METRIC REGULARITY FOR OPTIMAL TRANSPORT PLANS

FABIO CAVALLETTI AND MARTIN HUESMANN

Abstract. Let (X, d, m) be a geodesic non-branching metric measure space. Assuming just self-intersection of evolving sets, we prove a metric regularity results for optimal transport plans. We sharpen previously known shortening Lemmas in purely metric framework. For instance our result applies to spaces satisfying $\mathsf{CD}(K, N)$, both for $N \in \mathbb{N}$ and $N = \infty$.

1. Introduction

In the last 20 years the theory of optimal transport has developed in a wide and fascinating field with a huge variety of applications in different fields inside and outside of mathematics. One reason for this is the rich interplay of the structure of optimal transport plans with the underlying geometry. In order to better understand these structures it is fundamental to have a clear picture of regularity properties enjoyed by optimal transport plans and maps.

In this note we investigate the regularity property of the length map from intermediate times. More explicitly if $\{\mu_t\}_{t\in[0,1]}$ is a geodesic in the L^2-Wasserstein space over a geodesic non-branching metric measure space (X, d, m), we study the regularity properties of

$$\begin{array}{rcl} L_t : \mathrm{supp}(\mu_t) & \to & \mathbb{R} \\ x & \mapsto & L_t(x) := L(\gamma), \end{array}$$

where γ is the unique optimal geodesic passing through x at time t, $L(\gamma)$ denotes the length of γ and $t \in (0, 1)$.

Before stating our result, we recall the state of the art of this problem. In the smooth framework when X is a d-dimensional manifold, regularity of transport maps at intermediate times can be deduced from Mather's shortening Lemma (see Chapter 8 of [6] and references therein) that reads as follows: if $\gamma, \hat{\gamma} : [0, 1] \to X$ are two constant speed geodesics so that

$$d(\gamma_0, \gamma_1)^2 + d(\hat{\gamma}_0, \hat{\gamma}_1)^2 \le d(\hat{\gamma}_0, \gamma_1)^2 + d(\gamma_0, \hat{\gamma}_1)^2,$$

then for any $t_0 \in (0, 1)$

$$\sup_{t\in[0,1]} d(\gamma_t, \hat{\gamma}_t) \le \frac{C}{t_0(1 - t_0)} d(\gamma_{t_0}, \hat{\gamma}_{t_0}),$$

with C depending on the lengths $L(\gamma)$ and $L(\hat{\gamma})$. Clearly Mather's shortening Lemma implies Lipschitz regularity of the optimal transport map from any intermediate time $t_0 \in (0, 1)$ to $t = 1$ and therefore also of the length map L_t.

Dropping the smoothness of the ambient space X, the correct framework where to look for a generalization of Mather's shortening Lemma is given by Alexandrov spaces, where a lower bound on a generalized notion of sectional curvature holds. If (X, d) is an Alexandrov space with curvature bounded from below by $K \ge 0$, then Mather's shortening Lemma, in the version stated few lines above, has been proven by Alessio Figalli, see again Chapter 8 of [6] and references therein. If $K < 0$ the problem has not been solved and the question is still open.

Not using any comparison geometry and assuming (X, d) to be only a length space there is no hope to prove analogous estimates and the best you can hope is to control the length map. A 1/2-Hölder estimate holds and the result, called rough non-smooth shortening lemma, is due to Cedric Villani (see Theorem 8.22 of of [6]): if $\gamma, \hat{\gamma} : [0, 1] \to X$ are two constant speed geodesics so that

$$d(\gamma_0, \gamma_1)^2 + d(\hat{\gamma}_0, \hat{\gamma}_1)^2 \le d(\hat{\gamma}_0, \gamma_1)^2 + d(\gamma_0, \hat{\gamma}_1)^2,$$

MH gratefully acknowledges funding by SFB 611.

then for any $t_0 \in (0,1)$

$$|L(\gamma) - L(\hat{\gamma})| \leq \frac{C}{\sqrt{t_0(1-t_0)}} \sqrt{d(\gamma_{t_0}, \hat{\gamma}_{t_0})},$$

where C bounds the square root of the lengths $L(\gamma)$ and $L(\hat{\gamma})$.

In this paper we assume (X, d, m) to be a geodesic non-branching metric measure space and we prove that if the supports of the geodesic $\{\mu_t\}_{t\in[0,1]}$ have a self-intersection property, then

- the local Lipschitz constant of L_t is bounded by $C/(t_0(1-t_0))$;
- neglecting a set of geodesics of measure arbitrarily small, but not zero, the map L_t is locally Lipschitz with Lipschitz constant bounded by $C/(t_0(1-t_0))$,

improving for nice geodesics the rough non-smooth shortening lemma. For the definition of local Lipschitz constant see later. It is worth noticing that this self-intersection property is always satisfied if the ambient space verifies $\mathsf{CD}(K,\infty)$ or $\mathsf{CD}(K,N)$ or $\mathsf{MCP}(K,N)$.

We now describe the setting and the result in more detail.

1.1. **Setting and main result.** The following notations and definitions will be fixed throughout this note. Let (X, d, m) be a non-branching metric measure space. $\mathcal{P}(X)$ denotes the set of probability measures over the set X and $\mathcal{P}_2(X)$ denotes the set of probability measures with finite second moment over the metric space (X, d). The L^2-Wasserstein distance will be denoted by d_W.

We fix $\mu_0, \mu_1 \in \mathcal{P}_2(X)$. By classical optimal transport theory (e.g. Thm 5.10 in [6]), there exists a d^2-concave function φ such that the couple (φ, φ^d) is a couple of optimal Kantorovich potentials, namely any transference plan $\pi \in \mathcal{P}(X \times X)$ is optimal if and only if it gives measure one to the set

$$\Gamma := \{(x, y) \in X \times X : \varphi(x) + \varphi^d(y) = \frac{1}{2} d^2(x, y)\}.$$

We also fix a dynamical optimal transference plan $\boldsymbol{\gamma} \in \mathcal{P}(\mathcal{G}(X))$, i.e. if e_t is the evaluation map at $t \in [0,1]$, $e_t(\gamma) := \gamma_t$, then $\{(e_t)_\sharp \boldsymbol{\gamma}\}_{t\in[0,1]}$ is a geodesic in the L^2-Wasserstein space connecting μ_0 and μ_1 with $(e_0, e_1)_\sharp \boldsymbol{\gamma}$ being an optimal transference plan. The existence is ensured by Thm 7.21 in [6]. Let $G \subset \mathcal{G}(M)$ be the support of $\boldsymbol{\gamma}$, so $\boldsymbol{\gamma}(G) = 1$. Without loss of generality, we can assume G to be compact (using inner regularity of the Radon measure $\boldsymbol{\gamma}$).

We will write $\mu_t = (e_t)_\sharp \boldsymbol{\gamma}$ and for a geodesic γ we set $L(\gamma) = d(\gamma_0, \gamma_1)$. Fix $t \in (0,1)$. From the non-branching property of the space together with the d^2-monotonicity, the map

$$e_t(G) \ni x \mapsto e_t^{-1}(x) \in G$$

assigning to each point x of $e_t(G)$ the unique optimal geodesic passing through x at time t, is well defined. We will assume that a.e. geodesic stays in $e_t(G)$ for a positive amount of time with Lebesgue density 1 in t (see Assumption 1). The first main result of this note is

Theorem 1.1. *Provided Assumption 1 holds the length map*

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

has local Lipschitz constant bounded by $C/t(1-t)$, where C depends only on the length $L_t(x)$ and it is uniform if $L_t(x)$ is uniformly bounded.

Recall that for a map $f : X \to \mathbb{R}$ the *local Lipschitz constant* is defined by

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

The other main result of this note is

Theorem 1.2. *Provided Assumption 1 holds, for every $\eta > 0$ there exists a compact set $K \subset G$ so that*

- $\boldsymbol{\gamma}(\mathcal{G}(X) \setminus K) \leq \eta$;
- *for every $x \in e_t(K)$ there exists $U(x)$ open neighborhood of x so that*

$$|L_t(z) - L_t(w)| \leq \frac{2C}{t(1-t)}, \qquad \forall z, w \in U(x) \cap e_t(K),$$

where C depends on K and on the length $L_t(x)$.

The precise dependance of C on K and on $L_t(x)$ will be clarified in Theorem 5.1.

It is worth noticing that, as proved in Proposition 2.6, Assumption 1 holds in $\mathsf{CD}(K,N)$ and in $\mathsf{CD}(K,\infty)$ spaces. For a definition of these spaces we refer to [4, 5] or [6].

We start by recalling some facts about the time evolution of Kantorovich potentials and setting the framework of this note. The proof of the first main theorem is done in three steps. First we decompose the space into evolving level sets of the Kantorovich potential of "codimension" one which are nearly "orthogonal" to the geodesics used by the optimal transport plan. Then we prove the Lipschitz result on these level sets and on the geodesics separately. Finally, we will put these estimates together to get the desired result. In the last section we will adopt these ideas to prove the second main result.

2. Framework

Since we will consider Kantorovich potentials also for optimal transport between μ_0 and μ_t, and between μ_t and μ_1, we recall some results about their evolution in time. What follows is contained in [2].

Definition 2.1. Let $\varphi : X \to \mathbb{R} \cup \{\pm\infty\}$. Its *$d^2$-transform* $\varphi^d : X \to \mathbb{R} \cup \{-\infty\}$ is defined by

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

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 $\varphi = v^d$.

A *d^2-concave* function φ such that (φ, φ^d) 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.

The d^2-convex potentials evolve according to the Hopf-Lax evolution semigroup H_t^s via the following formula:

$$H_t^s(\psi)(x) := \begin{cases} \inf_{y \in X} \frac{1}{2} \frac{d^2(x,y)}{s-t} + \psi(y), & \text{if } t < s, \\ \psi(x), & \text{if } t = s, \\ \sup_{y \in X} \psi(y) - \frac{1}{2} \frac{d^2(x,y)}{t-s}, & \text{if } t > s. \end{cases} \tag{2.1}$$

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 [6] (Theorem 7.36) but here we report a slightly different version.

Theorem 2.2 ([1], 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 $c^{0,1}$-concave Kantorovich potential for (μ_1, μ_0), then $H_1^t(\phi)$ is a $c^{t,s}$-convex Kantorovich potential for (μ_t, μ_s), for any $t < s$.

This implies the following

Corollary 2.3. *Let φ be a d^2-concave Kantorovich potential for (μ_0, μ_1). Let $\varphi_t := -H_1^t(\varphi^d)$ be the $c^{t,1}$-concave Kantorovich potential for (μ_t, μ_1) and analogously let $\varphi_t^d := H_0^t(-\varphi)$ the $c^{0,t}$-concave Kantorovich potential for (μ_t, μ_0). Then:*

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

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

Since

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

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

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

Taking the infimum the claim follows. □

Also related to slopes of Kantorovich potentials is the following construction taken from [2]. Let $f : X \to \mathbb{R}$ be a Lipschitz function and consider the following maps

$$D^+(x,t) := \max\Big\{d(x,y) : y \in \operatorname{argmin}\Big\{y \mapsto f(y) + \frac{d^2(x,y)}{2t}\Big\}\Big\},$$
$$D^-(x,t) := \min\Big\{d(x,y) : y \in \operatorname{argmin}\Big\{y \mapsto f(y) + \frac{d^2(x,y)}{2t}\Big\}\Big\}.$$

We will use these maps with $f = -\varphi^d$ and $f = -\varphi$ in the proof of Proposition 3.4. Then a monotonicity result holds true.

Lemma 2.4 ([2], Proposition 3.1)**.** *For all $x \in X$ it holds,*

$$D^+(x,t) \leq D^-(x,s), \qquad 0 \leq t < s.$$

As a consequence, $D^+(x,\cdot)$ and $D^-(x,\cdot)$ are both nondecreasing, and they coincide with at most countably many exceptions in $[0,\infty)$.

We now prove the metric Brenier's Theorem for φ_t, with $t \in (0,1)$ without any additional assumption.

Proposition 2.5. *For every $t \in (0,1)$ and for every $\gamma \in G$*

$$\lim_{s\to 0} \frac{\varphi_t(\gamma_t) - \varphi_t(\gamma_{t+s})}{d(\gamma_t,\gamma_{t+s})} = d(\gamma_0,\gamma_1) = |D\varphi_t|(\gamma_t), \tag{2.2}$$

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

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

$$\operatorname{argmin}\Big\{y \mapsto \frac{d^2(\gamma_t,y)}{2(1-t)} - \varphi^d(y)\Big\},$$

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

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

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

$$\begin{aligned}
\frac{1}{2}d^2(\gamma_0,z) &\geq \varphi(\gamma_0) + \varphi^d(z) \\
&\geq \varphi(\gamma_0) - \varphi_t(\gamma_t) + \varphi_t(\gamma_t) + \varphi^d(z) \\
&= \frac{1}{2}\left(\frac{d^2(\gamma_0,\gamma_t)}{t} + \frac{d^2(\gamma_t,z)}{1-t}\right) \\
&\geq \frac{1}{2}d^2(\gamma_0,z),
\end{aligned}$$

then necessarily $d(\gamma_0,\gamma_1) = d(\gamma_0,\gamma_t) + d(\gamma_t, z)$. 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 [2]. 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

$$\begin{aligned}
\varphi_t(\gamma_t) - \varphi_t(\gamma_{t+s}) &= \varphi_t(\gamma_t) + \varphi^d(\gamma_1) - \varphi^d(\gamma_1) - \varphi_t(\gamma_{t+s}) \\
&\geq \frac{1}{2(1-t)}\left(d^2(\gamma_t,\gamma_1) - d^2(\gamma_{t+s},\gamma_1)\right) \\
&= \frac{1}{2(1-t)}\left(d(\gamma_t,\gamma_1) - d(\gamma_{t+s},\gamma_1)\right)\left(d(\gamma_t,\gamma_1) + d(\gamma_{t+s},\gamma_1)\right) \\
&= \frac{1}{2(1-t)} d(\gamma_t,\gamma_{t+s})\left(d(\gamma_t,\gamma_1) + d(\gamma_{t+s},\gamma_1)\right). \qquad (2.3)
\end{aligned}$$

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. □

The assumption that we make throughout the paper is the following.

Assumption 1. For every $t\in(0,1)$, for $\boldsymbol{\gamma}$-a.e. $\gamma\in G$ the point t is a point of Lebesgue density 1 for the set $I_t(\gamma) := \{\tau\in(0,1) : \gamma_\tau \in e_t(G)\}$. Namely for every $t\in(0,1)$

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

Proposition 2.6. *Let (X,d,m) be a non-branching metric measure space verifying $\mathsf{CD}(K,N)$, for $N\in\mathbb{N}\cup\{\infty\}$. If $N=\infty$ and $\mu_0,\mu_1 \ll m$ both with finite entropy, then for any dynamical optimal transference plan $\boldsymbol{\gamma}$, Assumption 1 holds. If $N<\infty$ and $\mu_0\ll m$, then for any $\mu_1\in\mathcal{P}_2(X)$ and any dynamical optimal transference plan $\boldsymbol{\gamma}$, Assumption 1 holds.*

Proof. Assume by contradiction Assumption 1 doesn't hold. Therefore there exists a set $H\subset G$, with $\boldsymbol{\gamma}(H)>0$ so that for all $\gamma\in H$ we have two possibilities: or

$$0 \leq \liminf_{\varepsilon\to 0} \frac{\mathcal{L}^1\big(I_t(\gamma)\cap(t-\varepsilon,t+\varepsilon)\big)}{2\varepsilon} < \limsup_{\varepsilon\to 0} \frac{\mathcal{L}^1\big(I_t(\gamma)\cap(t-\varepsilon,t+\varepsilon)\big)}{2\varepsilon} \leq 1.$$

either the limit exists but is not one:

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

In the first part of the proof we show that in the first case, the lim sup must be equal to 1, neglecting a set of zero $\boldsymbol{\gamma}$-measure. The same argument excludes immediately the second case.

Step 1. Suppose by contradiction the existence of a set $H\subset G$, with $\boldsymbol{\gamma}(H)>0$ so that for all $\gamma\in H$

$$\limsup_{\varepsilon\to 0} \frac{\mathcal{L}^1\big(I_t(\gamma)\cap(t-\varepsilon,t+\varepsilon)\big)}{2\varepsilon} < 1.$$

and therefore possibly restricting H,

$$\liminf_{\varepsilon\to 0} \frac{\mathcal{L}^1\big(I_t^c(\gamma)\cap(t-\varepsilon,t+\varepsilon)\big)}{2\varepsilon} > \alpha,$$

for some $\alpha>0$. Let

$$E := \{(\gamma,s)\in H\times(0,1) : t+s\in I_t(\gamma)^c\} = \{(\gamma,s)\in H\times(0,1) : d(\gamma_{t+s}, e_t(G))>0\}.$$

Then by Fubini's Theorem

$$\boldsymbol{\gamma}\otimes\mathcal{L}^1(E)=\int_{(0,1)}\mathcal{L}^1(E(\gamma))\boldsymbol{\gamma}(d\gamma),\qquad E(\gamma):=P_2\Big(E\cap\big(\{\gamma\}\times(0,1)\big)\Big),$$

where P_i denotes the projection map on the i-th component, for $i=1,2$. From Fatoú's Lemma

$$\liminf_{\varepsilon\to 0}\frac{\boldsymbol{\gamma}\otimes\mathcal{L}^1\Big(E\cap\big(H\times(t-\varepsilon,t+\varepsilon)\big)\Big)}{2\varepsilon}\geq\alpha\boldsymbol{\gamma}(H),$$

therefore

$$\liminf_{\varepsilon\to 0}\frac{1}{2\varepsilon}\int_{(t-\varepsilon,t+\varepsilon)}\boldsymbol{\gamma}(E(\tau))\mathcal{L}^1(ds)\geq\alpha\boldsymbol{\gamma}(H),\qquad E(\tau):=P_1\Big(E\cap\big(H\times\{\tau\}\big)\Big).$$

So there must be a sequence of $\{s_n\}_{n\in\mathbb{N}}$ converging to 0 so that $\boldsymbol{\gamma}(E(t+s_n))\geq C$, for some $C>0$. Then, since $e_{t+s_n}(G)$ converges to $e_t(G)$ in Hausdorff topology as s_n goes to 0, we have

$$m(e_t(G)^\varepsilon)\geq m(e_t(G)\cup e_{t+s_n}(E(t+s_n)))\geq m(e_t(G))+m(e_{t+s_n}(E(t+s_n))),$$

where $e_t(G)^\varepsilon:=\{z\in X: d(z,e_t(G))\leq\varepsilon\}$. Since $\varrho_t\leq M$ (from $\mathsf{CD}(K,N)$ or $\mathsf{CD}(K,\infty)$, provided both μ_0 and μ_1 are absolute continuous with respect to m) on $e_t(G)$ for all $t\in(0,1)$, it follows that $m(e_{t+s_n}(E(t+s_n)))$ remains uniformly strictly positive as s_n goes to 0. Since

$$m(e_t(G))\geq\limsup_{\varepsilon\to 0}m(e_t(G)^\varepsilon),$$

we have a contradiction. Hence we have shown that there exists H, $\boldsymbol{\gamma}$-negligible, so that

$$\limsup_{\varepsilon\to 0}\frac{\mathcal{L}^1\big(I_t(\gamma)\cap(t-\varepsilon,t+\varepsilon)\big)}{2\varepsilon}=1,$$

for all $\gamma\in G\setminus H$. Using the same reasoning we can also prove a stronger statement: for any sequence $\varepsilon_n\to 0$ there exists H, $\boldsymbol{\gamma}$-negligible and depending on the sequence ε_n, so that

$$\limsup_{n\to\infty}\frac{\mathcal{L}^1\big(I_t(\gamma)\cap(t-\varepsilon_n,t+\varepsilon_n)\big)}{2\varepsilon_n}=1,$$

for all $\gamma\in G\setminus H$. We now show that L^1-convergence holds.

Step 2. Consider any sequence ε_n converging to 0. Then from the equality

$$1-\frac{\mathcal{L}^1(I_t^c(\gamma)\cap(t-\varepsilon_n,t+\varepsilon_n))}{2\varepsilon_n}=\frac{\mathcal{L}^1(I_t(\gamma)\cap(t-\varepsilon_n,t+\varepsilon_n))}{2\varepsilon_n},$$

integrating over any set $K\subset G$ we get

$$\begin{aligned}\boldsymbol{\gamma}(K)-\limsup_{n\to\infty}&\int_K\frac{\mathcal{L}^1(I_t^c(\gamma)\cap(t-\varepsilon_n,t+\varepsilon_n))}{2\varepsilon_n}\boldsymbol{\gamma}(d\gamma)\\ &=\liminf_{n\to\infty}\int_K\frac{\mathcal{L}^1(I_t(\gamma)\cap(t-\varepsilon_n,t+\varepsilon_n))}{2\varepsilon_n}\boldsymbol{\gamma}(d\gamma).\end{aligned}$$

Hence, if there exists $K\subset G$ with $\boldsymbol{\gamma}(K)>0$ so that

$$\liminf_{n\to\infty}\int_K\frac{\mathcal{L}^1(I_t(\gamma)\cap(t-\varepsilon_n,t+\varepsilon_n))}{2\varepsilon_n}\boldsymbol{\gamma}(d\gamma)<\boldsymbol{\gamma}(K),$$

then

$$\limsup_{n\to\infty}\int_K\frac{\mathcal{L}^1(I_t^c(\gamma)\cap(t-\varepsilon_n,t+\varepsilon_n))}{2\varepsilon_n}\boldsymbol{\gamma}(d\gamma)>0.$$

Using Fubini's Theorem as in *Step 1.* we could find a sequence of slices of E, say $E(t+s_n)$, with uniformly positive $\boldsymbol{\gamma}$-measure and therefore we would get a contradiction. Reasoning in the same manner for the $\limsup$ we get that for any $K\subset G$,

$$\liminf_{n\to\infty}\int_K\frac{\mathcal{L}^1(I_t(\gamma)\cap(t-\varepsilon_n,t+\varepsilon_n))}{2\varepsilon_n}\boldsymbol{\gamma}(d\gamma)=\boldsymbol{\gamma}(K).$$

Since ε_n was arbitrarily chosen

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

We now show that the convergence is indeed point wise.

Step 3a. Suppose again by contradiction the existence of a set H, with $\boldsymbol{\gamma}(H) > 0$, so that for every $\gamma \in H$ there exists a sequence $\varepsilon_n(\gamma)$, converging to 0 as $n \to \infty$, such that

$$\liminf_{n\to\infty} \frac{\mathcal{L}^1(I_t(\gamma) \cap (t - \varepsilon_n(\gamma), t + \varepsilon_n(\gamma)))}{2\varepsilon_n(\gamma)} < 1.$$

Possibly restricting H, still of positive $\boldsymbol{\gamma}$-measure, we can assume

$$\limsup_{n\to\infty} \frac{\mathcal{L}^1(I_t^c(\gamma) \cap (t - \varepsilon_n(\gamma), t + \varepsilon_n(\gamma)))}{2\varepsilon_n(\gamma)} \geq c > 0, \qquad \forall \gamma \in H.$$

Consider the set $\mathcal{A} := \{\varepsilon_n(\gamma) : n \in \mathbb{N}, \gamma \in H\} \subset (0,1)$ and the family of open balls

$$\mathcal{F} := \{B_{\varepsilon_n^2(\gamma)}(\varepsilon_n(\gamma)) : n \in \mathbb{N}, \gamma \in H\}.$$

Clearly $\mathcal{F}$ is an open cover of $\mathcal{A}$. By Lindelöf's Theorem there exists a countable sub-cover of $\mathcal{A}$:

$$\mathcal{A} \subset \bigcup_{B \in \mathcal{F}_0} B, \qquad \mathcal{F}_0 := \{B_{\varepsilon^2_{n_i}(\gamma_i)}(\varepsilon_{n_j}(\gamma_i)) : i \in \mathbb{N}\}.$$

Hence we have a sequence $\{\alpha_i\}_{i\in\mathbb{N}}$ so that for any $\gamma \in H$ and any $n \in \mathbb{N}$ there exists $i_{n,\gamma} \in \mathbb{N}$ such that

$$|\varepsilon_n(\gamma) - \alpha_{i_{n,\gamma}}| \leq \left(\alpha_{i_{n,\gamma}}\right)^2.$$

Up to a reindexing we can assume $\alpha_{i+i} \leq \alpha_i$. Hence $\lim_i \alpha_i = 0$. We can consider the following covering of H: for $i, n \in \mathbb{N}$

$$\hat{\Lambda}_{i,n} := \{\gamma \in H : |\varepsilon_n(\gamma) - \alpha_i| \leq \alpha_i^2\}.$$

Clearly it could happen that for $i \neq j$ the set $\hat{\Lambda}_{i,n}$ and $\hat{\Lambda}_{j,n}$ are not disjoint. Using a standard argument we can find subsets $\Lambda_{i,n} \subset \hat{\Lambda}_{i,n}$ so that

$$\bigcup_{i=1}^{\infty} \Lambda_{i,n} = H, \qquad \Lambda_{i,n} \cap \Lambda_{j,n} = \emptyset, \quad i \neq j.$$

For fixed $n \in \mathbb{N}$, we look at the map $H \ni \gamma \mapsto \varepsilon_n(\gamma)$. Since the sequences $\varepsilon_n(\gamma)$ are just sequences so that

$$\limsup_{n\to\infty} \frac{\mathcal{L}^1(I_t^c(\gamma) \cap (t - \varepsilon_n(\gamma), t + \varepsilon_n(\gamma)))}{2\varepsilon_n(\gamma)} \geq c$$

it could be proved that the aforementioned map can be assumed Borel. Indeed if

$$H \times (0,1) \ni (\gamma, \varepsilon) \mapsto f(\gamma, \varepsilon) := \frac{\mathcal{L}^1(I_t^c(\gamma) \cap (t - \varepsilon, t + \varepsilon))}{2\varepsilon},$$

then f is continuous in ε and lower-semicontinuous in γ. Hence the set

$$K := \{(\gamma, \varepsilon) \in H \times (0,1) : f(\gamma, \varepsilon) \geq c, \varepsilon \leq \frac{1}{n}\}$$

is Borel and $K(\gamma) = P_1(K \cap (\{\gamma\} \times (0,1))) \neq \emptyset$, for all $\gamma \in H$. Using a standard selection theorem for each $n \in \mathbb{N}$ we can obtain sequences $\varepsilon_n(\gamma) \to 0$ as $n \to \infty$ so that for all $n \in \mathbb{N}$

$$\frac{\mathcal{L}^1(I_t^c(\gamma) \cap (t - \varepsilon_n(\gamma), t + \varepsilon_n(\gamma)))}{2\varepsilon_n(\gamma)} \geq c, \qquad \forall \gamma \in H,$$

and for fixed n the map $\gamma \mapsto \varepsilon_n(\gamma)$ is $\boldsymbol{\gamma}$-measurable on H.

Then thanks to Egorov's Theorem, we can find a subset of H, still of positive measure, so that $\varepsilon_n(\gamma) \to 0$, uniformly in γ. We denote this set still with H. Hence for any $\delta > 0$ there exists n_δ so that for all $n \geq n_\delta$

$$\varepsilon_n(\gamma) \leq \delta, \qquad \forall \gamma \in H.$$

Accordingly let i_δ be the smallest natural number so that

$$\alpha_{i_\delta} - \alpha_{i_\delta}^2 \leq \delta.$$

Hence, thanks to the monotonicity of α_i, for all $n \geq n_\delta$ the set $\Lambda_{i,n} = \emptyset$, whenever $i < i_\delta$. Letting $\delta \to 0$, being the sequence α_i not definitely equals to 0, it holds $i_\delta \to \infty$. This implies the existence of a sequence $\{i_n\}$, with $i_n \to \infty$ as $n \to \infty$ so that

$$H = \bigcup_{i=1}^{\infty} \Lambda_{i,n} = \bigcup_{i=i_n}^{\infty} \Lambda_{i,n}.$$

Step 3b. Now for fixed n

$$\begin{aligned}
&\int_H \left(\frac{\mathcal{L}^1(I_t^c(\gamma) \cap (t - \varepsilon_n(\gamma), t + \varepsilon_n(\gamma)))}{2\varepsilon_n(\gamma)} \right) \boldsymbol{\gamma}(d\gamma) \\
&= \sum_{i=i_n}^{\infty} \int_{\Lambda_{i,n}} \left(\frac{\mathcal{L}^1(I_t^c(\gamma) \cap (t - \varepsilon_n(\gamma), t + \varepsilon_n(\gamma)))}{2\varepsilon_n(\gamma)} \right) \boldsymbol{\gamma}(d\gamma) \\
&\leq \sum_{i=i_n}^{\infty} \int_{\Lambda_{i,n}} \left(\frac{\mathcal{L}^1(I_t^c(\gamma) \cap (t - (\alpha_i + \alpha_i^2), t + (\alpha_i + \alpha_i^2)))}{2(\alpha_i - \alpha_i^2)} \right) \boldsymbol{\gamma}(d\gamma) \\
&\leq \sum_{i=i_n}^{\infty} \boldsymbol{\gamma}(\Lambda_{i,n}) \left\| \frac{\mathcal{L}^1(I_t^c(\cdot) \cap (t - (\alpha_i + \alpha_i^2), t + (\alpha_i + \alpha_i^2)))}{2(\alpha_i - \alpha_i^2)} \right\|_{L^1(\boldsymbol{\gamma})} \\
&\leq \boldsymbol{\gamma}(H) \sup \left\{ \left\| \frac{\mathcal{L}^1(I_t^c(\cdot) \cap (t - (\alpha_i + \alpha_i^2), t + (\alpha_i + \alpha_i^2)))}{2(\alpha_i - \alpha_i^2)} \right\|_{L^1(\boldsymbol{\gamma})} : i \geq i_n \right\}.
\end{aligned} \tag{2.4}$$

By construction we have

$$c\boldsymbol{\gamma}(H) \leq \limsup_{n\to\infty} \int_H \left(\frac{\mathcal{L}^1(I_t^c(\gamma) \cap (t - \varepsilon_n(\gamma), t + \varepsilon_n(\gamma)))}{2\varepsilon_n(\gamma)} \right) \boldsymbol{\gamma}(d\gamma),$$

hence from (2.4)

$$\begin{aligned}
0 < c \leq &\liminf_{n\to\infty} \sup \left\{ \left\| \frac{\mathcal{L}^1(I_t^c(\cdot) \cap (t - (\alpha_i + \alpha_i^2), t + (\alpha_i + \alpha_i^2)))}{2(\alpha_i - \alpha_i^2)} \right\|_{L^1(\boldsymbol{\gamma})} : i \geq i_n \right\} \\
\leq &\limsup_{i\to\infty} \left\| \frac{\mathcal{L}^1(I_t^c(\cdot) \cap (t - (\alpha_i + \alpha_i^2), t + (\alpha_i + \alpha_i^2)))}{2(\alpha_i - \alpha_i^2)} \right\|_{L^1(\boldsymbol{\gamma})}.
\end{aligned} \tag{2.5}$$

In *Step 2.* we have already proved the L^1 convergence to 0 of the sequence of functions

$$\frac{\mathcal{L}^1(I_t^c(\cdot) \cap (t - (\alpha_i + \alpha_i^2), t + (\alpha_i + \alpha_i^2)))}{2(\alpha_i + \alpha_i^2)}.$$

Since $(\alpha_i + \alpha_i^2)/(\alpha_i - \alpha_i^2) \to 1$ the same convergence holds for the last right hand side of (2.5). Since $c > 0$ this is a contradiction and the claim is proved. □

Remark 2.7. As the last proof shows Assumption 1 is also satisfied if we replace in Proposition 2.6 the $\mathsf{CD}(K,N)$ condition by the slightly weaker measure contraction property $\mathsf{MCP}(K,N)$ (for a definition see [5] and [3]).

The following is a trivial consequence of Assumption 1 and Egorov's Theorem.

Corollary 2.8. *Fix $t \in (0,1)$ and let Assumption 1 holds. For every $\eta > 0$ there exists a compact set $K \subset G$ so that*

- *$\boldsymbol{\gamma}(G \setminus K) \leq \eta$;*
- *the limit $(2\varepsilon)^{-1}\mathcal{L}^1(\{I_t(\gamma) \cap (t-\varepsilon, t+\varepsilon)\})$ converges to 1 point wise as $\varepsilon \to 0$, uniformly in $\gamma \in K$.*

3. Split Regularity

Fix $t \in (0,1)$. In the first part of this section we prove the regularity property of the length map $L \circ e_t^{-1}$ restricted to two families of "orthogonal" sets. In order to have a shorter notation we set $L_t = L \circ e_t^{-1}$.

Consider the following sets depending on $a \in \mathbb{R}$

$$K_a := \{\gamma_t : \varphi(\gamma_0) = a\}, \qquad K_a^d := \{\gamma_t : \varphi^d(\gamma_1) = a\} \tag{3.1}$$

Proposition 3.1. *For every $a \in \varphi(\mathrm{supp}(\mu_0))$, the map L_t restricted to K_a is Lipschitz with Lipschitz constant equal to* 1*. The same holds replacing K_a with K_a^d and $a \in \varphi(\mathrm{supp}(\mu_0))$ with $a \in \varphi^d(\mathrm{supp}(\mu_1))$.*

Proof. Fix $a \in \varphi(\mathrm{supp}(\mu_0))$ and denote with G_a the set of geodesics in G leaving from $\varphi^{-1}(a)$, $G_a := G \cap e_0^{-1}(\varphi^{-1}(a))$. Then for any $\gamma \in G_a$ and any $z \in e_0(G_a)$

$$\frac{d^2(\gamma_0,\gamma_t)}{2t} = \varphi(\gamma_0) + \varphi_t^d(\gamma_t) = \varphi(z) + \varphi_t^d(\gamma_t) \le \frac{d^2(z,\gamma_t)}{2t}.$$

Hence for any $x \in K_a$,

$$L_t(x) = 1/t \cdot \min\big\{d(z,x) : z \in e_0(G_a)\big\},$$

therefore the length map L_t restricted to K_a is 1-Lipschitz. The proof for L_t restricted to K_a^d works the same. □

Remark 3.2. Note that the previous proposition also holds for $t = 0$.

With the same reasoning we obtain the following property.

Proposition 3.3. *For every $a \in \mathbb{R}$, the set $\Gamma_a := \{(\gamma_0,\gamma_1) : \gamma \in G, \varphi(\gamma_0) = a\}$ is d-cyclically monotone.*

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

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

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

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

and the claim follows. □

We now prove regularity of L_t in a direction transversal to K_a. Recall that $I_t(\gamma) := \{\tau \in [0,1] : \gamma_\tau \in e_t(G)\}$.

Proposition 3.4. *Let $\gamma \in G$ be any optimal geodesic satisfying Assumption 1. Then for every $t \in (0,1)$ and $s^-, s^+ > 0$ it holds that*

$$\frac{1-t-s^+}{1-t} \le \frac{L_t(\gamma_{t+s^+})}{L_t(\gamma_t)} \le \frac{t+s^+}{t}, \quad \textit{for } t+s^+ \in I_t(\gamma),$$

$$\frac{t-s^-}{t} \le \frac{L_t(\gamma_{t-s^-})}{L_t(\gamma_t)} \le \frac{1-t+s^-}{1-t}, \quad \textit{for } t-s^- \in I_t(\gamma). \tag{3.2}$$

Proof. Step 1. First consider the case $t+s^+$, so fix $t \in (0,1)$ and take s^+ such that $t+s^+ < 1$. For ease of notation we just write $s^+ = s$. Note that

$$L_t(\gamma_{t+s}) = \frac{1}{1-t}d(\gamma_{t+s},\gamma_1^s),$$

where $\gamma^s = e_t^{-1}(\gamma_{t+s})$ and

$$d(\gamma_{t+s},\gamma_1^s) = D(\gamma_{t+s}, 1-t),$$

where $D(x,t) = d(x,y)$ for $y \in \mathrm{argmin}\{y \mapsto -\varphi^d(y) + \frac{d^2(x,y)}{2t}\}$. The non-branching property of (X,d,m) implies that D is a well defined map. From Lemma 2.4, $D(x,\cdot)$ is nondecreasing, therefore

$$D(\gamma_{t+s}, 1-t) \ge D(\gamma_{t+s}, 1-(t+s)) = d(\gamma_{t+s},\gamma_1) = (1-(t+s))d(\gamma_0,\gamma_1).$$

Hence

$$L_t(\gamma_{t+s}) = \frac{1}{1-t}D(\gamma_{t+s},1-t) \ge \frac{1-(t+s)}{1-t}L_t(\gamma_t).$$

Step 2. Reasoning as before, we get the same inequality for $t - s^- > 0$. We write again $s^- = s$. Indeed since $L_t(\gamma_t) = D(\gamma_{t-s}, 1-t+s)$, from the monotonicity property of $D(x,\cdot)$ we have

$$L_t(\gamma_{t-s}) = \frac{1}{1-t} D(\gamma_{t-s}, 1-t) \le \frac{1}{1-t} D(\gamma_{t-s}, 1-(t-s)) = \frac{1}{1-t} d(\gamma_{t-s}, \gamma_1) = \frac{1-(t-s)}{1-t} L_t(\gamma_t).$$

To prove the remaining inequalities, just exchange the role of 0 and 1, i.e. use $\tilde{D}(x,s) = d(x,y)$ for $y \in \operatorname{argmin}\{y \mapsto -\varphi(y) + \frac{d^2(x,y)}{2s}\}$ together with the identity $(1-t)D(x,t) = t\tilde{D}(x, 1-t)$ for $t \in (0,1)$. □

4. General regularity

Throughout this section $t \in (0,1)$ is fixed. Consider the following real valued map defined for $x \in e_t(G)$:

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

Recall that by definition for $t+s \in I_t(\gamma)$ the map $L_t(\gamma_{t+s})$ is well defined. Hence, also the map $\Phi_t(\gamma_{t+s})$ is well defined for $t+s \in I_t(\gamma)$. Due to d^2-cyclical monotonicity, an equivalent definition is $\Phi_t(\gamma_t) := \varphi_0(\gamma_0)$ (see also Corollary 2.3).

Lemma 4.1. *The map Φ_t is continuous.*

Proof. For $t = 0$, $\Phi_0 = \varphi$ and there is nothing to prove. So assume $t \in (0,1)$. It is sufficient to prove that $e_t^{-1} : e_t(G) \to G$ is a continuous map. Note that e_t^{-1} is well defined again by d^2-cyclical monotonicity. So let $\{x_n\}_{n\in\mathbb{N}} \subset e_t(G)$ be a sequence converging to $x \in e_t(G)$ as $n \nearrow \infty$. If $\gamma^n = e_t^{-1}(x_n)$, then by compactness of G, there exists a subsequence $\{n_k\}_{k\in\mathbb{N}}$ such that γ^{n_k} will converge to $\hat{\gamma} \in G$. Since $\gamma_t^n = x_n$, necessarily $\hat{\gamma}_t = x$. By d^2-cyclical monotonicity, there is only one possible geodesic $\gamma \in G$ such that $\gamma_t = x$, hence the whole sequence $\{\gamma^n\}_{n\in\mathbb{N}}$ is converging to $e_t^{-1}(x)$. □

Lemma 4.2. *Let $\gamma \in G$ be fixed so that Assumption 1 holds. Then*

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

provided $t - s \in I_t(\gamma)$ for the first inequality, and $t + s \in I_t(\gamma)$ for the second one.

Proof. Step 1. We first prove the first inequality. Suppose by contradiction the existence of $s \in (0,1)$ such that $t - s \in I_t(\gamma)$ and $\Phi_t(\gamma_{t-s}) \le \Phi_t(\gamma_t)$. From Proposition 3.3, 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

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

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

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

and since $\hat{\gamma}_t = \gamma_{t-s}$ and $\hat{\gamma} \in G$, this is in contradiction with d^2-cyclical monotonicity.

Step 2. We now prove the other inequality. Suppose by contradiction the existence of $s \in (0,1)$ such that $t + s \in I_t(\gamma)$ and $\Phi_t(\gamma_{t+s}) \ge \Phi_t(\gamma_t)$. From Proposition 3.3, 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

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

So we can deduce

$$\frac{1}{2(t+s)} d^2(\gamma_0, \gamma_{t+s}) = \varphi(\gamma_0) + \varphi_{t+s}^d(\gamma_{t+s}) < \varphi(\hat{\gamma}_0) + \varphi_{t+s}^d(\gamma_{t+s}) \le \frac{1}{2(t+s)} d^2(\hat{\gamma}_0, \gamma_{t+s}),$$

and therefore $d(\gamma_0,\gamma_{t+s}) \le d(\hat\gamma_0,\gamma_{t+s})$. Hence

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

where the last inequality comes from $\hat\gamma_t = \gamma_{t+s}$. Since $\hat\gamma,\gamma \in G$, this is in contradiction with d^2-cyclical monotonicity. □

Lemma 4.3. *Fix $t \in (0,1)$. Then for $\boldsymbol{\gamma}$-a.e. $\gamma \in G$*

$$\limsup_{k\to\infty} \frac{d(\gamma_t,\gamma_{t+s_k^+})}{d(\gamma_t,\gamma_t^k)} \le \frac{1+t}{1-t},$$

for any sequence $\{\gamma^k\}_{k\in\mathbb{N}} \subset G$ with $\varphi(\gamma_0) > \varphi(\gamma_0^k)$ and $\Phi_t(\gamma_{t+s_k^+}) = \Phi_t(\gamma_t^k)$, where $\{s_k^+\}_{k\in\mathbb{N}}$ is any null-sequence of positive numbers so that $t+s_k^+ \in I_t(\gamma)$, and

$$\limsup_{k\to\infty} \frac{d(\gamma_t,\gamma_{t-s_k^-})}{d(\gamma_t,\gamma_t^k)} \le \frac{1+t}{1-t},$$

for any sequence $\{\gamma^k\}_{k\in\mathbb{N}} \subset G$ with $\varphi(\gamma_0) > \varphi(\gamma_0^k)$ and $\Phi_t(\gamma_{t+s_k^-}) = \Phi_t(\gamma_t^k)$, where $\{s_k^-\}_{k\in\mathbb{N}}$ is any null-sequence of positive numbers so that $t+s_k^- \in I_t(\gamma)$.

Proof. Let us start with the case $\varphi(\gamma_0) > \varphi(\gamma_0^k)$.

For $\gamma \in G$ satisfying Assumption 1 the following holds. From the hypothesis we have

$$\varphi_t(\gamma_{t+s_k^+}) + \frac{t}{2}L_t^2(\gamma_{t+s_k^+}) = \Phi_t(\gamma_{t+s_k^+}) = \varphi(\gamma_0^k) = \Phi_t(\gamma_t^k) = \varphi_t(\gamma_t^k) + \frac{t}{2}L_t^2(\gamma_t^k).$$

Hence using Proposition 3.1

$$\begin{aligned} \frac{\varphi_t(\gamma_t)-\varphi_t(\gamma_t^k)}{d(\gamma_t,\gamma_t^k)} &= \frac{\varphi_t(\gamma_t)-\varphi_t(\gamma_{t+s_k^+})}{d(\gamma_t,\gamma_{t+s_k^+})}\frac{d(\gamma_t,\gamma_{t+s_k^+})}{d(\gamma_t,\gamma_t^k)} + \frac{t}{2}\frac{L_t^2(\gamma_t^k)-L_t^2(\gamma_{t+s_k^+})}{d(\gamma_t,\gamma_t^k)} \\ &\ge \frac{\varphi_t(\gamma_t)-\varphi_t(\gamma_{t+s_k^+})}{d(\gamma_t,\gamma_{t+s_k^+})}\frac{d(\gamma_t,\gamma_{t+s_k^+})}{d(\gamma_t,\gamma_t^k)} - \frac{t}{2}(L_t(\gamma_t^k)+L_t(\gamma_{t+s_k^+}))\frac{d(\gamma_t^k,\gamma_{t+s_k^+})}{d(\gamma_t,\gamma_t^k)} \\ &\ge \frac{d(\gamma_t,\gamma_{t+s_k^+})}{d(\gamma_t,\gamma_t^k)}\left\{\frac{\varphi_t(\gamma_t)-\varphi_t(\gamma_{t+s_k^+})}{d(\gamma_t,\gamma_{t+s_k^+})} - \frac{t}{2}(L_t(\gamma_t^k)+L_t(\gamma_{t+s_k^+}))\right\} \\ &\quad - \frac{t}{2}(L_t(\gamma_t^k)+L_t(\gamma_{t+s_k^+}))\,, \end{aligned}$$

where we used the triangle inequality in the last step. From Proposition 2.5,

$$\lim_{k\to\infty}\frac{\varphi_t(\gamma_t)-\varphi_t(\gamma_{t+s_k^+})}{d(\gamma_t,\gamma_{t+s_k^+})} = L_t(\gamma_t).$$

Then taking the limsup it follows that

$$L_t(\gamma_t) \ge \limsup_{k\to\infty}\frac{d(\gamma_t,\gamma_{t+s_k^+})}{d(\gamma_t,\gamma_t^k)}L_t(\gamma_t)(1-t) - tL_t(\gamma_t)\,,$$

which directly implies the claim.

For the case $\varphi(\gamma_0) < \varphi(\gamma_0^k)$ we have to estimate

$$\frac{\varphi_t(\gamma_t^k)-\varphi_t(\gamma_t)}{d(\gamma_t,\gamma_t^k)}\,.$$

Using

$$\varphi_t(\gamma_{t-s_k^-}) + \frac{t}{2}L_t^2(\gamma_{t-s_k^-}) = \Phi_t(\gamma_{t-s_k^-}) = \varphi(\gamma_0^k) = \Phi_t(\gamma_t^k) = \varphi_t(\gamma_t^k) + \frac{t}{2}L_t^2(\gamma_t^k)\,,$$

the proof is completely the same as in the first case, proving the claim. □

We are now ready to prove the first main result of this note.

Theorem 4.4. *Fix $t \in (0,1)$ such that Assumption 1 holds true. Then the length map $L_t : e_t(G) \to \mathbb{R}$ has local Lipschitz constant bounded by $C/(t(1-t))$, that is for $\boldsymbol{\gamma}$-a.e. $\gamma \in G$*

$$\limsup_{k\to\infty} \frac{|L_t(\gamma_t^k) - L_t(\gamma_t)|}{d(\gamma_t^k, \gamma_t)} \leq \frac{C}{t(1-t)}$$

where $\{\gamma^k\} \subset G$ is converging to γ, C depends only on the length $L_t(\gamma_t)$ and it is uniform in G if $L_t(\gamma_t)$ is uniformly bounded.

Proof. Fix $t \in (0,1)$ and $\gamma \in G$ satisfying Assumption 1. Let $\{\gamma^k\}_{k\in\mathbb{N}} \subset G$ be converging to γ as $k \nearrow \infty$.

Step 1. Assume $\varphi(\gamma_0) > \varphi(\gamma_0^k)$. The function $s \mapsto \Phi_t(\gamma_{t+s})$ is defined just for $t+s \in I_t(\gamma)$. Since φ_t is Lipschitz, from Proposition 3.4 it follows that also the map $I_t(\gamma) \ni t+s \mapsto \Phi_t(\gamma_{t+s})$ satisfies (3.2).

Therefore we can find a continuous extension of Φ_t, say $\hat{\Phi}_t$, defined on an interval, say $I(\gamma)$ containing $I_t(\gamma)$, still verifying (3.2). Since $\Phi_t(\gamma_t) > \Phi_t(\gamma_{t+s})$, the extension $\hat{\Phi}_t$ can be taken strictly smaller than $\Phi_t(\gamma_t)$ as well. By continuity of $\hat{\Phi}_t$ for every $k \in \mathbb{N}$ sufficiently large there exists $s_k > 0$ such that

$$\Phi_t(\gamma_t^k) = \hat{\Phi}_t(\gamma_{t+s_k}).$$

Then it holds:

$$|\Phi_t(\gamma_t^k) - \Phi_t(\gamma_t)| = |\hat{\Phi}_t(\gamma_{t+s_k}) - \Phi_t(\gamma_t)| = |\hat{\Phi}_t(\gamma_{t+s_k}) - \hat{\Phi}_t(\gamma_t)| \leq \frac{C}{1-t} d(\gamma_t, \gamma_{t+s_k}),$$

where the last inequality follows from (3.2). For every $k \in \mathbb{N}$, consider the closest points to $t+s_k$ in $I_t(\gamma)$ from the left:

$$s_k^+ := \max\{s : t+s \in I_t(\gamma), t+s \leq t+s_k\}.$$

Since $I_t(\gamma)$ is a compact set, s_k^+ is well defined and $t+s_k^+ \in I_t(\gamma)$. Since t has density 1 in $I_t(\gamma)$, it follows that

$$\lim_{k\to\infty} \frac{d(\gamma_t, \gamma_{t+s_k^+})}{d(\gamma_t, \gamma_{t+s_k})} = 1,$$

and by construction $s_k^+ \leq s_k$: For every ε_k converging to zero, by Assumption 1

$$\lim_{k\to\infty} \frac{1}{\varepsilon_k} \mathcal{L}^1\Big(I_t(\gamma) \cap [t, t+\varepsilon_k]\Big) = 1.$$

Since $\mathcal{L}^1\Big(I_t(\gamma) \cap [t, t+s_k]\Big) \leq s_k^+$, taking $\varepsilon_k = s_k$

$$1 \leq \lim_{k\to\infty} \frac{s_k^+}{s_k} = \lim_{k\to\infty} \frac{d(\gamma_t, \gamma_{t+s_k^+})}{d(\gamma_t, \gamma_{t+s_k})} \leq 1.$$

Consequently $d(\gamma_{t+s_k^+}, \gamma_{t+s_k})/d(\gamma_t, \gamma_{t+s_k^+}) \to 0$ as $k \to \infty$.

Step 2. Put $\Phi_t(\gamma_{t+s_k^+}) =: \alpha_k$. Lemma 4.3 implies that if

$$\beta_k := \min\{d(\gamma_t, z) : \Phi_t(z) = \alpha_k\},$$

then

$$\limsup_{k\nearrow\infty} \frac{d(\gamma_t, \gamma_{t+s_k^+})}{\beta_k} \leq \frac{1+t}{1-t}.$$

Moreover since Φ_t is continuous,

$$\limsup_{k\nearrow\infty} \frac{\beta_k}{d(\gamma_t, \gamma_t^k)} \leq 1.$$

Than we can conclude as follows:

$$\begin{aligned} \frac{|\Phi_t(\gamma_t^k) - \Phi_t(\gamma_t)|}{d(\gamma_t, \gamma_t^k)} &\leq \frac{2C}{1-t}\left(\frac{d(\gamma_t, \gamma_{t+s_k^+})}{d(\gamma_t, \gamma_t^k)} + \frac{d(\gamma_{t+s_k^+}, \gamma_{t+s_k})}{d(\gamma_t, \gamma_t^k)}\right) \\ &= \frac{2C}{1-t}\left(\frac{d(\gamma_t, \gamma_{t+s_k^+})}{d(\gamma_t, \gamma_t^k)} + \frac{d(\gamma_{t+s_k^+}, \gamma_{t+s_k})}{d(\gamma_t, \gamma_{t+s_k^+})} \frac{d(\gamma_t, \gamma_{t+s_k^+})}{d(\gamma_t, \gamma_t^k)}\right) \end{aligned}$$

Passing to the limsup as $k \nearrow \infty$, we have

$$\limsup_{k\to\infty} \frac{|\Phi_t(\gamma_t^k) - \Phi_t(\gamma_t)|}{d(\gamma_t, \gamma_t^k)} \le \frac{2C(1+t)}{(1-t)^2}.$$

The case $\varphi(\gamma_0^k) < \varphi(\gamma_0)$ works the same.

Since $\Phi_t(\gamma_t) = \varphi_t(\gamma_t) + L_t^2(\gamma_t)t/2$, it is then fairly easy to obtain the estimate on L_t with the additional $1/t$ factor and with a different C that however can be taken uniform in γ, provided $L_t(\gamma_t)$ is uniformly bounded:

$$\limsup_{k\to\infty} \frac{|L_t(\gamma_t^k) - L_t(\gamma_t)|}{d(\gamma_t, \gamma_t^k)} \le \frac{C}{t(1-t)^2}.$$

Interchanging the role of 1 and 0, i.e. considering not the level sets of φ but the one of φ^d, doing the same calculations we obtain

$$\limsup_{k\to\infty} \frac{|L_t(\gamma_t^k) - L_t(\gamma_t)|}{d(\gamma_t, \gamma_t^k)} \le \frac{C}{t^2(1-t)}.$$

Observe now that

$$\min\left\{\frac{1}{t^2(1-t)}, \frac{1}{t(1-t)^2}\right\} \le \frac{2}{t(1-t)},$$

hence the claim follows. □

Remark 4.5. If $e_t(G)$ happens to be a length space and quasi-convex, that is for all $x, y \in e_t(G)$ there is a path in $e_t(G)$ joining x to y of length at most $Cd(x, y)$ for some uniform constant C, then the length map L_t is globally Lipschitz. Note that quasi-convexity holds for doubling metric measure spaces supporting a Poincaré inequality.

5. Additional regularity

Restricting the length map L_t to a compact subset of $e_t(G)$, paying a small error in terms of the measure $\boldsymbol{\gamma}$, we can strengthen the previous result and prove the local Lipschitz property. We will prove the following.

Theorem 5.1. *Fix $t \in (0, 1)$. For all $\eta > 0$ there exists a compact set $K \subset G$ so that $\boldsymbol{\gamma}(G \setminus K) \le \eta$ and for every $x \in e_t(K)$*

$$\limsup_{z,w\to x} \frac{|L_t(z) - L_t(w)|}{d(z, w)} \le \frac{C}{t(1-t)},$$

where $z, w \in e_t(H)$, C depends on the length $L_t(\gamma_t)$ and on the Lipschitz constant of φ_t on K. Moreover C is uniform in $x \in K$ if $L_t(\gamma_t)$ is bounded on K; it does not depend on K if the same holds for the Lipschitz constant of φ_t on K.

The proof technique is similar to the one of Theorem 4.4.

Proof. Let $\eta > 0$ be given. From Corollary 2.8 there exists a compact set K with $\boldsymbol{\gamma}(G \setminus K) \le \eta$ so that

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

uniformly for $\gamma \in K$.

Let $\gamma^k, \hat{\gamma}^k \in K$ so that they both converge to $\gamma \in K$. Assume that $\varphi(\hat{\gamma}_0^k) > \varphi(\gamma_0^k)$. As before, considering a continuous extension of the map $I_t(\hat{\gamma}^k) \ni \tau \mapsto \Phi_t(\hat{\gamma}_\tau^k)$, say $\hat{\Phi}_t^k$, to a whole interval $[0, 1]$, still verifying (3.2), we can find a sequence s_k so that $\Phi_t(\gamma_t^k) = \hat{\Phi}_t^k(\hat{\gamma}_{t+s_k}^k)$ and therefore from (3.2)

$$|\Phi_t(\gamma_t^k) - \Phi_t(\hat{\gamma}_t^k)| \le \frac{C}{1-t} d(\hat{\gamma}_t^k, \hat{\gamma}_{t+s_k}^k).$$

Since it could happen that $t + s_k \notin I_t(\hat{\gamma}^k)$ we are forced to consider

$$s_k^+ := \max\{s \in (0, 1) : t + s \in I_t(\hat{\gamma}^k), s \le s_k\}.$$

Thanks to the uniformity granted by K:

$$1 = \lim_{k\to\infty} \frac{\mathcal{L}^1\left(I_t(\hat{\gamma}^k) \cap [t, t+s_k]\right)}{s_k} \le \limsup_{k\to\infty} \frac{s_k^+}{s_k} \le 1.$$

Hence

$$\lim_{k\to\infty} \frac{d(\hat{\gamma}_t^k, \hat{\gamma}_{t+s_k^+}^k)}{d(\hat{\gamma}_t^k, \hat{\gamma}_{t+s_k}^k)} = 1, \qquad \lim_{k\to\infty} \frac{d(\hat{\gamma}_{t+s_k^+}^k, \hat{\gamma}_{t+s_k}^k)}{d(\hat{\gamma}_t^k, \hat{\gamma}_{t+s_k}^k)} = 0.$$

So the claim would be proved if

$$\limsup_{k\to\infty} \frac{d(\hat{\gamma}_t^k, \hat{\gamma}_{t+s_k^+}^k)}{d(\hat{\gamma}_t^k, \gamma_t^k)} \le C\frac{1+t}{1-t}, \tag{5.1}$$

that in turn would be implied by a uniform version of Lemma 4.3. We now prove (5.1) under an additional assumption and the general case will be then obtained just following the proof of Theorem 4.4. So assume $\Phi_t(\hat{\gamma}_{t+s_k^+}^k) = \Phi_t(\gamma_t^k)$, then

$$\varphi_t(\hat{\gamma}_{t+s_k^+}^k) + \frac{t}{2}L_t^2(\hat{\gamma}_{t+s_k^+}^k) = \varphi_t(\gamma_t^k) + \frac{t}{2}L_t^2(\gamma_t^k)$$

Hence

$$\varphi_t(\hat{\gamma}_t^k) - \varphi_t(\gamma_t^k) = \varphi_t(\hat{\gamma}_t^k) - \varphi_t(\hat{\gamma}_{t+s_k^+}^k) - \frac{t}{2}\left(L_t^2(\hat{\gamma}_{t+s_k^+}^k) - L_t^2(\gamma_t^k)\right)$$

It now follows from (2.3) and Proposition 3.1 that

$$\begin{aligned}
\varphi_t(\hat{\gamma}_t^k) - \varphi_t(\gamma_t^k) \ge{}& \frac{1}{2(1-t)} d(\hat{\gamma}_t^k, \hat{\gamma}_{t+s_k^+}^k)\left(d(\hat{\gamma}_t^k, \hat{\gamma}_1^k) + d(\hat{\gamma}_{t+s_k^+}^k, \hat{\gamma}_1^k)\right) \\
&- \frac{t}{2} d(\hat{\gamma}_{t+s_k^+}^k, \gamma_t^k)\left(L_t(\hat{\gamma}_{t+s_k^+}^k) + L_t(\gamma_t^k)\right) \\
\ge{}& \frac{1}{2(1-t)} d(\hat{\gamma}_t^k, \hat{\gamma}_{t+s_k^+}^k)\left(d(\hat{\gamma}_t^k, \hat{\gamma}_1^k) + d(\hat{\gamma}_{t+s_k^+}^k, \hat{\gamma}_1^k)\right) \\
&- \frac{t}{2} d(\hat{\gamma}_{t+s_k^+}^k, \hat{\gamma}_t^k)\left(L_t(\hat{\gamma}_{t+s_k^+}^k) + L_t(\gamma_t^k)\right) - \frac{t}{2} d(\hat{\gamma}_t^k, \gamma_t^k)\left(L_t(\hat{\gamma}_{t+s_k^+}^k) + L_t(\gamma_t^k)\right).
\end{aligned} \tag{5.2}$$

Dividing (5.2) by $d(\hat{\gamma}_t^k, \gamma_t^k)$, we get

$$lip_K(\varphi_t) \ge \limsup_{k\to\infty} \frac{d(\hat{\gamma}_t^k, \hat{\gamma}_{t+s_k^+}^k)}{d(\hat{\gamma}_t^k, \gamma_t^k)} L_t(\gamma_t)(1-t) - tL_t(\gamma_t),$$

where $lip_K(\varphi_t)$ denotes the Lipschitz constant of φ_t on K. Hence (5.1) holds and the claim is therefore proved. □

The following is a trivial consequence of Theorem 5.1.

Corollary 5.2. *For every $\eta > 0$ there exists a compact set $K \subset \mathcal{G}(X)$ so that*

- $\boldsymbol{\gamma}(\mathcal{G}(X) \setminus K) \le \eta$;
- *for every $x \in e_t(K)$ there exists $U(x)$ open neighborhood of x so that*

$$|L_t(z) - L_t(w)| \le \frac{2C}{t(1-t)}, \qquad \forall z, w \in U(x) \cap e_t(K),$$

where C is the one given from Theorem 5.1.

References

[1] L. Ambrosio and N. Gigli. A user's guide to optimal transport. preprint, 2011.
[2] 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.
[3] S.-I. Ohta. On the measure contraction property of metric measure spaces. *Comment. Math. Helv.*, 82:805–828, 2007.
[4] K.T. Sturm. On the geometry of metric measure spaces.I. *Acta Math.*, 196(1):65–131, 2006.
[5] K.T. Sturm. On the geometry of metric measure spaces.II. *Acta Math.*, 196(1):133–177, 2006.
[6] 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

Universität Bonn, Institut für angewandte Mathematik, Endenicher Allee 60, D-53115 Bonn (Germany)
E-mail address: huesmann@iam.uni-bonn.de