THE MONGE PROBLEM FOR DISTANCE COST IN GEODESIC SPACES

STEFANO BIANCHINI AND FABIO CAVALLETTI

Abstract. We address the Monge problem in metric spaces with a geodesic distance: (X, d) is a Polish space and d_L is a geodesic Borel distance which makes (X, d_L) a non branching geodesic space. We show that under the assumption that geodesics are d-continuous and locally compact, we can reduce the transport problem to 1-dimensional transport problems along geodesics.

We introduce two assumptions on the transport problem π which imply that the conditional probabilities of the first marginal on each geodesic are continuous or absolutely continuous w.r.t. the 1-dimensional Hausdorff distance induced by d_L. It is known that this regularity is sufficient for the construction of a transport map.

We study also the dynamics of transport along the geodesic, the stability of our conditions and show that in this setting d_L-cyclical monotonicity is not sufficient for optimality.

Contents

1. Introduction

This paper concerns the Monge transportation problem in geodesic spaces, i.e. metric spaces with a geodesic structure. Given two Borel probability measure $\mu, \nu \in \mathcal{P}(X)$, where (X, d) is a Polish space, we

---

Date: January 25, 2010.

study the minimization of the functional

$$\mathcal{I}(T) = \int d_L(x, T(x))\mu(dy)$$

where T varies over all Borel maps $T : X \to X$ such that $T_\sharp \mu = \nu$ and d_L is a Borel distance that makes (X, d_L) a non branching geodesic space.

Before giving an overview of the paper and of the existence result, we recall which are the main results concerning the Monge problem.

In the original formulation given by Monge in 1781 the problem was settled in $\mathbb{R}^n$, with the cost given by the Euclidean norm and the measures μ, ν were supposed to be absolutely continuous and supported on two disjoint compact sets. The original problem remained unsolved for a long time. In 1978 Sudakov [18] claimed to have a solution for any distance cost function induced by a norm: an essential ingredient in the proof was that if $\mu \ll \mathcal{L}^d$ and $\mathcal{L}^d$-a.e. $\mathbb{R}^d$ can be decomposed into convex sets of dimension k, then then the conditional probabilities are absolutely continuous with respect to the $\mathcal{H}^k$ measure of the correct dimension. But it turns out that when $d > 2$, $0 < k < d-1$ the property claimed by Sudakov is not true. An exemple with $d = 3$, $k = 1$ can be found in [13].

The Euclidean case has been correctly solved only during the last decade. L. C. Evans and W. Gangbo in [9] solved the problem under the assumptions that $\operatorname{spt}\mu \cap \operatorname{spt}\nu = \emptyset$, $\mu, \nu \ll \mathcal{L}^d$ and their densities are Lipschitz function with compact support. The first existence results for general absolutely continuous measures μ, ν with compact support have been independently obtained by L. Caffarelli, M. Feldman and R.J. McCann in [6] and by N. Trudinger and X.J. Wang in [19]. Afterwards M. Feldman and R.J. McCann [10] extended the results to manifolds with geodesic cost. The case of a general norm as cost function on $\mathbb{R}^d$, including also the case with non strictly convex unitary ball, has been solved first in the particular case of crystalline norm by L. Ambrosio, B. Kirchheim and A. Pratelli in [2], and then in fully generality independently by L. Caravenna in [7] and by T. Champion and L. De Pascale in [8].

1.1. Overview of the paper.

The presence of 1-dimensional sets (the geodesics) along which the cost is linear is a strong degeneracy for transport problems. This degeneracy is equivalent to the following problem in $\mathbb{R}$: if μ is concentrated on $(-\infty, 0]$, and ν is concentrated on $[0, +\infty)$, then every transference plan is optimal for the 1-dimensional distance cost $|\cdot|$. In fact, every $\pi \in \Pi(\mu, \nu)$ is supported on the set $(-\infty, 0] \times [0, +\infty)$, on which $|x - y| = y - x$ and thus

$$\int |x - y| \pi(dxdy) = -\int x\mu(dx) + \int y\nu(dy).$$

Nevertheless, for this easy case an explicit map $T : \mathbb{R} \to \mathbb{R}$ can be constructed if μ is continuous: the easiest choice is the monotone map, a minimizer of the quadratic cost $|\cdot|^2$.

The approach suggested by the above simple case is the following:

1. reduce the problem to transportation problems along distinct geodesics;
2. show that the disintegration of the marginal μ on each geodesic is continuous;
3. find a transport map on each geodesic and piece them together.

While the last point can be seen as an application of selection principles in Polish spaces, the first two points are more subtle.

The geodesics used by a given transference plan π to transport mass can be obtained from a set Γ on which π is concentrated. If π wants to be a minimizer, then it certainly chooses the shortest paths: however the metric space can be branching, i.e. geodesics can bifurcate. In this paper we assume that the space is non branching.

Under this assumption, a cyclically monotone plan π yields a natural partition R of the transport set $\mathcal{T}_e$, i.e. the set of the geodesics used by π:

- the set $\mathcal{T}$ made of inner points of geodesics,
- the set $a \cup b := \mathcal{T}_e \setminus \mathcal{T}$ of initial points a and end points b.

The non branching assumption and the cyclical monotonicity of Γ imply that the geodesics used by π are a partition on $\mathcal{T}$, but no other conditions can be obtained on $a \cup b$: one can think to the unit circle with $\mu = \delta_0$ and $\nu = \delta_\pi$. We note here that π gives also a direction along each component of R, as the one dimensional example above shows.

Even if we have a natural partition R in $\mathcal{T}$ and $\mu(a \cup b) = 0$, we cannot reduce the transport problem to one dimensional problems: a necessary and sufficient condition is that the disintegration of the measure μ is strongly consistent, which is equivalent to the fact that there exists a μ-measurable quotient map $f : \mathcal{T} \to \mathcal{T}$. In this case, one can write

$$m := f_\sharp \mu, \quad \mu = \int \mu_y m(dy), \quad \mu_y(f^{-1}(y)) = 1,$$

i.e. the conditional probabilities μ_y are concentrated on the counterimages $f^{-1}(y)$ (which are single geodesics). At this point we can obtain the one dimensional problems by partition π w.r.t. the partition $R \times (X \times X)$,

$$\pi = \int \pi_y m(dy), \quad \nu = \int \nu_y m(dy) \quad \nu_y := (P_2)_\sharp \pi_y,$$

and considering the one dimensional problems along the geodesic $R(y)$ with marginals μ_y, ν_y and cost $|\cdot|$, the length on the geodesic. At this point we can study the problem of the regularity of the conditional probabilities μ_y.

The fact that there exists a strongly consistent disintegration is a property of the geodesics of the metric space. In the setting considered in this paper, (X, d_L) is a non branching geodesic space, not necessarily Polish. To assure that standard measure theory can be used, there exists a second distance d on X which makes (X, d) Polish, and d_L is a Borel function on $X \times X$ with the metric $d \times d$. Note that we do not require d_L to be l.s.c., but we consider a d_L-cyclically monotone transference plan π. It is worth notice that we do not use the existence of optimal potentials (ϕ, ψ), as well as the optimality of π.

Thus, let π be a d_L-cyclically monotone transference plan. The strong consistency of the disintegration of μ along the geodesic used by π is a topological property of the geodesic of d_L as curves in (X, d): in fact we require that they are d-continuous and locally compact. Under this assumption, on $\mathcal{T}$ (the transport set without and points) it is possible to disintegrate μ. Moreover, a natural operation on sets can be considered: the translation along geodesics. If A is a subset of $\mathcal{T}$, we denote by A_t the set translated by t in the direction determined by π.

It turns out that the fact that $\mu(a \cup b) = 0$ and the measures μ_y are continuous depends on how the function $t \mapsto \mu(A_t)$ behaves. We can now state the main result.

Theorem 1.1 (Lemma 5.3 and Proposition 5.4)**.** *If $\sharp\{t > 0 : \mu(A_t) > 0\}$ is uncountable for all A Borel such that $\mu(A) > 0$, then $\mu(a \cup b) = 0$ and the conditional probabilities μ_y are continuous.*

This is sufficient to solve the Monge problem, i.e. to find a transport map which has the same cost as π. A second result concerns a stronger regularity assumption.

Theorem 1.2 (Theorem 5.7)**.** *Assume that $\mathcal{L}^1(\{t > 0 : \mu(A_t) > 0\}) > 0$ for all A Borel such that $\mu(A) > 0$. Then $\mu(a \cup b) = 0$ and μ_y is a.c. w.r.t. the 1-dimensional Hausdorff measure $\mathcal{H}^1_{d_L}$ induced by d_L.*

The assumption of the above theorem and the assumption $d_L \geq d$ allows to define a current in (X, d) which represents the vector field corresponding to the translation $A \mapsto A_t$, and moreover to solve the equation

$$\partial U = \mu - \nu$$

is the sense of current in metric space.

The final results of the paper are the stability of these conditions under Measure-Gromov-Hausdorff like convergence of structures $(X_n, d_n, d_{L,n})$. The conclusion is that a sort of uniform integrability condition on the conditional probability w.r.t. $\mathcal{H}^1_{d_{L,n}}$ passes to the limit, so that one can verify by approximation if Theorem 1.2 holds.

To conclude this introduction, we observe that it is probably possible to extend these results to the case $-d_L$ analytic function on $(X \times X, d \times d)$, see for example the analysis of [3].

We note moreover that in the case $d = d_L$ and (X, d, μ) geodesic measure space satisfying the $\mathrm{MCP}(K, N)$ condition for some K, N real numbers, $N \geq 1$, we can prove, via Theorem 8.6, that Assumption 2 holds. So if d is l.s.c. and (X, d, μ) satisfies $\mathrm{MCP}(K, N)$ we have a solution of the Monge problem.

For a complete view on the so colled Measure Contraction Property $\mathrm{MCP}(K, N)$ we refer to [14] and [12]. Observe that the Heisenberg group and Alexandrov spaces satisfy MCP.

1.2. **Structure of the paper.** The paper is organized as follows.

In Section 2, we recall the basic mathematical results we use. In Section 2.1 the fundamentals of projective set theory are listed. In Section 2.2 we recall the Disintegration Theorem, using the version of [3]. Next, the basic results of selection principles are in Section 2.3, and in Section 2.4 we define the geodesic structure (X, d, d_L) which is studied in this paper. Finally, Section 2.5 recalls some fundamental results in optimal transportation theory.

The next three sections are the key ones.

Section 3 shows how using only the d_L-cyclical monotonicity of a set Γ we can obtain a partial order relation $G \subset X \times X$ as follows (Lemma 3.3 and Proposition 3.7): xGy iff there exists $(w, z) \in \Gamma$ and a geodesic γ, passing trough w and z and with direction $w \to z$, such that x, y belongs to γ and $\gamma^{-1}(x) \leq \gamma^{-1}(y)$. This set G is analytic, and allows to define

- the transport ray set R (3.3),
- the transport sets $\mathcal{T}_e$, $\mathcal{T}$ (with and without and points) (3.4),
- the set of initial points a and final points b (3.7).

Moreover we show that $R\llcorner_{\mathcal{T}\times\mathcal{T}}$ is an equivalence relation (Proposition 3.7), we can assume that the set of final points b can be taken μ-negligible (Lemma 3.11), and in two final remarks we study what happens in the case more regularity on the cost d_L is assumed, Remark 3.12 and Remark 3.13.

Section 4 proves that the continuity and local compactness of geodesics imply that the disintegration induced by R on $\mathcal{T}$ is strongly consistent (Proposition 4.4). Using this fact, we can define an order preserving map g which allows to map our transport problem into a transport problem on $\mathcal{S} \times \mathbb{R}$, where $\mathcal{S}$ is a cross section of R (Proposition 4.6). Finally we show that under this assumption there exists a transference plan with the same cost of π which leaves the common mass $\mu \wedge \nu$ at the same place (note that in general this operation lowers the transference cost).

In Section 5 we prove Theorem 1.1 and Theorem 1.2. We first introduce the operation $A \mapsto A_t$, the translation along geodesic (5.1), and show that $t \mapsto \mu(A_t)$ is a universally measurable function if A is analytic (Lemma 5.2).

Next, we show that under the assumption

$$\mu(A) > 0 \quad \Longrightarrow \quad \sharp\big\{t > 0 : \mu(A_t) > 0\big\} > \aleph_0,$$

the set of initial points a is μ-negligible (Lemma 5.3) and the conditional probabilities μ_y are continuous w.r.t. the 1-dimensional Hausdorff measure $\mathcal{H}^1_{d_L}$.

Finally, we show that under the stronger assumption

$$\mu(A) > 0 \quad \Longrightarrow \quad \int_{\mathbb{R}^+} \mu(A_t) dt > 0, \tag{1.1}$$

the conditional probabilities μ_y are a.c. w.r.t. $\mathcal{H}^1_{d_L}$ (Theorem 5.7). A final result shows that actually Condition (1.1) yields that $t \mapsto \mu(A_t)$ has more regularity than just integrability (Proposition 5.8).

After the above results, the solution of the Monge problem is routine, and it is done in Theorem 6.2 of Section 6.

Under Condition 1.1 and $d \leq d_L$, in Section 7 we give a dynamic interpretation to the transport along geodesics. In Definition 7.1 we define the current $\dot{g}$ in (X, d), which represents the flow induced by the transference plan π. Not much can be said of this flow, unless some regularity assumptions are considered. These assumptions are the natural extensions of properties of transportation problems in finite dimensional spaces.

If there exists a background measure η whose disintegration along geodesic satisfies

$$\eta = \int q_y \mathcal{H}^1_{d_L} m(dy), \qquad q_y \in \mathrm{BV}, \quad \int \mathrm{Tot.Var.}(q_y) m(dy) < +\infty,$$

then $\dot{g}$ is a normal current, i.e. is boundary is a bounded measure on X (Lemma 7.2).

We can also consider the problem $\partial U = \mu - \nu$ in the sense of currents: Proposition 7.4 gives a solution, and in the case $q_y(t) > 0$ for $\mathcal{H}^1_{d_L}$-a.e. t we can write represent $U = \rho \dot{g}$, i.e. the flow $\dot{g}$ multiplied by a scalar density ρ (Corollary 7.6).

In Section 8 we address the stability of the assumptions under Measure-Gromov-Hausdorff-like convergence of structures $(X_n, d_n, d_{L,n})$. Under a uniform integrability condition of $\mu_{y,n}$ w.r.t. $\mathcal{H}^1_{d_{L,n}}$ for the approximating structures (Assumption 4 of Section 8.2), we show that the marginal μ can be represented

as the image of a measure $rm \otimes \mathcal{L}^1$ by a Borel function $h : \mathcal{T} \times \mathbb{R} \to \mathcal{T}_e$, with $r \in L^1(m \otimes \mathcal{L}^1)$ (Lemma 8.5). The key feature of h is that $t \mapsto h(y,t)$ is a geodesic of $\mathcal{T}$ for m-a.e. $y \in \mathcal{T}$.
Thus while $h(0,\mathcal{T})$ is not a cross section for R (in that case we would have finished the proof), in Proposition 8.3 we show which conditions on h imply that μ can be disintegrated with a.c. conditional probabilities, and we verify that this is our case in Theorem 8.6.
In two remarks we suggest how to pass also convex estimates on the disintegration on $(X_n, d_n, d_{L,n})$ to the transference problem in (X, d, d_L) (Remark 8.4 and Remark 8.7).

The last section contains two important examples. In Example 1 we show that if the geodesics are not locally compact, then in general the disintegration along transport rays is not strongly supported. In Example 2 we show that under our assumptions the c-monotonicity is not sufficient for optimality.

We end with a list of notations, Section A.

2. Preliminaries

In this section we recall some general facts about projective classes, the Disintegration Theorem for measure, measurable selection principles, geodesic spaces and optimal transportation problems.

2.1. Borel, projective and universally measurable sets.

The *projective class* $\Sigma^1_1(X)$ is the family of subsets A of the Polish space X for which there exists Y Polish and $B \in \mathcal{B}(X \times Y)$ such that $A = P_1(B)$. The *coprojective class* $\Pi^1_1(X)$ is the complement in X of the class $\Sigma^1_1(X)$. The class Σ^1_1 is called *the class of analytic sets*, and Π^1_1 are the *coanalytic sets*.

The *projective class* $\Sigma^1_{n+1}(X)$ is the family of subsets A of the Polish space X for which there exists Y Polish and $B \in \Pi^1_n(X \times Y)$ such that $A = P_1(B)$. The *coprojective class* $\Pi^1_{n+1}(X)$ is the complement in X of the class Σ^1_{n+1}.

If Σ^1_n, Π^1_n are the projective, coprojective pointclasses, then the following holds (Chapter 4 of [15]):

1. Σ^1_n, Π^1_n are closed under countable unions, intersections (in particular they are monotone classes);
2. Σ^1_n is closed w.r.t. projections, Π^1_n is closed w.r.t. coprojections;
3. if $A \in \Sigma^1_n$, then $X \setminus A \in \Pi^1_n$;
4. the *ambiguous class* $\Delta^1_n = \Sigma^1_n \cap \Pi^1_n$ is a σ-algebra and $\Sigma^1_n \cup \Pi^1_n \subset \Delta^1_{n+1}$.

We will denote by $\mathcal{A}$ the σ-algebra generated by Σ^1_1: clearly $\mathcal{B} = \Delta^1_1 \subset \mathcal{A} \subset \Delta^1_2$.

We recall that a subset of X Polish is *universally measurable* if it belongs to all completed σ-algebras of all Borel measures on X: it can be proved that every set in $\mathcal{A}$ is universally measurable.

Lemma 2.1. *If $f : X \to Y$ is universally measurable, then $f^{-1}(U)$ is universally measurable if U is.*

Proof. If $\mu \in \mathcal{M}(X)$, then $f_\sharp \mu \in \mathcal{M}(Y)$, so for $U \subset Y$ universally measurable there exist Borel sets B, B' such that $B \subset U \subset B'$ and

$$0 = (f_\sharp \mu)(B' \setminus B) = \mu\big(f^{-1}(B') \setminus f^{-1}(B)\big).$$

Since $f^{-1}(B), f^{-1}(B') \subset X$ are universally measurable, there exists Borel sets C, C' such that

$$C \subset f^{-1}(B) \subset f^{-1}(U) \subset f^{-1}(B') \subset C'$$

and $\mu(C' \setminus C) = 0$. The conclusion follows. □

2.2. Disintegration of measures.

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 $\mathscr{R}$:

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

which could be also 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 clearly measurable and inverse measure preserving.

Definition 2.2. 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

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

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 $m(A \vartriangle \hat{A}) = 0$.

We recall the following version of the disintegration theorem that can be found on [11], Section 452 (see [3] for a direct proof).

Theorem 2.3 (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 2.4. *If* $(S, \mathscr{S}) = (X, \mathcal{B}(X))$ *with* X *Polish space, then the disintegration is strongly consistent.*

2.3. **Selection principles.** Given a multivalued function $F : X \to Y$, X, Y metric spaces, the *graph* of F is the set

$$\operatorname{graph}(F) := \big\{(x, y) : y \in F(x)\big\}. \tag{2.1}$$

The *inverse image* of a set $S \subset Y$ is defined as:

$$F^{-1}(S) := \big\{x \in X\ :\ F(x) \cap S \neq \emptyset\big\}. \tag{2.2}$$

For $F \subset X \times Y$, we denote also the sets

$$F_x := F \cap \{x\} \times Y, \quad F^y := F \cap X \times \{y\}. \tag{2.3}$$

In particular, $F(x) = P_2(\operatorname{graph}(F)_x)$, $F^{-1}(y) = P_1(\operatorname{graph}(F)^y)$. We denote by F^{-1} the graph of the inverse function

$$F^{-1} := \big\{(x, y) : (y, x) \in F\big\}. \tag{2.4}$$

We say that F is $\mathcal{R}$*-measurable* if $F^{-1}(B) \in \mathcal{R}$ for all B open. We say that F is *strongly Borel measurable* if inverse images of closed sets are Borel. A multivalued function is called *upper-semicontinuous* if the preimage of every closed set is closed: in particular u.s.c. maps are strongly Borel measurable.

In the following we will not distinguish between a multifunction and its graph. Note that the *domain of* F (i.e. the set $P_1(F)$) is in general a subset of X. The same convention will be used for functions, in the sense that their domain may be a subset of X.

Given $F \subset X \times Y$, a *section* u *of* F is a function from $P_1(F)$ to Y such that $\operatorname{graph}(u) \subset F$. We recall the following selection principle, Theorem 5.5.2 of [15], page 198.

Theorem 2.5. *Let* X *and* Y *be Polish spaces,* $F \subset X \times Y$ *analytic, and* $\mathcal{A}$ *the* σ*-algebra generated by the analytic subsets of* X*. Then there is an* $\mathcal{A}$*-measurable section* $u : P_1(F) \to Y$ *of* F*.*

A *cross-section of the equivalence relation* E is a set $S \subset E$ such that the intersection of S with each equivalence class is a singleton. We recall that a set $A \subset X$ is saturated for the equivalence relation $E \subset X \times X$ if $A = \cup_{x\in A} E(x)$.

The next result is taken from [15], Theorem 5.2.1.

Theorem 2.6. *Let Y be a Polish space, X a nonempty set, and $\mathcal{L}$ an algebra of subset of X. Denote with $\mathcal{L}_\sigma$ the σ-algebra generated by $\mathcal{L}$. Every $\mathcal{L}_\sigma$-measurable, closed value multifunction $F : X \to Y$ admits an $\mathcal{L}_\sigma$-measurable section.*

A standard corollary of the above selection principle is that if the disintegration is strongly consistent in a Polish space, then up to a saturated set of negligible measure there exists a Borel cross-section.

In particular, we will use the following corollary.

Corollary 2.7. *Let $F \subset X \times X$ be $\mathcal{A}$-measurable, X Polish, such that F_x is closed and define the equivalence relation $x \sim y \ \Leftrightarrow\ F(x) = F(y)$. Then there exists a $\mathcal{A}$-section $f : P_1(F) \to X$ such that $(x, f(x)) \in F$ and $f(x) = f(y)$ if $x \sim y$.*

Proof. For all open sets $G \subset X$, consider the sets $F^{-1}(G) = P_1(F \cap X \times G) \in \mathcal{A}$, and let $\mathcal{R}$ be the σ-algebra generated by $F^{-1}(G)$. Clearly $\mathcal{R} \subset \mathcal{A}$.

If $x \sim y$, then

$$x \in F^{-1}(G) \quad \Longleftrightarrow \quad y \in F^{-1}(G),$$

so that each equivalence class is contained in an atom of $\mathcal{R}$, and moreover by construction $x \mapsto F(x)$ is $\mathcal{R}$-measurable.

We thus conclude by using Theorem 2.6 that there exists an $\mathcal{R}$-measurable section f: this measurability condition implies that f is constant on atoms, in particular on equivalence classes. □

2.4. **Metric setting.** In this section we refer to [5].

Definition 2.8. A *length structure* on a topological space X is a class $\mathtt{A}$ of admissible paths, which is a subset of all continuous paths in X, together with a map $L : \mathtt{A} \to [0, +\infty]$: the map L is called *length of path.* The class $\mathtt{A}$ satisfies the following assumptions:

- **closure under restrictions:** if $\gamma : [a, b] \to X$ is admissible and $a \leq c \leq d \leq b$, then $\gamma\llcorner_{[c,d]}$ is also admissible.
- **closure under concatenations of paths:** if $\gamma : [a, b] \to X$ is such that its restrictions γ_1, γ_2 to $[a, c]$ and $[c, b]$ are both admissible, then so is γ.
- **closure under admissible reparametrizations:** for an admissible path $\gamma : [a, b] \to X$ and a for $\varphi : [c, d] \to [a, b]$, $\varphi \in B$, with B class of admissible homeomorphisms that includes the linear one, the composition $\gamma(\varphi(t))$ is also admissible.

The map L satisfies the following properties:

- **additivity:** $L(\gamma\llcorner_{[a,b]}) = L(\gamma\llcorner_{[a,c]}) + L(\gamma\llcorner_{[c,b]})$ for any $c \in [a, b]$.
- **continuity:** $L(\gamma\llcorner_{[a,t]})$ is a continuous function of t.
- **invariance:** The length is invariant under admissible reparametrizations.
- **topology:** Length structure agrees with the topology of X in the following sense: for a neighborhood U_x of a point $x \in X$, the length of paths connecting x with points of the complement of U_x is separated from zero:

$$\inf \big\{ L(\gamma) : \gamma(a) = x, \gamma(b) \in X \setminus U_x \big\} > 0.$$

Given a length structure, we can define a distance

$$d_L(x, y) = \inf \Big\{ L(\gamma) : \gamma : [a, b] \to X, \gamma \in \mathtt{A}, \gamma(a) = x, \gamma(b) = y \Big\},$$

that makes (X, d_L) a metric space (allowing d_L to be $+\infty$). The metric d_L is called *intrinsic.* Observe that for this metric, if $\gamma : [0, 1] \to X$ is d_L-Lipschitz and $|\dot\gamma|_L$ is the metric derivative, then

$$L(\gamma) = \int_0^1 |\dot\gamma|_L(t) dt.$$

Definition 2.9. A length structure is said to be *complete* if for every two points x, y there exists an admissible path joining them whose length $L(\gamma)$ is equal to $d_L(x, y)$.

In other words, a length structure is complete if there exists a shortest path between two points.

Intrinsic metrics associated with complete length structure are said to be *strictly intrinsic.* The metric space (X, d_L) with d_L strictly intrinsic is called a *geodesic space.* A curve whose length equals the distance between its end points is called *geodesic.*

Let $B_{r,L}(x) = \{y : d_L(x, y) < r\}$ denote the usual open ball centred in x with radius r.

Definition 2.10. Let (X, d_L) be a metric space. The distance d_L is said to be *strictly convex* if, for all $r \geq 0$, $d_L(x, y) = r/2$ implies that $B_{r,L}(x) \cap B_{r/2,L}(x)$ is a singleton.

The definition can be restated in geodesics spaces as: geodesics cannot bifurcate in the interior, i.e. *the geodesic space* (X, d_L) *is not branching.* An equivalent requirement is that if $\gamma_1 \neq \gamma_2$ and $\gamma_1(0) = \gamma_2(0)$, $\gamma_1(1) = \gamma_2(1)$, then $\gamma_1((0, 1)) \cap \gamma_2((0, 1)) = \emptyset$ and such geodesics do not admit a geodesic extension i.e. they are not a part of a longer geodesic.

From now on we assume the following:

1. (X, d) Polish space;
2. $d_L : X \times X \to [0, +\infty]$ Borel distance;
3. (X, d_L) is a non-branching geodesic space;
4. geodesics are continuous w.r.t. d;
5. geodesics are locally compact in (X, d): if γ is a geodesic for (X, d_L), then for each $x \in \gamma$ there exists r such that $\gamma^{-1}(\bar{B}_r(x))$ is compact in $\mathbb{R}$.

Since we have two metric structures on X, we denote the quantities relating to d_L with the subscript L: for example

$$B_r(x) = \big\{y : d(x, y) < r\big\}, \quad B_{r,L}(x) = \big\{y : d_L(x, y) < r\big\}.$$

In particular we will use the notation

$$D_L(x) = \big\{y : d_L(x, y) < +\infty\big\},$$

$(\mathcal{K}, d_H)$ for the compact sets of (X, d) with the Hausdorff distance d_H and $(\mathcal{K}_L, d_{H,L})$ for the compact sets of (X, d_L) with the Hausdorff distance $d_{H,L}$. We recall that $(\mathcal{K}, d_H)$ is Polish.

We write

$$\gamma_{[x,y]} := \Big\{\gamma \in \mathrm{Lip}_{d_L}([0, 1]; X) : \gamma(0) = x, \gamma(1) = 1, L(\gamma) = d_L(x, y)\Big\}. \tag{2.5}$$

With a slight abuse of notation, we will write

$$\gamma_{(x,y)} = \bigcup_{\gamma \in \gamma_{[x,y]}} \gamma((0, 1)), \quad \gamma_{[x,y]} = \bigcup_{\gamma \in \gamma_{[x,y]}} \gamma([0, 1]). \tag{2.6}$$

We will also use the following definition.

Definition 2.11. We say that $A \subset X$ is *geodesically convex* if for all $x, y \in A$ the minimizing geodesic $\gamma_{[x,y]}$ between x and y is contained in A:

$$\Big\{\gamma((0, 1)) : \gamma(0) = x, \gamma(1) = y, L(\gamma) = d(x, y), x, y \in A\Big\} \subset A.$$

Lemma 2.12. *If A is analytic in (X, d), then $\{x : d_L(A, x) < \epsilon\}$ is analytic for all $\epsilon > 0$.*

Proof. Observe that

$$\big\{x : d_L(A, x) < \epsilon\big\} = P_1\Big(X \times A \cap \big\{(x, y) : d_L(x, y) < \epsilon\big\}\Big),$$

so that the conclusion follows from the invariance of the class Σ^1_1 w.r.t. projections. □

In particular, $\overline{A}^{d_L}$, the closure of A w.r.t. d_L, is analytic if A is analytic.

Remark 2.13. A simple computation shows that assuming $d_L(x, y) \geq d(x, y)$ then the following holds

1. d_L-compact/closed sets are d-compact/closed;
2. d-Lipschitz functions are d_L-Lipschitz with the same costant.

Whenever more regularity is required we will use the following setting:

- (2') $d_L : X \times X \to [0, +\infty]$ l.s.c. distance,
- (4') $d_L(x, y) \geq d(x, y)$,
- (5') $\cup_{x \in K_1, y \in K_2} \gamma_{[x,y]}$ is compact if K_1, K_2 are d-compact, $d_{L \llcorner K_1 \times K_2}$ uniformly bounded.

2.5. **General facts about optimal transportation.** Let $(X,\mathcal{B},\mu)$ and $(Y,\mathcal{B},\nu)$ be two Polish probability spaces and $c : X \times Y \to \mathbb{R}$ be a Borel measurable function. Consider the set of *transference plans*

$$\Pi(\mu,\nu) := \Big\{\pi \in \mathcal{P}(X \times Y) : (P_1)_\sharp \pi = \mu, (P_2)_\sharp \pi = \nu\Big\}.$$

Define the functional

$$\begin{array}{rcccl} \mathcal{I} & : & \Pi(\mu,\nu) & \to & \mathbb{R}^+ \\ & & \pi & \mapsto & \mathcal{I}(\pi) := \int c\pi. \end{array} \tag{2.7}$$

The *Monge-Kantorovich minimization problem* is to find the minimum of $\mathcal{I}$ over all transference plans.

If we consider a *transport map* $T : X \to Y$ such that $T_\sharp \mu = \nu$, the functional (2.7) becomes

$$\mathcal{I}(T) := \mathcal{I}\big((Id \times T)_\sharp \mu\big) = \int c(x,T(x))\mu(dx).$$

The minimum problem over all T is called *Monge minimization problem.*

The Kantorovich problem admits a (pre) dual formulation.

Definition 2.14. A map $\varphi : X \to \mathbb{R} \cup \{-\infty\}$ is said to be *c-concave* if it is not identically $-\infty$ and there exists $\psi : Y \to \mathbb{R} \cup \{-\infty\}$, $\psi \not\equiv -\infty$, such that

$$\varphi(x) = \inf_{y \in Y} \big\{c(x,y) - \psi(y)\big\}.$$

The *c-transform* of φ is the function

$$\varphi^c(y) := \inf_{x \in X} \{c(x,y) - \varphi(x)\}. \tag{2.8}$$

The *c-superdifferential* $\partial^c\varphi$ of φ is the subset of $X \times Y$ defined by

$$\partial^c\varphi := \Big\{(x,y) : c(x,y) - \varphi(x) \le c(z,y) - \varphi(z)\ \forall z \in X\Big\} \subset X \times Y. \tag{2.9}$$

Definition 2.15. A set $\Gamma \subset X \times Y$ is said to be *c-cyclically monotone* if, for any $n \in \mathbb{N}$ and for any family $(x_0,y_0),\dots,(x_n,y_n)$ of points of Γ, the following inequality holds:

$$\sum_{i=0}^n c(x_i,y_i) \le \sum_{i=0}^n c(x_{i+1},y_i),$$

where $x_{n+1} = x_0$.

A transfence plan is said to be *c-cyclically monotone* if it is concentrated on a c-cyclically monotone set.

Consider the set

$$\Phi_c := \Big\{(\varphi,\psi) \in L^1(\mu) \times L^1(\nu) : \varphi(x) + \psi(y) \le c(x,y)\Big\}. \tag{2.10}$$

Define for all $(\varphi,\psi) \in \Phi_c$ the functional

$$J(\varphi,\psi) := \int \varphi\mu + \int \psi\nu. \tag{2.11}$$

The following is a well known result (see Theorem 5.10 of [20]).

Theorem 2.16 (Kantorovich Duality)**.** *Let X and Y be Polish spaces, let $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$, and let $c : X \times Y \to [0,+\infty]$ be lower semicontinuous. Then the following holds:*

1. *Kantorovich duality:*

$$\inf_{\pi \in \Pi[\mu,\nu]} \mathcal{I}(\pi) = \sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi).$$

Moreover, the infimum on the left-hand side is attained and the right-hand side is also equal to

$$\sup_{(\varphi,\psi)\in\Phi_c \cap C_b} J(\varphi,\psi),$$

where $C_b = C_b(X,\mathbb{R}) \times C_b(Y,\mathbb{R})$.

2. *If $-c$ is real valued and the optimal cost is finite, then there is a measurable c-cyclically monotone set $\Gamma \subset X \times Y$, closed if c is continuous, such that for any $\pi \in \Pi(\mu,\nu)$ the following statements are equivalent:*

   (a) π *is optimal;*
   (b) π *is c-cyclically monotone;*
   (c) π *is concentrated on* Γ*;*
   (d) *there exists a c-concave function* φ *such that* π*-a.s.* $\varphi(x) + \varphi^c(y) = c(x, y)$.
(3) *If moreover*

$$c(x, y) \leq c_X(x) + c_Y(y), \qquad c_X \ \mu\text{-integrable}, \ c_Y \ \nu\text{-integrable},$$

then the supremum is attained:

$$\sup_{\Phi_c} J = J(\varphi, \varphi^c) = \inf_{\pi \in \Pi(\mu,\nu)} \mathcal{I}(\pi).$$

We recall also that if $-c$ is analytic, then every optimal transference plan π is concentrated on a c-cyclically monotone set [3].

3. OPTIMAL TRANSPORTATION IN GEODESIC SPACES

Let $\mu, \nu \in \mathcal{P}(X)$ and consider the transportation problem with cost $c(x, y) = d_L(x, y)$, and let $\pi \in \Pi(\mu, \nu)$ be a d_L-cyclically monotone transference plan with finite cost. By inner regularity, we can assume that the optimal transference plan is concentrated on a σ-compact d_L-cyclically monotone set $\Gamma \subset \{d_L(x, y) < +\infty\}$. By Lusin Theorem, we can require also that $d_{L \llcorner \Gamma}$ is σ-continuous:

$$\Gamma = \cup_n \Gamma_n, \ \Gamma_n \subset \Gamma_{n+1} \text{ compact}, \quad d_{L \llcorner \Gamma_n} \text{ continuous}.$$

Consider the set

$$\begin{aligned} \Gamma' := \bigg\{ &(x, y) : \exists I \in \mathbb{N}_0, (w_i, z_i) \in \Gamma \text{ for } i = 0, \dots, I, \ z_I = y \\ &w_{I+1} = w_0 = x, \ \sum_{i=0}^{I} d_L(w_{i+1}, z_i) - d_L(w_i, z_i) = 0 \bigg\}. \end{aligned} \tag{3.1}$$

In other words, we concatenate points $(x, z), (w, y) \in \Gamma$ if they are initial and final point of a cycle with total cost 0.

Lemma 3.1. *The following holds:*
(1) $\Gamma \subset \Gamma' \subset \{d_L(x, y) < +\infty\}$*;*
(2) *if* Γ *is analytic, so is* Γ'*;*
(3) *if* Γ *is* d_L*-cyclically monotone, so is* Γ'*.*

Proof. For the first point, set $I = 0$ and $(w_{n,0}, z_{n,0}) = (x, y)$ for the first inclusion. If $d_L(x, y) = +\infty$, then $(x, y) \notin \Gamma$ and all finite set of points in Γ are bounded.

For the second point, observe that

$$\begin{aligned} \Gamma' &= \bigcup_{I \in \mathbb{N}_0} P_{12}(A_I) \\ &= \bigcup_{I \in \mathbb{N}_0} P_{12}\bigg(\prod_{i=0}^{I} \Gamma \cap \bigg\{ \prod_{i=1}^{I} (w_i, z_i) : \sum_{i=0}^{I} d_L(w_{i+1}, z_i) - d_L(w_i, z_i) = 0, w_{I+1} = w_0 \bigg\} \bigg). \end{aligned}$$

For each $I \in \mathbb{N}_0$, since d_L is Borel, it follows that

$$\bigg\{ \prod_{i=1}^{I} (w_i, z_i) : \sum_{i=0}^{I} d_L(w_{i+1}, z_i) - d_L(w_i, z_i) = 0, w_{I+1} = w_0 \bigg\}$$

is Borel in $\prod_{i=0}^{I}(X \times X)$, so that for Γ analytic each set $A_{n,I}$ is analytic. Hence $P_{12}(A_I)$ is analytic, and since the class Σ^1_1 is closed under countable unions and intersections it follows that Γ' is analytic.

For the third point, observe that for all $(x_j, y_j) \in \Gamma'$, $j = 0, \dots, J$, there are $(w_{j,i}, z_{j,i}) \in \Gamma$, $i = 0, \dots, I_j$, such that

$$d_L(x_j, y_j) + \sum_{i=0}^{I_j - 1} d_L(w_{j,i+1}, z_{j,i}) - \sum_{i=0}^{I_j} d_L(w_{j,i}, z_{j,i}) = 0.$$

Hence we can write for $x_{J+1} = x_0$, $w_{j,I_j+1} = w_{j+1,0}$, $w_{J+1,0} = w_{0,0}$

$$\sum_{j=0}^{J} d_L(x_{j+1}, y_j) - d_L(x_j, y_j) = \sum_{j=0}^{J} \sum_{i=0}^{I_j} d_L(w_{j,i+1}, z_{j,i}) - d_L(w_{j,i}, z_{j,i}) \geq 0,$$

using the d_L-cyclical monotonicity of Γ. □

Definition 3.2 (Transport rays)**.** Define the *set of oriented transport rays*

$$G := \Big\{(x,y) : \exists (w,z) \in \Gamma', d_L(w,x) + d_L(x,y) + d_L(y,z) = d_L(w,z)\Big\}. \tag{3.2}$$

For $x \in X$, the *outgoing transport rays from* x is the set $G(x)$ and the *incoming transport rays in* x is the set $G^{-1}(x)$. Define the *set of transport rays* as the set

$$R := G \cup G^{-1}. \tag{3.3}$$

Lemma 3.3. *The following holds:*

1. *G is d_L-cyclically monotone;*
2. *$\Gamma' \subset G \subset \{d_L(x,y) < +\infty\}$;*
3. *the sets G, $R := G \cup G^{-1}$ are analytic.*

Proof. The second point follows by the definition: if $(x,y) \in \Gamma'$, just take $(w,z) = (x,y)$ in the r.h.s. of (3.2).

The third point is consequence of the fact that

$$G = P_{34}\Big(\big(\Gamma' \times X \times X\big) \cap \Big\{(w,z,x,y) : d_L(w,x) + d_L(x,y) + d_L(y,z) = d_L(w,z)\Big\}\Big),$$

and the result follows from the properties of analytic sets.

The first point follows from the following observation: if $(x_i, y_i) \in \gamma_{[w_i,z_i]}$, then from triangle inequality

$$\begin{aligned}
d_L(x_{i+1}, y_i) - d_L(x_i, y_i) + d_L(x_i, y_{i-1}) &\geq d_L(x_{i+1}, z_i) - d_L(z_i, y_i) - d_L(x_i, y_i) + d_L(x_i, y_{i-1}) \\
&= d_L(x_{i+1}, z_i) - d_L(x_i, z_i) + d_L(x_i, y_{i-1}) \\
&\geq d_L(x_{i+1}, z_i) - d_L(x_i, z_i) + d_L(w_i, y_{i-1}) - d_L(w_i, x_i) \\
&= d_L(x_{i+1}, z_i) - d_L(w_i, z_i) + d_L(w_i, y_{i-1}).
\end{aligned}$$

Repeating the above inequality finitely many times one obtain

$$\sum_i d_L(x_{i+1}, y_i) - d_L(x_i, y_i) \geq \sum_i d_L(w_{i+1}, z_i) - d_L(w_i, z_i) \geq 0.$$

Hence the set G is d_L-cyclically monotone. □

Definition 3.4. Define the *transport sets*

$$\mathcal{T} := P_1\big(\mathrm{graph}(G^{-1}) \setminus \{x = y\}\big) \cap P_1\big(\mathrm{graph}(G) \setminus \{x = y\}\big), \tag{3.4a}$$

$$\mathcal{T}_e := P_1\big(\mathrm{graph}(G^{-1}) \setminus \{x = y\}\big) \cup P_1\big(\mathrm{graph}(G) \setminus \{x = y\}\big). \tag{3.4b}$$

From the definition of G it is fairly easy to prove that $\mathcal{T}$, $\mathcal{T}_e$ are analytic sets. The subscript e refers to the endpoints of the geodesics: clearly we have

$$\mathcal{T}_e = P_1(R \setminus \{x = y\}). \tag{3.5}$$

The follwing lemma shows that we have only to study the Monge problem in $\mathcal{T}_e$.

Lemma 3.5. *It holds $\pi(\mathcal{T}_e \times \mathcal{T}_e \cup \{x = y\}) = 1$.*

Proof. If $x \in P_1(\Gamma \setminus \{x = y\})$, then $x \in G^{-1}(y) \setminus \{y\}$. Similarly, $y \in P_2(\Gamma \setminus \{x = y\})$ implies that $y \in G(x) \setminus \{x\}$. Hence $\Gamma \setminus \mathcal{T}_e \times \mathcal{T}_e \subset \{x = y\}$. □

As a consequence, $\mu(\mathcal{T}_e) = \nu(\mathcal{T}_e)$ and any maps T such that for $\nu \llcorner_{\mathcal{T}_e} = T_\sharp \mu \llcorner_{\mathcal{T}_e}$ can be extended to a map T' such that $\nu = T_\sharp \mu$ with the same cost by setting

$$T'(x) = \begin{cases} T(x) & x \in \mathcal{T}_e \\ x & x \notin \mathcal{T}_e \end{cases} \tag{3.6}$$

We now use the non branching assumption.

Lemma 3.6. *If $x \in \mathcal{T}$, then $R(x)$ is a single geodesic.*

Proof. Since $x \in \mathcal{T}$, there exists $(w,x),(x,z) \in G \setminus \{x=y\}$: from the d_L-cyclical monotonicity and triangular inequality, it follows that

$$d_L(w,z) = d_L(w,x) + d_L(x,z),$$

so that $(w,z) \in G$ and $x \in \gamma_{(w,z)}$. Hence from the non branching assumption the set

$$R(x) = \bigcup_{y \in G(x)} \gamma_{[x,y]} \cup \bigcup_{z \in G^{-1}(x)} \gamma_{[z,x]}$$

is a single geodesic. □

Proposition 3.7. *The set $R \cap \mathcal{T} \times \mathcal{T}$ is an equivalence relation on $\mathcal{T}$. The set G is a partial order relation on $\mathcal{T}_e$.*

Proof. Using the definition of R, one has in $\mathcal{T}$:

1. $x \in \mathcal{T}$ implies that
$$\exists y \in G(x) \setminus \{x=y\},$$
so that from the definition of G it follows $(x,x) \in G$;
2. if $y \in R(x)$, $x,y \in \mathcal{T}$, then from Lemma 3.6 there exists $(w,z) \in G$ such that $x,y \in \gamma_{(w,z)}$. Hence $x \in R(y)$;
3. if $y \in R(x)$, $z \in R(y)$, $x,y,z \in \mathcal{T}$, then from Lemma 3.6 it follows again there exists $(w,z) \in G$ such that $x,y,z \in \gamma_{(w,z)}$. Hence $z \in R(x)$.

The second part follows similarly:

1. $x \in \mathcal{T}_e$ implies that
$$\exists (x,y) \in \big(G \setminus \{x=y\}\big) \cup \big(G^{-1} \setminus \{x=y\}\big),$$
so that in both cases $(x,x) \in G$;
2. as in Lemma 3.6, $(x,y),(y,z) \in G \setminus \{x=y\}$ implies by d_L-cyclical monotonicity that $(x,z) \in G$.

□

Remark 3.8. Note that $G \cup \{x=y\}$ is an order relation on X.

Definition 3.9. Define the multivalued *endpoint graphs* by:

$$a := \big\{(x,y) \in G^{-1} : G^{-1}(y) \setminus \{y\} = \emptyset\big\}, \tag{3.7a}$$

$$b := \big\{(x,y) \in G : G(y) \setminus \{y\} = \emptyset\big\}. \tag{3.7b}$$

We call $P_2(a)$ the set of *initial points* and $P_2(b)$ the set of *final points.*

Even if a, b are not in the analytic class, still they belong to the σ-algebra $\mathcal{A}$.

Proposition 3.10. *The following holds:*

1. *the sets*
$$a, b \subset X \times X, \quad a(A), b(A) \subset X,$$
belong to the $\mathcal{A}$-class if A analytic;
2. *$a \cap b \cap \mathcal{T}_e \times X = \emptyset$;*
3. *$a(x)$, $b(x)$ are singleton or empty when $x \in \mathcal{T}$;*
4. *$a(\mathcal{T}) = a(\mathcal{T}_e)$, $b(\mathcal{T}) = b(\mathcal{T}_e)$;*
5. *$\mathcal{T}_e = \mathcal{T} \cup a(\mathcal{T}) \cup b(\mathcal{T})$, $\mathcal{T} \cap (a(\mathcal{T}) \cup b(\mathcal{T})) = \emptyset$.*

Proof. Define

$$C := \big\{(x,y,z) \in \mathcal{T}_e \times \mathcal{T}_e \times \mathcal{T}_e : y \in G(x), z \in G(y)\big\} = (G \times X) \cap (X \times G) \cap \mathcal{T}_e \times \mathcal{T}_e \times \mathcal{T}_e,$$

that is clearly analytic. Then

$$b = \big\{(x,y) \in G : y \in G(x), G(y) \setminus \{y\} = \emptyset\big\} = G \setminus P_{1,2}(C \setminus X \times \{y=z\}),$$

$$b(A) = \big\{y : y \in G(x), G(y) \setminus \{y\} = \emptyset, x \in A\big\} = P_2(G \cap A \times X) \setminus P_2(C \setminus X \times \{y=z\}).$$

A similar computation holds for a:

$$a = G^{-1} \setminus P_{23}(C \setminus \{x = y\} \times X), \quad a(A) = P_1(G \cap X \times A) \setminus P_1(C \setminus \{x = y\} \times X)$$

Hence $a, b \in \mathcal{A}(X \times X)$, $a(A), b(A) \in \mathcal{A}(X)$, being the intersection of an analytic set with a coanalytic one.

If $x \in \mathcal{T}$, then from Lemma 3.6 it follows that $a(x)$, $b(x)$ are singleton and $a(x) \neq b(x)$. If $x \in \mathcal{T}_e \setminus \mathcal{T}$, then it follows that the geodesic $\gamma_{[w,z]}$, $(w, z) \in G$, to which x belongs cannot be prolonged in at least one direction: hence $x \in a(x) \cup b(x)$.

The other point follows easily. □

We finally show that we can assume that the μ-measure of final points and the ν-measure of the initial points are 0.

Lemma 3.11. *The sets $G \cap b(\mathcal{T}) \times X$, $G \cap X \times a(\mathcal{T})$ is a subset of the graph of the identity map.*

Proof. From the definition of b one has that

$$x \in b(\mathcal{T}) \quad \Longrightarrow \quad G(x) \setminus \{x\} = \emptyset,$$

A similar computation holds for a. □

Hence we conclude that

$$\pi(b(\mathcal{T}) \times X) = \pi(G \cap b(\mathcal{T}) \times X) = \pi(\{x = y\}),$$

and following (3.6) we can assume that

$$\mu(b(\mathcal{T})) = \nu(a(\mathcal{T})) = 0.$$

Remark 3.12. In the case considered in Remark 2.13, it is possible to show that Γ' is σ-compact: in fact, if one restrict to Γ_n, then the set of cycles of order I is compact, and thus

$$\Gamma'_{n,\bar{I}} := \bigg\{ (x, y) : \exists I \in \{0, \dots, \bar{I}\}, (w_i, z_i) \in \Gamma_n \text{ for } i = 0, \dots, I, \ z_I = y$$
$$w_{I+1} = w_0 = x, \ \sum_{i=0}^{I} d_L(w_{i+1}, z_i) - d_L(w_i, z_i) = 0 \bigg\}$$

is compact. Finally $\Gamma' = \cup_{n,I} \Gamma'_{n,I}$.

Moreover, $d_L \llcorner_{\Gamma'_{n,I}}$ is continuous. If $(x_n, y_n) \to (x, y)$, then from the l.s.c. and

$$\sum_{i=0}^{I} d_L(w_{n,i+1}, z_{n,i}) = \sum_{i=0}^{I} d_L(w_{n,i}, z_{n,i}), \quad w_{n,I+1} = w_{n,0} = x_n, \ z_{n,I} = y_n,$$

it follows also that each $d_L(w_{n,i+1}, z_{n,i})$ is continuous.

Similarly the sets G, R, a, b are σ-compact: assumption (5') and the above computation in fact shows that

$$G_{n,I} := \Big\{ (x, y) : \exists (w, z) \in \Gamma'_{n,I}, d_L(w, x) + d_L(x, y) + d_L(y, z) = d_L(w, z) \Big\}$$

is compact. For a, b, one uses the fact that projection of σ-compact sets is σ-compact.

Remark 3.13. Many simplifications occur in the case the disintegration w.r.t. the partition $\{D_L(x)\}_{x \in X}$ is strongly consistent. Let

$$\pi = \int_0^1 \pi_\alpha m(d\alpha), \quad \mu = \int_0^1 \mu_\alpha m(d\alpha), \ \nu = \int_0^1 \nu_\alpha m(d\alpha)$$

be strongly consistent disintegrations such that

$$\mu_\alpha(D_L(x_\alpha)) = \nu_\alpha(D_L(x_\alpha)) = 1, \quad \pi_\alpha \in \Pi(\mu_\alpha, \nu_\alpha).$$

We have used the fact that the partition $\{D_L(x) \times D_L(x)\}_{x \in X}$ has the crosswise structure, and then we can apply the results of [3].

1) Optimality of π_α. Since π is d_L-cyclically monotone, also the π_α are d_L-cyclically monotone: precisely they are concentrated on the sets

$$\Gamma_\alpha = \Gamma \cap D_L(x_\alpha) \times D_L(x_\alpha),$$

if Γ is d_L-cyclically monotone and $\pi(\Gamma) = 1$.

Using the fact that $(D_L(x_\alpha), d_L)$ is a metric space, then we can construct a potential $\varphi(x, x_\alpha)$ using the formula

$$\varphi(x, x_\alpha) = \inf \bigg\{ \sum_{i=0}^{I} d_L(x_{i+1}, y_i) - d_L(x_i, y_i), (x_i, y_i) \in \Gamma_\alpha, x_{I+1} = x, (x_0, y_0) = (x_\alpha, x_\alpha) \bigg\}.$$

and since this is bounded on $(D_L(x_\alpha), d_L)$, we see that π_α and hence π are optimal.

2. Potential for π. Extend $\varphi(x, x_\alpha)$ to X by setting $\varphi(x, x_\alpha) = +\infty$ if $x \notin D_L(x_\alpha)$. If $\{(x_\alpha, x_\alpha)\}_{\alpha \in [0,1]}$ is a Borel section, then the function

$$\varphi(x) = \inf_\alpha \{\varphi(x, \alpha)\}$$

is easily seen to be analytic. This function is clearly a potential for π. In particular, it follows again from [3] that π is optimal if it is d_L-cyclically monotone.

3.Transport set. We can then define the set of oriented transport rays as the set

$$G = \Big\{ (x, y) \in X \times X : \varphi(x) - \varphi(y) = d_L(x, y) \Big\}.$$

In general, this sets is larger than the one of definition 3.2.

4. PARTITION OF THE TRANSPORT SET $\mathcal{T}$

Let $\{x_i\}_{i \in \mathbb{N}}$ be a dense sequence in (X, d).

Lemma 4.1. *The sets*

$$W_{ijk} := \Big\{ x \in \mathcal{T} \cap \bar{B}_{2^{-j}}(x_i) : L(G(x)), L(G^{-1}(x)) \geq 2^{2-k}, L\big(R(x) \cap \bar{B}_{2^{1-j}}(x_i)\big) \leq 2^{-k} \Big\}$$

form a countable covering of $\mathcal{T}$ *of class* $\mathcal{A}$.

Proof. We first prove the measurability. We consider separately the conditions defining W_{ijk}.

Point 1. The set

$$A_{ij} := \mathcal{T} \cap \bar{B}_{2^{-j}}(x_i)$$

is clearly analytic.

Point 2. The set

$$B_k := \big\{ x \in \mathcal{T} : L(G(x)) \geq 2^{2-k} \big\} = P_1\Big(G \cap \big\{ d_L(x, y) \geq 2^{2-k} \big\} \Big)$$

is again analytic, being the projection of an analytic set. Similarly, the set

$$C_k := \big\{ x \in \mathcal{T} : L(G^{-1}(x)) \geq 2^{2-k} \big\} = P_1\Big(G^{-1} \cap \big\{ d_L(x, y) \geq 2^{2-k} \big\} \Big)$$

is again analytic.

Point 3. The set

$$\begin{aligned} D_{jk} &:= \big\{ x \in \mathcal{T} : L\big(R(x) \cap \bar{B}_{2^{-j}}(x_i)\big) \leq 2^{-k} \big\} \\ &= \mathcal{T} \setminus P_1\Big(R \cap \big(\{(x, y) : d(x_i, y) \leq 2^{1-j}\} \cap \{d_L(x, y) > 2^{-k}\} \big) \Big) \end{aligned}$$

is in the $\mathcal{A}$-class, being the difference of two analytic sets.

We finally can write

$$W_{ijk} = A_{ij} \cap B_k \cap C_k \cap D_{jk},$$

and the fact that $\mathcal{A}$ is a σ-algebra proves that $W_{ijk} \in \mathcal{A}$.

To show that it is a covering, notice that for all $x \in \mathcal{T}$ it holds

$$\min \big\{ L(G(x)), L(G^{-1}(x)) \big\} \geq 2^{2-\bar{k}}$$

for some $\bar{k} \in \mathbb{N}$.

From the local compactness of geodesics, Assumption (5) of page 8, it follows that if $\gamma^{-1}(\bar{B}_r(x))$ is compact, then the continuity of γ implies that $\gamma^{-1}(\bar{B}_{r'}(x))$ is also compact for all $r' \leq r$, and $\operatorname{diam}_{d_L}(\gamma \cap \bar{B}_{r'}(x)) \to 0$ and $r' \to 0$. In particular there exists $\bar{j} \in \mathbb{N}$ such that

$$L\big(R(x) \cap \bar{B}_{2^{1-\bar{j}}}(x)\big) \leq 2^{-\bar{k}},$$

with $\bar{k}$ the one chosen above.

Finally, one choose $x_{\bar{i}}$ such that $d(x, x_{\bar{i}}) < 2^{-1-\bar{j}}$, so that $x \in \bar{B}_{2^{-\bar{j}}}(x_{\bar{i}}) \subset \bar{B}_{2^{1-\bar{j}}}(x)$ and thus

$$L\big(R(x) \cap \bar{B}_{2^{-\bar{j}}}(x_{\bar{i}})\big) \leq 2^{-\bar{k}}.$$

□

Lemma 4.2. *There exist μ-negligible sets $N_{ijk} \subset W_{ijk}$ such that the family of sets*

$$\mathcal{T}_{ijk} = R^{-1}(W_{ijk} \setminus N_{ijk})$$

is a countable covering of $\mathcal{T}_e \setminus \cup_{ijk} N_{ijk}$ into saturated analytic sets.

Proof. First of all, since $W_{ijk} \in \mathcal{A}$, then there exists μ-negligible set $N_{ijk} \subset W_{ijk}$ such that $W_{ijk} \setminus N_{ijk} \in \mathcal{B}(X)$. Hence $\{W_{ijk} \setminus N_{ijk}\}_{i,j,k \in \mathbb{N}}$ is a countable covering of $\mathcal{T} \setminus \cup_{ijk} N_{ijk}$. It follows immediately that $\{\mathcal{T}_{ijk}\}_{i,j,k \in \mathbb{N}}$ satisfies the lemma. □

Remark 4.3. Observe that $\bar{B}_{2^{-j}}(x_i) \cap R(x)$ is compact for all $x \in \mathcal{T}_{ijk}$: in fact, during the proof of Lemma 4.1 we have already shown that $\gamma^{-1}(\bar{B}_{2^{-j}}(x_i))$ is compact.

From any analytic countable covering, we can find a countable partition into $\mathcal{A}$-class saturated sets by defining

$$\mathcal{Z}_{m,e} := \mathcal{T}_{i_m j_m k_m} \setminus \bigcup_{m'=1}^{m-1} \mathcal{T}_{i_{m'} j_{m'} k_{m'}}, \quad \mathcal{Z}_{0,e} := \mathcal{T}_e \setminus \bigcup_{m \in \mathbb{N}} \mathcal{Z}_{m,e}, \tag{4.1}$$

where

$$\mathbb{N} \ni m \mapsto (i_m, j_m, k_m) \in \mathbb{N}^3$$

is a bijective map. Intersecting the above sets with $\mathcal{T}$, we obtain the countable partition of $\mathcal{T}$ in $\mathcal{A}$-sets

$$\mathcal{Z}_m := \mathcal{Z}_{m,e} \cap \mathcal{T}, \quad m \in \mathbb{N}_0. \tag{4.2}$$

Since R is an equivalence relation on $\mathcal{T}$, we use this partition to prove the strong consistency.

On $\mathcal{Z}_m$, $m > 0$, we define the closed values map

$$\mathcal{Z}_m \ni x \mapsto F(x) := R(x) \cap \bar{B}_{2^{-j_m}}(x_{i_m}) \in \mathcal{K}\big(\bar{B}_{2^{-j_m}}(x_{i_m})\big), \tag{4.3}$$

where $\mathcal{K}(\bar{B}_{2^{-j_m}}(x_{i_m}))$ is the space of compact subsets of $\bar{B}_{2^{-j_m}}(x_{i_m})$.

Proposition 4.4. *There exists a μ-measurable cross section $f : \mathcal{T} \to \mathcal{T}$ for the equivalence relation R.*

Proof. First we show that F is $\mathcal{A}$-measurable: for $\delta > 0$,

$$\begin{aligned} F^{-1}(B_\delta(y)) &= \Big\{x \in \mathcal{Z}_m : R(x) \cap B_\delta(y) \cap \bar{B}_{2^{-j_m}}(x_{i_m}) \neq \emptyset\Big\} \\ &= \mathcal{Z}_m \cap P_1\Big(R \cap \big(X \times B_\delta(y) \cap \bar{B}_{2^{-j_m}}(x_{i_m})\big)\Big). \end{aligned}$$

Being the intersection of two $\mathcal{A}$-class sets, $F^{-1}(B_\delta(y))$ is in $\mathcal{A}$.

By Corollary 2.7 there exists a $\mathcal{A}$-class section $f_m : \mathcal{Z}_m \to \bar{B}_{2^{-j_m}}(x_{i_m})$. The proposition follows by setting $f\llcorner_{\mathcal{Z}_m} = f_m$ on $\cup_m \mathcal{Z}_m$, and defining it arbitrarily on $\mathcal{T} \setminus \cup_m \mathcal{Z}_m$: the latter being negligible, f is μ-measurable. □

Up to a μ-negligible saturated set $\mathcal{T}_N$, we can assume it to have σ-compact range: just let $S \subset f(\mathcal{T})$ be a σ-compact set where $f_\sharp \mu$ is concentrated, and set

$$\mathcal{T}_S := R^{-1}(S) \cap \mathcal{T}, \quad \mathcal{T}_N := \mathcal{T} \setminus \mathcal{T}_S, \quad \mu(\mathcal{T}_N) = 0. \tag{4.4}$$

Having a measurable cross-section

$$\mathcal{S} := S \cup f(\mathcal{T}_N) = (\text{Borel}) \cup (f(\mu\text{-negligible})),$$

we can define the parametrization of $\mathcal{T}$, $\mathcal{T}_e$ by geodesics.

Definition 4.5 (Ray map)**.** Define the *ray map* g by the formula

$$\begin{aligned} g :=& \Big\{ (y,t,x) : y \in \mathcal{S}, t \in [0,+\infty), x \in G(y) \cap \{ d_L(x,y) = t\} \Big\} \\ & \cup \Big\{ (y,t,x) : y \in \mathcal{S}, t \in (-\infty,0), x \in G^{-1}(y) \cap \{ d_L(x,y) = -t\} \Big\} \\ =& \, g^+ \cup g^-. \end{aligned}$$

Proposition 4.6. *The following holds.*

1. *The set $g \cap S \times \mathbb{R} \times X$ is analytic.*
2. *It is the graph of a map with range $\mathcal{T}_e$.*
3. *$t \mapsto g(y,t)$ is a d_L 1-Lipschitz G-order preserving for $y \in \mathcal{T}$.*
4. *$(t,y) \mapsto g(y,t)$ is bijective on $\mathcal{T}$, and its inverse is*
$$x \mapsto g^{-1}(x) = \big(f(y), \pm d_L(x, f(y))\big)$$
where f is the quotient map of Proposition 4.4 and the positive/negative sign depends on $x \in G(f(y))/x \in G^{-1}(f(y))$.

Proof. For the first point just observe that

$$\begin{aligned} g^+ =& \Big\{ (y,t,x) : y \in S, t \in \mathbb{R}^+, x \in G(y) \cap \{ d_L(x,y) = t\} \Big\} \\ =& \, S \times R^+ \times X \cap \{ (y,t,x) : (y,x) \in G\} \cap \{ (y,t,x) : d_L(x,y) = t\} \in \Sigma^1_1. \end{aligned}$$

Similarly

$$g^- = \Big\{ (y,t,x) : y \in S, t \in \mathbb{R}^-, x \in G^{-1}(y) \cap \{ d_L(x,y) = -t\} \Big\} \in \Sigma^1_1.$$

Using the fact that $\mathcal{S} \subset \mathcal{T}$ and $R(y)$ is a subset of a single geodesic for $y \in \mathcal{S} \subset \mathcal{T}$, the second part of the proposition follows. The third point is a direct consequence of the definition.

The fourth point follows by substitution. □

We finally prove the following property of c-cyclically monotone transference plans π.

Proposition 4.7. *There exists a d_L-cyclically monotone transference plan $\tilde{\pi}$ with the same cost of π such that it coincides the identity on $\mu \wedge \nu$.*

We will use the disintegration technique exploited also in the next section. We observe that another proof can be the direct composition of the transference plan with itself, using the fact that the mass moves along geodesics and the disintegration makes the problem one dimensional.

Proof. We have already shown that we can take

$$\mu(P_2(b)) = \nu(P_2(a)) = 0,$$

so that $\mu \wedge \nu$ is concentrated on $\mathcal{T}_S$.

Step 1. On $\mathcal{T}$ we can use the Disintegration Theorem to write

$$\mu \llcorner_{\mathcal{T}} = \int_S \mu_y m(dy), \quad m = f_\sharp(\mu \llcorner_{\mathcal{T}}), \ \mu_y \in \mathcal{P}(R(y) \cap \mathcal{T}). \tag{4.5}$$

In fact, the existence of a Borel section is equivalent to the strong consistency of the disintegration. Since $\{R(y) \times X\}_{y \in \mathcal{T}}$ is also a partition on $\mathcal{T} \times X$, we can similarly write

$$\pi \llcorner_{\mathcal{T} \times X} = \int_S \pi_y m(dy), \quad \pi_y(R(y) \times R(y)) = 1.$$

We write moreover

$$\nu_y := (P_2)_\sharp(\pi \llcorner_{\mathcal{T} \times X}), \quad \tilde{\nu} := \int_S \nu_y m(dy) = \int_S (P_2)_\sharp \pi_y m(dy). \tag{4.6}$$

Clearly the rest of the mass starts from $a(\mathcal{T})$, so we have just to show how to rearrange the transference plan in $\mathcal{T}$ in order to obtain $\mu \perp \nu$. Using g, we can reduce the problem to a transport problem on $S \times \mathbb{R}$ with cost

$$c((y,t),(y',t')) = \begin{cases} |t-t'| & y = y' \\ +\infty & y \neq y' \end{cases}$$

By standard regularity argument, we can assume that $S \ni y \mapsto \pi_y \in \mathcal{P}(R(y) \times R(y))$ is σ-continuous, i.e. its graph is σ-compact.

Step 2. Using the fact that $(\mu, \nu) \mapsto \mu \wedge \nu$ is Borel w.r.t. the weak topology [3], we can assume that $S \ni y \mapsto \mu_y \wedge \nu_y \in \mathcal{P}(R(y))$ is σ-continuous, so that also the map

$$S \ni y \mapsto (\mu_y - \mu_y \wedge \nu_y, \nu_y - \mu_y \wedge \nu_y) \in \mathcal{P}(R(y)) \times \mathcal{P}(R(y))$$

is σ-continuous.

Step 3. Since in each $R(y)$ the problem is one dimensional, one can take the unique transference plan

$$\tilde{\pi}_y \in \Pi\big(\mu_y - \mu_y \wedge \nu_y, \nu_y - \mu_y \wedge \nu_y\big)$$

concentrated on a monotone set: clearly

$$\int d_L \tilde{\pi}_y = \int d_L \pi_y.$$

Step 4. If we define the left-continuous distribution functions

$$H(y,s) := \big(\mu_y - \mu_y \wedge \nu_y\big)(-\infty, s), \quad F(y,t) := \big(\nu_y - \mu_y \wedge \nu_y\big)(-\infty, t),$$

and

$$G(y,s,t) := \tilde{\pi}_y\big((-\infty, s) \times (-\infty, t)\big),$$

then the measure $\tilde{\pi}_y$ is uniquely determined by $G(y,s,t) = \min\{H(y,s), F(y,t)\}$.

The σ-continuity of $y \mapsto (\mu_y - \mu_y \wedge \nu_y, \nu_y - \mu_y \wedge \nu_y)$ yields that H, F are again σ-l.s.c., so that G is Borel, and finally $y \mapsto \tilde{\pi}_y$ is σ-continuous up to a $f_\sharp\mu$-negligible set.

Step 5. Define

$$\hat{\pi}_y := \tilde{\pi}_y + (\mathbb{I}, \mathbb{I})_\sharp(\mu_y \wedge \nu_y) \in \Pi(\mu_y, \nu_y).$$

The above steps show that $\hat{\pi}$ is m-measurable, and thus we can define the measure

$$\hat{\pi} := \pi_{\llcorner(\mathcal{T}_e \setminus \mathcal{T}) \times X} + \int \hat{\pi}_y m(dy).$$

It is routine to check that $\hat{\pi}$ has the required properties. □

5. Regularity of the disintegration

Let μ be a probability measure on (X, d). This section is divided in two parts.

In the first one we consider the translation of Borel sets by the optimal geodesic flow, we introduce a first regularity assumption (Assumption 1) on the measure μ and we show that an immediate consequence is that the set of initial points is negligible. A second consequence is that the disintegration of η w.r.t. the R has continuous conditions probabilities.

In the second part we consider a stronger regularity assumption (Assumption 2) which gives that the conditional probabilities are absolutely continuous with respect to $\mathcal{H}^1$ along geodesics.

5.1. Evolution of Borel sets.

Let $A \subset \mathcal{T}_e$ be an analytic set and define for $t \in \mathbb{R}$ the *t-evolution A_t of A* by

$$A_t := g\big(g^{-1}(A) + (0, t)\big). \tag{5.1}$$

Lemma 5.1. *The set $A_t \cap g(S \times \mathbb{R})$ is analytic, and A_t is μ-measurable for $t \geq 0$.*

Proof. Divide A into two parts:

$$A_S := A \cap g(S \times \mathbb{R}) \quad \text{and} \quad A_N := A \setminus A_S.$$

From Point (1) of Proposition 4.6 it follows that A_S is analytic. We consider the evolution of the two sets separately.

Again by Point (1) of Proposition 4.6, the set $(A_S)_t$ is analytic, hence universally measurable for all $t \in \mathbb{R}$.

Since $\mathcal{T}_N$ is μ-negligible (see (4.4)), it follows that $(A_N)_t$ is μ-negligible for all $t > 0$, and by the assumptions it is clearly measurable for $t = 0$. □

We can show that $t \mapsto \mu(A_t)$ is measurable.

Lemma 5.2. *Let A be analytic. The function $t \mapsto \mu(A_t)$ is $\mathcal{A}$-measurable for $t \geq 0$. If $A \subset g(S \times \mathbb{R})$, then $t \mapsto \mu(A_t)$ is $\mathcal{A}$-measurable for $t \in \mathbb{R}$.*

Proof. As before, we split the A into the sets

$$A_S := A \cap g(S \times \mathbb{R}) \quad \text{and} \quad A_N := A \setminus A_S.$$

The function

$$t \mapsto \mu(A_{N,t}) = \begin{cases} \mu(A_N) & t = 0 \\ 0 & t > 0 \end{cases}$$

is clearly Borel. Obserbe that since $\mathcal{T}_N \subset \mathcal{T}$ and μ-measure of final points is 0, the value of $\mu(A_{N,t})$ is known only for $t > 0$.

Since A_S is analytic, then $g^{-1}(A_S)$ is analytic, and the set

$$\tilde{A}_S := \big\{(y, \tau, t) : (y, \tau - t) \in g^{-1}(A_S)\big\}$$

is easily seen to be again analytic. Define the analytic set $\hat{A}_S \subset X \times \mathbb{R}$ by

$$\hat{A}_S := (g, \mathbb{I})(\tilde{A}_S).$$

Clearly $(A_S)_t = \hat{A}_S(t)$. We now show in two steps that the function $t \mapsto \mu((A_S)_t)$ is analytic.

Step 1. Define the closed set in $\mathcal{P}(X \times [0,1])$

$$\Pi(\mu) := \big\{\pi \in \mathcal{P}(X \times [0,1]) : (P_1)_\sharp(\pi) = \mu\big\}$$

and let $B \subset X \times \mathbb{R} \times [0,1]$ a Borel set such that $P_{12}(B) = \hat{A}_S$.

Consider the function

$$\mathbb{R} \times \Pi(\mu) \ni (t, \pi) \mapsto \pi(B(t)).$$

A slight modification of Lemma 4.12 in [3] shows that this function is Borel.

Step 2. Since supremum of Borel function are $\mathcal{A}$-measurable, pag. 134 of [15], the proof is concluded once we show that

$$\mu(A_t) = \mu(\hat{A}_S(t)) = \sup_{\pi \in \Pi(\mu)} \pi(B(t)).$$

From the Disintegration Theorem, for all $\pi \in \Pi(\mu)$ we have

$$\pi(B(t)) = \int \pi_x(B(t)) \mu(dx) \leq \int_{P_1(B(t))} \mu(dx) = \mu(\hat{A}_S(t)).$$

On the other hand from Theorem 2.5, there exists an $\mathcal{A}$-measurable section $u : \hat{A}_S(t) \to B(t)$. Clearly for $\pi_u = (\mathbb{I}, u)_\sharp(\mu)$ it holds $\pi_u(B(t)) = \mu(\hat{A}_S(t))$. □

The next assumption is the first fundamental assumption of the paper.

Assumption 1 (Non-degeneracy assumption)**.** For all Borel sets A such that $\mu(A) > 0$ the set $\{t \in \mathbb{R}^+ : \mu(A_t) > 0\}$ has cardinality $> \aleph_0$.

By inner regularity, it is clearly enough to verifies Assumption 1 only for compact sets.

An immediate consequence of the Assumption 1 is that the measure μ is concentrated on $\mathcal{T}$.

Lemma 5.3. *If μ satisfies Assumption 1 then*

$$\mu(\mathcal{T}_e \setminus \mathcal{T}) = 0.$$

Proof. If $A \subset a(X)$, then $A_t \cap A_s = \emptyset$ for $0 \leq s < t$. Hence

$$\sharp\big\{t \in \mathbb{R}^+ : \mu(A_t) > 0\big\} \leq \aleph_0,$$

because of the boundedness of μ. This contradicts the assumptions. □

Once we know that $\mu(\mathcal{T}) = 1$, we can use the Disintegration Theorem 2.3 to write

$$\mu = \int_S \mu_y m(dy), \quad m = f_\sharp \mu, \; \mu_y \in \mathcal{P}(R(y)). \tag{5.2}$$

The disintegration is strongly consistent since the quotient map $f : \mathcal{T} \to \mathcal{T}$ is μ-measurable and $(\mathcal{T}, \mathcal{B}(\mathcal{T}))$ is countably generated.

The second consequence of Assumption 1 is that μ_y is continuous, i.e. $\mu_y(\{x\}) = 0$ for all $x \in X$.

Proposition 5.4. *The conditional probabilities μ_y are continuous for m-a.e. $y \in S$.*

Proof. From the regularity of the disintegration and the fact that $m(S) = 1$, we can assume that the map $y \mapsto \mu_y$ is weakly continuous on a compact set $K \subset S$ of comeasure $< \epsilon$ such that $L(R(y)) > \epsilon$ for all $y \in K$. It is enough to prove the proposition on K.

Step 1. From the continuity of $K \ni y \mapsto \mu_y \in \mathcal{P}(X)$ w.r.t. the weak topology, it follows that the map

$$y \mapsto A(y) := \big\{x \in R(y) : \mu_y(\{x\}) > 0\big\} = \cup_n \big\{x \in R(y) : \mu_y(\{x\}) \geq 2^{-n}\big\}$$

is σ-closed: in fact, if $(y_m, x_m) \to (y, x)$ and $\mu_{y_m}(\{x_m\}) \geq 2^{-n}$, then $\mu_y(\{x\}) \geq 2^{-n}$ by u.s.c. on compact sets.

Hence it is Borel, and by Lusin Theorem (Theorem 5.8.11 of [15]) it is the countable union of Borel graphs: setting in case $c_i(y) = 0$, we can consider them as Borel functions on S and order w.r.t. G,

$$\mu_{y,\text{atomic}} = \sum_{i \in \mathbb{Z}} c_i(y) \delta_{x_i(y)}, \quad x_{i+1}(y) \in G(x_i(y)),\ i \in \mathbb{Z}.$$

Step 2. Define the sets

$$S_{ij}(t) := \big\{y \in K : x_i(y) = g\big(g^{-1}(x_j(y)) + t\big)\big\} \cap \mathcal{T}.$$

Since $K \subset S$, to define S_{ij} we are using the graph $g \cap S \times \mathbb{R} \times \mathcal{T}$, which is analytic: hence $S_{ij} \in \Sigma^1_1$.

For $A_j := \{x_j(y), y \in K\}$ and $t \in \mathbb{R}^+$ we have that

$$\begin{aligned}\mu((A_j)_t) &= \int_K \mu_y((A_j)_t) m(dy) = \int_K \mu_{y,\text{atomic}}((A_j)_t) m(dy) \\ &= \sum_{i \in \mathbb{Z}} \int_K c_i(y) \delta_{x_i(y)}\big(g(g^{-1}(x_j(y)) + t)\big) m(dy) = \sum_{i \in \mathbb{Z}} \int_{S_{ij}(t)} c_i(y) m(dy).\end{aligned}$$

We have used the fact that $A_j \cap R(y)$ is a singleton.

Step 3. For fixed $i, j \in \mathbb{N}$, again from the fact that $A_j \cap R(y)$ is a singleton

$$S_{ij}(t) \cap S_{ij}(t') = \begin{cases} S_{ij}(t) & t = t' \\ \emptyset & t \neq t' \end{cases}$$

so that

$$\sharp\big\{t : m(S_{ij}(t)) > 0\big\} \leq \aleph_0.$$

Finally

$$\mu((A_j)_t) > 0 \quad \Longrightarrow \quad t \in \bigcup_i \big\{t : m(S_{ij}(t)) > 0\big\},$$

whose cardinality is $\leq \aleph_0$, contradicting Assumption 1. ☐

5.2. **Absolute continuity.** We next assume a stronger regularity assumption.

Assumption 2 (Absolute continuity assumption)**.** For every Borel set $A \subset \mathcal{T}_e$

$$\mu(A) > 0 \quad \Longrightarrow \quad \int_0^{+\infty} \mu(A_t) dt > 0.$$

Again by inner regularity, Assumption 2 can be verified only for compact sets. Note that the condition is meaningful by Lemma 5.2. Observe moreover that Assumption 2 implies Assumption 1, so that in the following we will restrict the map g to the set $g^{-1}(\mathcal{T})$, where it is analytic. Moreover, we can consider shift $t \mapsto A_t$ for $t \in \mathbb{R}$, because of Lemma 5.2.

Remark 5.5. An equivalent form of the Assumption 2 is the following:

$$\mu(A) > 0 \quad \Longrightarrow \quad \int_{t,s \geq 0} \mu(A_t \cap A_s) dt ds > 0.$$

In fact, due to $\mu(X) = 1$, in the set $I_n := \{t : \mu(A_t) > 2^{-n}\}$ the set $\{s \in I_n : \mu(A_s \cap A_t) = 0, t \in I_n\}$ has cardinality at most 2^{-n}. Hence, since for some n $\mathcal{L}^1(I_n) > 0$ by Assumption 2, it follows that

$$\mathcal{L}^2(I_n \times I_n) = \big(\mathcal{L}^1(I_n)\big)^2 > 0.$$

The opposite implication is a consequence of Fubini theorem.

The next results show regularity of the Radon-Nikodym derivative of μ_y w.r.t. $(\mathcal{H}^1_L)\llcorner f^{-1}(y)$, where $\mathcal{H}^1_L$ is the 1-dimensional Hausdorff measure w.r.t. the d_L-distance. Note that along d_L 1-Lipschitz geodesics, $\mathcal{H}^1_L$ is equivalence to $g(y,\cdot)_\sharp\mathcal{L}^1$: in the following we will use both notations.

Lemma 5.6. *Let μ be a Radon measure and*

$$\mu_y = r(y,\cdot)g(y,\cdot)_\sharp\mathcal{L}^1 + \omega_y, \quad \omega_y \perp g(y,\cdot)_\sharp\mathcal{L}^1$$

be the Radon-Nikodym decomposition of μ_y w.r.t. $g(y,\cdot)_\sharp\mathcal{L}^1$. Then there exists a Borel set C such that

$$\mathcal{L}^1\big(g^{-1}(C)\cap(\{y\}\times\mathbb{R})\big)\big) = 0$$

and $\omega_y = \mu_{y\llcorner C}$ for m-a.e. $y\in[0,1]$.

Proof. Consider the measure

$$\lambda = g_\sharp(m\otimes\mathcal{L}^1),$$

and compute the Radon-Nikodym decomposition

$$\mu = \frac{D\mu}{D\lambda}\lambda + \omega.$$

Then there exists a Borel set C such that $\omega = \mu\llcorner_C$ and $\lambda(C) = 0$. The set C proves the Lemma. Indeed $C = \cup_{y\in[0,1]}C_y$ where $C_y = C\cap f^{-1}(y)$ is such that $\mu_{y\llcorner C_y} = \omega_y$ and $g(y,\cdot)_\sharp\mathcal{L}^1(C_y) = 0$ for m-a.e. $y\in[0,1]$. □

Theorem 5.7. *If μ satisfies Assumption 2, then for m-a.e. $y\in[0,1]$ the conditional probabilities μ_y are absolutely continuous w.r.t. $g(y,\cdot)_\sharp\mathcal{L}^1$.*

The proof is based on the following simple observation.

Let η be a Radon measure on $\mathbb{R}$. Suppose that for all $A\subset\mathbb{R}$ Borel with $\eta(A)>0$ it holds

$$\int_{\mathbb{R}^+}\eta(A+t)dt = \eta\otimes\mathcal{L}^1\big(\{(x,t) : t\geq 0, x-t\in A\}\big) > 0.$$

Then $\eta \ll \mathcal{L}^1$.

Proof. The proof will use Lemma 5.6: take C the set constructed in Lemma 5.6 and suppose by contradiction that

$$\mu(C) > 0 \quad \text{and} \quad m\otimes\mathcal{L}^1(g^{-1}(C)) = 0.$$

In particular, for all $t\in\mathbb{R}$ it follows that

$$m\otimes\mathcal{L}^1(g^{-1}(C_t)) = m\otimes\mathcal{L}^1(g^{-1}(C)+(0,t)) = 0.$$

By Fubini-Tonelli Theorem

$$\begin{aligned}
0 < \int_{\mathbb{R}^+}\mu(C_t)dt &= \int_{\mathbb{R}^+}\bigg(\int_{g^{-1}(C_t)}(g^{-1})_\sharp\mu(dyd\tau)\bigg)dt \\
&= \big((g^{-1})_\sharp\mu\otimes\mathcal{L}^1\big)\Big(\Big\{(y,\tau,t) : (y,\tau)\in g^{-1}(\mathcal{T}), (y,\tau-t)\in g^{-1}(C)\Big\}\Big) \\
&\leq \int_{S\times\mathbb{R}}\mathcal{L}^1\big(\{\tau - g^{-1}(C\cap f^{-1}(y))\}\big)(g^{-1})_\sharp\mu(dyd\tau) \\
&= \int_{S\times\mathbb{R}}\mathcal{L}^1\big(g^{-1}(C\cap f^{-1}(y))\big)(g^{-1})_\sharp\mu(dyd\tau) \\
&= \int_S\mathcal{L}^1\big(g^{-1}(C\cap f^{-1}(y))\big)m(dy) = 0.
\end{aligned}$$

That gives a contradiction. □

Now we will study the regularity of the map $t\mapsto\mu(A_t)$ under Assumption 2. We will use the following notation:

$$\mu(A) = \int_S\mu_y(A)m(dy) = \int_S\bigg(\int_{g(y,\cdot)^{-1}(A)}r(y,\tau)d\tau\bigg)m(dy) = g_\sharp(rm\otimes\mathcal{L}^1).$$

Proposition 5.8. *μ satisfies Assumption 2 if and only if for all A Borel $t\mapsto\mu(A_t)$ is continuous. Moreover if A is geodesically convex then $\mu(A_t)$ is absolutely continuous.*

Proof. It is enough to prove the continuity for $t = 0$. Since

$$\mu(A_t) = \int_S \bigg(\int_{g(y,\cdot)^{-1}(A_t)} r(y,\tau) d\tau \bigg) m(dy),$$

its continuity is a direct consequence of Lebesgue dominated convergence theorem applied to the function:

$$t \mapsto \mu_y(A_t) = \int_{g(y,\cdot)^{-1}(A_t)} r(y,\tau) d\tau.$$

Suppose now A geodesically convex. Each $g(y,\cdot)^{-1}(A)$ is an interval $(\alpha(y), \omega(y))$, so that the map

$$t \mapsto \int_{g(y,\cdot)^{-1}(A_t)} r(y,\tau) d\tau$$

is absolutely continuous with derivative

$$h(y,t) = r(y, \omega(y) + t) - r(y, \alpha(y) + t).$$

Since $h(y,t) \in L^1(m \otimes \mathcal{L}^1)$ the result follows by a standard computation. □

6. Solution to the Monge problem

In this section we show that Theorem 5.7 allows to construct an optimal map T. We recall the one dimensional result for the Monge problem [20].

Theorem 6.1. *Let μ, ν be probability measures on $\mathbb{R}$, μ continuous, and let*

$$H(s) := \mu((-\infty, s)), \quad F(t) := \nu((-\infty, t)),$$

be the left-continuous distribution functions of μ and ν respectively. Then the following holds.

1. *The non decreasing function $T : \mathbb{R} \to \mathbb{R} \cup [-\infty, +\infty)$ defined by*
$$T(s) := \sup \big\{ t \in \mathbb{R} : F(t) \leq H(s) \big\}$$
maps μ to ν. Moreover any other non decreasing map T' such that $T'_\sharp \mu = \nu$ coincides with T on the support of μ up to a countable set.
2. *If $\phi : [0, +\infty] \to \mathbb{R}$ is non decreasing and convex, then T is an optimal transport relative to the cost $c(s,t) = \phi(|s-t|)$. Moreover T is the unique optimal transference map if ϕ is strictly convex.*

Assume that μ satisfies Assumption 1. Then we can disintegrate μ and π respect to the ray equivalence relation R and $R \times X$ as in (5.2),

$$\mu = \int \mu_y m(dy), \ \pi = \int \pi_y m(dy), \quad \mu_y \text{ continuous}, \ (P_1)_\sharp \pi_y = \mu_y. \tag{6.1}$$

We write moreover

$$\nu = \int \nu_y m(dy) = \int (P_2)_\sharp \pi_y m(dy). \tag{6.2}$$

Note that $\pi_y \in \Pi(\mu_y, \nu_y)$ is d_L-cyclically monotone (and hence optimal, because $R(y)$ is one dimensional) for m-a.e. y. If $\nu(\mathcal{T}) = 1$, then (6.2) is the disintegration of ν w.r.t. R.

Theorem 6.2. *Let $\pi \in \Pi(\mu,\nu)$ be a d_L-cyclically monotone transference plan, and assume that Assumption 1 holds. Then there exists a Borel map $T : X \to X$ with the same transport cost as π.*

Proof. By means of the map g^{-1}, we reduce to a transport problem on $S \times \mathbb{R}$, with cost

$$c((y,s),(y',t)) = \begin{cases} |t-s| & y = y' \\ +\infty & y \neq y' \end{cases}$$

It is enough to prove the theorem in this setting under the following assumptions: S compact and $S \ni y \mapsto (\mu_y, \nu_y)$ weakly continuous. We consider here the probabilities μ_y, ν_y on $\mathbb{R}$.

Step 1. From the weak continuity of the map $y \mapsto (\mu_y, \nu_y)$, it follows that for all t the map

$$(y,t) \mapsto H(y,t) := \mu_y((-\infty, t)),$$

is continuous in t and l.s.c. in y, hence l.s.c.. Similarly, the map

$$(y,t) \mapsto F(y,t) := \nu_y((-\infty, t))$$

is easily seen to be l.s.c.. Both are clearly increasing in t.

Step 2. The map T defined as Theorem 6.1 by

$$T(y,s) := \Big(y, \sup\big\{t : F(y,t) \leq H(y,s)\big\}\Big)$$

is Borel. In fact, for A Borel,

$$T^{-1}(A \times [t,+\infty)) = \big\{(y,s) : y \in A, H(y,s) \geq F(y,t)\big\} \in \mathcal{B}(S \times \mathbb{R}).$$

Step 3. By the definition of the set G, it follows that along each geodesic $\mu_y(g(y,(-\infty,t))) \geq \nu_y(g(y,(-\infty,t)))$, because in the opposite case G is not d_L-cyclically monotone. Hence $T(s) \geq s$, and $c((y,s),T(y,s)) = P_2(T(y,s)) - s$. □

7. Dynamic interpretation

In this section we show how the regularity of the disintegration yields a correct definition of the current $\dot{g}$ representing the flow along the geodesics of an optimal transference plan. This allows to solve the PDE

$$\partial U = \mu - \nu$$

in the sense of currents in metric spaces. In particular, under additional regularity assumptions, one can prove that the boundary $\partial\dot{g}$ is well defined and satisfies an ODE along geodesics. This gives a dynamic interpretation to the transport problem.

The setting here is slightly different from the previous sections:

1. $d(x,y) \leq d_L(x,y)$;
2. there exists a probability measure η, such that it (or more precisely $\eta \llcorner_{\mathcal{T}_e}$) satisfies Assumption 2 along the transport rays of the transportation problem with marginals μ, ν;
3. $\mu \ll \eta$, so that also μ satisfies Assumption 2.

In particular, $\mathrm{Lip}(X) \subset \mathrm{Lip}_{d_L}(X)$.

The main reference for this chapter is [1].

7.1. Definition of $\dot{g}$.

For any Lipschitz function $\omega : X \to \mathbb{R}$ we can define the derivative $\partial_t\omega$ along the geodesic $g(t,y)$ for a.e. $t \in \mathbb{R}$,

$$\partial_t\omega(g(y,t)) := \frac{d}{dt}\omega(g(t,y))$$

and using the disintegration formula

$$\eta\llcorner_{\mathcal{T}} = \int (g(y,\cdot))_\sharp(q(y,\cdot)\mathcal{L}^1)m(dy) = g_\sharp(qm \otimes \mathcal{L}^1),$$

we can define the measure $\partial_t\omega\eta$ as

$$\int \phi(x)(\partial_t\omega\eta)(dx) := \int_S \int_{\mathbb{R}} \phi(g(y,t))\partial_t\omega(g(y,t))q(y,t)dtm(dy).$$

where $\phi \in C_b(X,\mathbb{R})$.

Definition 7.1. We define the *flow* $\dot{g}$ as the current

$$\langle \dot{g}, (h,\omega)\rangle = \int_{S\times\mathbb{R}} h(g(y,t))\partial_t\omega(g(y,t))q(y,t)dtm(dy)$$

where h, ω are Lipschitz functions of (X,d) with h bounded.

It is fairly easy to see that $\dot{g}$ is a current: in fact,

1. $\dot{g}$ has finite mass, namely
$$\big|\langle \dot{g}, (h,\omega)\rangle\big| \leq \mathrm{Lip}(\omega)\int h\eta;$$
2. $\dot{g}$ is linear in h, ω;
3. if $\omega_n \to \omega$ pointwise in X with uniformly bounded Lipschitz constant, then by Lebesgue Dominated Convergence Theorem if follows that
$$\lim_{n\to+\infty}\langle \dot{g}, (h_n,\omega_n)\rangle = \langle \dot{g}, (h,\omega)\rangle;$$
4. $\langle \dot{g}, (h,\omega)\rangle = 0$ if ω is constant in $\{h \neq 0\}$.

In general, $\dot g$ is only a current, with boundary $\partial \dot g$ defined by the duality formula

$$\langle \partial \dot g, \omega \rangle = \langle \dot g, (1,\omega)\rangle. \tag{7.1}$$

Under additional assumptions, the current $\dot g$ is a normal current, i.e. $\partial \dot g$ is also a scalar current, in particular it is a bounded measure on (X,d).

Lemma 7.2. *Assume that* $q(y,\cdot) : \mathbb{R} \to \mathbb{R}$ *belongs to* $\mathrm{BV}(\mathbb{R})$ *for* m*-a.e.* y *and*

$$\sigma_y := -\frac{d}{dt} q(y,t), \qquad \int_S |\sigma_y(\mathbb{R})| m(dy) = \int_S \mathrm{Tot.Var.}(q(y,\cdot) m(dy) < +\infty.$$

Then $\dot g$ *is a normal current and its boundary is given by*

$$\langle \partial \dot g, \omega \rangle = \int_S \int_{\mathbb{R}} \omega(g(y,t)) \sigma_y(dt) m(dy).$$

Note that in the above formula we cannot restrict σ_y to $g^{-1}(\mathcal{T})$: in fact, in general

$$\int_S (g(y,\cdot)_\sharp \sigma_y)(\mathcal{T}_e \setminus \mathcal{T}) m(dy) > 0.$$

Proof. First of all, by using the formula $q(y,t) = \sigma_y((t,+\infty))$, it follows that σ_y is m-measurable, i.e. for all $\phi \in C_b(X,\mathbb{R})$ the integral

$$\int \bigg(\int \phi(g(y,t)) \sigma_y(dt) \bigg) m(dy)$$

is meaningful and then

$$\int \big(g(y,\cdot)_\sharp \sigma_y \big) m(dy)$$

is a finite measure on (X,d).

A direct computation yields

$$\langle \partial \dot g, \omega \rangle = \langle \dot g, (1,\omega)\rangle = \int_S \int_{\mathbb{R}} \partial_t \omega(g(t,y)) \sigma_y((t,+\infty)) dt m(dy) = \int_S \int_{\mathbb{R}} \omega(g(t,y)) \sigma_y(dt) m(dy).$$

□

Remark 7.3. In many cases the measure $\int (g(y,\cdot)_\sharp \sigma_y)\llcorner_{\mathcal{T}} m(dy)$ is absolutely continuous w.r.t. η, i.e. for m-a.e. $y \in [0,1]$

$$\sigma_y \llcorner_{\mathcal{T}} = h(g(t,y)) q(y,t) \mathcal{L}^1.$$

for some $\mathcal{L}^1(\eta)$ function h. In that case we obtain that

$$\begin{aligned} \langle \partial \dot g, \omega \rangle = & \int \omega(b(y)) \sigma_y \big(P_2(\{g^{-1}(b(y))\}) \big) m(dy) \\ & - \int \omega(a(y)) \sigma_y \big(P_2(\{g^{-1}(a(y))\}) \big) m(dy) + \int \omega(x) h(x) \eta(dx). \end{aligned}$$

7.2. **Transport equation.** We now consider the problem $\partial U = \mu - \nu$ in the sense of currents:

$$\langle U, (1,\omega)\rangle = \langle \mu - \nu, \omega \rangle = \int \omega(x)(\mu - \nu)(dx).$$

Using the disintegration formula and (6.1), (6.2) we can write

$$\langle U, (1,\omega)\rangle = \int_S \bigg\{ \int_{\mathbb{R}} \omega(g(y,t)) (g^{-1}(y,\cdot)_\sharp \mu_y)(dt) - \int_{\mathbb{R}} \omega(g(y,t)) (g^{-1}(y,\cdot)_\sharp \nu_y)(dt) \bigg\} m(dy).$$

By integrating by parts we obtain

$$\int_{\mathbb{R}} \omega(g(y,t)) (g^{-1}(y,\cdot)_\sharp \mu_y)(dt) = -\int_{\mathbb{R}} \mu_y(g(y,(-\infty,t))) \partial_t \omega(g(y,t)) dt = -\int_{\mathbb{R}} H(y,t) \partial_t \omega(g(y,t)) dt,$$

$$\int_{\mathbb{R}} \omega(g(y,t)) (g^{-1}(y,\cdot)_\sharp \nu_y)(dt) = -\int_{\mathbb{R}} \nu_y(g(y,(-\infty,t))) \partial_t \omega(g(y,t)) dt = -\int_{\mathbb{R}} F(y,t) \partial_t \omega(g(y,t)) dt.$$

Observe that the map

$$S \times \mathbb{R} \ni (y,t) \mapsto F(y,t) - H(y,t) \in \mathbb{R}$$

is in $L^1(m \otimes \mathcal{L}^1)$ if the transport cost $\mathcal{I}(\pi)$ is finite: in fact, using the fact that $F(y,t) \leq H(y,t)$ and integrating by parts,

$$\int_{\mathbb{R}} H(y,t) - F(y,t) dt = \int_{\mathbb{R}} (g^{-1}(y,\cdot)_\sharp \mu_y - g^{-1}(y,\cdot)_\sharp \nu_y)(-\infty,t)(dt) = \int_{\mathbb{R}^2} (t-s)\tilde{\pi}(ds,dt), \tag{7.2}$$

where $\tilde{\pi}$ is the monotone rearrangement.

We deduce the following proposition.

Proposition 7.4. *Under Assumption 1, a solution to $\partial U = \mu - \nu$ is given by the current U defined as*

$$\langle U, (h,\omega)\rangle = \int_S \bigg(\int_{\mathbb{R}} (F(y,t) - H(y,t)) h(g(y,t)) \partial_t \omega(g(y,t)) dt \bigg) m(dy).$$

In general, the solution is not unique: just add a boundary free current to our solution.

Some further assumptions allow to represent our solution U as the product of a scalar ρ with the current $\dot{g}$.

Proposition 7.5. *Assume that $q(y,t) > 0$ whenever $H(y,t) - F(y,t) > 0$. Then $R = \rho\dot{g}$, where*

$$\rho(g(y,t)) = \frac{F(y,t) - H(y,t)}{q(y,t)}.$$

Proof. It is enough to observe that

$$\begin{aligned} \int_{S\times\mathbb{R}} F(y,t) - H(y,t) dt m(dy) &= \int_{S\times\mathbb{R}} \frac{F(y,t) - H(y,t)}{q(y,t)} q(y,t) dt m(dy) \\ &= \int_{S\times\mathbb{R}} \rho(g(y,t)) q(y,t) dt m(dy) = \int_X \rho(x)\eta(dx), \end{aligned}$$

and from (7.2) we conclude that $\rho \in L^1(\eta)$. □

Corollary 7.6. *If $q(y,t) \neq 0$ for $m \otimes \mathcal{L}^1$-a.e. $(y,t) \in g^{-1}(\mathcal{T})$, then there exists a scalar function ρ such that $\partial(\rho\dot{g}) = \mu - \nu$.*

8. Stability of the non degeneracy condition

In this section we prove a general approximation theorem, which will be then applied to two important situations: the Measure-Gromov-Hausdorff (MGH) convergence and the approximations of the transport problem by δ-measures. In both cases, if a uniform estimate holds for the disintegration in the approximating spaces, we deduce the regularity of the disintegration also in the limit.

8.1. A general stability result.

We consider the following setting:

(1) μ_n is a sequence of measure converging to μ weakly;
(2) there exists functions $g_n : S_n \times \mathbb{R} \to X$, $S_n \subset X$ Borel, and measures $r_n m_n \otimes \mathcal{L}^1 \in \mathcal{P}(S_n \times \mathbb{R})$ such that

$$\mu_n = (g_n)_\sharp (r_n m_n \otimes \mathcal{L}^1).$$

The following is the basic tool for our stability result.

Proposition 8.1. *Let $\{\xi_n\}_{n\in\mathbb{N}} \subset \mathcal{P}(Y) \times \mathcal{M}(X)$, Y, X Polish and X locally compact, such that $\xi_n \rightharpoonup \xi$. Consider $\{r_n\}_{n\in\mathbb{N}}$, $r_n \geq 0$, such that $r_n\xi_n \in \mathcal{P}(Y \times X)$ is tight and the following equintegrability condition holds:*

$$\forall \delta > 0\ \exists \varepsilon > 0 \bigg(\forall A \in \mathcal{B}, \xi_n(A) < \varepsilon \quad \Longrightarrow \quad \int_A r_n \xi_n < \delta \bigg).$$

Then, up to subsequences, there exists $r \in L^1(\xi)$ such that $r_n\xi_n \rightharpoonup r\xi$.

Proof. From the tightness it follows that there exists $\omega \in \mathcal{P}(Y \times X)$ such that $r_n\xi_n \rightharpoonup \omega$. We will show that $\omega(B) = 0$ for all B such that $\xi(B) = 0$. Clearly by inner and outer regularity, it is enough to prove the following statement:

$$\forall \delta > 0\ \exists \varepsilon > 0 \bigg(1 \geq \phi \geq 0,\ \int \phi\xi < \varepsilon \quad \Longrightarrow \int \phi\omega < \delta \bigg).$$

Fix $\delta > 0$ and take the corresponding ε given by the equintegrability condition on r_n. Consider $\phi \in C_b(Y)$ positive such that

$$\int \phi\xi \leq \varepsilon^2/2.$$

From the weak convergence for n great enough

$$\int \phi\xi_n \leq \varepsilon^2,$$

so that we can estimate

$$\int \phi r_n \xi_n \leq \int_{\phi > \varepsilon} r_n \xi_n + \varepsilon < \delta + \varepsilon.$$

Hence $\int \phi\omega < \delta + \varepsilon$. □

Theorem 8.2. *Assume that* $r_n \in L^1(m_n \otimes \mathcal{L}^1)$ *is equintegrable and*

$$(\mathbb{I}, \mathbb{I}, g_n)_\sharp \big(r_n m_n \otimes \mathcal{L}^1\big) \rightharpoonup (\mathbb{I}, \mathbb{I}, g)_\sharp n$$

with $n \in \mathcal{P}(S \times \mathbb{R})$ *and* g *being the ray map (Definition 4.5). Then* $n = rm \otimes \mathcal{L}^1$ *for some function* $r \in L^1(m \otimes \mathcal{L}^1)$*, measure* $m \in \mathcal{P}(S)$ *and the disintegration of* μ *is a.c. w.r.t.* $\mathcal{L}^1$ *on each geodesic.*

Proof. The fact that $g_\sharp n$ is a disintegration is a consequence of the a.c. of n along each geodesic: in this case the initial points have n-measure 0, and moreover $g_\sharp n = \mu$. We conclude by using Proposition 8.1. □

In general the convergence of the graph of g_n is too strong: the next result considers a more general case.

Proposition 8.3. *Assume that* $\tilde{n} \in \Pi(rm \otimes \mathcal{L}^1, \mu)$ *is concentrated on the graph of a Borel function* $h : \mathcal{T} \times \mathbb{R} \to \mathcal{T}_e$ *such that*

1. $(y, t) \mapsto e(y) := f(h(y, t)) \in S$ *is constant w.r.t.* t*,*
2. *it holds*

$$h(y, \cdot)_\sharp \big(r(y, \cdot)\mathcal{L}^1\big) \ll \mathcal{H}^1 \llcorner_{g(e(y),\mathbb{R})}.$$

Then the disintegration w.r.t. g *has absolutely continuous conditional probability.*

Proof. We can disintegrate the measure m as follows:

$$m = \int_S m_z (e_\sharp m)(dz),$$

and by the second assumption

$$h(y, \cdot)_\sharp (r(y, \cdot)\mathcal{L}^1) = g(e(y), \cdot)_\sharp (\tilde{r}(y, \cdot)\mathcal{L}^1),$$

for m-a.e. $y \in \mathcal{T}$. Hence by explicit computation,

$$\begin{aligned} \mu &= \int_S h(y, \cdot)_\sharp (r(y, \cdot)\mathcal{L}^1) m(dy) = \int_S g(e(y), \cdot)_\sharp (\tilde{r}(y, \cdot)\mathcal{L}^1) m(dy) \\ &= \int_S \bigg(\int_{e^{-1}(z)} g(z, \cdot)_\sharp (\tilde{r}(y, \cdot)\mathcal{L}^1) m_z(dy) \bigg) e_\sharp m(dz). \end{aligned}$$

To conclude the proof observe that

$$\int_{e^{-1}(z)} g(z, \cdot)_\sharp (\tilde{r}(y, \cdot)\mathcal{L}^1) m_z(dy) \ll g(z, \cdot)_\sharp \mathcal{L}^1.$$

□

Remark 8.4. A corollary of the above results is that if r_n belongs to some closed convex set of L^1 for all n, then this condition passes to the limit r. Just observe that if $\phi \in C(S \times \mathbb{R})$ then

$$\forall n \int \phi r_n m_n \otimes \mathcal{L}^1 \geq k \quad \Longrightarrow \quad \int \phi r m \otimes \mathcal{L}^1 \geq k.$$

In relation with the previous section, we consider the following cases:

(1) for all $\varepsilon > 0$ there exists $k_\varepsilon > 0$ such that

$$(r_n \circ g^{-1})\llcorner_{\{x \in \mathcal{T} : d_L(x, a(x) \cup b(x)) > \varepsilon\}} \geq k_\epsilon;$$

(2) for all $\varepsilon > 0$, $y \in S$ there exists $L_\varepsilon > 0$ such that

$$(r_n \circ g^{-1})\llcorner_{\{x \in R(y) : d_L(x, a(y) \cup b(y)) > \varepsilon\}} \in \mathrm{Lip}_{L_\varepsilon}(\mathbb{R});$$

(3) for all $\varepsilon > 0$, $y \in S$ there exists $M_\varepsilon > 0$ such that

$$\mathrm{Tot.Var.}\Big((r_n \circ g^{-1})\llcorner_{\{x \in R(y) : d_L(a(y), b(y)) > \varepsilon\}} \Big) \leq M_\varepsilon.$$

The first condition yields that the assumptions of Corollary 7.6 holds.

The second and third conditions imply that we are under the conditions for Remark 7.3, if the rays are of uniformly positive length.

8.2. **Approximations by metric spaces.** In this section we explain a procedure to verify if the transport problem under consideration satisfied Assumption 2. The basic references for this sections are [16, 17].

We consider the following setting:

(1) $(X_n, d_n, d_{L,n})$, $n \in \mathbb{N}$, is a family of metric structures satisfying the assumptions of page 8 and Remark 2.13;
(2) $\mu_n, \nu_n \in \mathcal{P}(X_n)$, $\mu_n \perp \nu_n$;
(3) $\pi_n \in \Pi(\mu_n, \nu_n)$ $d_{L,n}$-cyclically monotone transference plan with finite cost.

We assume that the structure $(X_n, d_n, d_{L,n}, \mu_n, \nu_n, \pi_n)$ converges to $(X, d, d_L, \mu, \nu, \pi)$ in the following sense: for all $n \in \mathbb{N}$ there exist Borel sets $A_n \subset X_n$ and Borel maps $\ell_n : A_n \to X$ such that

$$(\ell_n, \ell_n)_\sharp \pi_n \llcorner_{A_n \times A_n} \rightharpoonup \pi, \quad \big| d(\ell_n(x), \ell_n(y)) - d_n(x, y) \big| \leq 2^{-n}, \big| d_L(\ell_n(x), \ell_n(y)) - d_{L,n}(x, y) \big| \leq 2^{-n}, \tag{8.1}$$

and if $(\ell_n(x_n), \ell_n(y_n)) \to (x, y)$, then

$$d_L(x, y) = \lim_n d_{L,n}(x_n, y_n). \tag{8.2}$$

By σ-additivity of measures, Remark 3.12 and up to subsequences, we can reduce the problem as follows:

(1) Γ, G are compact subsets of $X \times X$;
(2) A_n compact, $\ell_n(A_n) \subset \ell_{n+1}(A_{n+1})$ and ℓ_n continuous;
(3) the compact sets $\ell_n(A_n)$ converge in Hausdorff distance to a compact set K on which μ and ν are concentrated;
(4) $\pi_n(A_n) = 1$ upon renormalization and $(\ell_n, \ell_n)_\sharp \pi_n$ is concentrated in

$$\big\{ (x, y) : d(\bar{x}, \bar{y}) > 8r, d(x, \bar{x}), d(y, \bar{y}) \leq r \big\};$$

(5) the section S_n of Proposition 4.4 is a compact subset of $X_n \setminus B_{2r}(\bar{x}) \cup B_{2r}(\bar{y})$;
(6) the compact sets $\ell_n(S_n)$ converge to a compact subset of $\Gamma \setminus (B_{2r}(\bar{x}) \cup B_{2r}(\bar{y}))$;
(7) there exists a sequence of $d_{L,n}$-cyclically monotone sets $\Gamma_n \subset X_n$ such that $(\ell_n, \ell_n)(\Gamma_n) \to \Gamma$.

For the transport problems in (X_n, d_n) with measures μ_n, ν_n, we assume the following.

Assumption 3. The $d_{L,n}$-cyclically monotone plan π_n satisfies Assumption 2 for all $n \in \mathbb{N}$.

It follows that for all $n \in \mathbb{N}$ we can write the disintegration of μ_n along the geodesics as

$$\mu_n = (g_n)_\sharp (r_n m \otimes \mathcal{L}^1) = \int g_n(y, \cdot)_\sharp (r(y, \cdot) \mathcal{L}^1) m(dy),$$

with $r_n \in L^1(m \otimes \mathcal{L}^1)$. The next assumption is fundamental.

Assumption 4 (Equintegrability)**.** The map $(y, t) \mapsto r_n(y, t)$ is equintegrable w.r.t. the measure $m_n \otimes \mathcal{L}^1$: $r_n m_n \otimes \mathcal{L}^1$ is tight and there exists $\delta > 0$ such that if $(m_n \otimes \mathcal{L}^1)(A) < \delta$ then

$$\int_A r_n(y, t) dt m_n(dy) < \varepsilon.$$

Note that since $\int r_n(y,t)dt = 1$, then the equintegrability is essentially a condition on $\mathcal{L}^1$: in fact, in the disintegration theorem no regularity on m is required.

Let $g_n : S_n \times \mathbb{R} \to X_n$ be the transport maps for the approximating problems, and write

$$\begin{array}{rccc} \tilde{g}_n(y,t) & : & S_n \times \mathbb{R} & \to & X \times \mathbb{R} \\ & & (y,t) & \mapsto & \tilde{g}_n(y,t) := \big(\ell_n(y), t\big) \end{array}$$

$$\tilde{r}_n \tilde{m}_n \otimes \mathcal{L}^1 := (\tilde{g}_n)_\sharp (r_n m_n \otimes \mathcal{L}^1) \in \mathcal{P}(X \times \mathbb{R}).$$

This allows us to consider the space X as a section. Define finally

$$n_n := (\tilde{g}_n, \ell_n \circ g_n)_\sharp (r_n m_n \otimes \mathcal{L}^1) \in \Pi\Big(\tilde{r}_n \tilde{m}_n \otimes \mathcal{L}^1, (\ell_n \circ g_n)_\sharp (r_n m_n \otimes \mathcal{L}^1)\Big).$$

This measure is not supported on a graph: it is obtained by considering the measure $(\mathbb{I}, \ell_n \circ g_n)_\sharp (r_n m_n \otimes \mathcal{L}^1)$ and taking the image measure of the first marginal by $\tilde{g}_n$.

Lemma 8.5. *Up to subsequences, $n_n \rightharpoonup n$, where $n \in \Pi(rm \otimes \mathcal{L}^1, \mu)$ is supported on a Borel graph $h : \mathcal{T} \times \mathbb{R} \to \mathcal{T}_e$ such that $t \mapsto h(y,t)$ is the d_L 1-Lipschitz curve $R(y)$ for m-a.e. $y \in X$.*

Proof. The convergence to the correct marginals is a consequence of the fact that

$$(P_2)_\sharp \tilde{n}_n = (\ell_n \circ g_n)_\sharp (r_n m_n \otimes \mathcal{L}^1) = (\ell_n)_\sharp \mu_n \rightharpoonup \mu$$

by (8.1), and by Proposition 8.1 and Assumption 4

$$\tilde{r}_n \tilde{m}_n \otimes \mathcal{L}^1 \rightharpoonup rm \otimes \mathcal{L}^1.$$

We next observe that from (8.1) in A_n

$$\Big| d_L\big(\ell_n \circ g_n(y_n, 0), \ell_n \circ g_n(y_n, t)\big) - t \Big| \leq 2^{-n},$$

so that if $\ell_n(y_n,) \to y$ then by (8.2) $\ell_n \circ g_n(y_n, \cdot)$ converges to a d_L 1-Lipschitz geodesic curve θ_y in K passing through y.

Define the multivalued map

$$\begin{array}{rccc} h_n & : & \ell_n(S_n) \times \mathbb{R} & \to & X \\ & & (y,t) & \mapsto & \{\ell_n(g_n(z,t)) : \ell_n(z) = y\}. \end{array}$$

W.l.o.g. we can assume g, g_n continuous therefore, by assumptions on A_n and ℓ_n, the graph(h_n) is locally compact. It follows that up to subsequence graph$(h_n) \to$ graph(h) in Hausdorff metric where h is a multivalued function with locally compact graph. Since $n_n(\text{graph}(h_n)) = 1$ we can conclude that $n(\text{graph}(h)) = 1$.

By Point 6 it follows that

$$\big\{y : \exists (t,x), (y,t,x) \in \text{graph}(h)\big\} \subset \mathcal{T}$$

then h is a Borel function $h : \mathcal{T} \times \mathbb{R} \to X$.

We conclude by showing that the curve $t \mapsto h(y,t)$ is the d_L 1-Lipschitz parametrization of $R(y)$. First, from the $d_{L,n}$-cyclical monotonicity of π_n and the estimate (8.1), it follows that

$$\Big\{(x,x') \in K \times K : x = h(y,s), x' = h(y,s'), s \geq s'\Big\}$$

is d_L-cyclically monotone: in fact for all sequences $(y_i, t_i) \in K$, $i = 0, \dots, I$,

$$\sum_{i=0}^{I} d_L\big(h(y_i,t_i), h(y_{i+1}, s_{i+1})\big) - d_L\big(h(y_i,t_i), h(y_i,s_i)\big) \geq -2^{-n} I +$$
$$\sum_{i=0}^{I} d_{L,n}\big(g_n(y_i^n,t_i), g_n(y_{i+1}^n, s_{i+1})\big) - d_{L,n}\big(g_n(y_i^n,t_i), g_n(y_i^n,s_i)\big) \geq -2^{-n} I.$$

Hence on K the geodesics θ_y can intersects only at the endpoints, which have measure 0 w.r.t. $\mu = h_\sharp(rm \otimes \mathcal{L}^1)$.

Finally, Point 7 implies that each θ_y coincides with $R(y)$. In fact, if $\ell_n \circ g_n(y_n, t) \to h(y, t)$ and $x'_n \in R(g_n(y_n))$ is such that $\ell_n(x'_n) \in B_{2r}(\bar{y})$, then

$$\begin{aligned} d_L(h(y,t), x') =& \lim_{n \to +\infty} d_L\big(\ell_n \circ g_n(y_n, t), \ell_n(x'_n)\big) \\ =& \lim_{n \to +\infty} d_{L,n}(g_n(y_n, 0), x'_n) + t = d_L(y, x') + t. \end{aligned}$$

The conclusion follows. $\square$

Theorem 8.6. *The transport plan π satisfy Assumption 2.*

Proof. By Lemma 8.5, it is enough to verify the hypothesis of Proposition 8.3. First of all, from d_L-cyclical monotonicity or from the proof of Lemma 8.5, it follows that $(t, y) \mapsto e(y) := f(h(y, t))$ does not depend on t.

Next, the d_L 1-Lipschitz dependence implies that if $r(y, \cdot) \in L^1(\mathbb{R})$, then also $h(y, \cdot)_\sharp(r(y, \cdot)\mathcal{L}^1)$ is a.c. w.r.t. $\mathcal{L}^1$. $\square$

Remark 8.7. As in Remark 8.4, if we know more regularity of the disintegrations for the approximating problems, we can pass them to the limit. Here the key observation is that geodesics converges to geodesic, so that continuous functions on them converge pointwise to continuous functions.

A special case is when $d_L = d$: a natural approximation is by transport plans where ν is atomic, with a finite number of atoms. Using techniques similar to [4], one can prove the next proposition.

Proposition 8.8. *Let $\{x_n\}_{n \in \mathbb{N}}$ be a dense sequence in X, and assume that for all transport problems*

$$\inf \bigg\{ \int d(x, y)\pi, \pi \in \Pi(\mu, \delta_{x_n}) \bigg\}$$

the disintegration along transport rays $\mu = \int \mu_{n,y} m_n(dy)$ satisfies

$$\mu_{n,y} \in \big[h(d(x, x_n), L(R(y))), H(d(x, x_n), L(R(y)))\big]\mathcal{L}^1$$

for some strictly positive continuous functions $h, H : (\mathbb{R}^+)^2 \mapsto \mathbb{R}$. Then for all $\nu \in \mathcal{P}(X)$ the disintegration $\mu = \int \mu_y m(dy)$ satisfies the same bounds.

Proof. Observe that for all $a > 0$ along rays of length in $[a, b]$ and for $d(x, x_n) \leq a/2$, the densities $\mu_{n,y}$ are equintegrable. $\square$

Proposition 8.8 ensure that there exists a solution for the Monge minimization problem in all the spaces that satisfies the Measure Contraction Property, like Alexandrov spaces and the d-dimensional Heisenberg group, see [12]. The reference for the measure contraction property is [14].

9. EXAMPLES

We end this paper with some examples which shows how the different hypotheses of Section 2.4 enter into the analysis. In the following we denote the standard Euclidean scalar product in $\mathbb{R}^d$ as $\cdot$ and the standard distance in $\mathbb{T}^d$ by $|\cdot|$. We will also denote points by $p = (x, y, z, \dots) \in \mathbb{R}^d$, and α a fixed constant in $[0, 1] \setminus \mathbb{Q}$.

Example 1 (Non strongly consistent disintegration along rays)**.** Consider the metric space

$$(X, d) = \big(\mathbb{T}^2, |\cdot|\big)$$

and the l.s.c. distance in the local chart $X = \{(x, y) : 0 \leq x, y < 1\}$

$$d_L(p_1, p_2) := \begin{cases} |x_1 - x_2 + i| & y_1 - y_2 = \alpha(x_1 - x_2) + i\alpha + n \\ +\infty & \text{otherwise} \end{cases}$$

for $i, n \in \mathbb{Z}$. The sets D_L are given by

$$D_L(p_1) = \Big\{(x, y) : y = y_1 + \alpha(x - x_1 + i) \mod 1, i \in \mathbb{N}\Big\},$$

so that it is easy to see that the partition $\{D_L(p)\}_{p \in X}$ does not yield a strongly consistent disintegration. Since $t \mapsto (t \mod 1, \alpha t \mod 1)$ is a continuous not locally compact geodesic, Condition (5) is not verified in this system.

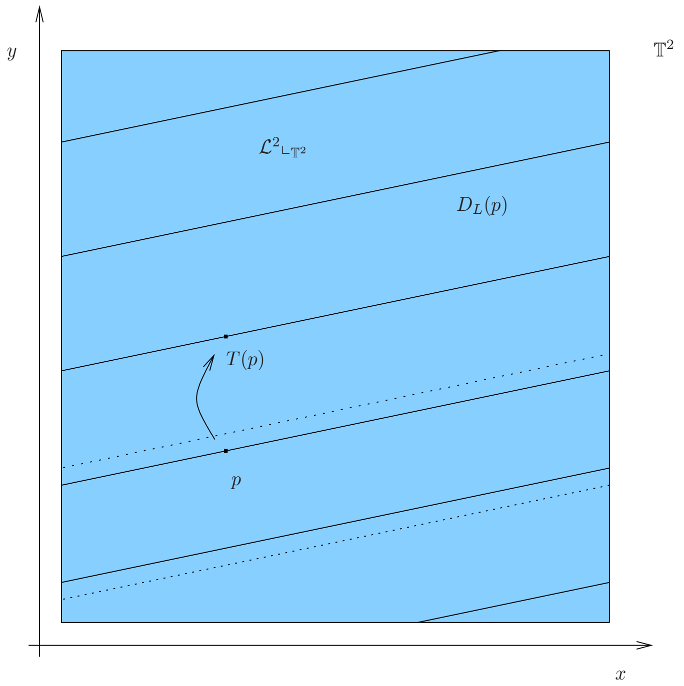

FIGURE 1. The metric space of Example 1

Consider the measures $\mu = \mathcal{L}^2\llcorner_{\mathbb{T}}$ and the map $T : (x, y) \mapsto (x, y + \alpha \mod 1)$: being μ invariant w.r.t. translations, one has $T_\sharp \mu = \mu$, and moreover

$$\int d_L(x, T(x))\mu(dx) = 1.$$

If we consider points $(p_i, (x_i, y_i + \alpha \mod 1))$, $i = 1, \dots, I$, then the only case for which $d_L(p_{i+1}, p_i) < +\infty$ is when $p_{i+1} = (x_i + t \mod 1, y_i + \alpha t \mod 1)$ for some $t \in \mathbb{R}$, i.e. they belong to the geodesic

$$\mathbb{R} \ni t \mapsto (x_i + t \mod 1, y_i + \alpha t \mod 1) \in X.$$

Hence, to prove d_L-cyclical monotonicity, it is sufficient to consider path which belongs to a single geodesic, where d_L reduces to the the one dimensional length:

$$d_L\big((x, y), (x + t \mod 1, y + \alpha t \mod 1)\big) = |t|.$$

Since translations in $\mathbb{R}$ are cyclically monotone w.r.t. the absolute value, we conclude that T is d_L-cyclically monotone.

The fact that the optimal rays coincide with the sets D_L yields that the disintegration is not strongly consistent, in particular there is not a Borel section up to a saturated negligible set. Note that every transference plan which leaves the common mass in the same place has cost 0, so that this example shows the necessity of Condition (5) for Proposition 4.7.

Example 2 (Non optimality of transport map)**.** Consider countable copies of the manifolds $\mathbb{T}^2$: we denote them in local coordinates by

$$C := \big\{(x, y) : 0 \leq x, y < 1\big\}, \quad C^i := \big\{(x^i, y^i) : 0 \leq x^i, y^i < 1\big\},\ i \in \mathbb{Z} \setminus \{0\}.$$

With this fixed choice of coordinates, identify the points $(x, 0) \equiv (x^i, 0)$ if $0 \leq x = x^i < 1$. In other words, we glue the sets C, C^i, $i \in \mathbb{Z} \setminus \{0\}$, along a maximal circle S, which will be written in local coordinates by

$$S = \big\{\theta : 0 \leq \theta < 1\big\}.$$

The space X is the set obtained with this procedure.

In the following points we need to divide C into two parts: with the same coordinates as above, we set

$$C^- := \big\{(x, y) \in C : 0 \leq x < 1, 0 \leq y \leq 1/2\big\}, \quad C^+ := \big\{(x, y) \in C : 0 \leq x < 1, 1/2 < y < 1\big\}.$$

Definition of d and d_L. The distance d is defined as follows:

$$d(p_1,p_2)=\min\Big\{|p_1-p_2|,|(x_1,y_1)-(\theta,0)|+|(x_2,y_2)-(\theta,0)|,\theta\in S^1\Big\}.$$

Note that p_1-p_2 can be computed only when the points belong to the same component. It is fairly easy to see that (X,d) is a compact set, in particular Polish.

The distance d_L is defined as follows: if $\gamma:[0,1]\to X$ is a d-Lipschitz map, then set

$$L(\gamma):=\int_0^1\omega(\gamma(t),\dot\gamma(t))dt,\quad \omega(p,v):=\begin{cases}|\dot\gamma| & p\in C^-,\dot\gamma\cdot(-1,\alpha)=0\\ 4|\dot\gamma| & p\in C^+\setminus S,\dot\gamma\cdot(-1,\alpha)=0\\ |\dot\gamma| & p\in C^i\setminus S,\dot\gamma\cdot(-1,i\alpha)=0\\ +\infty & \text{otherwise}\end{cases}$$

In otherwords, the Lipschitz path with finite length are a countable union of segments in C or C^i, $i\in\mathbb{Z}\setminus\{0\}$ with slope $(\alpha,1)$, $(i\alpha,1)$, respectively. The distance d_L is define then by

$$d_L(p_1,p_2):=\inf\Big\{L(\gamma):\gamma\in\mathrm{Lip}([0,1],X),\gamma(0)=p_1,\gamma(1)=p_2\Big\}.$$

Study of the distance d_L. To study the distance d_L, observe that it is enough to analyze the induced distance on S^1. We consider the length of the return map on S depending on which sets we are moving on:

1. if we take the path $\theta\to\theta+\alpha$ along C, then its length is $\frac{5}{2}\sqrt{1+\alpha^2}$;
2. if we take the path $\theta\to\theta+i\alpha$ along C^i, then its length is $\sqrt{1+(i\alpha)^2}$.

In particular, geodesics starting from S and ending in some C^i, $i\in\mathbb{Z}\setminus\{0\}$, never take values in $C\setminus S$. Due to the invariance w.r.t. translations $(x,y)\mapsto(x+\alpha\mod 1,y)$, it is sufficient to study the structure the metric space $(D_L((0,0)),d_L)$.

The set $D_L((0,0))$ is the set $\{y=\alpha x+z\alpha\mod 1,z\in\mathbb{Z}\}$ in each component C, C^i, $i\in\mathbb{Z}\setminus\{0\}$. The metric $d_{L\llcorner D_L((0,0))}$ is obtained as follows: given two points (p_1,p_2), we can connect them using a path on the same component or by connecting each of them to points θ_1, θ_2 of S, and using one of the C^i to connect these last points.

It follows that $(D_L((0,0)),d_L)$ is geodesic, and a more careful analysis shows that d_L is actually l.s.c.. Moreover, the fact that $\sqrt{\cdot}$ is subadditive yields that there are not geodesic of infinite length: in particular all the assumptions listed on Page 8 are satisfied.

Transport problem. Define the sets

$$A:=\Big\{(x,y)\in C:0\le x<1,\frac{1}{2}<y<\frac{5}{8}\Big\},\quad B:=\Big\{(x,y)\in C:0\le x<1,\frac{7}{8}<y<1\Big\},$$

and the measures $\mu:=\mathcal{L}^2\llcorner_A$, $\nu:=\mathcal{L}^2\llcorner_B$. Consider the two maps defined on A

$$\begin{array}{cccc} T^+ : & A & \to & B\\ & (x,y) & \mapsto & T^+(x,y):=\Big(x+\frac{3}{8}\alpha\mod 1,y+\frac{3}{8}\Big)\end{array}$$

$$\begin{array}{cccc} T^- : & A & \to & B\\ & (x,y) & \mapsto & T^-(x,y):=\Big(x-\frac{5}{8}\alpha\mod 1,y+\frac{3}{8}\Big)\end{array}$$

It is standard to show that $T^\pm_\sharp\mu=\nu$.

Let $(p_i,T^+(p_i))\in\mathrm{graph}(T^+)$, $i=1,\dots,I$: from the definition of d_L, $d_L(p_{i+1},T^+(p_i))$ can be either equal to $d_L(p_i,T^+(p_i))$ or greater than $3\sqrt{1+\alpha^2}$ (by taking the path along $C^-\cup C^1$). Since $d_L(p_i,T^+(p_i))=\frac{3}{2}\sqrt{1+\alpha^2}$, it follows that T^+ is d_L-cyclically monotone.

However, one has

$$\int d_L(x,T^-(x))\mu(dx)=\frac{1}{8}\sqrt{1+\alpha^2}<\frac{3}{2}\frac{1}{8}\sqrt{1+\alpha^2}=\int d_L(x,T^+(x))\mu(dx).$$

Hence the d_L-cyclical monotonicity is not sufficient for optimality. Note that Assumption 2 is verified.

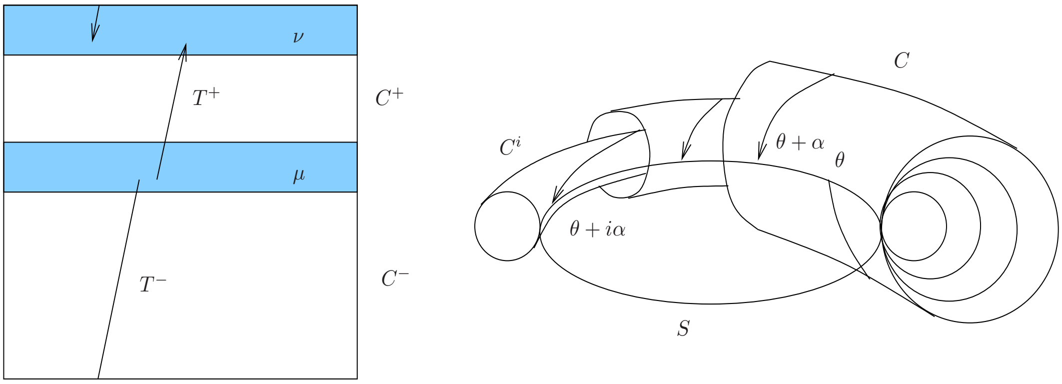

Figure 2. The metric space of Example 2.

Appendix A. Notation

$P_{i_1\dots i_I}$	projection of $x \in \Pi_{k=1,\dots,K} X_k$ into its $(i_1,\dots,i_I)$ coordinates, keeping order
$\mathcal{P}(X)$ or $\mathcal{P}(X,\Omega)$	probability measures on a measurable space (X,Ω)
$\mathcal{M}(X)$ or $\mathcal{M}(X,\Omega)$	signed measures on a measurable space (X,Ω)
$f_{\llcorner A}$	the restriction of the function f to A
$\mu_{\llcorner A}$	the restriction of the measure μ to the σ-algebra $A\cap\Sigma$
$\mathcal{L}^d$	Lebesgue measure on $\mathbb{R}^d$
$\mathcal{H}^k$	k-dimensional Hausdorff measure
$\Pi(\mu_1,\dots,\mu_I)$	$\pi \in \mathcal{P}(\Pi_{i=1}^I X_i, \otimes_{i=1}^I \Sigma_i)$ with marginals $(P_i)_\sharp \pi = \mu_i \in \mathcal{P}(X_i)$
$\mathcal{I}(\pi)$	cost functional (2.7)
c	cost function : $X \times Y \mapsto [0,+\infty$
$\mathcal{I}$	transportation cost (2.7)
ϕ^c	c-transform of a function ϕ (2.8)
$\partial^c\varphi$	d-subdifferential of φ (2.9)
Φ_c	subset of $L^1(\mu)\times L^1(\nu)$ defined in (2.10)
$J(\phi,\psi)$	functional defined in (2.11)
C_b or $C_b(X,\mathbb{R})$	continuous bounded functions on a topological space X
(X,d)	Polish space
(X,d_L)	non-branching geodesic separable metric space
$D_L(x)$	the set $\{y : d_L(x,y) < +\infty\}$
$L(\gamma)$	length of the Lipschitz curve γ, Definition 2.8
$\gamma_{[x,y]}(t)$	geodesics $\gamma:[0,1]\to X$ such that $\gamma(0)=x$, $\gamma(1)=y$
$\gamma_{(x,y)}$, $\gamma_{[x,y]}$	open, closed geodesics (2.6)
$B_r(x)$	open ball of center x and radius r in (X,d)
$B_{r,L}(x)$	open ball of center x and radius r in (X,d_L)
$\mathcal{K}(X)$	space of compact subsets of X
$d_H(A,B)$	Hausdorff distance of A, B w.r.t. the distance d
A_x, A^y	x, y section of $A\subset X\times Y$ (2.3)
$\mathcal{B}$, $\mathcal{B}(X)$	Borel σ-algebra of X Polish
Σ^1_1, $\Sigma^1_1(X)$	the pointclass of analytic subsets of Polish space X, i.e. projection of Borel sets
Π^1_1	the pointclass of coanalytic sets, i.e. complementary of Σ^1_1
Σ^1_n, Π^1_n	the pointclass of projections of Π^1_{n-1}-sets, its complementary
Δ^1_n	the ambiguous class $\Sigma^1_n\cap\Pi^1_n$
$\mathcal{A}$	σ-algebra generated by Σ^1_1
$\mathcal{A}$-function	$f : X\to\mathbb{R}$ such that $f^{-1}((t,+\infty])$ belongs to $\mathcal{A}$
$h_\sharp\mu$	push forward of the measure μ through h, $h_\sharp\mu(A) = \mu(h^{-1}(A))$
graph(F)	graph of a multifunction F (2.1)
F^{-1}	inverse image of multifunction F (2.2)
F_x, F^y	sections of the multifunction F (2.3)

$\mathrm{Lip}_1(X)$	Lipschitz functions with Lipschitz constant 1
Γ'	transport set (3.1)
G, G^{-1}	outgoing, incoming transport ray, Definition 3.2
R	set of transport rays (3.3)
$\mathcal{T}$, $\mathcal{T}_e$	transport sets (3.4)
$a, b : \mathcal{T}_e \to \mathcal{T}_e$	endpoint maps (3.7)
$\mathcal{Z}_{m,e}$, $\mathcal{Z}_m$	partition of the transport set Γ (4.1), (4.2)
$\mathcal{S}$	cross-section of $R\llcorner_{\mathcal{T}\times\mathcal{T}}$
$g = g^+ \cup g^-$	ray map, Definition 4.5
A_t	evolution of $A \subset \mathcal{Z}_{k,i,j}$ along geodesics (5.1)
$\dot g$	current on (X, d) corresponding to the flow along geodesics, Definition 7.1
$\partial \dot g$	boundary of the current $\dot\gamma$ (7.1)

REFERENCES

[1] L. Ambrosio and B. Kirchheim. Currents in metric spaces. *Acta Mathematica*, 185(1):1–80, 2000.

[2] L. Ambrosio, B. Kirchheim, and A. Pratelli. Existence of optimal transport maps for crystalline norms. *Duke Math. J.*, 125(2):207–241, 2004.

[3] S. Bianchini and L. Caravenna. On the extremality, uniqueness and optimality of transference plans. preprint.

[4] S. Bianchini and M. Gloyer. The Euler-Lagrange equation for a singular variational problem. *Math. Z.*, 2009.

[5] D. Burago, Y. Burago, and S. Ivanov. *A course in metric geometry*. Graduate studies in mathematics. AMS, 2001.

[6] L. Caffarelli, M. Feldman, and R.J. McCann. Constructing optimal maps for Monge's transport problem as a limit of strictly convex costs. *J. Amer. Math. Soc.*, 15:1–26, 2002.

[7] L. Caravenna. An existence result for the Monge problem in $\mathbb{R}^n$ with norm cost function. *preprint*, 2009.

[8] T. Champion and L. De Pascale. The Monge problem in $\mathbb{R}^d$. *preprint*, 2009.

[9] L.C. Evans and W. Gangbo. Differential equations methods for the Monge-Kantorovich mass transfer problem. *Current Developments in Mathematics*, pages 65–126, 1997.

[10] M. Feldman and R. McCann. Monge's transport problem on a Riemannian manifold. *Trans. Amer. Math. Soc.*, 354:1667–1697, 2002.

[11] D. H. Fremlin. *Measure Theory*, volume 4. Torres Fremlin, 2002.

[12] N. Juillet. Ricci curvature bounds and geometric inequalities in the Heisenberg group. *preprint*, 2009.

[13] D.G. Larman. A compact set of disjoint line segments in $\mathbb{R}^3$ whose end set has positive measure. *Mathematika*, 18:112–125, 1971.

[14] S.-I. Ohta. On the measure contraction property of metric measure spaces. *Comment. Math. Helv.*, 82:805–828, 2007.

[15] A. M. Srivastava. *A course on Borel sets*. Springer, 1998.

[16] K.T. Sturm. On the geometry of metric measure spaces. *Acta Mathematica*, 2004.

[17] K.T. Sturm. On the geometry of metric measure spaces ii. *Acta Mathematica*, 2004.

[18] V.N. Sudakov. Geometric problems in the theory of dimensional distributions. *Proc. Steklov Inst. Math.*, 141:1–178, 1979.

[19] N. Trudinger and X.J. Wang. On the Monge mass transfer problem. *Calc. Var. PDE*, 13:19–31, 2001.

[20] C. Villani. *Optimal transport, old and new*. Springer, 2008.

SISSA, VIA BEIRUT 2, IT-34014 TRIESTE (ITALY)
E-mail address: `bianchin@sissa.it`

SISSA, VIA BEIRUT 2, IT-34014 TRIESTE (ITALY)
E-mail address: `cavallet@sissa.it`