := \liminf_{\varepsilon \downarrow 0} \frac{|D(g + \varepsilon f)|_w^2 - |Dg|_w^2}{2\varepsilon},$$

$$D^- f(\nabla g) := \limsup_{\varepsilon \uparrow 0} \frac{|D(g + \varepsilon f)|_w^2 - |Dg|_w^2}{2\varepsilon}.$$

are well defined.

Spaces where the two differentials coincide are called *infinitesimally strictly convex*, i.e. (X, d, m) is said to be infinitesimally strictly convex provided

$$(2.9) \qquad \int D^+ f(\nabla g)(x) m(dx) = \int D^- f(\nabla g)(x) m(dx), \qquad \forall f, g \in S^2(X, d, m).$$

It is proven in [15] that (2.9) is equivalent to the point-wise one:

$$D^+ f(\nabla g) = D^- f(\nabla g), \quad m - a.e., \quad \forall f, g \in S^2_{loc}(X, d, m).$$

If the space is infinitesimally strictly convex, we can denote by $Df(\nabla g)$ the common value and $Df(\nabla g)$ is linear in f and 1-homogeneous and continuous in g.

There is a strong link between differentials and derivation along families of curves. For $\boldsymbol{\gamma} \in \mathcal{P}(C([0, 1], X))$, define the norm $\|\boldsymbol{\gamma}\|_2 \in [0, \infty]$ of $\boldsymbol{\gamma}$ by

$$\|\boldsymbol{\gamma}\|_2^2 := \limsup_{t \downarrow 0} \frac{1}{t} \int \int_0^t |\dot\gamma_s|^2 ds \boldsymbol{\gamma}(d\gamma),$$

if $\gamma \in \mathcal{P}(AC([0, 1], X))$ and $+\infty$ otherwise.

Definition 2.12. Let $g \in S^2(X, d, m)$. We say that $\boldsymbol{\gamma} \in \mathcal{P}(C([0, 1], X))$ represents ∇g if $\boldsymbol{\gamma}$ is of bounded compression, $\|\boldsymbol{\gamma}\|_2 < \infty$, and it holds

$$(2.10) \qquad \liminf_{t \downarrow 0} \int \frac{g(\gamma_t) - g(\gamma_0)}{t} \boldsymbol{\gamma}(d\gamma) \geq \frac{1}{2}\big(\||Dg|_w\|_{L^2(X, e_0 \sharp \boldsymbol{\gamma})}^2 + \|\boldsymbol{\gamma}\|_2^2\big).$$

A straightforward consequence of (2.10) is that if $\boldsymbol{\gamma}$ represents ∇g, then the whole limit in the lefthand-side of (2.10) exists and verifies

$$\lim_{t \downarrow 0} \int \frac{g(\gamma_t) - g(\gamma_0)}{t} \boldsymbol{\gamma}(d\gamma) = \frac{1}{2}\big(\||Dg|_w\|_{L^2(X, e_0 \sharp \boldsymbol{\gamma})}^2 + \|\boldsymbol{\gamma}\|_2^2\big).$$

Theorem 2.13 ([15], Theorem 3.10)**.** *Let $f, g \in S^2(X, d, m)$. For every $\boldsymbol{\gamma} \in \mathcal{P}(C([0, 1], M))$ representing ∇g it holds*

$$\int D^+ f(\nabla g) e_0 \sharp \boldsymbol{\gamma} \geq \limsup_{t \downarrow 0} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \boldsymbol{\gamma}(d\gamma)$$

$$\geq \liminf_{t \downarrow 0} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \boldsymbol{\gamma}(d\gamma) \geq \int D^- f(\nabla g) e_0 \sharp \boldsymbol{\gamma}.$$

2.4. **Hopf-Lax formula for Kantorovich potentials.** What follows is contained in [5].

The definitions below make sense for a general Borel and real valued cost but we will only consider the $d^2/2$ case, for this reason c has to be interpret as $d^2/2$.

Definition 2.14. Let $\varphi : X \to \mathbb{R} \cup \{\pm\infty\}$. Its d^2-*transform* $\varphi^c : X \to \mathbb{R} \cup \{-\infty\}$ is defined by

$$\varphi^c(y) := \inf_{x \in X} \frac{1}{2} d^2(x, y) - \varphi(x).$$

Accordingly $\varphi : X \mapsto \mathbb{R} \cup \{\pm\infty\}$ is d^2-*concave* if there exists $v : X \to \mathbb{R} \cup \{-\infty\}$ such that $\varphi = v^c$.

A d^2-*concave* function φ such that (φ, φ^c) is a maximizing pair for the dual Kantorovich problem between μ_0, μ_1 is called a d^2-*concave Kantorovich potential* for the couple (μ_0, μ_1). A function φ is called a d^2-*convex Kantorovich potential* if $-\varphi$ is a d^2-concave Kantorovich potential.

We are interested in the evolution of potentials. They evolve accordingly to the Hopf-Lax evolution semigroup H_t^s via the following formula:

$$(2.11) \qquad H_t^s(\psi)(x) := \begin{cases} \inf_{y \in X} \dfrac{1}{2} \dfrac{d^2(x, y)}{s - t} + \psi(y), & \text{if } t < s, \\ \psi(x), & \text{if } t = s, \\ \sup_{y \in X} \psi(y) - \dfrac{1}{2} \dfrac{d^2(x, y)}{t - s}, & \text{if } t > s. \end{cases}$$

We also introduce the rescaled cost $c^{t,s}$ defined by

$$c^{t,s}(x,y) := \frac{1}{2}\frac{d^2(x,y)}{s-t}, \qquad \forall t < s, \ x,y \in X.$$

Observe that for $t < r < s$

$$c^{t,r}(x,y) + c^{r,s}(y,z) \geq c^{t,s}(x,z), \qquad \forall x,y,z \in X,$$

and equality holds if and only if there is a constant speed geodesic $\gamma : [t,s] \to X$ such that $x = \gamma_t$, $y = \gamma_r$ and $z = \gamma_s$. The following result is taken from [22] (Theorem 7.30 and Theorem 7.36) but here we report a different version.

Theorem 2.15 ([4], Theorem 2.18). *Let $(\mu_t) \subset \mathcal{P}_2(X)$ be a constant speed geodesic in $(\mathcal{P}_2(X,d), d_W)$ and ψ a $c^{0,1}$-convex Kantorovich potential for the couple (μ_0, μ_1). Then $\psi_s := H_0^s(\psi)$ is a $c^{t,s}$-concave Kantorovich potential for (μ_s, μ_t), for any $t < s$.*

Similarly, if ϕ is a c-concave Kantorovich potential for (μ_1, μ_0), then H_1^t is a $c^{t,s}$-convex Kantorovich potential for (μ_t, μ_s), for any $t < s$.

The following is an easy consequence.

Corollary 2.16. *Let φ be a d^2-concave Kantorovich potential for (μ_0, μ_1). Let $\varphi_t := -H_1^t(\varphi^c)$ be a $c^{t,1}$-concave Kantorovich potential for (μ_t, μ_1) and analogously let $\varphi_t^c := H_0^t(-\varphi)$ a $c^{0,t}$-concave Kantorovich potential for (μ_t, μ_0). Then:*

$$\varphi_t(\gamma_t) = \varphi(\gamma_0) - \frac{t}{2}d^2(\gamma_0, \gamma_1), \qquad \varphi_t^c(\gamma_t) = \varphi^c(\gamma_1) - \frac{1-t}{2}d^2(\gamma_0, \gamma_1), \qquad \boldsymbol{\gamma} - a.e.\ \gamma.$$

Proof. Since the proofs of the statements for φ_t and for φ_t^c are the same, we prefer to present only the one for φ_t.

Since

$$\varphi_t(x) = -H_1^t(\varphi^c)(x) = \inf_{y \in X} \frac{1}{2}\frac{d^2(x,y)}{1-t} - \varphi^c(y).$$

for $\boldsymbol{\gamma}$-a.e. γ

$$\varphi_t(\gamma_t) \leq \frac{1}{2}\frac{d^2(\gamma_t, \gamma_1)}{1-t} + \varphi(\gamma_0) - \frac{1}{2}d^2(\gamma_0, \gamma_1) = \varphi(\gamma_0) - \frac{t}{2}d^2(\gamma_0, \gamma_1).$$

To prove the opposite inequality: observe that

$$\frac{d^2(\gamma_0, \gamma_t)}{t} + \frac{d^2(\gamma_t, y)}{1-t} \geq d^2(\gamma_0, y),$$

therefore for $\boldsymbol{\gamma}$-a.e. γ

$$\frac{1}{2}\frac{d^2(\gamma_t, y)}{1-t} - \varphi^c(y) \geq \frac{1}{2}\frac{d^2(\gamma_t, y)}{1-t} - \frac{1}{2}d^2(\gamma_0, y) + \varphi(\gamma_0) \geq \varphi(\gamma_0) - \frac{1}{2}\frac{d^2(\gamma_0, \gamma_t)}{t} = \varphi(\gamma_0) - \frac{t}{2}d^2(\gamma_0, \gamma_1).$$

Taking the infimum the claim follows. $\square$

2.5. **Disintegration of measures.** We conclude this introductory part with a short review on disintegration theory. What follows is taken from [8].

Given a measurable space $(R, \mathscr{R})$ and a function $r : R \to S$, with S generic set, we can endow S with the *push forward σ-algebra* $\mathscr{S}$ of $\mathscr{R}$:

$$Q \in \mathscr{S} \iff r^{-1}(Q) \in \mathscr{R},$$

which could also be defined as the biggest σ-algebra on S such that r is measurable. Moreover given a measure space $(R, \mathscr{R}, \rho)$, the *push forward measure* η is then defined as $\eta := (r_\sharp \rho)$.

Consider a probability space $(R, \mathscr{R}, \rho)$ and its push forward measure space $(S, \mathscr{S}, \eta)$ induced by a map r. From the above definition the map r is measurable.

Definition 2.17. A *disintegration* of ρ consistent with r is a map $\rho : \mathscr{R} \times S \to [0,1]$ such that

    (1) $\rho_s(\cdot)$ is a probability measure on $(R, \mathscr{R})$ for all $s \in S$,
    (2) $\rho_\cdot(B)$ is η-measurable for all $B \in \mathscr{R}$,

and satisfies for all $B \in \mathscr{R}, C \in \mathscr{S}$ the consistency condition

$$\rho \left(B \cap r^{-1}(C) \right) = \int_C \rho_s(B) \eta(ds).$$

A disintegration is *strongly consistent with respect to* r if for all s we have $\rho_s(r^{-1}(s)) = 1$.

The measures ρ_s are called *conditional probabilities*.

We say that a σ-algebra $\mathcal{H}$ is *essentially countably generated* with respect to a measure m if there exists a countably generated σ-algebra $\hat{\mathcal{H}}$ such that for all $A \in \mathcal{H}$ there exists $\hat{A} \in \hat{\mathcal{H}}$ such that $m(A \triangle \hat{A}) = 0$.

We recall the following version of the disintegration theorem that can be found on [13], Section 452 (see [8] for a direct proof).

Theorem 2.18 (Disintegration of measures)**.** *Assume that $(R, \mathscr{R}, \rho)$ is a countably generated probability space, $R = \cup_{s \in S} R_s$ a partition of R, $r : R \to S$ the quotient map and $(S, \mathscr{S}, \eta)$ the quotient measure space. Then $\mathscr{S}$ is essentially countably generated w.r.t. η and there exists a unique disintegration $s \mapsto \rho_s$ in the following sense: if ρ_1, ρ_2 are two consistent disintegration then $\rho_{1,s}(\cdot) = \rho_{2,s}(\cdot)$ for η-a.e. s.*

If $\{S_n\}_{n \in \mathbb{N}}$ is a family essentially generating $\mathscr{S}$ define the equivalence relation:

$$s \sim s' \iff \{s \in S_n \iff s' \in S_n, \ \forall n \in \mathbb{N}\}.$$

Denoting with p the quotient map associated to the above equivalence relation and with $(L, \mathscr{L}, \lambda)$ the quotient measure space, the following properties hold:

- *$R_l := \cup_{s \in p^{-1}(l)} R_s = (p \circ r)^{-1}(l)$ is ρ-measurable and $R = \cup_{l \in L} R_l$;*
- *the disintegration $\rho = \int_L \rho_l \lambda(dl)$ satisfies $\rho_l(R_l) = 1$, for λ-a.e. l. In particular there exists a strongly consistent disintegration w.r.t. $p \circ r$;*
- *the disintegration $\rho = \int_S \rho_s \eta(ds)$ satisfies $\rho_s = \rho_{p(s)}$ for η-a.e. s.*

In particular we will use the following corollary.

Corollary 2.19. *If $(S, \mathscr{S}) = (X, \mathcal{B}(X))$ with X Polish space, then the disintegration is strongly consistent.*

3. SETTING

We fix here the objects, notations and hypothesis that will be used throughout this note.

(X, d, m) will be a non-branching metric measure space verifying $\mathsf{CD}_{loc}(K, N)$ or equivalently $\mathsf{CD}^*(K, N)$. The marginal measure $\mu_0, \mu_1 \in \mathcal{P}_2(X, d, m)$ are fixed together with $\pi \in \Pi(\mu_0, \mu_1)$ the optimal coupling and $\gamma \in \mathcal{P}(\mathcal{G}(X))$ the associated optimal dynamical transference plan such that

$$[0,1] \ni t \mapsto (e_t)_\sharp \gamma = \mu_t, \qquad (e_0, e_1)_\sharp \gamma = \pi,$$

with μ_t geodesic in the L^2-Wasserstein space and e_t is the evaluation map at time t: for any geodesic $\gamma \in \mathcal{G}(X)$, $e_t(\gamma) = \gamma_t$. $\mathcal{P}(\mathcal{G}(X))$ denotes the space of probability measures over $\mathcal{G}(X)$, the space of geodesic in X endowed with the uniform topology inherited as a subset of $C([0,1], X)$. The support of γ will be denoted with G. The evaluation map e without subscript is defined on $[0,1] \times G$ by $e(s, \gamma) = \gamma_s$.

Moreover

$$\mu_t = \varrho_t m, \qquad \forall \in t \in [0,1].$$

Thanks to recent results on existence and uniqueness of optimal maps, see [16], only one geodesic in G has a given couple of points as initial and final points, that is for $\gamma \in G$

$$(e_0, e_1)^{-1}\{(\gamma_0, \gamma_1)\} = \{\gamma\}.$$

Moreover by inner regularity of compact sets we can assume without loss of generality that G is compact,

$$\varrho_t \leq M, \qquad \forall \in t \in [0,1],$$

and metric Brenier's Theorem holds for all $\gamma \in G$, that is

(3.1) $$d(\gamma_0, \gamma_1) = |D\varphi|_w(\gamma_0).$$

A d^2-concave Kantorovich potential for (μ_0, μ_1) is φ and φ_t will be the d^2-concave Kantorovich potential for (μ_t, μ_1) obtained through Theorem 2.15. When it will be needed, we will prefer the notation φ_0 to φ.

Thanks to compactness of G we can also assume φ to be Lipschitz. Its $d^2/2$-transform will be denoted by φ^c. From Corollary 2.16 it follows that $\varphi_1 = -\varphi^c$ μ_1-a.e. and

$$(3.2) \qquad \varphi_t(\gamma_t) = (1-t)\varphi_0(\gamma_0) + t\varphi_1(\gamma_1).$$

We will also use the following notation

$$\varphi_t(\mu_t) = \varphi_t(\text{supp}[\mu_t]), \qquad \forall t \in [0,1].$$

Since we will make an extensive use of the following sets, we fix their names once for all:

$$(3.3) \qquad \Gamma := \left\{ (x,y) \in X \times X : \varphi(x) + \varphi^c(y) = \frac{d^2(x,y)}{2} \right\},$$

contains the support of π and the transportation set for (μ_t, μ_1) is

$$(3.4) \qquad \Gamma_t : \left\{ (x,y) \in X \times X : \varphi_t(x) + \varphi^c(y) = \frac{d^2(x,y)}{2(1-t)} \right\}.$$

and again $(e_t, e_1)_\sharp \gamma(\Gamma_t) = 1$. Fix also the set of curves with starting point in $\varphi^{-1}(a)$:

$$(3.5) \qquad G_a := \left\{ \gamma \in G : \varphi(\gamma_0) = a \right\}.$$

and the corresponding subset of Γ

$$(3.6) \qquad \Gamma_a = \left\{ (x,y) \in X \times X : \varphi(x) + \varphi^c(y) = \frac{d^2(x,y)}{2}, \varphi(x) = a \right\} = \Gamma \cap \left(\varphi^{-1}(a) \times X \right).$$

In Section 5.2 and Section 6.2 to disintegrate the reference measure m in the direction of evolution, for $r \in [0,1]$ we will use the "closed" and "open" evolution sets:

$$(3.7) \qquad \bar{\Gamma}_a(r) := e\left([0,r] \times G_a\right), \qquad \Gamma_a(r) := e\left([0,r) \times G_a\right).$$

As it will be proved in Proposition 4.1, the set Γ_a is d-cyclically monotone. We will denote with ϕ_a a Kantorovich potential associated to it, that is ϕ_a is 1-Lipschitz function such that

$$\Gamma_a \subset \{(x,y) \in X \times X : \phi_a(x) - \phi_a(y) = d(x,y)\}.$$

The d-monotone set associated to ϕ_a will be used again so we will denote it with K_a:

$$(3.8) \qquad K_a := \{(x,y) \in X \times X : \phi_a(x) - \phi_a(y) = d(x,y)\}.$$

A relevant function for the analysis is the length map at time t: for $t \in [0,1]$ the map $L_t : e_t(G) \to (0,\infty)$ is defined by

$$L_t(x) := L(e_t^{-1}(x)).$$

Again by inner regularity of compact sets, we can assume that there exists a positive constant C such that

$$\frac{1}{C} < L(\gamma) < C, \qquad \forall \gamma \in G.$$

In order to study the behavior of the evolution after time t of the level sets of φ, i.e. $\{\gamma_t : \gamma \in G, \varphi(\gamma_0) = a\}$ for $a \in \mathbb{R}$, is convenient to see them as level set of a particular function. As it will be proven during this note this particular function is defined by

$$(3.9) \qquad e_t(G) \ni \gamma_t \mapsto \Phi_t(\gamma_t) := \varphi_t(\gamma_t) + \frac{t}{2} L_t^2(\gamma_t),$$

where in the definition of Φ_t we used that, for $t \in (0,1]$, for every $x \in e_t(G)$ there exists only one geodesic $\gamma \in G$ with $\gamma_t = x$. This property for $t = 1$ holds only if $\mu_1 \ll m$. Another possible definition is $\Phi_t(\gamma_t) := \varphi(\gamma_0)$, see (3.2).

The map Φ_t enjoys the next monotonicity property.

Lemma 3.1. *Let $\gamma \in G$ be fixed. Then for every $s > 0$ it holds that*

$$\Phi_t(\gamma_{t-s}) > \Phi_t(\gamma_t) > \Phi_t(\gamma_{t+s}),$$

provided $\gamma_{t-s} \in e_t(G)$ for the first inequality and $\gamma_{t+s} \in e_t(G)$ for the second one.

Proof. We first prove the first inequality. Suppose by contradiction the existence of $s > 0$ such that $\gamma_{t-s} \in e_t(G)$ and $\Phi_t(\gamma_{t-s}) \leq \Phi_t(\gamma_t)$. From Proposition 4.1, necessarily $\Phi_t(\gamma_{t-s}) < \Phi_t(\gamma_t)$. So let $\hat{\gamma} := e_t^{-1}(\gamma_{t-s})$, then the previous inequality reads as

$$\varphi(\hat{\gamma}_0) < \varphi(\gamma_0).$$

So we can deduce

$$\frac{1}{2t}d^2(\hat{\gamma}_0, \gamma_{t-s}) = \varphi(\hat{\gamma}_0) + \varphi_t^c(\gamma_{t-s}) < \varphi(\gamma_0) + \varphi_t^c(\gamma_{t-s}) \leq \frac{1}{2t}d^2(\gamma_0, \gamma_{t-s}),$$

and therefore $d(\hat{\gamma}_0, \gamma_{t-s}) < d(\gamma_0, \gamma_{t-s})$. Hence

$$
\begin{aligned}
d^2(\gamma_0, \gamma_{t-s}) + d^2(\hat{\gamma}_0, \gamma_t) &\leq d^2(\gamma_0, \gamma_{t-s}) + \big(d(\hat{\gamma}_0, \gamma_{t-s}) + d(\gamma_{t-s}, \gamma_t)\big)^2 \\
&= d^2(\gamma_0, \gamma_{t-s}) + d^2(\hat{\gamma}_0, \gamma_{t-s}) + d^2(\gamma_{t-s}, \gamma_t) + 2d(\hat{\gamma}_0, \gamma_{t-s})d(\gamma_{t-s}, \gamma_t) \\
&< d^2(\gamma_0, \gamma_{t-s}) + d^2(\hat{\gamma}_0, \gamma_{t-s}) + d^2(\gamma_{t-s}, \gamma_t) + 2d(\gamma_0, \gamma_{t-s})d(\gamma_{t-s}, \gamma_t) \\
&= d^2(\gamma_0, \gamma_t) + d^2(\hat{\gamma}_0, \gamma_{t-s}),
\end{aligned}
$$

and since $\hat{\gamma}_t = \gamma_{t-s}$ and $\hat{\gamma} \in G$, this is in contradiction with d^2-cyclical monotonicity. The proof of the other inequality follows in the same way. $\square$

Another important set for our analysis is the following one: for $\gamma \in G$ and $t \in (0,1)$

(3.10) $$I_t(\gamma) := \{\tau \in (0,1) : \gamma_\tau \in e_t(G)\},$$

that is the set of τ for which γ_τ belongs to $e_t(G)$. A priori one can only say that t belongs to $I_t(\gamma)$ but actually the set $I_t(\gamma)$ has sufficiently many points in a neighborhood of t. The following Lemma proves a density result and it has been obtained in collaboration with Martin Huesmann in [10].

Lemma 3.2. *For $\mathcal{L}^1$-a.e. $t \in [0,1]$,*

$$\lim_{\varepsilon \to 0} \frac{1}{2\varepsilon}\mathcal{L}^1(I_t(\gamma) \cap (t-\varepsilon, t+\varepsilon)) = 1, \qquad in\ L^1(G, \gamma).$$

That is, the point $\tau = t$ is a point of Lebesgue density (in L^1 sense) 1 for the set $I_t(\gamma) := \{\tau \in [0,1] : \gamma_\tau \in e_t(G)\}$.

4. ON THE METRIC STRUCTURE OF OPTIMAL TRANSPORTATION

Only for this Section the setting will be more general than the one specified in Section 3. Here we drop all the assumption on the curvature of the space. So (X, d, m) is a geodesic, non branching and separable metric measure space, μ_t is geodesic in the L^2-Wasserstein space together with a family of Kantorovich potential φ_t for $t \in [0,1]$ associated to it. We will use the notation of Section 3 for everything and related to this objects.

Fix $a \in \varphi(\mu_0)$. We will prove that Γ_a is d-cyclically monotone. Recall that

$$\Gamma_a := \Gamma \cap \varphi^{-1}(a) \times X,$$

with Γ transport set for (μ_0, μ_1) as from (3.3).

Proposition 4.1. *The set Γ_a is d-cyclically monotone.*

Proof. Let $(x_i, y_i) \in \Gamma_a$ for $i = 1, \ldots, n$ and observe that

$$\frac{1}{2}d^2(x_i, y_i) = \varphi(x_i) + \varphi^c(y_i) = \varphi(x_{i-1}) + \varphi^c(y_i) \leq \frac{1}{2}d^2(x_{i-1}, y_i).$$

Hence $d(x_i, y_i) \leq d(x_{i-1}, y_i)$ and therefore

$$\sum_{i=1}^n d(x_i, y_i) \leq \sum_{i=1}^n d(x_i, y_{i+1})$$

and the claim follows. $\square$

The main consequence of Proposition 4.1 is that two distinct geodesic of G, starting from the same level set of φ, can meet only for $t = 0$ or $t = 1$, provided the metric Brenier's Theorem holds (in the sense of (3.1)). Recall the definition of

$$\bar{\Gamma}_a(1) = e([0,1] \times G_a).$$

already introduced in (3.7) and G_a the set of geodesics starting from the level set a of f, see (3.5).

Lemma 4.2. *If the metric Brenier's Theorem holds, then the family $\{e_t(G_a)\}_{t \in [0,1]}$ is a partition of $\bar{\Gamma}_a(1)$.*

Proof. By construction the family covers $\bar{\Gamma}_a(1)$, so we have only to show that overlapping doesn't occur. Assume by contradiction the existence of $\hat{\gamma}, \tilde{\gamma} \in G_a$, $\hat{\gamma} \neq \tilde{\gamma}$ such that $\hat{\gamma}_s = \tilde{\gamma}_t = z$ with, say, $s < t$.

Then d-cyclical monotonicity implies that $\hat{\gamma}$ and $\tilde{\gamma}$ form a cycle of zero cost and then non-branching property of (X, d, m) implies that they are contained in a longer geodesic: if $\hat{\gamma}_0 = x_0, \hat{\gamma}_1 = y_0$ and $\tilde{\gamma}_0 = x_1, \tilde{\gamma}_1 = y_1$ then

$$d(x_0, y_1) + d(x_1, y_0) = d(x_0, y_0) + d(x_1, y_1).$$

There are two possible cases: or $d(x_1, y_0) \leq d(x_0, y_0)$ or $d(x_0, y_1) \leq d(x_1, y_1)$, indeed if both were false we would have a contradiction with the previous identity. In the first case

$$\frac{1}{2}d^2(x_0, y_0) = \varphi(x_0) + \varphi^c(y_0) = \varphi(x_1) + \varphi^c(y_0) \leq \frac{1}{2}d^2(x_1, y_0) \leq \frac{1}{2}d^2(x_0, y_0).$$

Therefore $d(x_0, y_0) = d(x_1, y_0)$ and since they lie on the same geodesic $x_0 = x_1$. In the second case

$$\frac{1}{2}d^2(x_1, y_1) = \varphi(x_1) + \varphi^c(y_1) = \varphi(x_0) + \varphi^c(y_1) \leq \frac{1}{2}d^2(x_0, y_1) \leq \frac{1}{2}d^2(x_1, y_1),$$

and the same conclusion holds true: $x_0 = x_1$.

Hence we have $(x_0, y_0), (x_0, y_1) \in \Gamma_a$. It follows from metric Brenier's Theorem (Proposition 2.10) that for all $\gamma \in G$

$$|D\varphi|_w(x) = d(\gamma_0, \gamma_1).$$

Therefore necessarily $y_0 = y_1$. Since $\hat{\gamma}, \tilde{\gamma}$ have also an inner common point, they must coincide implying a contradiction. $\qquad\square$

The next is a simple consequence of Lemma 4.2.

Corollary 4.3. *For each $a \in \mathbb{R}$, the map $e : [0,1] \times G_a \to X$ defined by*

$$e(s, \gamma) := \gamma_s$$

is a measurable isomorphism.

The following is, to our knowledge, a new result and it proves that for $t \in (0,1)$ the Kantorovich potentials φ_t, obtained with the Hopf-Lax formula from any Kantorovich potential φ_0, verifies a property similar to the point wise metric Brenier's Theorem.

Proposition 4.4. *For every $t \in (0,1)$ and for every $\gamma \in G$*

(4.1) $$\lim_{s \to 0} \frac{\varphi_t(\gamma_t) - \varphi_t(\gamma_{t+s})}{d(\gamma_t, \gamma_{t+s})} = d(\gamma_0, \gamma_1) = |D\varphi_t|(\gamma_t),$$

where $|D\varphi_t|$ denotes the local Lipschitz constant of φ_t.

Proof. Step 1. Fix $\gamma \in G$. Observe that the set

$$\text{argmin}\left\{ y \mapsto \frac{d^2(\gamma_t, y)}{2(1-t)} - \varphi^c(y) \right\},$$

is single valued and contains only γ_1. Indeed suppose by contradiction the contrary. Then there exists $z \in X$ different from γ_1 so that

$$\varphi_t(\gamma_t) + \varphi^c(z) = \frac{d^2(\gamma_t, z)}{2(1-t)},$$

then since $\varphi_t = -\varphi_t^c$ we have

$$\frac{1}{2}d^2(\gamma_0, z) \geq \varphi(\gamma_0) + \varphi^c(z)$$

$$= \varphi(\gamma_0) - \varphi_t(\gamma_t) + \varphi_t(\gamma_t) + \varphi^c(z)$$

$$= \frac{1}{2}\left(\frac{d^2(\gamma_0, \gamma_t)}{t} + \frac{d^2(\gamma_t, z)}{1-t}\right)$$

$$\geq \frac{1}{2}d^2(\gamma_0, z),$$

then necessarily $d(\gamma_0, z) = d(\gamma_0, \gamma_t) + d(\gamma_t, z)$. But then non-branching property of (X, d, m) implies a contradiction and then $z = \gamma_1$.

Step 2. Then by Hopf-Lax formula for Hamilton-Jacobi equations on length spaces

$$\limsup_{y \to \gamma_t} \frac{|\varphi_t(y) - \varphi_t(\gamma_t)|}{d(y, \gamma_t)} = \frac{D^+(\gamma_t, 1-t)}{1-t}$$

see Proposition 3.6 in [5]. Hence from *Step 1.* it follows that

$$\limsup_{y \to \gamma_t} \frac{|\varphi_t(y) - \varphi_t(\gamma_t)|}{d(y, \gamma_t)} = d(\gamma_0, \gamma_1).$$

To conclude the proof observe that

$$\varphi_t(\gamma_t) - \varphi_t(\gamma_{t+s}) = \varphi_t(\gamma_t) + \varphi^c(\gamma_1) - \varphi^c(\gamma_1) - \varphi_t(\gamma_{t+s})$$

$$\geq \frac{1}{2(1-t)}\left(d^2(\gamma_t, \gamma_1) - d^2(\gamma_{t+s}, \gamma_1)\right)$$

$$= \frac{1}{2(1-t)}\left(d(\gamma_t, \gamma_1) - d(\gamma_{t+s}, \gamma_1)\right)\left(d(\gamma_t, \gamma_1) + d(\gamma_{t+s}, \gamma_1)\right)$$

(4.2)
$$= \frac{1}{2(1-t)}d(\gamma_t, \gamma_{t+s})\left(d(\gamma_t, \gamma_1) + d(\gamma_{t+s}, \gamma_1)\right).$$

Hence

$$\liminf_{s \to 0} \frac{\varphi_t(\gamma_t) - \varphi_t(\gamma_{t+s})}{d(\gamma_t, \gamma_{t+s})} \geq \frac{d(\gamma_t, \gamma_1)}{1-t} = d(\gamma_0, \gamma_1)$$

and the claim follows. $\square$

As a consequence of Proposition 4.4, the construction presented so far is purely metric. Indeed instead of analyzing the geometric properties of the Wasserstein geodesic $[0,1] \ni t \mapsto \mu_t$ one could restrict the domain of μ_t to $[\varepsilon, 1-\varepsilon]$ for any $\varepsilon > 0$ and Lemma 4.2 is true without assuming any curvature bound on the space (X, d, m).

4.1. **From L^2-geodesics to L^1-geodesics.** Thanks to the properties proved so far we can construct a link between L^2 Wasserstein geodesics and the linear structure of d-cyclically monotone sets. Since the distance is finite, from d-monotonicity of Γ_a we deduce the existence a 1-Lipschitz function $\phi_a : X \to \mathbb{R}$ so that

$$\Gamma_a \subset K_a := \{(x, y) \in X \times X : \phi_a(x) - \phi_a(y) = d(x, y)\}.$$

Note that also the following inclusion holds

$$\{(\gamma_s, \gamma_t) : \gamma \in G_a, s \leq t\} \subset K_a.$$

Remark 4.5. Even if an explicit expression of ϕ_a is not strictly needed to our analysis, for the sake of completeness, a possible choice of ϕ_a is the following one:

$$e([0,1], G_a) = \bar{\Gamma}_a(1) \ni \gamma_s \mapsto \phi_a(\gamma_s) = a - d(\gamma_0, \gamma_s) = a - sL(\gamma).$$

Indeed all the geodesics in G_a follows at time 0 the direction of $\nabla\varphi$ and therefore they are somehow orthogonal to $\varphi^{-1}(a)$. The same geodesics of G_a follows also the steepest descent direction of ϕ_a and therefore they have the same direction of $\nabla\phi_a$. Hence one would expect that at time 0 the set $\varphi^{-1}(a)$ is a level set also for ϕ_a: therefore one could expect $\phi_a(\gamma_s) = a - d(\gamma_0, \gamma_s)$.

We now prove that this heuristic motivation make sense. If $0 \leq s \leq t \leq 1$ and $\gamma \in G_a$

$$\phi_a(\gamma_s) - \phi_a(\gamma_t) = (t - s)L(\gamma) = d(\gamma_s, \gamma_t),$$

and therefore

$$(\gamma_s, \gamma_t) : \gamma \in G_a, 0 \leq s \leq t \leq 1\} \subset K_a.$$

If $s, t \in [0, 1]$ and $\gamma, \hat{\gamma} \in G_a$

$$\phi_a(\gamma_s) - \phi_a(\hat{\gamma}_t) = d(\hat{\gamma}_0, \hat{\gamma}_t) - d(\gamma_0, \gamma_s),$$

and since

$$\frac{1}{2}d(\hat{\gamma}_0, \hat{\gamma}_t)^2 = \varphi(\hat{\gamma}_0) + \varphi_t(\hat{\gamma}_t) = \varphi(\gamma_0) + \varphi_t(\hat{\gamma}_t) \leq \frac{1}{2}d(\gamma_0, \hat{\gamma}_t)^2,$$

it follows that

$$\phi_a(\gamma_s) - \phi_a(\hat{\gamma}_t) \leq d(\gamma_0, \hat{\gamma}_t) - d(\gamma_0, \gamma_s) \leq d(\gamma_s, \hat{\gamma}_t).$$

Hence ϕ_a is 1-Lipschitz and therefore it is a good L^1-Kantorovich potential for the d-monotone set $\{(\gamma_s, \gamma_t) : \gamma \in G_a, s \leq t\}$. Note moreover that the calculations above proves that

$$\bar{\Gamma}_a(1) \ni \gamma_s \mapsto d(\gamma_0, \gamma_s)$$

is 1-Lipschitz and coincides with $\gamma_s \mapsto \min\{d(\hat{\gamma}_0, \gamma_s) : \hat{\gamma} \in G_a\}$.

The following holds.

Lemma 4.6. *Let $\Delta \subset K_a$ be any set so that:*

$$(x_0, y_0), (x_1, y_1) \in \Delta \quad \Rightarrow \quad (\phi_a(y_1) - \phi_a(y_0)) \cdot (\phi_a(x_1) - \phi_a(x_0)) \geq 0.$$

Then Δ is d^2-cyclically monotone.

Proof. It follows directly from the hypothesis of the Lemma that the set

$$\{(\phi_a(x), \phi_a(y)) : (x, y) \in \Delta\} \subset \mathbb{R} \times \mathbb{R}$$

is $|\cdot|^2$-cyclically monotone, where $|\cdot|$ denotes the modulus. Then for $\{(x_i, y_i)\}_{i \leq N} \subset \Delta$, since $\Delta \subset K_a$, it holds

$$\sum_{i=1}^{N} d^2(x_i, y_i) = \sum_{i=1}^{N} |\phi_a(x_i) - \phi_a(y_i)|^2$$

$$\leq \sum_{i=1}^{N} |\phi_a(x_i) - \phi_a(y_{i+1})|^2$$

$$\leq \sum_{i=1}^{N} d^2(x_i, y_{i+1}),$$

where the last inequality is given by the 1-Lipschitz regularity of ϕ_a. The claim follows. $\square$

Fix an interval $(a_0, b_0) \subset \mathbb{R}$ and for any γ so that $(a_0, b_0) \subset \phi_a(\gamma_{[0,1]})$ we can define $R_0^\gamma, L_0^\gamma \subset [0, 1]$ so that

$$\phi_a \circ \gamma\left((R_0^\gamma, R_0^\gamma + L_0^\gamma)\right) = (a_0, b_0)$$

that is equivalent to say

$$\phi_a \circ \gamma(R_0^\gamma) = b_0, \qquad \phi_a \circ \gamma(R_0^\gamma + L_0^\gamma) = a_0.$$

In the same manner, for another interval $(a_1, b_1) \subset \mathbb{R}$ we can associate to any γ so that $(a_1, b_1) \subset \phi_a(\gamma_{[0,1]})$ the corresponding time interval $(R_1^\gamma, R_1^\gamma + L_1^\gamma)$ Accordingly for all $t \in [0, 1]$ we define $R_t^\gamma := (1-t)R_0^\gamma + tR_1^\gamma$ and $L_t^\gamma := (1-t)L_0^\gamma + tL_1^\gamma$.

We use these coefficients to construct an L^2-Wasserstein geodesic.

Proposition 4.7. *Let $H \subset G_a$ be so that for all $\gamma \in H$ both $(a_0, b_0), (a_1, b_1) \subset \phi_a(\gamma_{(0,1)})$ with $b_0 > a_0 > b_1 > a_1$. Define the curve*

$$(4.3) \qquad [0, 1] \ni t \mapsto \nu_t := \int_H \frac{1}{L_t^\gamma}\mathcal{L}^1 \llcorner_{[R_t^\gamma, R_t^\gamma + L_t^\gamma]}\eta(d\gamma) \in \mathcal{P}([0, 1] \times G_a),$$

with η probability measure on $\mathcal{G}(X)$ so that $\eta(H) = 1$. Then $[0, 1] \ni t \mapsto (e)_\sharp\nu_t$ is a W_2-geodesic.

Proof. First note that for any fixed $s \in [0,1]$ the value $\phi_a(\gamma_{R_0^\gamma + sL_0^\gamma})$ do not depend on $\gamma \in H$. Indeed

$$
\begin{aligned}
\phi_a(\gamma_{R_0^\gamma + sL_0^\gamma}) &= \phi_a(\gamma_{R_0^\gamma}) - d(\gamma_{R_0^\gamma}, \gamma_{R_0^\gamma + sL_0^\gamma}) \\
&= \phi_a(\gamma_{R_0^\gamma}) - s d(\gamma_{R_0^\gamma}, \gamma_{R_0^\gamma + L_0^\gamma}) \\
&= \phi_a(\gamma_{R_0^\gamma}) - s \left(\phi_a(\gamma_{R_0^\gamma}) - \phi_a(\gamma_{R_0^\gamma + L_0^\gamma}) \right) \\
&= b_0 - s(b_0 - a_0),
\end{aligned}
$$

and the same applies for $\phi_a(\gamma_{R_1^\gamma + sL_1^\gamma})$. It follows that the set

$$
\left\{ (\gamma_{R_0^\gamma + sL_0^\gamma}, \gamma_{R_1^\gamma + sL_1^\gamma}) : \gamma \in H, s \in [0,1] \right\}
$$

is d^2-cyclically monotone. Indeed, using Lemma 4.6, we have only to show that for any $\hat{\gamma}, \gamma \in H$ and $\hat{s}, s \in [0,1]$:

$$
(\phi_a(\hat{\gamma}_{R_1^{\hat\gamma} + \hat s L_1^{\hat\gamma}}) - \phi_a(\gamma_{R_1^\gamma + sL_1^\gamma})) \cdot (\phi_a(\hat{\gamma}_{R_0^{\hat\gamma} + \hat s L_0^{\hat\gamma}}) - \phi_a(\gamma_{R_0^\gamma + sL_0^\gamma})) \geq 0.
$$

But as observed few lines above

$$
\phi_a(\gamma_{R_1^\gamma + sL_1^\gamma}) = \phi_a(\hat{\gamma}_{R_1^{\hat\gamma} + sL_1^{\hat\gamma}}), \qquad \phi_a(\gamma_{R_0^\gamma + sL_0^\gamma}) = \phi_a(\hat{\gamma}_{R_0^{\hat\gamma} + sL_0^{\hat\gamma}}).
$$

Hence the claim is equivalent to

$$
(\phi_a(\hat{\gamma}_{R_1^{\hat\gamma} + \hat s L_1^{\hat\gamma}}) - \phi_a(\hat{\gamma}_{R_1^{\hat\gamma} + sL_1^{\hat\gamma}})) \cdot (\phi_a(\hat{\gamma}_{R_0^{\hat\gamma} + \hat s L_0^{\hat\gamma}}) - \phi_a(\hat{\gamma}_{R_0^{\hat\gamma} + sL_0^{\hat\gamma}})) \geq 0,
$$

that in turn is equivalent to

$$
(s - \hat s)(b_1 - a_1) \cdot (s - \hat s)(b_0 - a_0) = (s - \hat s)^2 (b_1 - a_1)(b_0 - a_1) \geq 0,
$$

hence the claim follows. $\square$

Hence, an optimal transport is achieved by not changing the "angular" parts and coupling radial parts according to optimal coupling on $\mathbb{R}$. Since in the radial (or linear) part of the coupling is linear, one is allowed to rescale the radial speed and gain one degree of freedom.

In the next Sections we will use regularity properties of CD_{loc}-spaces to properly apply the constructions of this Section to improve the curvature estimates and to study the globalization problem.

5. Dimension reduction for a class of optimal transportations

In this Section we start our general analysis in the particular case of optimal transport plan with lengths of geodesics depending only on the level set of φ from where they start, that is

$$
L(\gamma) = f(\varphi(\gamma_0)), \qquad \forall \gamma \in G,
$$

and the level sets of φ maintain their order during the evolution, that is

$$
\varphi(\mu_0) \ni a \mapsto a - \frac{1}{2} f^2(a) \in \mathbb{R},
$$

is non decreasing. In what follows we will denote with $F(a)$ the function $a - f^2(a)/2$. Thanks to Luzin's Theorem, we can also assume the map $e_0(G) \ni \gamma_0 \mapsto f(\varphi(\gamma_0))$ to be continuous.

Under this particular assumption, the transportation enjoys nice properties. In the following Lemma we prove that level sets are moved by $\boldsymbol{\gamma}$ in a monotone way. Recall that

$$
e_t(G_a) = \{ \gamma_t : \gamma \in G, \varphi(\gamma_0) = a \}.
$$

Lemma 5.1. *Assume that $L(\gamma) = f(\varphi(\gamma_0))$, then it holds*

$$
e_t(G_a) = \varphi_t^{-1} \left(a - \frac{t}{2} f^2(a) \right) \cap e_t(G).
$$

for $a \in \varphi(\mu_0)$.

Proof. The first inclusion follows immediately from Corollary 2.16: if $\varphi(\gamma_0) = a$ then

$$\varphi_t(\gamma_t) = a - \frac{t}{2}L^2(\gamma) = a - \frac{t}{2}f^2(a).$$

To prove the other inclusion we observe that the evolution at time t of two different level sets of φ_0 cannot be contained in the same level set of φ_t. Indeed if $a > b \in \varphi(\mu_0)$ then for all $\gamma \in G_a$ and $\bar{\gamma} \in G_b$ it holds

$$\varphi_1(\gamma_1) = -\varphi^c(\gamma_1) = F(a) \geq F(b) = -\varphi^c(\bar{\gamma}_1) = \varphi_1(\bar{\gamma}_1).$$

Hence for all $t \in (0,1)$

$$\varphi_t(\gamma_t) = (1-t)\varphi(\gamma_0) + t\varphi_1(\gamma_1) > (1-t)\varphi(\bar{\gamma}_0) + t\varphi_1(\bar{\gamma}_1) = \varphi_t(\bar{\gamma}_t).$$

The claim follows. $\square$

5.1. **Level sets of Kantorovich potentials.** On the set $e_0(G)$ we will consider the partition given by the saturated sets of φ, i.e. $\{\varphi^{-1}(a)\}_{a \in \mathbb{R}}$. Disintegration Theorem implies that

$$m_{\llcorner e_0(G)} = \int_{\varphi(\mu_0)} \tilde{m}_a q(da), \qquad \varphi_\sharp\left(m_{\llcorner e_0(G)}\right) = q,$$

with $\tilde{m}_a(\varphi^{-1}(a)^c) = 0$ for q-a.e. $a \in \varphi(\mu_0)$, where, in order to have a shorter notation, we have denoted by $\varphi(\mu_0)$ the set $\varphi(\mathrm{supp}[\mu_0])$.

Proposition 5.2. *The measure $q = \varphi_\sharp\left(m_{\llcorner e_0(G)}\right)$ is absolute continuous w.r.t. $\mathcal{L}^1$. Moreover for q-a.e. $a \in \varphi(\mu_0)$ it holds*

$$\tilde{m}_a \ll \mathcal{S}^h,$$

where $\mathcal{S}^h$ is the spherical Hausdorff measure of codimension one.

Proof. Step 1. Recall that on G the point wise metric Brenier's Theorem holds true: $|D\varphi|_w(\gamma_0) = d(\gamma_0, \gamma_1)$, for all $\gamma \in G$. Define the map

$$e_0(G) \ni x \mapsto \hat{\varphi}(x) := \inf_{y \in e_1(G)} \left\{ \frac{d^2(x,y)}{2} - \varphi^c(y) \right\}.$$

Since G is compact, $e_0(G)$ and $e_1(G)$ are bounded and $\hat{\varphi}$ is obtained as the infimum of Lipschitz maps with uniformly bounded Lipschitz constant. Therefore $\hat{\varphi}$ is Lipschitz and coincide with $\varphi(x)$. Extend $\hat{\varphi}$ to the whole space keeping the same Lipschitz constant.

We can use the Coarea formula (see Section 2.2) in the particular case of Lipschitz maps: for any $B \subset X$ Borel

$$(5.1) \qquad \int_{-\infty}^{+\infty} P(\{\hat{\varphi} > a\}, B) da \geq c_0 \int_B \|\nabla \hat{\varphi}\|(x) m(dx),$$

where c_0 is a strictly positive constant.

Step 2. For $(x,y) \in (e_0, e_1)(G)$, $\|\nabla \hat{\varphi}\|(x) \geq d(x,y)$. Indeed fix $(x,y) \in (e_0, e_1)(G)$, then $\hat{\varphi}(x) + \varphi^c(y) = d^2(x,y)/2$ and

$$\hat{\varphi}(x) - \hat{\varphi}(z) \geq \frac{1}{2}(d^2(x,y) - d^2(z,y)) = \frac{1}{2}(d(x,y) - d(z,y))(d(x,y) + d(z,y))$$

Select a minimizing sequence $\rho_n \to 0$ for $\|\nabla\hat{\varphi}\|(x)$ and z_n on the geodesic connecting x to y at distance ρ_n from x. Then

$$\frac{1}{\rho_n} \sup_{z \in B_{\rho_n}(x)} |\hat{\varphi}(z) - \hat{\varphi}(x)| \geq \frac{1}{2}\frac{1}{\rho_n}(d(x,y) - d(z_n, y))(d(x,y) + d(z_n, y)) = \frac{1}{2}(d(x,y) + d(z_n, y)).$$

Passing to the limit we have $\|\nabla\hat{\varphi}\|(x) \geq d(x,y)$.

Let $E \subset \mathbb{R}$ with $\mathcal{L}^1(E) = 0$, then from (5.1) it follows that

$$(5.2) \quad \int_{\varphi^{-1}(E) \cap e_0(G)} \|\nabla\hat{\varphi}\|(x) m(dx) = \int_{\hat{\varphi}^{-1}(E) \cap e_0(G)} \|\nabla\hat{\varphi}\|(x) m(dx) \leq \frac{1}{c_0} \int_E P(\{\hat{\varphi} > a\}, e_0(G)) da = 0.$$

But on $e_0(G)$ the gradient of $\hat{\varphi}$ is strictly positive, it follows that $m(\varphi^{-1}(E) \cap e_0(G)) = 0$ and therefore the first part of the claim is proved. Moreover from (5.2) it follows that

$$\tilde{m}_a \leq P(\{\hat{\varphi} > a\}, \cdot).$$

Being the latter absolutely continuous with respect to $\mathcal{S}^h$, also the second part of the statement follows.
□

Remark 5.3. Proposition 5.2 proves a property of disintegration at time $t = 0$ where the particular shape of the optimal transportation or of the Kantorovich potentials do not play any role and indeed the proof is done without using any particular assumption. Hence the result will be used also in the general case.

So Proposition 5.2 implies the following decomposition for $t = 0$:

$$m_{\llcorner e_0(G)} = \int_{\varphi(\mu_0)} \tilde{m}_a q(a) \mathcal{L}^1(da) = \int_{\varphi(\mu_0)} \hat{m}_a \mathcal{L}^1(da),$$

with clearly again $\hat{m}_a \ll \mathcal{S}^h$.

For $t \in [0,1)$ an analogous partition can be considered also on the support of μ_t, $e_t(G)$. Indeed the d^2-cyclical monotonicity of Γ implies that the family

$$\{\gamma_t : \gamma \in G, \varphi(\gamma_0) = a\}_{a \in \varphi(\mu_0)} = \{e_t(G_a)\}_{a \in \varphi(\mu_0)}$$

is a disjoint family and a partition of $e_t(G)$. Therefore we consider the disintegration of $m_{\llcorner e_t(G)}$ w.r.t. the aforementioned family. Since for every $t \in [0,1)$

$$\mu_0(\varphi^{-1}(A)) = \mu_t(\{\gamma_t : \varphi(\gamma_0) \in A\}),$$

the quotient measures of μ_0 and μ_t are the same measure. We can conclude that the quotient measures of $m_{\llcorner e_t(G)}$ and of $m_{\llcorner e_0(G)}$ are equivalent and

(5.3) $$m_{\llcorner e_t(G)} = \int_{\varphi(\mu_0)} \tilde{m}_{a,t} f_t(a) \mathcal{L}^1(da) = \int_{\varphi(\mu_0)} \hat{m}_{a,t} \mathcal{L}^1(da), \qquad \hat{m}_{a,t}(\{\gamma_t : \varphi(\gamma_0) = a\}^c) = 0.$$

To keep notation consistent, we will denote also the conditional probabilities for $t = 0$ with $\hat{m}_{a,0}$.

For $t = 1$ only if μ_1 is absolute continuous with respect to m we can do the same disintegration. Indeed if this is the case, from Theorem 2.7 of [16], for m-a.e. $x \in e_1(G)$ there is only one geodesic γ in G so that $\gamma_1 = x$ and the family

$$\{\gamma_1 : \gamma \in G, \varphi(\gamma_0) = a\}_{a \in \varphi(\mu_0)}$$

is again partition of $e_1(G)$. Since we are assuming both μ_0 and μ_1 absolute continuous with respect to m, we have

$$m_{\llcorner e_t(G)} = \int_{\varphi(\mu_0)} \hat{m}_{a,t} \mathcal{L}^1(da), \qquad \hat{m}_{a,t}(e_t(G_a)) = \|\hat{m}_{a,t}\|,$$

for all $t \in [0,1]$.

Lemma 5.4. *For every $t \in [0,1]$ and $\mathcal{L}^1$-a.e. $a \in \varphi(\mu_0)$, with the exceptional set depending on t, it holds*

$$\hat{m}_{a,t} \ll \mathcal{S}^h,$$

where $\mathcal{S}^h$ is the spherical Hausdorff measure of codimension one defined in (2.5).

Proof. For $t = 0$ the claim has been already obtained in Proposition 5.2. For $t \in (0,1]$ we observe that from Lemma 5.1 the partition of $e_t(G)$

$$\{\gamma_t : \gamma \in G_a\}_{a \in \varphi(\mu_0)},$$

can be equivalently written as

$$\{\varphi_t^{-1}(a)\}_{a \in \varphi_t(\mu_t)}.$$

Then using coarea formula as in Proposition 5.2 the claim follows.   □

Since level sets of φ_0 are moved after time t to level sets of φ_t, the monotone map $F(a) = a - f^2(a)/2$ is the optimal map between the quotient measures.

Lemma 5.5. *For each $t \in [0,1]$, consider the map $F_t(a) := a - tf^2(a)/2$ defined on $\varphi(\mu_0)$. Then for each $t \in [0,1]$, F_t is the optimal map for between*

$$(\varphi_0)_{\sharp}\mu_0, \qquad (\varphi_t)_{\sharp}\mu_t$$

and f is locally Lipschitz.

Proof. Just note that

$$(5.4) \qquad F_t(\varphi(\gamma_0)) = \varphi(\gamma_0) - \frac{t}{2}L^2(\gamma) = \varphi_t(\gamma_t).$$

Since g_1 is monotone by assumption and, thanks to Proposition 5.2, $(\varphi_i)_\sharp\mu_i$ are absolute continuous w.r.t. to $\mathcal{L}^1$ for $i = 0, 1$, the claim follows. □

5.2. **Disintegration in the direction of motion.** As already motivated in the Introduction, $\hat{m}_{a,t}$ is not the right reference measure to improve the curvature estimate to a "codimension 1"-like estimate. So we consider the evolution in time of a single level set as a whole subset of X, that is the set $e([0,1] \times G_a)$, and we disintegrate the reference measure m with respect to the family $\{e_t(G_a)\}_{t \in [0,1]}$. In this way the quotient space of the disintegration will be the time variable and as t moves the conditional probabilities will move in the same direction of the optimal transportation.

Recall $\bar{\Gamma}_a(1) := e([0,1] \times G_a)$. Thanks to Lemma 4.2, we can consider the disintegration of $m\llcorner_{\bar{\Gamma}_a(1)}$ w.r.t. the family of sets $\{e_t(G_a)\}_{t \in [0,1]}$:

$$m\llcorner_{\bar{\Gamma}_a(1)} = \int_{[0,1]} \bar{m}_{a,t}q(dt), \qquad q \in \mathcal{P}([0,1]), \quad q(I) = m(e(I \times G_a)).$$

Observe that any $\gamma \in G_a$ can be taken as quotient set, therefore Corollary 2.19 implies the strong consistency of the disintegration, i.e. for q-a.e. $t \in [0,1]$ $\bar{m}_{a,t}$ is concentrated on $e_t(G_a)$.

Proposition 5.6. *The quotient measure q_a is absolute continuous with respect to $\mathcal{L}^1$.*

Proof. Since Γ_a is d-cyclically monotone, we can consider another partition of $\bar{\Gamma}_a(1)$. Consider the family of sets $\{\gamma_s : s \in [0,1]\}_{\gamma \in G_a}$. By Lemma 4.2, we have that

- the following disintegration holds true:

$$m\llcorner_{\bar{\Gamma}_a(1)} = \int \eta_y q_a(dy),$$

  where the quotient measure q_a is concentrated on $\{\gamma_{1/2} : \gamma \in G_a\}$ and q_a-a.e. conditional probability η_y is concentrated on $\{\gamma_s : s \in [0,1], \gamma \in e_{1/2}^{-1}(y) \cap G_a\}$;

- Since $\mathsf{CD}_{loc}(K,N)$ implies $\mathsf{MCP}(K,N)$, from Theorem 9.5 of [9] we have that $\eta_y = g(y,\cdot)\mathcal{L}^1\llcorner_{[0,1]}$ for q_a-a.e. y, and for $r \leq R$

$$(5.5) \qquad \left(\frac{\sin\left(\frac{r}{R}d(\gamma_0,\gamma_R)\sqrt{K/(N-1)}\right)}{\sin\left(d(\gamma_0,\gamma_R)\sqrt{K/(N-1)}\right)}\right)^{N-1} \leq \frac{g(y,r)}{g(y,R)} \leq \left(\frac{\sin\left(\frac{r}{R}d(\gamma_r,\gamma_1)\sqrt{K/(N-1)}\right)}{\sin\left(d(\gamma_r,\gamma_1)\sqrt{K/(N-1)}\right)}\right)^{N-1},$$

  where $\gamma = e_{1/2}^{-1}(y) \cap G_a$, and and the measure $g(y,\cdot)\mathcal{L}^1\llcorner_{[0,1]}$ has to be intended as $(\gamma)_\sharp(g(y,\cdot)\mathcal{L}^1\llcorner_{[0,1]})$, with γ the unique element of G_a so that $\gamma_{1/2} = y$.

To prove the claim it is enough to observe that the two disintegration proposed for $m\llcorner_{\bar{\Gamma}_a(1)}$ are the same. Use Fubini's Theorem to get

$$\int_{[0,1]} \bar{m}_{a,t}q(dt) = m\llcorner_{\bar{\Gamma}_a(1)} = \int g(y,\cdot)\mathcal{L}^1(dt)q_a(dy) = \int_{[0,1]} g(\cdot,t)q_a(dy)dt,$$

therefore from uniqueness of disintegration,

$$\bar{m}_{a,t} = g(\cdot,t)q_a\left(\int g(y,t)q_a(dy)\right)^{-1}, \qquad q = \left(\int g(y,t)q_a(dy)\right)\mathcal{L}^1,$$

and the claim follows. □

Hence if $dq/d\mathcal{L}^1$ denotes the density of q with respect to $\mathcal{L}^1$, posing $m_{a,t} := (dq/d\mathcal{L}^1)\bar{m}_{a,t}$, we have

$$(5.6) \qquad m\llcorner_{\bar{\Gamma}_a(1)} = \int_{[0,1]} m_{a,t}dt.$$

Note that Proposition 5.6, and therefore (5.6), has been obtained without using the assumption of constant speed of geodesics along the level set of φ. So we will use it also in the general case without any need of prove it again.

Proposition 5.7. *For $\mathcal{L}^2$-a.e. $(a,t) \in \varphi(\mu_0) \times [0,1]$ it holds*

$$m_{a,t} \ll \mathcal{S}^h,$$

where $\mathcal{S}^h$ is the spherical Hausdorff measure of codimension one defined in (2.5).

Proof. Following the proof of Lemma 5.4, the claim will be proved if we write the family of sets

$$\{\gamma_t : \gamma \in G_a\}_{t \in [0,1]} = \{e_t(G_a)\}_{t \in [0,1]},$$

as a family of level sets $\{\Lambda^{-1}(t)\}_{t \in [0,1]}$ for some locally Lipschitz $\Lambda : \bar{\Gamma}_a(1) \to \mathbb{R}$ with non zero gradient.

Step 1. Consider the evaluation map $e : [0,1] \times G_a \to \bar{\Gamma}_a(1)$ as $e(s,\gamma) := \gamma_s$ and define the following function

$$\bar{\Gamma}_a(1) \ni x \mapsto \Lambda(x) := P_1(e^{-1}(x)).$$

Hence $\Lambda(x)$ is the unique t for which there exists $\gamma \in G_a$ so that $\gamma_t = x$. From its definition, Λ is clearly measurable and

$$\Lambda^{-1}(t) = \{\gamma_t : \gamma \in G_a\}.$$

Its derivative in the direction of $s \mapsto \gamma_{t+s}$ is 1 for any $t \in (0,1)$ and $\gamma \in G_a$. We now show that Λ is locally Lipschitz. Note that for $s < t$ and any $\gamma, \hat{\gamma} \in G_a$

$$\Lambda(\gamma_t) - \Lambda(\hat{\gamma}_s) = t - s = \frac{1}{L(\hat{\gamma})} d(\hat{\gamma}_s, \hat{\gamma}_t).$$

On the other hand from Lemma 5.1 $\varphi_s(\gamma_s) = \varphi_s(\hat{\gamma}_s)$ and therefore

$$\frac{1}{(t-s)^2} d^2(\hat{\gamma}_s, \hat{\gamma}_t) = \varphi_s(\hat{\gamma}_s) + \varphi^c_{1-t}(\hat{\gamma}_t) = \varphi_s(\hat{\gamma}_s) + \varphi^c_{1-t}(\gamma_t) \leq \frac{1}{(t-s)^2} d^2(\hat{\gamma}_s, \gamma_t).$$

Hence

$$|\Lambda(\gamma_t) - \Lambda(\hat{\gamma}_s)| \leq \frac{1}{C} d(\gamma_t, \hat{\gamma}_s).$$

and therefore the claim is proved.   □

So the results obtained in this Section are: assuming that $L(\gamma) = f(\varphi(\gamma_0))$, and $a - \frac{1}{2}f^2(a)$ is non decreasing, we have

$$m \llcorner e_t(G) = \int_{\varphi(\mu_0)} \hat{m}_{a,t} \mathcal{L}^1(da), \qquad m \llcorner \bar{\Gamma}_a(1) = \int_{[0,1]} m_{a,t} \mathcal{L}^1(dt)$$

for all $t \in [0,1]$ and $a \in \varphi(\mu_0)$ and $\hat{m}_{a,t}, m_{a,t} \ll \mathcal{S}^h \llcorner e_t(G_a)$ for $\mathcal{L}^1$-a.e. $a, t \in [0,1]$.

6. DIMENSION REDUCTION FOR THE GENERAL TRANSPORTATION

In this Section we obtain the result of Section 5 dropping the assumption of constant length on level sets of φ but assuming few regularity properties for γ. In particular we will assume a regularity property of the length map that has been already introduce in Section 3: for $t \in (0,1)$

$$e_t(G) \ni x \mapsto L_t(x) := L(e_t^{-1}(x)) \in (0, \infty).$$

Assumption 1. For all $t \in (0,1)$ the map L_t is locally Lipschitz: for μ_t-a.e. $x \in e_t(G)$ there exists an open neighborhood $U(x)$ of x and a positive constant C so that

$$|L_t(z) - L_t(w)| \leq Cd(z,w), \qquad \forall z,w \in U(x).$$

We can now introduce the function Φ_t defined on $e_t(G)$:

$$\Phi_t(x) := \varphi_t(x) + \frac{t}{2} L_t^2(x).$$

As already pointed out in Section 3, the relevance of Φ_t is explained by the following equivalent identities:

$$\Phi_t(\gamma_t) = \varphi(\gamma_0), \qquad \Phi_t^{-1}(a) = \{\gamma_t : \gamma \in G, \varphi(\gamma_0) = a\}.$$

It follows from Assumption 1 that also Φ_t is locally Lipschitz. Moreover almost by definition

$$(\Phi_t)_\sharp m \llcorner e_t(G) \ll \mathcal{L}^1,$$

indeed since $\Phi_t \circ e_t = \varphi \circ e_0$ it follows that $(\Phi_t)_\sharp \mu_t = (\varphi)_\sharp \mu_0$ and therefore

$$(\Phi_t)_\sharp \varrho_t m \llcorner e_t(G) = (\varphi)_\sharp \varrho_0 m \llcorner e_0(G).$$

Since $\varrho_t > 0$ on $e_t(G)$, also the converse is true, that is

$$\mathcal{L}^1 \llcorner_{\varphi(\mu_0)} \ll (\Phi_t)_\sharp m \llcorner_{e_t(G)}.$$

Anyway this property is not sufficient to guarantee that its metric gradient do not vanish. See [1] for a counter example to this property (constructed on $\mathbb{R}^2$). One of the first steps we have to do is prove that the reference measures of codimension one are all absolute continuous with respect to the spherical Hausdorff measure $\mathcal{S}^h$, and, as in the proof of Proposition 5.2, we will use Coarea formula and we will apply it to the function Φ_t. Since Coarea formula brings information only where the gradient is non zero we have to ask for the following property to hold.

Assumption 2. For all $t \in [0,1]$ for $\boldsymbol{\gamma}$-a.e. $\gamma \in G$ the following holds

$$\lim_{s \to 0} \frac{\Phi_t(\gamma_t) - \Phi_t(\gamma_{t+s})}{d(\gamma_t, \gamma_{t+s})} \in (0, \infty).$$

Remark 6.1. Assumption 1 is verified in the hypothesis of Section 5 that is:

$$L(\gamma) = f(\varphi(\gamma_0)), \qquad \forall \gamma \in G,$$

with f so that $a \mapsto a - f^2(a)/2$ is non-decreasing. Indeed as proved in Lemma 5.5 f is locally Lipschitz. Moreover if $F_t : \varphi_0(\mu_0) \to \varphi_t(\mu_t)$ is the locally bi-Lipschitz function of (5.4), then

$$e_t(G) \ni \gamma_t \mapsto f \circ F_t^{-1} \circ \varphi_t(\gamma_t)$$

is locally Lipschitz and coincides with $L_t(\gamma_t)$. Noticing that $\Phi_t = F_t^{-1} \circ \varphi_t$ Assumption 2 is straightforward.

Before showing how Assumption 1 and 2 are used in the metric framework, we prove that if X is a Riemannian manifold with d geodesic distance induced by a Riemannian tensor g and m is the volume measure, then Assumption 1 and 2 are verified.

Proposition 6.2. *Assume (X, d, m) has a Riemannian structure, that is (X, d) is a Riemannian manifold with metric g and m is the volume measure. Then Assumption 1 and 2 are both verified.*

Proof. Assumption 1 follows from the Monge-Mather shortening principle, see [22] Theorem 8.5.

Actually Theorem 8.5 of [22] proves Lipschitz regularity on compact sets of the transport map from intermediate times: if

$$T_t : e_0(G) \to e_t(G), \qquad (T_t)_\sharp \mu_0 = \mu_t$$

then for any $t \in (0,1)$ the map T_t^{-1} is Lipschitz and in particular m-almost every where differentiable: Since $\Phi_t = \varphi \circ T_t^{-1}$, it follows that Assumption 2 is equivalent to prove that

$$g_{\gamma_0}(\nabla \varphi(\gamma_0), DT_t^{-1} \nabla \varphi_t(\gamma_t)) > 0.$$

In order to compute the previous quantity is convenient to consider the expression of DT_t^{-1} proved in [12], see Theorem 4.2:

$$DT_t^{-1} = Y(H - tHess_x \varphi_t^c),$$

where Y is the differential of the exponential map in $(\gamma_t, -t\nabla \varphi_t^c) \in X \times T_{\gamma_t} X$, H is the Hessian of the squared of distance function and φ_t^c has been introduced in Section 2.4 and minus its gradient composed with the exponential maps gives the optimal transport from μ_t to μ_0. Since

$$Y \nabla \varphi_t(\gamma_t) = \nabla \varphi(\gamma_0)$$

it follows from Gauss Lemma, see [14] Theorem 3.70, that

$$(6.1) \qquad g_{\gamma_0}(\nabla \varphi(\gamma_0), DT_t^{-1} \nabla \varphi_t(\gamma_t)) = g_{\gamma_t}(\nabla \varphi(\gamma_t), (H - tHess_x \varphi_t^c) \nabla \varphi_t(\gamma_t)).$$

Since $(H - tHess_x \varphi_t^c)$ is symmetric with strictly positive determinant μ_t-almost everywhere, the claim follows. $\qquad \square$

6.1. **Level sets of Kantorovich potentials.**

Proposition 6.3. *For every $t \in [0,1)$ and $\mathcal{L}^1$-a.e. $a \in \varphi(\mu_0)$, with the exceptional set depending on t, it holds*

$$\hat{m}_{a,t} \ll \mathcal{S}^h,$$

where $\mathcal{S}^h$ is the spherical Hausdorff measure of codimension one defined in (2.5).

Proof. Step 1. For $t = 0$ the claim has been already proved in Proposition 5.2, see Remark 5.3.

As a consequence of Assumption 1, Φ_t is locally Lipschitz on $e_t(G)$. Since we are proving a local property, possibly taking a compact subset of G, we can assume without loss of generality that Φ_t is Lipschitz on the whole $e_t(G)$. Denote with $\hat{\Phi}_t$ its Lipschitz extension to X. Coarea formula for Lipschitz maps applies (see Section 2.2 and references therein): for any measurable $A \subset X$

$$\int_{-\infty}^{+\infty} P(\{\hat{\Phi}_t > a\}, A)\mathcal{L}^1(da) \geq c_0 \int_A \|\nabla\hat{\Phi}_t\|(x)m(dx),$$

where c_0 is a strictly positive constant.

Step 2. Since $\hat{\Phi}_t$ is Lipschitz,

$$P(\{\hat{\Phi}_t > a\}, \cdot) \leq c\mathcal{S}^h.$$

where c is a positive constant depending on K and N. So we have

$$\int \|\nabla\hat{\Phi}_t\|\hat{m}_{a,t}\mathcal{L}^1(da) \leq \int_{-\infty}^{+\infty} P(\{\hat{\Phi}_t > a\}, \cdot)\mathcal{L}^1(da),$$

which in turn gives

$$\|\nabla\hat{\Phi}_t\|\hat{m}_{a,t} \leq c\mathcal{S}^h.$$

From Assumption 2 it follows that $\|\nabla\hat{\Phi}_t\| > 0$ on $e_t(G)$ and the claim follows. $\square$

6.2. **Disintegration in the direction of motion.** As in Section 5.2, we disintegrate $m \llcorner_{\bar{\Gamma}_a(1)}$ in with respect to the partition $\{e_t(G_a)\}_{t \in [0,1]}$. From Proposition 5.6 we have

$$m\llcorner_{\bar{\Gamma}_a(1)} = \int_{[0,1]} m_{a,t}\mathcal{L}^1(dt).$$

We now prove a regularity property for the conditional measures $m_{a,t}$. Recall that we are considering optimal transportation with uniformly positive and bounded lengths: there exists $C > 0$ so that

$$\frac{1}{C} < L(\gamma) < C, \qquad \forall \gamma \in G.$$

Lemma 6.4. *For $\mathcal{L}^2$-a.e. $(a,t) \in \varphi(\mu_0) \times [0,1]$ it holds*

$$m_{a,t} \ll \mathcal{S}^h,$$

where $\mathcal{S}^h$ is the spherical Hausdorff measure of codimension one defined in (2.5).

Proof. The idea of the proof is exactly the same as Proposition 5.7.

Step 1. Define the map

$$\bar{\Gamma}_a(1) \ni x \mapsto \Lambda(x) := P_1(e^{-1}(x)),$$

hence $\Lambda(x)$ is the unique t for which there exists $\gamma \in G_a$ so that $\gamma_t = x$. From its definition, Λ is clearly measurable, $\Lambda^{-1}(t) = \{\gamma_t : \gamma \in G_a\}$ and its derivative in the direction of $s \mapsto \gamma_{t+s}$ is 1 for any $t \in (0,1)$ and $\gamma \in G_a$.

Step 2. We now show that Λ is locally Lipschitz. Consider the map

$$G_a \times G_a \times [0,1]^2 \ni (\bar{\gamma}, \hat{\gamma}, s, t) \mapsto Y(\bar{\gamma}, \hat{\gamma}, s, t) = d(\bar{\gamma}_s, \hat{\gamma}_t) - \frac{1}{C}|t - s|.$$

Fix $\gamma \in G_a$ and note that for any $s, t \in [0,1]$:

$$Y(\gamma, \gamma, s, t) = |t - s|\left(L(\gamma) - \frac{1}{C}\right) > 0.$$

By continuity, there exists an open set U in $G_a \times G_a \times [0,1]^2$ so that $Y(U) \subset (0,\infty)$ and

$$K := \{(\gamma, \gamma, s, t) : s, t \in [0,1]\} \subset U.$$

Since K is compact, there exists $\varepsilon > 0$ so that $K^\varepsilon \subset U$ where K^ε is the ε-neighborhood of K in the metric space $G_a \times G_a \times [0,1]^2$. Consider $\bar{\gamma}, \hat{\gamma} \in G_a$ so that

$$d_\infty(\bar{\gamma}, \gamma) < \frac{\varepsilon}{2}, \qquad d_\infty(\hat{\gamma}, \gamma) < \frac{\varepsilon}{2},$$

where d_∞ the metric on $\mathcal{G}(X)$. Then $(\hat{\gamma}, \bar{\gamma}, s, t) \in K^\varepsilon$ for any $s, t \in [0,1]$. Therefore

$$d(\hat{\gamma}_s, \bar{\gamma}_t) > \frac{1}{C}|s - t|, \qquad \forall\, s, t \in [0,1].$$

Hence we have shown that for any $\gamma \in G_a$ there exists $\varepsilon > 0$ so that the map

$$e\left([0,1] \times B_\varepsilon(\gamma)\right) \ni x \mapsto \Lambda(x)$$

is Lipschitz indeed for $x, y \in e\left([0,1] \times B_\varepsilon(\gamma)\right)$ with say $x = \bar{\gamma}_s$ and $y = \hat{\gamma}_t$ it holds

$$|\Lambda(x) - \Lambda(y)| = |s - t| \le Cd(\bar{\gamma}_s, \hat{\gamma}_t).$$

Step 3. Repeating the proof of Proposition 6.3 with coarea formula we obtain that

$$m_{a,t} \llcorner e([0,1] \times B_\varepsilon(\gamma)) \ll \mathcal{S}^h,$$

for $\mathcal{L}^1$-a.e. $t \in [0,1]$. Since G_a is compact the claim on the whole $\bar{\Gamma}_a(1)$ follows by a covering argument. $\square$

7. UNIQUENESS OF CONDITIONAL MEASURES

This Section is devoted to find a relation and possibly a comparison between $m_{a,t}$ and $\hat{m}_{a,t}$. Find a comparison between this two different reference measure of codimension one is fundamental. Indeed disintegrate $\boldsymbol{\gamma}$ w.r.t. $\{e_0^{-1}(\varphi^{-1}(a))\}_{a \in \mathbb{R}}$ that is the set of geodesic starting from a given level set of φ:

$$\boldsymbol{\gamma} = \int_{\varphi(\mu_0)} \boldsymbol{\gamma}_a q(a) \mathcal{L}^1(da), \qquad \boldsymbol{\gamma}_a\left((\varphi \circ e_0)^{-1}(a)\right) = 1.$$

Clearly this disintegration is just the lift for each t of the disintegration of μ_t w.r.t. $\{e_t(G_a)\}_{a \in \varphi(\mu_0)}$. Therefore the quotient measure $q(a)\mathcal{L}^1(da)$ is the same quotient measure of μ_t for every $t \in [0,1]$. Then necessarily,

$$\int_{\varphi(\mu_0)} \varrho_t \hat{m}_{a,t} \mathcal{L}^1(da) = \mu_t = (e_t)_\sharp \boldsymbol{\gamma} = \int_{\varphi(\mu_0)} (e_t)_\sharp \boldsymbol{\gamma}_a q(a) \mathcal{L}^1(da),$$

and from uniqueness of disintegration,

$$(e_t)_\sharp \boldsymbol{\gamma}_a = \left(\int \varrho_t(z) \hat{m}_{a,t}(dz)\right)^{-1} \varrho_t \hat{m}_{a,t}.$$

Hence if we want to express the geodesic of codimension one $(e_t)_\sharp \boldsymbol{\gamma}_a$ in terms of the reference measure $m_{a,t}$ moving in the same direction of the optimal transportation, we have to prove that $(e_t)_\sharp \boldsymbol{\gamma}_a \ll m_{a,t}$. To do that we will prove that $\hat{m}_{a,t} \ll m_{a,t}$.

Remark 7.1. Here we want to stress the differences between $\hat{m}_{a,t}$ and $m_{a,t}$. It is worth underlining again that both measures are concentrated on $e_t(G_a)$. Also they are both obtained as conditional measures of m or, otherwise stated, they belong to the range of two different disintegration maps of m:

$$m \llcorner e_t(G) = \int_{\varphi(\mu_0)} \hat{m}_{a,t} \mathcal{L}^1(da), \qquad m \llcorner \bar{\Gamma}_a(1) = \int_{[0,1]} m_{a,t} \mathcal{L}^1(dt).$$

Since in both disintegrations the quotient measure is $\mathcal{L}^1$, conditional measures can be interpret as the "derivative" with respect to the parameter in the quotient space, a in the first case and t in the second one, of m. Even if m and $e_t(G_a)$ are fixed, what do matters, and implies $\hat{m}_{a,t} \ne m_{a,t}$, is the difference between $e_{t+\varepsilon}(G_a)$ and $e_t(G_{a+\varepsilon})$. The difference can be observed in Figure 1.

Lemma 7.2. *For every* $a \in \varphi(\mu_0)$,

$$\lim_{s \to 0} \frac{1}{s} \int_{(t, t+s)} m_{a,\tau} \mathcal{L}^1(d\tau) = m_{a,t},$$

for $\mathcal{L}^1$-*a.e.* $t \in [0,1]$, *where the convergence is in the weak sense.*
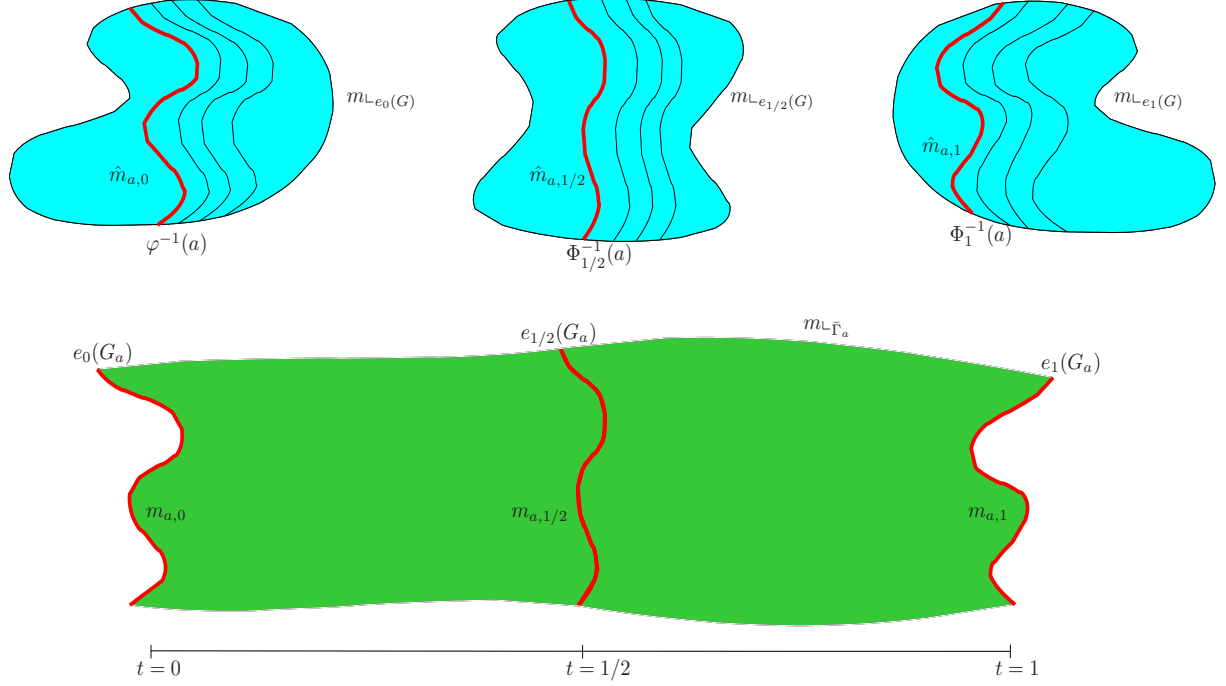
FIGURE 1. Above and below the disintegration with conditional $\hat{m}_{a,t}$ and $m_{a,t}$, respectively.

Proof. Since (X, d, m) is locally compact, the space of real valued continuous and bounded functions $C_b(X)$ is separable. Let $\{f_k\}_{k \in \mathbb{N}} \subset C_b(X)$ be a dense family.

Fix $a \in \varphi(\mu_0)$. The Lebesgue differentiation theorem implies that for every $k \in \mathbb{N}$

$$\frac{1}{s} \int_{(t,t+s)} \left(\int f_k(z) m_{a,\tau}(dz) \right) \mathcal{L}^1(d\tau) \to \int f_k(z) m_{a,t}(dz), \quad \text{as } s \searrow 0,$$

as real numbers, for all $t \in [0, 1] \setminus E_{a,k}$ with $\mathcal{L}^1(E_{a,k}) = 0$. Hence $E_a := \cup_{m \in \mathbb{N}} E_{a,k}$ is $\mathcal{L}^1$-negligible. Take $f \in C_b(X)$ and chose $\{f_{k_h}\}_{h \in \mathbb{N}}$ approximating f in the uniform norm. Using f_{k_h}, it is then fairly easy to show that

$$\lim_{s \to 0} \frac{1}{s} \int_{(t,t+s)} \left(\int f(z) m_{a,\tau}(dz) \right) \mathcal{L}^1(d\tau) = \int f(z) m_{a,t}(dz)$$

for all $t \in [0, 1] \setminus E_a$. □

The analogous statement of Lemma 7.2 is true for the conditional measures $\hat{m}_{a,t}$ of (5.3): fix $t \in [0, 1]$, then

$$\lim_{b \to 0} \frac{1}{b} \int_{(a,a+b)} \hat{m}_{\alpha,t} \mathcal{L}^1(d\alpha) = \hat{m}_{a,t},$$

for $\mathcal{L}^1$-a.e. $a \in \varphi(\mu_0)$, where the convergence is in the weak sense.

7.1. **Comparison between conditional measures.** The next one is the main technical statement of the Section.

Proposition 7.3. *For $\mathcal{L}^1$-a.e. $a \in \varphi(\mu_0)$ and every sequence $\varepsilon_n \to 0^+$ there exists a subsequence ε_{n_k} so that:*

$$\lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \cdot m_{\llcorner \Phi_t^{-1}([a-\varepsilon,a])} = \lim_{k \to \infty} \frac{1}{\varepsilon_{n_k}} \cdot m_{\llcorner \Phi_t^{-1}([a-\varepsilon_{n_k},a]) \cap \bar{\Gamma}_a(1)}$$

for $\mathcal{L}^1$-a.e. $t \in [0, 1]$, where the exceptional set depends on the subsequence ε_{n_k} and the limit is in the weak topology.

Proof. Step 1. We show that for every $a \in \varphi(\mu_0)$,

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_{(0,1)} m\big(\Phi_t^{-1}([a-\varepsilon,a]) \setminus \bar{\Gamma}_a(1)\big) dt = 0.$$

Suppose by contradiction the existence of $a \in \varphi(\mu_0)$ and of a sequence $\varepsilon_n \to 0$ such that

$$\lim_{n \to \infty} \frac{1}{\varepsilon_n} \int_{(0,1)} m\big(\Phi_t^{-1}([a - \varepsilon_n, a]) \setminus \bar{\Gamma}_a(1)\big) dt \geq \alpha.$$

Then, since in Lemma 5.4 we have proved that $\|\nabla \hat{\Phi}_t\| \hat{m}_{b,t} \leq c \mathcal{S}^h \llcorner_{e_t(G_b)}$ for $\mathcal{L}^1$-a.e. $b \in \varphi(\mu_0)$, with $\|\nabla \hat{\Phi}_t\|$ positive m-a.e., it follows that

$$\lim_{n \to \infty} \frac{1}{\varepsilon_n} \int_{(0,1)} \int_{[a - \varepsilon_n, a]} \mathcal{S}^h(e_t(G_b) \setminus \bar{\Gamma}_a(1)) \mathcal{L}^1(db) \mathcal{L}^1(dt) \geq \alpha.$$

Then by Fubini's Theorem

$$\lim_{n \to \infty} \frac{1}{\varepsilon_n} \int_{[a - \varepsilon_n, a]} \int_{(0,1)} \mathcal{S}^h(e_t(G_b) \setminus \bar{\Gamma}_a(1)) \mathcal{L}^1(dt) \mathcal{L}^1(db) \geq \alpha.$$

Hence there exists a sequence a_k converging to a from below such that

$$\int_{(0,1)} \mathcal{S}^h(e_t(G_{a_k}) \setminus \bar{\Gamma}_a(1)) \mathcal{L}^1(dt) \geq \alpha.$$

for all $k \in \mathbb{N}$.

Step 2. It follows from Lemma 6.4 and Proposition 2.8 that, since $m \llcorner_{\bar{\Gamma}_{a_k}(1)} = \int m_{a_k,t} dt$:

$$\liminf_{k \to 0} m\big(\bar{\Gamma}_{a_k}(1) \setminus \bar{\Gamma}_a(1)\big) \geq \alpha.$$

Since as $k \to \infty$ the sequence a_k is converging to a, the sequence of compact sets of geodesics $\{G_{a_k}\}_{k \in \mathbb{N}}$ is converging in Hausdorff topology to a subset of G_a, hence the same happens for the sequence of compact sets $\{\bar{\Gamma}_{a_k}(1)\}_{k \in \mathbb{N}}$. Then just observe that

$$m(\bar{\Gamma}_a(1)) = \lim_{\delta \to 0} m(\bar{\Gamma}_a(1)^\delta) \geq m(\bar{\Gamma}_a(1)) + \liminf_{k \to \infty} m(\bar{\Gamma}_{a_k}(1) \setminus \bar{\Gamma}_a(1)) \geq m(\bar{\Gamma}_a(1)) + \alpha,$$

where $\bar{\Gamma}_a(1)^\delta = \{z \in X : d(z, \bar{\Gamma}_a(1)) \leq \delta\}$ is a neighborhood of $\bar{\Gamma}_a(1)$ and the first inequality follows from the definition of Hausdorff convergence. Since $\alpha > 0$ we have a contradiction and therefore for each $a \in \varphi(\mu_0)$

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_{(0,1)} m\big(\Phi_t^{-1}([a - \varepsilon, a]) \setminus \bar{\Gamma}_a(1)\big) dt = 0.$$

So for each sequence $\varepsilon_n \to 0$ there exists a subsequence ε_{n_k} such that

$$\lim_{k \to \infty} \frac{1}{\varepsilon_{n_k}} m\big(\Phi_t^{-1}([a - \varepsilon_{n_k}, a]) \setminus \bar{\Gamma}_a(1)\big) = 0,$$

for $\mathcal{L}^1$-a.e. $t \in [0,1]$.

Step 3. Let $\{f_h\}_{h \in \mathbb{N}} \subset C_b(X)$ be a dense family. Then for each f_k

$$\lim_{k \to \infty} \left(\frac{1}{\varepsilon_{n_k}} \int_{[a - \varepsilon_{n_k}, a]} f_h \hat{m}_{b,t} \mathcal{L}^1(db) - \frac{1}{\varepsilon_{n_k}} \int_{\Phi_t^{-1}([a - \varepsilon_{n_k}, a]) \cap \bar{\Gamma}_a(1)} f_h m \right) = 0$$

for all $t \in [0,1]$ minus a set of measure zero. Reasoning as Lemma 7.2, we have the existence of a set $E \subset [0,1]$ with $\mathcal{L}^1(E) = 0$ such that for all $f \in C_b(X)$ it holds

$$\lim_{k \to \infty} \left(\frac{1}{\varepsilon_{n_k}} \int_{[a - \varepsilon_{n_k}, a]} f \hat{m}_{b,t} \mathcal{L}^1(db) - \frac{1}{\varepsilon_{n_k}} \int_{\Phi_t^{-1}([a - \varepsilon_{n_k}, a]) \cap \bar{\Gamma}_a(1)} f m \right) = 0$$

for all $t \in [0,1] \setminus E$. Then again from Lemma 7.2 applied to $\hat{m}_{a,t}$ we have the claim. $\square$

The proof of the next Corollary follows from Lemma 7.2 and Proposition 7.3.

Corollary 7.4. *For $\mathcal{L}^1$-a.e. $a \in \varphi(\mu_0)$ the following holds: for every sequence $\varepsilon_n \to 0$ there exists a subsequence $\varepsilon_{n_k} \to 0$ so that*

$$\lim_{k \to \infty} \frac{1}{\varepsilon_{n_k}} \cdot m \llcorner_{\Phi_t^{-1}([a - \varepsilon_{n_k}, a]) \cap \bar{\Gamma}_a(1)} = \hat{m}_{a,t}$$

for $\mathcal{L}^1$-a.e. $t \in [0,1]$, where the exceptional set depends on the subsequence ε_{n_k} and the limit is in the weak topology.

We now prove that $\hat{m}_{a,t} \ll m_{a,t}$. Let us recall the disintegration formula for m as constructed in Proposition 5.6: for each $a \in \varphi(\mu_0)$ since the geodesics in G_a are disjoint even for different times it holds

$$(7.1) \qquad m_{\bar{\Gamma}_a(1)} = \int_{e_{1/2}(G_a)} g(y, \cdot)\mathcal{L}^1 \llcorner_{[0,1]} q_a(dy).$$

where g satisfies (5.5), q_a is the quotient measure satisfying for $I \subset e_{1/2}(G_a)$

$$q_a(I) = m(\{\gamma_t : \gamma \in G_a, \gamma_{1/2} \in I\})$$

and the measure $g(y, \cdot)\mathcal{L}^1\llcorner_{[0,1]}$ has to be intended as $(\gamma)_\sharp(g(y, \cdot)\mathcal{L}^1\llcorner_{[0,1]})$, with γ the unique element of G_a so that $\gamma_{1/2} = y$. Being the evaluation map $e_{1/2}$ a Borel isomorphism between G_a and $e_{1/2}(G)$ the measure q_a can be also interpret as a measure on G_a.

Proposition 7.5. *For $\mathcal{L}^1$-a.e. $a \in \varphi(\mu_0)$*

$$\hat{m}_{a,t} \ll m_{a,t}$$

for $\mathcal{L}^1$-a.e. $t \in [0, 1]$. Equivalently $(e_{1/2})_\sharp \gamma_a \ll q_a$.

Proof. Consider $a \in \varphi(\mu_0)$ and a subsequence ε_{n_k} so that Corollary 7.4 holds.

Step 1. Consider the evaluation map $e : [0, 1] \times e_{1/2}(G_a) \to \bar{\Gamma}_a(1)$ defined as usual by

$$e(s, y) = e_s \circ e_{1/2}^{-1}(y).$$

Note that it is continuous, surjective and its inverse is continuous as well. Hence $\bar{\Gamma}_a(1)$ and $[0,1] \times e_{1/2}(G_a)$ are homeomorphic.

Take I compact subset of $e_{1/2}(G_a)$ with $q_a(I) = 0$. Since q_a is a regular finite measure on $e_{1/2}(G_a)$, by outer regularity there exists a sequence $\{A_i\}_{i \in \mathbb{N}}$ with $A_i \subset e_{1/2}(G_a)$ and open in the subspace topology of $e_{1/2}(G_a)$ so that

$$I \subset A_i, \qquad q_a(A_i) \leq \frac{1}{i}.$$

Take now any open set $U \subset [0, 1]$ neighborhood of $1/2$. Then $e(U \times A_i)$ will be an open set in $\bar{\Gamma}_a(1)$ for each $i \in \mathbb{N}$.

Step 2. Then

$$\frac{1}{\varepsilon_{n_k}} m\llcorner_{\Phi_t^{-1}([a-\varepsilon_{n_k},a]) \cap \bar{\Gamma}_a(1)} \big(e(U \times A_i)\big)$$

$$= \frac{1}{\varepsilon_{n_k}} \int_{A_i} (g(y, \cdot)\mathcal{L}^1) \left(U \cap \{\tau \in [t, 1] : \Phi_t(\gamma_\tau) \in [a - \varepsilon_{n_k}, a]\}\right) q_a(dy).$$

Let $s_k \in (0, 1)$ be such that

$$s_k = \max\{s : \Phi_t(\gamma_{t+s}) \geq a - \varepsilon_{n_k}\}.$$

Then $s_k \geq \mathcal{L}^1\left(\{\tau \in [t, 1] : \Phi_t(\gamma_\tau) \in [a - \varepsilon_{n_k}, a]\}\right)$ and

$$\varepsilon_{n_k} = a - (a - \varepsilon_{n_k}) \geq \Phi_t(\gamma_t) - \Phi_t(\gamma_{t+s_k})$$

Therefore

$$(7.2) \qquad \lim_{k \to \infty} \frac{1}{\varepsilon_{n_k}} \mathcal{L}^1\left(\{\tau \in [0, 1] : \Phi_t(\gamma_\tau) \in [a - \varepsilon_{n_k}, a]\}\right) \leq \lim_{k \to \infty} \frac{s_k}{\Phi_t(\gamma_t) - \Phi_t(\gamma_{t+s_k})}$$

and the last term by Assumption 2 is bounded. Since g is uniformly bounded as well, it follows that

$$\frac{1}{\varepsilon_{n_k}} m\llcorner_{\Phi_t^{-1}([a-\varepsilon_{n_k},a]) \cap \bar{\Gamma}_a(1)} \big(e(U \times A_i)\big) \leq C q_a(A_i),$$

for some positive constant C not depending on k.

Step 3. We now observe that $\bar{\Gamma}_a(1)$ is a compact set. Hence any function $f \in C_b(X)$ can be extended, by Tiezte's Theorem, to a bounded and continuous function on the whole space, say $\tilde{f}$. It follows that

$$\lim_{k \to \infty} \frac{1}{\varepsilon_{n_k}} \cdot m\llcorner_{\Phi_t^{-1}([a-\varepsilon_{n_k},a]) \cap \bar{\Gamma}_a(1)} = \hat{m}_{a,t}$$

holds also in the weak topology of $\mathcal{P}(\bar{\Gamma}_a(1))$. So we can use lower semicontinuity on open sets of weakly converging measures, also for open sets in the trace topology of $\bar{\Gamma}_a(1)$. Therefore

$$\hat{m}_{a,t}(e(U \times A_i)) \leq \liminf_{k \to \infty} \frac{1}{\varepsilon_{n_k}} m {\llcorner}_{\Phi_t^{-1}([a-\varepsilon_{n_k},a]) \cap \bar{\Gamma}_a(1)}(e(U \times A_i)) \leq C\frac{1}{i}.$$

By outer regularity, $\hat{m}_{a,t}(I) = 0$ and the claim follows. $\qquad\square$

Direct consequence of Proposition 7.5 is that for $\mathcal{L}^1$-a.e. $a \in \varphi(\mu_0)$ we have $\hat{m}_{a,t} = \theta_{a,t}q_a$ for $\mathcal{L}^1$-a.e. $t \in [0,1]$, that is

$$\hat{m}_{a,t}(K) = \int_{e_{1/2}(e_t^{-1}(K))} \theta_{a,t}(y)q_a(dy),$$

for all $K \subset e_t(G_a)$.

7.2. **A formula for the density.** We now derive an explicit expression for the density of $\hat{m}_{a,t}$ with respect to $m_{a,t}$.

Lemma 7.6. *For $\mathcal{L}^1$-a.e. $a \in \varphi(\mu_0)$ and every sequence $\varepsilon_n \to 0^+$, there exists a subsequence ε_{n_k} such that the limit*

$$(7.3) \qquad \lim_{k \to \infty} \frac{1}{\varepsilon_{n_k}} \mathcal{L}^1\Big(\big\{\tau \in (0,1) : \Phi_t(\gamma_\tau) \in [a - \varepsilon_{n_k}, a]\big\}\Big)$$

exists for $\boldsymbol{\gamma}_a$-a.e. $\gamma \in G_a$ and $\mathcal{L}^1$-a.e. $t \in [0,1]$. If we denote by $\lambda_t(\gamma_t)$ its value, then

$$\hat{m}_{a,t} = \lambda_t m_{a,t}.$$

Proof. Consider $a \in \varphi(\mu_0)$ and ε_{n_k} so that Corollary 7.4 and Proposition 7.5 holds. Then we have

$$\lim_{k \to \infty} \int \frac{1}{\varepsilon_{n_k}} \Big(g(y,\cdot)\mathcal{L}^1\Big)\Big(\big\{_{\tau \in (0,1):\Phi_t(\gamma_\tau) \in [a-\varepsilon_{n_k},a]}\big\}\Big) q_a(dy) = \theta_{a,t}q_a,$$

again for $\mathcal{L}^1$-a.e. $t \in [0,1]$ with the exceptional set depending on the subsequence and where the convergence is in the weak topology. Using a localization argument on the support of q_a, it follows that there exists another subsequence that we will call again ε_{n_k} so that

$$\lim_{k \to \infty} \frac{1}{\varepsilon_{n_k}} \Big(g(y,\cdot)\mathcal{L}^1\Big)\Big(\big\{_{\tau \in (0,1):\Phi_t(\gamma_\tau) \in [a-\varepsilon_{n_k},a]}\big\}\Big) = \theta_{a,t}\delta_y$$

for $\mathcal{L}^1$-a.e. $t \in [0,1]$ and q_a-a.e. $y \in e_{1/2}(G_a)$. Then by continuity of $t \mapsto g(y,t)$ for q_a-a.e. y, it follows that

$$\lim_{k \to \infty} \frac{1}{\mathcal{L}^1\Big(\big\{\tau \in (0,1) : \Phi_t(\gamma_\tau) \in [a - \varepsilon_{n_k}, a]\big\}\Big)}$$
$$\cdot \Big(g(y,\cdot)\mathcal{L}^1\Big)\Big(\big\{_{\tau \in (0,1):\Phi_t(\gamma_\tau) \in [a-\varepsilon_{n_k},a]}\big\}\Big) = g(y,t)\delta_y$$

for $\mathcal{L}^1$-a.e. $t \in [0,1]$ and q_a-a.e. $y \in e_{1/2}(G_a)$. Then necessarily

$$\lim_{k \to \infty} \frac{1}{\varepsilon_{n_k}} \mathcal{L}^1\Big(\big\{\tau \in (0,1) : \Phi_t(\gamma_\tau) \in [a - \varepsilon_{n_k}, a]\big\}\Big)$$

exists $\mathcal{L}^1$-a.e. $t \in [0,1]$ and q_a-a.e. $y \in e_{1/2}(G_a)$. By uniqueness of the limit

$$\theta_{a,t} = g(y,t) \cdot \lim_{k \to \infty} \frac{1}{\varepsilon_{n_k}} \mathcal{L}^1\Big(\big\{\tau \in (0,1) : \Phi_t(\gamma_\tau) \in [a - \varepsilon_{n_k}, a]\big\}\Big).$$

Hence if we define $\lambda_t(y) = \theta_{a,t}(y)/g(y,t)$ then

$$\lambda_t(y) = \lim_{k \to \infty} \frac{1}{\varepsilon_{n_k}} \mathcal{L}^1\Big(\big\{\tau \in (0,1) : \Phi_t(\gamma_\tau) \in [a - \varepsilon_{n_k}, a]\big\}\Big)$$

and since $m_{a,t} = g(\cdot,t)q_a$ and $\hat{m}_{a,t} = \theta_{a,t}q_a$ it follows that

$$\hat{m}_{a,t} = \lambda_t m_{a,t},$$

and therefore the claim. $\qquad\square$

At the beginning of this section we observed that

$$e_{t\,\sharp}\boldsymbol{\gamma}_a = \left(\int \varrho_t(z)\hat{m}_{a,t}(dz)\right)^{-1}\varrho_t\hat{m}_{a,t},$$

so now Proposition 7.5 and Lemma 7.6 implies the next corollary.

Corollary 7.7. *The measure $(e_t)_\sharp\boldsymbol{\gamma}_a$ is absolute continuous with respect to the surface measure $m_{a,t}$.*

Let $\hat{h}_{a,t}$ be such that $(e_t)_\sharp\boldsymbol{\gamma}_a = \hat{h}_{a,t}m_{a,t}$. We prefer to think of $\hat{h}_{a,t}$ as a function defined on G_a rather than on $e_t(G_a)$, hence define $h_{a,r} : G_a \to [0,\infty]$ by $h_{a,r}(\gamma) := \hat{h}_{a,r}(\gamma_r)$. So we have found a decomposition of ϱ_t:

$$\varrho_t(\gamma_t) = \left(\int \varrho_t(z)\hat{m}_{a,t}(dz)\right)\frac{1}{\lambda_t(\gamma_t)}h_{a,t}(\gamma),$$

where $a = \varphi(\gamma_0)$. We now deduce a more convenient expression for λ_t. Recall the definition

$$I_t(\gamma) = \{\tau \in [0,1] : \gamma_\tau \in e_t(G)\} = \{\tau \in [0,1] : d(\gamma_\tau, e_t(G)) = 0\},$$

Theorem 7.8. *For $\mathcal{L}^1$-a.e $t \in [0,1]$*

$$(7.4) \qquad\qquad \frac{1}{\lambda_t(\gamma_t)} = \lim_{s\to 0}\frac{\Phi_t(\gamma_t) - \Phi_t(\gamma_{t+s})}{s},$$

point wise for $\boldsymbol{\gamma}$-a.e. $\gamma \in G$.

Proof. Step 1. From Lemma 7.6 for $\mathcal{L}^1$-a.e. $a \in \varphi(\mu_0)$, for every $\varepsilon_n \to 0^+$ there exists a subsequence ε_{n_k} such that

$$\lambda_t(\gamma_t) = \lim_{k\to\infty}\frac{1}{\varepsilon_{n_k}}\mathcal{L}^1\Big(\big\{\tau \in (0,1) : \Phi_t(\gamma_\tau) \in [a - \varepsilon_{n_k}, a]\big\}\Big),$$

point wise $\boldsymbol{\gamma}_a \otimes \mathcal{L}^1$-a.e. in $G_a \times [0,1]$. An equivalent expression of $\lambda_t(\gamma_t)$ is:

$$\lambda_t(\gamma_t) = \lim_{k\to\infty}\frac{\big((\Phi_t \circ \gamma)_\sharp\mathcal{L}^1\big)([a - \varepsilon_{n_k}, a])}{\mathcal{L}^1([a - \varepsilon_{n_k}, a])}.$$

Using Assumption 2 λ_t can be written in terms of the same limit above substituting $\Phi_t \circ \gamma$, that is defined only on $I_t(\gamma)$, with an extension of Φ_t to a neighborhood of t.

Since for each $\gamma \in G$ the set $I_t(\gamma)$ is compact, we can extend Φ_t by linearity on each geodesic of G_a. By d-monotonicity this will create no problem in the definition. More specifically: for $\delta > 0$ fixed, for each $\tau \in [t-\delta, t+\delta]$ and $\gamma \in G_a$ there exists

$$\tau_m = \max\{s \in I_t(\gamma) : s \leq \tau\}, \qquad \tau_M = \min\{s \in I_t(\gamma) : \tau \leq s\}.$$

Clearly τ_m and τ_M depends on γ and if $\tau \in I_t(\gamma)$ they all coincide $\tau = \tau_m = \tau_M$. Then we define the extension map $\hat{\Phi}_t$ by linearity

$$\hat{\Phi}_t(\gamma_\tau) = \Phi_t(\gamma_{\tau_m}) + (\tau - \tau_m)\frac{\Phi_t(\gamma_{\tau_M}) - \Phi_t(\gamma_{\tau_m})}{\tau_M - \tau_m}.$$

Since by d-cyclical monotonicity $\gamma_t \neq \bar{\gamma}_s$ for all $t,s \in [0,1]$ if $\gamma, \bar{\gamma} \in G_a$ with $\gamma \neq \bar{\gamma}$, the map $\hat{\Phi}_t$ is well defined on $e([t-\delta, t+\delta] \times G_a)$ and is measurable. Moreover by Assumption 2, on each line the map

$$[t-\delta, t+\delta] \ni \tau \mapsto \hat{\Phi}_t(\gamma_\tau)$$

is differentiable in t with strictly negative derivative and is Lipschitz in the whole interval $[t-\delta, t+\delta]$.

Consider now $\tau_k = \max\{\tau \in [t, t+\delta] : \hat{\Phi}_t(\gamma_\tau) \geq a - \varepsilon_{n_k}\}$, then

$$\mathcal{L}^1\Big(\big(\hat{\Phi}_t \circ \gamma\big)^{-1}[a - \varepsilon_{n_k}, a]\Big) \leq \tau_k - t.$$

Since by construction

$$\hat{\Phi}_t(\gamma_t) - \hat{\Phi}_t(\gamma_{\tau_k}) \geq \frac{1}{c}(\tau_k - t),$$

for some positive constant C, we have $(\tau_k - t) \leq c\varepsilon_{n_k}$ implying that

$$\int_{G_a} \left| \frac{\left((\Phi_t \circ \gamma)_\sharp \mathcal{L}^1\right)([a - \varepsilon_{n_k}, a])}{\varepsilon_{n_k}} - \frac{\left((\hat{\Phi}_t \circ \gamma)_\sharp \mathcal{L}^1\right)([a - \varepsilon_{n_k}, a])}{\varepsilon_{n_k}} \right| \gamma_a(d\gamma)$$

$$\leq \int_{G_a} \frac{\left((\hat{\Phi}_t \circ \gamma)_\sharp \mathcal{L}^1\right)([a - \varepsilon_{n_k}, a] \cap I_t(\gamma)^c)}{\varepsilon_{n_k}} \gamma_a(d\gamma)$$

$$\leq \int_{G_a} \frac{\mathcal{L}^1 \left((t - c\varepsilon_{n_k}, t + c\varepsilon_{n_k}) \cap I_t(\gamma)^c\right)}{\varepsilon_{n_k}} \gamma_a(d\gamma).$$

By Lemma 3.2 the last integral converges to 0 as $k \to \infty$. We have therefore proved that for $\mathcal{L}^1$-a.e. $a \in \varphi(\mu_0)$, for every $\varepsilon_n \to 0$ there exists a subsequence ε_{n_k} such that

$$\lambda_t(\gamma_t) = \lim_{k \to \infty} \frac{1}{\varepsilon_{n_k}} \mathcal{L}^1 \left(\left\{ \tau \in (0,1) : \hat{\Phi}_t(\gamma_\tau) \in [a - \varepsilon_{n_k}, a] \right\} \right),$$

for $\gamma_a \otimes \mathcal{L}^1$-a.e. in $G_a \times [0,1]$.

Step 2. Now we take advantage from the fact that $\hat{\Phi}_t \circ \gamma$ is defined on a connected set and invertible. For any ε sufficiently small the following identity holds:

$$\frac{\left((\hat{\Phi}_t \circ \gamma)_\sharp \mathcal{L}^1\right)([a - \varepsilon, a])}{\varepsilon} = \frac{s_\varepsilon}{\hat{\Phi}_t(\gamma_t) - \hat{\Phi}_t(\gamma_{t+s_\varepsilon})},$$

where s_ε is the unique $s \in [t, t + \delta]$ such that

$$\hat{\Phi}_t(\gamma_{t+s_\varepsilon}) = a - \varepsilon.$$

It follows that

(7.5) $$\frac{1}{\lambda_t(\gamma_t)} = \lim_{s \to 0} \frac{\hat{\Phi}_t(\gamma_t) - \hat{\Phi}_t(\gamma_{t+s})}{s},$$

for $\gamma_a \otimes \mathcal{L}^1$-a.e. $(\gamma, t) \in G_a \times [0,1]$. Restricting s to $I_t(\gamma)$ the claim follows. $\qquad\square$

8. Global estimates and main theorems

So far we have proved that in a metric measure space (X, d, m) verifying $\mathsf{CD}_{loc}(K, N)$ or $\mathsf{CD}^*(K, N)$ (actually $\mathsf{MCP}(K, N)$ would be enough), given a geodesic $\mu_t = \varrho_t m$ in the L^2-Wasserstein space with some regularity, the following decomposition holds:

$$\varrho_t(\gamma_t) = \left(\int \varrho_t(z) \hat{m}_{a,t}(dz) \right) \frac{1}{\lambda_t(\gamma_t)} h_{a,t}(\gamma),$$

where $a = \varphi(\gamma_0)$ and the functions involved in the decomposition are determined by the following identities:

$$h_{a,t} m_{a,t} = (e_t)_\sharp \gamma_a = \left(\int \varrho_t(z) \hat{m}_{a,t}(dz) \right)^{-1} \varrho_t \hat{m}_{a,t},$$

$$\frac{1}{\lambda_t(\gamma_t)} = \lim_{s \to 0} \frac{\Phi_t(\gamma_t) - \Phi_t(\gamma_{t+s})}{s}.$$

From Assumption 2, $\lambda_t(\gamma_t) > 0$ γ-a.e. and the above expression make sense.

To give a complete meaning to this decomposition we have to prove additional properties for both $h_{a,t}$ and $\lambda_{a,t}$. In this Section we will consider this function $h_{a,t}$ and λ_t in the perspective of lower curvature bounds. In particular, thanks to the metric results proved in Section 4, we prove that $h_{a,t}$ verifies $\mathsf{CD}^*(K, N-1)$.

As already observed (see (7.1)) a disintegration of $m_{\llcorner \bar{\Gamma}_a(1)}$ is given by the next expression:

(8.1) $$m_{\bar{\Gamma}_a(1)} = \int_{e_{1/2}(G_a)} \left(g(y, \cdot) \mathcal{L}^1_{\llcorner [0,1]} \right) q_a(dy),$$

where $g(y, \cdot)\mathcal{L}^1_{\llcorner [0,1]}$ has to be intended as a measure on $\gamma_{[0,1]} \subset X$, the image of γ where $\gamma = e_{1/2}^{-1}(y)$.

Since $\lambda_t(\gamma_t) > 0$ also for $\boldsymbol{\gamma}_a$-a.e $\gamma \in G_a$, it follows from Proposition 7.5 that $(e_{1/2})_\sharp\boldsymbol{\gamma}_a$ can be taken to be the quotient measure in (8.1), at the price of changing the value of g:

$$(8.2) \qquad m_{\bar{\Gamma}_a(1)} = \int_{G_a} \big(g(\gamma_{1/2}, \cdot)\mathcal{L}^1 \llcorner_{[0,1]} \big) \boldsymbol{\gamma}_a(d\gamma),$$

with the change of the value constant in t and therefore the new g still verifies (5.5). For ease of notation in what follows we will just denote with $g(\gamma, t)$ instead of $g(\gamma_{1/2}, t)$. The new densities g enjoy the following property.

Lemma 8.1. *For $\boldsymbol{\gamma}_a$-a.e. $\gamma \in G_a$*

$$h_{a,t}(\gamma)g(\gamma, t) = 1, \qquad \mathcal{L}^1 - a.e. \ t \in [0,1].$$

Proof. The function $\hat{h}_{a,t}$ has been introduced after Corollary 7.7. For any measurable sets $H \subset G_a$, $I \subset [0,1]$ the following identities hold:

$$\begin{aligned}
\boldsymbol{\gamma}_a(H)\mathcal{L}^1(I) &= \int_I (h_{a,t}m_{a,t})(e_t(H))dt = \int_{\{\gamma_t : \gamma \in H, t \in I\}} \hat{h}_{a,t}(z)m_{a,t}(dz)dt \\
&= \int_{\{\gamma_t : \gamma \in H, t \in I\}} \hat{h}_{a,t}(z)m(dz) \\
&= \int_H \left(\int_I h_{a,t}(\gamma)g(\gamma, t)dt \right) \boldsymbol{\gamma}_a(d\gamma),
\end{aligned}$$

where passing from the second to the third line we used (8.2) and $\hat{h}_{a,t}$ was introduced after Corollary 7.7. The claim follows from the arbitrariness of H and I. $\qquad \square$

8.1. **Gain of one degree of freedom.** As proved in Section 4, for any $a_1 < b_1 < a_0 < b_0$ and for any $\gamma \in G_a$ so that $(a_0, b_1) \subset \phi_a(\gamma_{[0,1]})$ we can define $R_0^\gamma, L_0^\gamma \subset [0,1]$ and $R_1^\gamma, L_1^\gamma \subset [0,1]$ so that

$$\phi_a \circ \gamma\left((R_0^\gamma, R_0^\gamma + L_0^\gamma)\right) = (a_0, b_0), \qquad \phi_a \circ \gamma\left((R_1^\gamma, R_1^\gamma + L_1^\gamma)\right) = (a_1, b_1),$$

where ϕ_a is a Kantorovich potential associated to the d-monotone set $\{(\gamma_s, \gamma_t) : \gamma \in G_a, s \leq t\}$. The previous equations are equivalent to

$$\phi_a \circ \gamma(R_0^\gamma) = b_0, \qquad \phi_a \circ \gamma(R_0^\gamma + L_0^\gamma) = a_0.$$

and

$$\phi_a \circ \gamma(R_1^\gamma) = b_1, \qquad \phi_a \circ \gamma(R_1^\gamma + L_0^\gamma) = a_1.$$

Accordingly for all $t \in [0,1]$ we define

$$R_t^\gamma := (1-t)R_0^\gamma + tR_1^\gamma, \qquad L_t^\gamma := (1-t)L_0^\gamma + tL_1^\gamma.$$

Let $H \subset G_a$ be so that for all $\gamma \in H$ both $(a_0, b_0), (a_1, b_1) \subset \phi_a(\gamma_{(0,1)})$ with $a_1 < b_1 < a_0 < b_0$. The Proposition 4.7 implies if we define

$$(8.3) \qquad [0,1] \ni t \mapsto \nu_t := \frac{1}{\boldsymbol{\gamma}_a(H)} \int_H \frac{1}{L_t^\gamma} \mathcal{L}^1 \llcorner_{[R_t^\gamma, R_t^\gamma + L_t^\gamma]} \boldsymbol{\gamma}_a(d\gamma) \in \mathcal{P}([0,1] \times G_a),$$

then $[0,1] \ni t \mapsto (e_\sharp)\nu_t$ is a W_2-geodesic.

Moreover from Lemma 8.1 we can deduce that for each $t \in [0,1]$ the density $p_t(x)$ of $(e)_\sharp\nu_t$ w.r.t. m is given by

$$(8.4) \qquad p_t(\gamma_\tau) = \begin{cases} \dfrac{1}{\boldsymbol{\gamma}_a(H)L_t^\gamma} h_{a,\tau}(\gamma), & \tau \in [R_t^\gamma, R_t^\gamma + L_t^\gamma], \\ 0, & \text{otherwise.} \end{cases}$$

The dynamical optimal plan associated to ν_t can be obtained as follows: consider the following map

$$\begin{aligned}
\Theta : \mathcal{G}(X) \times [0,1] &\rightarrow \mathcal{G}(X) \\
(\gamma, s) &\mapsto t \mapsto \eta_t = \gamma_{(1-t)(R_0^\gamma + sL_0^\gamma) + t(R_1^\gamma + sL_1^\gamma)}
\end{aligned}$$

Then if we pose

$$(8.5) \qquad \tilde{\boldsymbol{\gamma}}_a := \Theta_\sharp\left(\frac{1}{\boldsymbol{\gamma}_a(H)} \boldsymbol{\gamma}_a \llcorner_H \otimes \mathcal{L}^1 \llcorner_{[0,1]} \right),$$

it follows that $(e)_\sharp \nu_t = (e_t)_\sharp \tilde{\gamma}_a$.

Theorem 8.2. *For γ_a-a.e. $\gamma \in G_a$ and for any $0 \leq \tau_0 < \tau_1 \leq 1$ the following inequality holds true:*

$$(8.6) \qquad h_{a,\tau_{1/2}}^{-\frac{1}{N-1}}(\gamma) \geq \sigma_{K,N-1}^{(1/2)}\big((\tau_1 - \tau_0)L(\gamma)\big)\left\{ h_{a,\tau_0}^{-\frac{1}{N-1}}(\gamma) + h_{a,\tau_1}^{-\frac{1}{N-1}}(\gamma) \right\},$$

where $\tau_{1/2} = (\tau_0 + \tau_1)/2$.

Proof. As a preliminary step, we note that in order to prove the claim is sufficient to prove (8.6) locally, i.e. for R_0 and R_1 sufficiently close. As proved in [7], reduced curvature dimension condition enjoys the globalization property.

Step 1. Since ϕ_a is 1-Lipschitz and G_a is compact, there exist real numbers α_i, β_i for $i = 0, 1$ so that

$$\phi_a \circ e_0(G_a) \subset [\alpha_0, \alpha_1], \quad \phi_a \circ e_1(G_a) \subset [\beta_0, \beta_1].$$

For any $n \in \mathbb{N}$ and $\mathbb{N} \ni k \leq n-1$ we can consider the following family of curves

$$E_{k,n} := (\phi_a \circ e_0)^{-1}\left(\left[\alpha_0 + \frac{k}{n}\alpha_1, \alpha_0 + \frac{k+1}{n}\alpha_1\right]\right), \qquad D_{k,n} := (\phi_a \circ e_1)^{-1}\left(\left[\beta_0 + \frac{k}{n}\beta_1, \beta_0 + \frac{k+1}{n}\beta_1\right]\right),$$

where the maps $\phi_a \circ e_i$, for $i = 0, 1$, has to be considered as defined only on G_a. Then we define the family of compact sets

$$M_{h,k,n} := E_{h,n} \cap D_{k,n}.$$

For any $n \in \mathbb{N}$, as h and k vary from 0 to $n-1$ the sets $M_{h,k,n}$ cover G_a. In particular we will consider this covering for n so that

$$\frac{1}{n} \ll \min\{L(\gamma), \gamma \in G_a\}, |\alpha_1 - \alpha_0|, |\beta_1 - \beta_0|.$$

Under the previous condition

$$\min\{\phi_a(\gamma_0) : \gamma \in M_{h,k,n}\} \gg \max\{\phi_a(\gamma_1) : \gamma \in M_{h,k,n}\}.$$

Then for any $a > b$ real numbers so that

$$\min\{\phi_a(\gamma_0) : \gamma \in M_{h,k,n}\} > a > b > \max\{\phi_a(\gamma_1) : \gamma \in M_{h,k,n}\},$$

for any $\gamma \in M_{h,k,n}$ the image $\phi_a(\gamma_{[0,1]})$ contains $[b, a]$. Therefore we are under the hypothesis of Proposition 4.7.

Step 2. Fix a compact set $H \subset M_{h,k,n}$ and a, b such that the curvature dimension condition $\mathsf{CD}(K, N)$ holds true for all measures supported in

$$\phi_a^{-1}([b, a]) \cap \{\gamma_{[0,1]} : \gamma \in H\}.$$

Chose now a_0, b_0 and a_1, b_1 so that $(b_0, a_0), (b_1, a_1) \subset [b, a]$. In the same manner as Proposition 4.7 consider $R_0^\gamma, R_1^\gamma, L_0^\gamma$ and L_1^γ. Finally define $\{(e)_\sharp \nu_t\}_{t \in [0,1]}$ as before in (8.3) and the associated dynamical optimal plan $\tilde{\gamma}_a$ as in (8.5). Note that since $M_{h,k,n}$ is a covering of G_a we can always assume $\gamma_a(M_{h,k,n}) > 0$ and therefore $\gamma_a(H) > 0$.

Condition $\mathsf{CD}_{loc}(K, N)$ for $t = 1/2$ imply that for $\tilde{\gamma}_a$-a.e. $\eta \in \mathcal{G}(X)$

$$p_{1/2}^{-1/N}(\eta_{1/2}) \geq \tau_{K,N}^{(1/2)}(d(\eta_0, \eta_1))\left\{ p_0^{-1/N}(\eta_0) + p_1^{-1/N}(\eta_1) \right\},$$

that can be formulated also in the following way: for $\mathcal{L}^1$-a.e. $s \in [0, 1]$ and γ_a-a.e. $\gamma \in H$

$$p_{1/2}^{-1/N}(\gamma_{R_{1/2}^\gamma + sL_{1/2}^\gamma}) \geq \tau_{K,N}^{(1/2)}\big((R_1^\gamma - R_0^\gamma + s|L_1^\gamma - L_0^\gamma|)L(\gamma)\big)\left\{ p_0^{-1/N}(\gamma_{R_0^\gamma + sL_0^\gamma}) + p_1^{-1/N}(\gamma_{R_1^\gamma + sL_1^\gamma}) \right\}.$$

Then using (8.4) and the continuity of $r \mapsto h_r(\gamma)$ (Lemma 8.1), letting $s \searrow 0$, it follows that

$$(L_0^\gamma + L_1^\gamma)^{1/N} h_{a,R_{1/2}^\gamma}^{-1/N}(\gamma)$$

$$(8.7) \qquad \geq \sigma_{K,N-1}^{(1/2)}\big((R_1^\gamma - R_0^\gamma)L(\gamma)\big)^{\frac{N-1}{N}}\left\{ (L_0^\gamma)^{1/N} h_{a,R_0^\gamma}^{-1/N}(\gamma) + (L_1^\gamma)^{1/N} h_{a,R_1^\gamma}^{-1/N}(\gamma) \right\},$$

for γ_a-a.e. $\gamma \in H$, with exceptional set depending on a_0, b_0, a_1, b_1.

Step 3. Note that all the involved quantities in (8.7) are continuous w.r.t. $R_0^\gamma, L_0^\gamma, R_1^\gamma, L_1^\gamma$, that in turn are continuous functions of a_0, b_0, a_1, b_1 respectively. Therefore there exists a common exceptional set $H' \subset H$ of zero γ_a-measure such that (8.7) holds true for all for all $a_0 > a_1 \in (a, b)$, and all b_0, b_1 so that

$a_0 - b_0, a_1 - b_1$ are sufficiently small and all $\gamma \in H \setminus H'$. Then for fixed fixed $\gamma \in H \setminus H'$, varying L_0^γ, L_1^γ in (8.7) yields

$$h_{a,R_{1/2}^\gamma}^{-\frac{1}{N-1}}(\gamma) \geq \sigma_{K,N-1}^{(1/2)}\big((R_1^\gamma - R_0^\gamma)L(\gamma)\big)\left\{h_{a,R_0^\gamma}^{-\frac{1}{N-1}}(\gamma) + h_{a,R_1^\gamma}^{-\frac{1}{N-1}}(\gamma)\right\}.$$

Indeed the optimal choice is

$$L_0^\gamma = L\frac{h_{a,R_0^\gamma}^{-1/(N-1)}(\gamma)}{h_{a,R_0^\gamma}^{-1/(N-1)}(\gamma) + h_{a,R_1^\gamma}^{-1/(N-1)}(\gamma)}, \qquad L_1^\gamma = L\frac{h_{a,R_1^\gamma}^{-1/(N-1)}(\gamma)}{h_{a,R_0^\gamma}^{-1/(N-1)}(\gamma) + h_{a,R_1^\gamma}^{-1/(N-1)}(\gamma)}$$

for sufficiently small $L > 0$.

Using the same argument of [11], we prove the global (8.6) for $\tau_0^\gamma, \tau_1^\gamma$ so that

$$\phi_a(\gamma_{\tau_0^\gamma}) \leq a, \qquad \phi_a(\gamma_{\tau_1^\gamma}) \geq b,$$

for $\boldsymbol{\gamma}$-a.e. $\gamma \in M_{h,k,n}$. Since n can be as big as we want, τ_0^γ and τ_0^γ can be taken 0 and 1 respectively. Therefore we obtain the claim. $\qquad\square$

We have therefore proved one of the main results of this note.

Theorem 8.3. *Let (X, d, m) be a non-branching metric measure space verifying $\mathsf{CD}_{loc}(K, N)$ or $\mathsf{CD}^*(K, N)$ and let $\{\mu_t\}_{t\in[0,1]} \subset \mathcal{P}_2(X, d, m)$ be a geodesic with $\mu_t = \varrho_t m$. Assume moreover Assumption 1 and Assumption 2. Then*

$$\varrho_t(\gamma_t) = C(a)\frac{1}{\lambda_t(\gamma_t)}h_{a,t}(\gamma), \qquad \boldsymbol{\gamma} - a.e. \ \gamma \in G,$$

where $a = \varphi(\gamma_0)$ and $C(a) = \|\boldsymbol{\gamma}_a\|$ is a constant depending only on a. The map $[0,1] \ni t \mapsto h_{a,t}(\gamma)$ verifies $\mathsf{CD}^*(K, N-1)$ for $\boldsymbol{\gamma}$-a.e. $\gamma \in G$ and

$$\frac{1}{\lambda_t(\gamma_t)} = \lim_{s\to0}\frac{\Phi_t(\gamma_t) - \Phi_t(\gamma_{t+s})}{s}.$$

8.2. **Globalization for a class of optimal transportation.** In order to prove globalization theorem of CD_{loc} it now necessary to show concavity in time of $\lambda_t(\gamma_t)$. We will do that in the framework of Section 5: $L(\gamma)$ depends only on $\varphi(\gamma_0)$, i.e.

$$L(\gamma) = f(\varphi(\gamma_0)), \qquad \boldsymbol{\gamma} - a.e. \ \gamma \in G,$$

for some $f : \varphi(\mu_0) \to (0, \infty)$ such that $\varphi(\mu_0) \ni a \mapsto a - f^2/a$ is non increasing.

Proposition 8.4. *Assume the following: Then for $\boldsymbol{\gamma}$-a.e. $\gamma \in G$ the following holds true*

$$\lambda_t(\gamma_t) = (1-t)\lambda_0(\gamma_0) + t\lambda_1(\gamma_1),$$

for every $t \in [0,1]$.

Proof. Since $\Phi_t = F_t^{-1} \circ \varphi_t$, where $F_t(a) = a - tf^2/2$ and

$$\lim_{s\to0}\frac{\varphi_t(\gamma_t) - \varphi_t(\gamma_{t+s})}{s} = L^2(\gamma),$$

for all $t \in (0, 1)$, it follows that from Theorem 7.8 that

$$\lambda_t(\gamma_t) = (\partial_a F_t)(F_t^{-1}(\varphi_t(\gamma_t)))\frac{1}{L^2(\gamma)}$$

$$= (\partial_a F_t)(\varphi(\gamma_0))\frac{1}{L^2(\gamma)}.$$

Since $(\partial_a F_t)g_t$ is linear in t the claim follows. $\qquad\square$

Using the results proved so far, we can now state the following.

Theorem 8.5. *Let (X, d, m) be a non-branching metric measure space verifying $\mathsf{CD}_{loc}(K, N)$ or $\mathsf{CD}^*(K, N)$ and let $\{\mu_t\}_{t \in [0,1]} \subset \mathcal{P}_2(X, d, m)$ be a geodesic with $\mu_t = \varrho_t m$. Assume moreover that*

$$L(\gamma) = f(\varphi(\gamma_0)),$$

for some $f : \varphi(\mu_0) \to (0, \infty)$ such that $\varphi(\mu_0) \ni a \mapsto a - f^2/a$ is non increasing. Then

$$\varrho_t(\gamma_t)^{-1/N} \geq \varrho_0(\gamma_0)^{-1/N} \tau_{K,N}^{(1-t)}(d(\gamma_0, \gamma_1)) + \varrho_1(\gamma_1)^{-1/N} \tau_{K,N}^{(s)}(d(\gamma_0, \gamma_1)),$$

for every $t \in [0, 1]$ and for $\boldsymbol{\gamma}$-a.e. $\gamma \in G$.

Proof. From Remark 6.1,

$$\varrho_t(\gamma_t) = \left(\int \varrho_t(z) \hat{m}_{a,t}(dz) \right) \frac{h_{a,t}(\gamma)}{\lambda_t(\gamma_t)},$$

where the integral is constant in t and therefore in order to prove the claim we can assume

$$\varrho_t(\gamma_t) = \frac{1}{\lambda_t(\gamma_t)} h_{a,t}(\gamma).$$

Then from Theorem 8.2 and Proposition 8.4

$$
\begin{aligned}
\varrho_t^{-1/N}(\gamma_t) &= \left(\frac{1}{\lambda_t(\gamma_t)} h_{a,t}(\gamma) \right)^{-1/N} \\
&= \left((1-t)\lambda_0(\gamma_0) + t\lambda_1(\gamma_1) \right)^{\frac{1}{N}} \left(h_{a,t}^{-1/(N-1)}(\gamma) \right)^{\frac{N-1}{N}} \\
&\geq \left((1-t)\lambda_0(\gamma_0) \right)^{1/N} \left(\sigma_{K,N-1}^{(1-t)}(d(\gamma_0, \gamma_1)) h_{a,0}^{-\frac{1}{N-1}}(\gamma) \right)^{\frac{N-1}{N}} \\
&\quad + \left(t\lambda_1(\gamma_1) \right)^{1/N} \left(\sigma_{K,N-1}^{(t)}(d(\gamma_0, \gamma_1)) h_{a,1}^{-\frac{1}{N-1}}(\gamma) \right)^{\frac{N-1}{N}} \\
&= \varrho_0^{-1/N}(\gamma_0) \tau_{K,N}^{(1-t)}(d(\gamma_0, \gamma_1)) + \varrho_1^{-1/N}(\gamma_1) \tau_{K,N}^{(t)}(d(\gamma_0, \gamma_1)).
\end{aligned}
$$

The claim follows. $\qquad\square$

9. More on the one-dimensional component

Assuming the metric measure space (X, d, m) to be *infinitesimally strictly convex*, see Subsection 2.3, we can give an more explicit expression for λ_t.

Define the restriction map as follows. For any $t \in (0, 1)$ let $restr_{[t,1]} : \mathcal{G}(X) \to \mathcal{G}(X)$ be defined as follows $restr_{[t,1]}(\gamma)_s = \gamma_{(1-s)t+s}$. Denote by $\boldsymbol{\gamma}_{[t,1]}$ the measure $restr_{[t,1]\sharp} \boldsymbol{\gamma}$.

Lemma 9.1. *For all $t \in [0, 1)$ the measure $\boldsymbol{\gamma}_{[t,1]}$ represents $\nabla(1-t)(-\varphi_t)$.*

The notion of test plans representing gradients has been introduced in Definition 2.12.

Proof. First observe that $\varphi_t \in S^2(e_t(G), d, m)$. Indeed from Proposition 2.10, since φ_t is a Kantorovich potential for (μ_t, μ_1), it follows that

$$|D\varphi_t|_w(\gamma_t) = \frac{d(\gamma_t, \gamma_1)}{1-t} = d(\gamma_0, \gamma_1), \qquad \text{for } \boldsymbol{\gamma} - a.e.\gamma,$$

and therefore $|D\varphi_t|_w \in L^2(e_t(G), m)$. We know that $\boldsymbol{\gamma}_{[t,1]}$ is the optimal dynamical transference plan between μ_t and μ_1 and $(1-t)\varphi_t$ is the Kantorovich potential for the d^2 cost, hence Proposition 2.10 implies that

$$\lim_{t \downarrow 0} \int \frac{\varphi_t(\gamma_0) - \varphi_t(\gamma_\tau)}{\tau} \boldsymbol{\gamma}_{[t,1]}(d\gamma) = \int \frac{d^2(\gamma_0, \gamma_1)}{1-t} \boldsymbol{\gamma}_{[t,1]}(d\gamma).$$

Since $\|\boldsymbol{\gamma}_{[t,1]}\|_2^2 = \int d^2(\gamma_0, \gamma_1) \boldsymbol{\gamma}_{[t,1]}(d\gamma)$,

$$(1-t) \lim_{t \downarrow 0} \int \frac{\varphi_t(\gamma_0) - \varphi_t(\gamma_\tau)}{\tau} \boldsymbol{\gamma}_{[t,1]}(d\gamma) = \|\boldsymbol{\gamma}_{[t,1]}\|_2^2$$

and the claim follows. $\qquad\square$

Using Theorem 2.13 and Theorem 7.8 we can now write λ_t in a differential expression.

Proposition 9.2. *Let (X, d, m) be infinitesimally strictly convex. Then λ_t verifies the following identity: for every $t \in [0, 1)$*

$$\frac{1}{\lambda_t(\gamma_t)} = D\Phi_t(\nabla\varphi_t)(\gamma_t), \qquad \gamma - a.e.\gamma,$$

where the exceptional set depends on t.

Proof. Since (X, d, m) is infinitesimally strictly convex, and $\gamma_{[t,1]}$ represents $\nabla(1-t)(-\varphi_t)$, from Theorem 2.13 it follows that

$$\lim_{\tau\downarrow 0} \int_{restr_{[t,1]}(G)} \frac{\Phi_t(\gamma_0) - \Phi_t(\gamma_\tau)}{\tau} \gamma_{[t,1]}(d\gamma) = (1-t)\int D\Phi_t(\nabla\varphi_t)(x)\mu_t(dx)$$

$$= (1-t)\int_{restr_{[t,1]}(G)} D\Phi_t(\nabla\varphi_t)(\gamma_0)\gamma_{[t,1]}(d\gamma).$$

Since the previous identity holds true even if we restrict to a subset of $restr_{[t,1]}(G)$, it follows that it holds point-wise: for $\gamma_{[t,1]}$-a.e. γ

$$\lim_{\tau\downarrow 0} \frac{\Phi_t(\gamma_0) - \Phi_t(\gamma_\tau)}{\tau} = (1-t)D\Phi_t(\nabla\varphi_t)(\gamma_0).$$

So fix $\hat{\gamma}$ in the support of $\gamma_{[t,1]}$ such that the limit exists and consider γ in the support of γ such that $\hat{\gamma}_\tau = \gamma_{(1-\tau)t+\tau}$, then we have

$$\frac{\Phi_t(\hat{\gamma}_0) - \Phi_t(\hat{\gamma}_\tau)}{\tau} = \frac{\Phi_t(\gamma_t) - \Phi_t(\gamma_{(1-\tau)t+\tau})}{\tau} = \frac{\Phi_t(\gamma_t) - \Phi_t(\gamma_{(1-\tau)t+\tau})}{\tau(1-t)}(1-t),$$

and therefore the claim follows from Theorem 7.8. $\qquad\qquad\square$

Under the infinitesimally strictly convexity assumption, we have therefore the following decomposition:

$$\frac{1}{c(\varphi(\gamma_0))}\varrho_t(\gamma_t) = D\Phi_t(\nabla\varphi_t)(\gamma_t)h_{a,t}(\gamma_t),$$

where $c(a) = \int \varrho_t(z)\hat{m}_{a,t}(dz)$ is independent of t, and h verifies $\mathsf{CD}^*(K, N-1)$.

9.1. **A formal computation.** We conclude this note with a formal calculation in order to show a formal expression of $D\Phi_t(\nabla\varphi_t)(\gamma_t)$ in a smooth framework.

So let us assume X be the Euclidean space with distance given by the euclidean distance and m any measure absolute continuous with respect to the Lebesgue measure of the right dimension. Let $\mu_t = \varrho_t m$ be the usual geodesic in the L^2-Wasserstein space over X and let $T_t, T_{t,1} : X \to X$ be optimal maps such that

$$(T_t)_\sharp\mu_0 = \mu_t \qquad (T_{t,1})_\sharp\mu_t = \mu_1.$$

Hence

$$T_t = Id - t\nabla\varphi_0, \qquad T_{t,1} = Id - (1-t)\nabla\varphi_t,$$

with φ_0 a Kantorovich potential associated to μ_0, μ_1 and φ_t the usual evolution at time t of φ_0. Then the standard identity holds:

$$\varphi_t(\gamma_t) = (1-t)\varphi_0(\gamma_0) + t\varphi_1(\gamma_1).$$

Clearly $\gamma_0 = T_t^{-1}(\gamma_t)$ and $\gamma_1 = T_{t,1}(\gamma_t)$. Then one can differentiate the standard identity in the direction $s \mapsto \gamma_{t+s}$. Then we get

$$\|\nabla\varphi_t\|^2(\gamma_t) = (1-t)\langle\nabla\varphi_0(\gamma_0), DT_t^{-1}(\gamma_t)\nabla\varphi_t(\gamma_t)\rangle + t\langle\nabla\varphi_1(\gamma_1), DT_{t,1}(\gamma_t)\nabla\varphi_t(\gamma_t)\rangle.$$

Moreover one can write Φ_t in a more convenient way:

$$\Phi_t = \varphi_0 \circ T_t^{-1}$$

and then compute λ_t using Proposition 9.2

$$\frac{1}{\lambda_t(\gamma_t)} = \langle(DT_t^{-1})^t(\gamma_0)\nabla\varphi_0(\gamma_0), \nabla\varphi_t(\gamma_t)\rangle$$

$$= \langle\nabla\varphi_0(\gamma_0), DT_t^{-1}(\gamma_t)\nabla\varphi_t(\gamma_t)\rangle$$

Then using what calculated before

$$\frac{1}{\lambda_t(\gamma_t)} = \frac{1}{1-t}\|\nabla\varphi_t(\gamma_t)\|^2 - \frac{t}{1-t}\langle\nabla\varphi_1(\gamma_1), DT_{t,1}(\gamma_t)\nabla\varphi_t(\gamma_t)\rangle$$
$$= \|\nabla\varphi_t(\gamma_t)\|^2 + t\langle H\varphi_t(\gamma_t)\nabla\varphi_t(\gamma_t), \nabla\varphi_t(\gamma_t)\rangle,$$

where $H\varphi_t$ is the Hessian of φ_t. Clearly the effect of curvature would change the expression of $DT_{t,1}$. Hence on a linear space

$$\frac{1}{\lambda_t(\gamma_t)} = \langle(Id + tH\varphi_t(\gamma_t))\nabla\varphi_t(\gamma_t), \nabla\varphi_t(\gamma_t)\rangle.$$

As a final comment, by Corollary 2.16, it holds that $\varphi^t = -\varphi_t^c$ and since

$$\varphi_t^c(x) = H_0^t(-\varphi) = \inf_{y\in X}\frac{1}{2t}d^2(x,y) - \varphi(y),$$

it follows by semi-concavity that $Id - tH\varphi_t^c \geq 0$, in the sense of symmetric matrices. Note that we have derived in a different way the same expression for λ_t obtained in (6.1) from the decomposition of the differential of optimal transport map on manifold of [12]. Again from [12] it follows that

$$Id - tH\varphi_t^c > 0,$$

showing again consistency with Assumption 2.

References

[1] G. Alberti, S. Bianchini, and G. Crippa. Structure of level sets and sard-type properties of lipschitz maps. *to appear on Ann. Sc. Norm. Super. Pisa, Cl. Sci.*, 2011.

[2] L. Ambrosio. Some fine properties of sets of finite perimeter in ahlfors regular metric measure spaces. *Advances in Mathematics*, 159:51–67, 2001.

[3] L. Ambrosio. Fine properties of sets of finite perimeter in doubling metric measure spaces. *Set Valued Analysis*, 10:111–128, 2002.

[4] L. Ambrosio and N. Gigli. A user's guide to optimal transport. preprint, 2011.

[5] L. Ambrosio, N. Gigli, and G. Savarè. Calculus and heat flow in metric measure spaces and application to spaces with Ricci curvature buonded from below. Preprint, arXiv:1106.2090.

[6] L. Ambrosio, M. Miranda Jr., and D. Pallara. Special functions of bounded variation in doubling metric measure spaces. *Calculus of variations: topics from the mathematical heritage of E. De Giorgi*, pages 1–45, 2004.

[7] K. Bacher and K.T. Sturm. Localization and tensorization properties of the Curvature-Dimension condition for metric measure spaces. *J. Funct. Anal.*, 259(1):28–56, 2010.

[8] S. Bianchini and L. Caravenna. On the extremality, uniqueness and optimality of transference plans. *Bull. Inst. Math. Acad. Sin.(N.S.)*, 4(4):353–454, 2009.

[9] S. Bianchini and F. Cavalletti. The Monge problem for distance cost in geodesic spaces. *Comm. Math. Phys*, 318:615 – 673, 2013.

[10] F. Cavalletti and M. Huesmann. Self-intersection of optimal geodesics. preprint arXiv:1211.6547, 2012.

[11] F. Cavalletti and K.-T. Sturm. Local curvature-dimension condition implies measure-contraction property. *J. Funct. Anal.*, 262:5110 – 5127, 2012.

[12] D. Cordero-Erausquin, R. J. McCann, and M. Schmuckenshläger. A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. *Invent. Math.*, 146:219–257, 2011.

[13] D. H. Fremlin. *Measure Theory*, volume 4. Torres Fremlin, 2002.

[14] S. Gallot, D. Hulin, and J. Lafontaine. *Riemannian Geometry*. Springer-Verlag, 1987.

[15] N. Gigli. On the differential structure of metric measure spaces and applications. preprint, arXiv:1205.6622, 2012.

[16] N. Gigli. Optimal maps in non branching spaces with ricci curvature bounded from below. *Geom. Funct. Anal.*, 22:990–999, 2012.

[17] M. Miranda Jr. Functions of bounded variation on "good" metric spaces. *J. Math. Pures Appl.*, 82:975–1004, 2003.

[18] J. Lott and C. Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math.*, 169(3):903–991, 2009.

[19] T. Rajala and K.-T. Sturm. Non-branching geodesics and optimal maps in strong $\mathsf{CD}(k,\infty)$-spaces. arXiv 1207.6754, 2012.

[20] K.T. Sturm. On the geometry of metric measure spaces.I. *Acta Math.*, 196(1):65–131, 2006.

[21] K.T. Sturm. On the geometry of metric measure spaces.II. *Acta Math.*, 196(1):133–177, 2006.

[22] C. Villani. *Optimal transport, old and new*. Springer, 2008.

RWTH, Department of Mathematics, Templergraben 64, D-52062 Aachen (Germany)
E-mail address: cavalletti@instmath.rwth-aachen.de