Geometric properties of disintegration of measures

RENATA POSSOBON⬤ and CHRISTIAN S. RODRIGUES⬤

Institute of Mathematics, Department of Applied Mathematics,
Universidade Estadual de Campinas, 13.083-859 Campinas, SP, Brazil
(e-mail: re.possobon@gmail.com, rodrigues@ime.unicamp.br)

Abstract. In this paper, we study a connection between disintegration of measures and geometric properties of probability spaces. We prove a disintegration theorem, addressing disintegration from the perspective of an optimal transport problem. We look at the disintegration of transport plans, which are used to define and study disintegration maps. Using these objects, we study the regularity and absolute continuity of disintegration of measures. In particular, we exhibit conditions for which the disintegration map is weakly continuous and one can obtain a path of measures given by this map. We show a rigidity condition for the disintegration of measures to be given into absolutely continuous measures.

Key words: disintegration theorem, probability, optimal transport, ergodic theory, dynamical systems
2020 Mathematics Subject Classification: 37C40, 49K45, 49N60 (Primary); 37H10, 37C05 (Secondary)

1. Introduction

The disintegration of a measure over a partition of the space on which it is defined is a way to rewrite this measure as a combination of probability measures, which are concentrated on the elements of the partition. As an example, consider the probability space $(X, \mathcal{F}, \mu)$ and its partition into a finite number of measurable subsets $P_1, \ldots, P_n$ with positive measure. A disintegration of μ with respect to (w.r.t.) this partition is a family of probabilities $\{\mu_1, \ldots \mu_n\}$ on X such that for $i = 1, \ldots, n$, we have $\mu_i(P_i) = 1$ and, for every measurable set $E \subset X$, the conditional measures are given by $\mu_i(E) = \mu(E \cap P_i)/\mu(P_i)$. It is possible to write the original measure as a combination of the conditional ones

$$\mu(E) = \sum_{i=1}^{n} \mu(P_i)\mu_i(E) = \sum_{i=1}^{n} \mu(P_i)\frac{\mu(E \cap P_i)}{\mu(P_i)}.$$

More generally, consider a probability space $(X, \mathcal{F}, \mu)$ and a partition $\mathcal{P}$ of X into measurable subsets. Let ϱ be the natural projection that associates each point $x \in X$ to the element $P \in \mathcal{P}$ which contains x. The measurable function ϱ can be used to induce a probability $\hat{\mu}$ on $\mathcal{P}$. A subset $B \subset \mathcal{P}$ is measurable if and only if $\varrho^{-1}(B)$ is a measurable subset of X. Then, the family $\hat{\mathcal{B}}(\mathcal{P})$ of measurable subsets is a σ-algebra on $\mathcal{P}$. Let $\hat{\mu}$ denote the measure given by

$$\hat{\mu}(B) = \varrho_* \mu(B) := \mu \circ \varrho^{-1}(B) = \mu(\{x \in X : \varrho(x) \in B\})$$

for every $B \in \hat{\mathcal{B}}(\mathcal{P})$. In this case, $\hat{\mu}$ is called the *law* of ϱ, denoted by law(ϱ). A disintegration of μ w.r.t. $\mathcal{P}$ into conditional measures is a family $\{\mu_P : P \in \mathcal{P}\}$ of probability measures on X such that for every $E \in \mathcal{F}$:

(1) $\mu_P(P) = 1$ for $\hat{\mu}$-almost every (a.e.) $P \in \mathcal{P}$;
(2) $P \mapsto \mu_P(E)$ is measurable;
(3) $\mu(E) = \int \mu_P(E) \, d\hat{\mu}(P)$.

There are several reasons why one may wish to study such possible combinations of measures. In ergodic theory, for example, the disintegration of a measure is directly related to the ergodic decomposition of invariant measures, which are crucial objects encoding the asymptotic behaviour of dynamical systems [OV14]. The concept of disintegration, however, appears in much broader context in areas, such as probability [CP97, Par67] and geometry [Stu06], among others [BM17, GL20, Var16, Vil03].

The idea of disintegrating a measure was devised by Von Neumann in 1932 [Von32]. Since then, different versions of disintegration theorems have been presented and used, for example, in [AGS05, AP03, DM78, Tue75], to name a few. In particular, the well-known Rokhlin disintegration theorem shows that there exists a disintegration of μ relative to $\mathcal{P}$ if X is a complete separable metric space and $\mathcal{P}$ is a measurable partition. By measurable partition, we mean that there exists some measurable set $X_0 \subset X$ such that $\mu(X) = \mu(X_0)$ and $\mathcal{P} = \bigvee_{n=1}^{\infty} \mathcal{P}_n = \{P_1 \cap P_2 \cap \cdots : P_n \in \mathcal{P}_n \text{ for all } n \geq 1\}$ restricted to X_0, for an increasing sequence $\mathcal{P}_1 \prec \mathcal{P}_2 \prec \cdots \prec \mathcal{P}_n \prec \cdots$ of countable partitions [Rok52].

Recently, Simmons has proposed a more general and subtle formulation of Rokhlin's disintegration theorem, where he has considered any universally measurable space $(X, \mathcal{B}, \mu)$ and a measure space Y for which there exists an injective map $Y \to \{0, 1\}^{\mathbb{N}}$. That is, Y is any subspace of the standard Borel space. He has shown that there is a (unique) system of conditional measures $(\mu_y)_{y \in Y}$, a disintegration of μ [Sim12]. Then, his formulation is further developed to address σ-finite measure spaces with absolutely continuous morphisms. One of the facts standing out in Simmons' formulation is a fibre-wise perspective, which we wish to further explore.

Even though geometric properties of the space where a measure is defined and statistical properties obtained via disintegration theorems seem to be strictly connected, for instance, in foliated manifolds, very little geometric information is taken into account while studying disintegration of measures. In particular, intrinsic geometric properties of probability spaces are very often neglected. The purpose of this paper is to advertise the viewpoint of tackling disintegration of measures taking into consideration intrinsic structures of probability spaces obtained from optimal transport theory. To this end, we will formulate

disintegration of probability measures in terms of a transportation problem to explore a fibre-wise formulation of a disintegration and its consequences.

1.1. *Main results.* As a first result in this paper, we prove a disintegration theorem, Theorem A, and we introduce a fibre-wise perspective on disintegration. Using this disintegration theorem, we study conditions in which one can obtain a path of conditional measures in the space of probability measures. In particular, in Propositions 5.2, 5.4 and 5.8, we investigate the weak continuity of the disintegration map, which parametrizes the conditional measures. The last result in this paper is Theorem B. Its first part shows how to construct a path of conditional measures in the space of probabilities. Its second part gives us a sort of rigidity result for disintegration of measures. Namely, we show that if one of the measures in this path is absolutely continuous, then all measures in the associated path must also be absolutely continuous. The third part is a particular case of absolute continuity in which the disintegration map is an isometry.

The paper is organized as follows. Section 2 contains the main concepts from optimal transport theory to be used throughout the paper. In §3, a disintegration theorem, stated as Theorem A, is proved. In §4, Theorem A is used to define and study what we called disintegration maps. These crucial objects are used in §5. There we proved a series of propositions about a disintegration map: in Proposition 5.2, we show that this map is nearly weakly continuous, under some assumptions on the reference measure ν. In Propositions 5.4 and 5.8, we study hypotheses about the disintegration for which this map is weakly continuous. Afterwards, we prove our main result in this section, Theorem B, about paths of measures given via disintegration maps and a rigidity condition establishing absolute continuity of measures in these paths.

2. *Spaces of probability measures, Wasserstein spaces and optimal transport*

In this section, we set up some notation to be used throughout the text. We also introduce some basic terminology from optimal transport theory, which is meant for readers not familiar with this area. Those who are skilled on the topic may wish to skip this section. Our main references for this section are [**AGS05**, **Amb00**, **Vil09**].

The cornerstone for the theory of optimal transportation is considered to be a logistic problem addressed by Gaspar Monge in 1781. The main idea is to transport masses from a given location to another one at minimal cost. To state it in a modern formulation, let $\mathscr{P}(X)$ be the set of all Borel probability measures on X. The problem amounts to the following.

Monge transport problem. Let X, Y be Radon spaces. Given measures $\mu \in \mathscr{P}(X)$, $\nu \in \mathscr{P}(Y)$ and a fixed Borel cost function $c : X \times Y \to [0, \infty]$, minimize

$$T \mapsto \int_X c(x, T(x)) \, d\mu$$

among all maps T such that $T_*\mu = \nu$.

The maps T fulfilling $T_*\mu = \nu$ are called *transport maps*. The Monge transport problem actually may be ill-posed and such a map does not need to exist. That is the case, for

example, when one of the measures is a Dirac mass and the other one is not. A way around is given by a different formulation as proposed by Kantorovich.

Monge–Kantorovich transport problem. Let X, Y be Radon spaces. Given measures $\mu \in \mathscr{P}(X)$, $\nu \in \mathscr{P}(Y)$ and a fixed Borel cost function $c : X \times Y \to [0, \infty]$, minimize

$$\gamma \mapsto \int_{X \times Y} c(x, y) \, d\gamma(x, y) \tag{1}$$

among all measures $\gamma \in \mathscr{P}(X \times Y)$ with *marginals* μ and ν, i.e. γ satisfying $(\text{proj}_X)_* \gamma = \mu$ and $(\text{proj}_Y)_* \gamma = \nu$, where proj_X and proj_Y are the canonical projections $(x, y) \mapsto x$ and $(x, y) \mapsto y$, respectively.

The measures $\gamma \in \mathscr{P}(X \times Y)$ are called *transport plans*. We denote the set of all transport plans with marginals μ and ν by $\Pi(\mu, \nu)$. The value $C(\mu, \nu) = \inf_{\gamma \in \Pi(\mu, \nu)} \int_{X \times Y} c(x, y) \, d\gamma(x, y)$ is called *optimal cost*. Note that $(\text{proj}_X)_* \gamma = \mu$ is equivalent to $\gamma[A \times Y] = \mu[A]$ for every $A \in \mathcal{B}(X)$, and $(\text{proj}_Y)_* \gamma = \nu$ is equivalent to $\gamma[X \times B] = \nu[B]$ for every $B \in \mathcal{B}(Y)$. Moreover, it is possible to describe this problem in terms of the coupling of measures, in the following sense.

Definition 2.1. Let (X, μ) and (Y, ν) be probability spaces. A *coupling* of (μ, ν) is a pair $(\mathcal{X}, \mathcal{Y})$ of measurable functions in a probability space $(\Omega, \mathbb{P})$ such that $\text{law}(\mathcal{X}) = \mu$ and $\text{law}(\mathcal{Y}) = \nu$.

Considering $\Omega = X \times Y$, coupling μ and ν means to construct $\gamma \in \mathscr{P}(X \times Y)$ with marginals μ and ν. Then, the Monge–Kantorovich transport problem can be understood as the minimization of the total cost, over all possible couplings of (μ, ν).

Whenever these problems are stated in metric spaces, we may choose the cost function to be the distance function itself, which in turn allows us to introduce a distance function between measures.

Definition 2.2. (Wasserstein distance) Let (X, d) be a separable complete metric space. Consider probability measures μ and ν on X and $p \in [1, \infty)$. The *Wasserstein distance* of order p between μ and ν is given by

$$W_p(\mu, \nu) := \left(\inf_{\gamma \in \Pi(\mu, \nu)} \int d(x_1, x_2)^p \, d\gamma(x_1, x_2) \right)^{1/p} .$$

In general, W_p is not a distance in the strict sense, because it can take the value $+\infty$. To rule this situation out, it is natural to constrain W_p to a subset in which it takes finite values.

Definition 2.3. (Wasserstein space) The *Wasserstein space* of order p is defined by

$$\mathscr{P}_p(X) := \left\{ \mu \in \mathscr{P}(X) : \int d(x, \tilde{x})^p \mu(dx) < +\infty \right\}$$

for $\tilde{x} \in X$ arbitrary.

Therefore, W_p sets a (finite) distance on $\mathscr{P}_p(X)$. It turns out that if (X, d) is a complete separable metric space and d is bounded, then the p-Wasserstein distance metrizes the weak topology over $\mathscr{P}(X)$ [**Vil09**, Corollary 6.13]. Furthermore, if X is a complete metric space, then so is $\mathscr{P}(X)$ with the p-Wasserstein distance [**Vil09**, Theorem 6.18].

Although the Wasserstein distance is defined for every $p \geq 1$, in this paper, we will chose either $p = 1$ or $p = 2$, as stated later on. The reason is that for W_1, we can explicitly compute distance bounds, while for W_2, the space $\mathscr{P}_2(X)$ inherits geometric properties of the space X. The choice is always indicated throughout the text.

We finish this section recalling two well-known results for later use. The first one is regarding measurable and continuous functions, the Lusin theorem in a specific form. The other one gives us a 'recipe' to glue different couplings.

THEOREM 2.4. [**Fed69**, Theorem 2.3.5] *Let M be a locally compact metric space and N a separable metric space. Consider μ a Borel measure on M, $A \subset M$ a measurable set with finite measure and $f : M \to N$ a measurable map. Then, for each $\delta > 0$, there is a closed set $K \subset A$, with $\mu(A \backslash K) < \delta$, such that the restriction of f to K is continuous.*

LEMMA 2.5. (Gluing lemma) [**Vil09**, Ch. 1] *Let (X_i, μ_i), $i = 1, 2, 3$, be complete separable metric probability spaces. If $(\mathcal{X}_1, \mathcal{X}_2)$ is a coupling of (μ_1, μ_2) and $(\mathcal{Y}_2, \mathcal{Y}_3)$ is a coupling of (μ_2, μ_3), then one can construct a triple of random variables $(\mathcal{Z}_1, \mathcal{Z}_2, \mathcal{Z}_3)$ such that $(\mathcal{Z}_1, \mathcal{Z}_2)$ has the same law as $(\mathcal{X}_1, \mathcal{X}_2)$, and $(\mathcal{Z}_2, \mathcal{Z}_3)$ has the same law as $(\mathcal{Y}_2, \mathcal{Y}_3)$. If μ_{12} stands for the law of $(\mathcal{X}_1, \mathcal{X}_2)$ on $X_1 \times X_2$ and μ_{23} stands for the law of $(\mathcal{X}_2, \mathcal{X}_3)$ on $X_2 \times X_3$, then to construct the joint law μ_{123} of $(\mathcal{Z}_1, \mathcal{Z}_2, \mathcal{Z}_3)$, one just has to glue μ_{12} and μ_{23} along their common marginal μ_2.*

3. Disintegration of measures

To grasp some of the properties of the probability spaces while studying disintegration of measures, we would like to associate the latter with the optimal transport theory. Before doing so, we prove the following disintegration theorem. Our proof is based on the idea of choosing a dense subset of a vector space using separability. Then, we extend a tailor-made linear functional to the whole space, which is implicitly used in the proof of [**DM78**, III-70], although our conditions are different.

THEOREM A. *Let X and Y be locally compact and separable metric spaces. Let $\pi : X \to Y$ be a Borel map and take $\mu \in \mathcal{M}_+(X)$, where $\mathcal{M}_+(X)$ is the set of all positive and finite Radon measures on X. Define $\nu = \pi_* \mu$ in $\mathcal{M}_+(Y)$. Then, there exist measures $\mu_y \in \mathcal{M}_+(X)$ such that:*

(1)   *$y \mapsto \mu_y$ is a Borel map and $\mu_y \in \mathscr{P}(X)$ for ν-a.e. $y \in Y$;*
(2)   *$\mu = \nu \otimes \mu_y$, that is, $\mu(A) = \int_Y \mu_y(A) \, d\nu(y)$ for every $A \in \mathcal{B}(X)$;*
(3)   *μ_y is concentrated on $\pi^{-1}(y)$ for ν-a.e. $y \in Y$.*

Proof. We shall first consider the disintegration of measures on compact metric spaces. Then, we tackle the general case as it is stated.

Step 1: To get started, consider X to be a compact metric space with its Borel σ-algebra $\mathcal{B}(X)$ and let μ be a Radon measure on X. If $(Y, \mathcal{E})$ is a measurable space, we define

a measurable map $q : \mathcal{B}(X) \to \mathcal{E}$ so that we set $\nu = q_* \mu$. Let $C(X)$ be the set of all continuous real functions $\omega : X \to \mathbb{R}$. Then, for each $\omega \in C(X)$, we associate a measure λ given by

$$\lambda(A) = \int_{q^{-1}(A)} \omega(x) \, d\mu(x)$$

for every $A \in \mathcal{E}$. The measure λ is absolutely continuous w.r.t. ν. Indeed, for every $A \in \mathcal{E}$, we have that $\nu(A) = \mu(q^{-1}(A)) = 0$ implies $\lambda(A) = 0$. Therefore, since ν and λ are positive measures and $\lambda \ll \nu$, by the Radon–Nikodym theorem, there exists $h : Y \to [0, \infty]$, the density of λ w.r.t. ν, such that $\lambda(A) = \int_A h \, d\nu$ for $A \in \mathcal{E}$. Thus,

$$\int_A h(y) \, d\nu(y) = \int_{q^{-1}(A)} \omega(x) \, d\mu(x). \tag{2}$$

Recall that $C(X)$ is a separable space with the supremum norm. Let $\mathcal{H} = \{\omega_1 \equiv 1, \omega_2, \omega_3, \ldots\}$ be a dense subset of $C(X)$. Suppose, without loss of generality, that $\mathcal{H}$ is a vector space over $\mathbb{Q}$. Then, for each $n \in \mathbb{N}$, we consider the Radon–Nikodym density h_n associated with ω_n given by equation (2) so that for each $A \in \mathcal{E}$,

$$\int_A h_n(y) \, d\nu(y) = \int_{q^{-1}(A)} \omega_n(x) \, d\mu(x). \tag{3}$$

Note that $h_n \geq 0$ almost always (a.a.), if $\omega_n \geq 0$, since $\nu = q_* \mu$.

Step 2: We will use the associated densities to construct a linear functional which will be extended using the Hahn–Banach theorem. To do so, we denote by $\mathbb{A}$ the set of all $y \in Y$, such that, if $\omega_i = \alpha \omega_j + \beta \omega_k$, then we have for their associated densities that the relation $h_i(y) = \alpha h_j(y) + \beta h_k(y)$ holds true, where $\alpha, \beta \in \mathbb{Q}$ and the associated density to ω_1 is set to $h_1(y) = 1$. The set $\mathbb{A}$ is measurable and $\nu(\mathbb{A}) = 1$. Indeed,

$$\int_{q^{-1}(Y)} \omega_i(x) \, d\mu = \int_{q^{-1}(Y)} (\alpha \omega_j + \beta \omega_k)(x) \, d\mu$$

$$= \alpha \left(\int_{q^{-1}(Y)} \omega_j(x) \, d\mu \right) + \beta \left(\int_{q^{-1}(Y)} \omega_k(x) \, d\mu \right), \text{ which by equation (3)},$$

$$= \alpha \left(\int_Y h_j(y) \, d\nu \right) + \beta \left(\int_Y h_k(y) \, d\nu \right)$$

$$= \int_Y \alpha h_j(y) + \beta h_k(y) \, d\nu,$$

so $\int h_i(y) \, d\nu = \int \alpha h_j(y) + \beta h_k(y) \, d\nu$. Consequently, $h_i(y) = \alpha h_j(y) + \beta h_k(y) \nu$-a.a. Therefore, whenever ω_n is a linear combination of elements of $\mathcal{H}$, then the associated densities defined by equation (3) can also be written point-wise as linear combinations of Radon–Nikodym densities.

Step 3: For each $y \in \mathbb{A}$, we define the functional $\tilde{\varphi}_y : \mathcal{H} \to \mathbb{R}$, given by $\tilde{\varphi}_y(\omega_n) := h_n(y)$. Note that $\tilde{\varphi}_y : \mathcal{H} \to \mathbb{R}$ is $\mathbb{Q}$-linear with $\|\tilde{\varphi}_y\| \leq 1$ and, since $\tilde{\varphi}_y(1) = 1$, we have actually that $\|\tilde{\varphi}_y\| = 1$. Therefore, by the Hahn–Banach theorem, $\tilde{\varphi}_y$ can be

extended to a continuous positive linear functional $\varphi_y : C(X) \to \mathbb{R}$, with $\|\varphi_y\| = 1$. Furthermore, the Riesz–Markov–Kakutani representation theorem assures that there exists a unique Radon measure μ_y on X such that $\varphi_y(\omega) = \int \omega \, d\mu_y$ for every $\omega \in C(X)$ and $\varphi_y(1) = 1$. Note that $\int \omega \, d\mu_y \leq 1$, since $\|\varphi_y\| = 1$. Thus, we conclude that μ_y is a probability measure. Note also that μ_y is supported on $q^{-1}\{y \in \mathbb{A}\}$. For $y \notin \mathbb{A}$, consider $\mu_y = 0$.

Step 4: Observe that $y \mapsto \int_X \omega_n \, d\mu_y$ is $\mathcal{E}$-measurable for every $\omega_n \in \mathcal{H}$ and

$$\int_Y \int_{q^{-1}(y)} \omega_n(x) \, d\mu_y \, d\nu = \int_Y \varphi_y(\omega_n) \, d\nu = \int_Y h_n(y) \, d\nu = \int_X \omega_n(x) \, d\mu.$$

By the definition of $\mathcal{H}$, we have that for each $\omega \in C(X)$, there exists a sequence $(\omega_i)_i$, with $\omega_i \to \omega$ uniformly. So, by uniform convergence, we have that $y \mapsto \int_X \omega \, d\mu_y$ is $\mathcal{E}$-measurable for every $\omega \in C(X)$. Furthermore,

$$\int_Y \int_{q^{-1}(y)} \omega(x) \, d\mu_y \, d\nu = \int_X \omega(x) \, d\mu.$$

The same holds true for any bounded and $\mathcal{B}(X)$-measurable function ω. Indeed, denote by $\mathscr{C}$ the class of functions such that $y \mapsto \int_X \omega \, d\mu_y$ is $\mathcal{E}$-measurable and $\int_Y \int_{q^{-1}(y)} \omega(x) \, d\mu_y \, d\nu = \int_X \omega(x) \, d\mu$. Note that $C(X) \subset \mathscr{C}$, from what was shown before. If $A \subset X$ is an open set, then $\mathbb{1}_A \in \mathscr{C}$. So, let $\mathcal{D}$ be the set of all characteristic functions in $\mathscr{C}$. If $\mathbb{1}_{A_n} \in \mathcal{D}$ for $n \in \mathbb{N}$, we have that $\mathbb{1}_{\cup A_n} \in \mathcal{D}$. If $\mathbb{1}_A \in \mathcal{D}$, we have that $\mathbb{1}_{A^c} \in \mathcal{D}$. Thus, the class of measurable sets whose characteristic functions are in $\mathcal{D}$ is $\mathcal{B}(X)$. Therefore, the result follows by monotone convergence. This completes the proof of Theorem A for compact spaces.

Step 5: Let X be a locally compact and separable metric space. Then, the set of continuous real-valued functions with compact support on X, denoted by $C_c(X)$, is a vector space. Such a vector space can be seen as the union of the spaces $C_c(K_i)$ of continuous functions with support on compact sets K_i. Since μ is a Radon measure on X, the map $\varphi : C_c(X) \to \mathbb{R}$, such that, $\omega \mapsto \int_X \omega(x) \, d\mu$, is a continuous positive linear map. Note also that μ is supported on a set $\tilde{\mathcal{K}}$, which is a countable union of compact subsets $K_i \subset X$. So, we can imbed X into a compact metric space $\mathcal{K}$ and identify μ with a measure on $\mathcal{K}$ with support on $\tilde{\mathcal{K}}$, and construct the measures μ_y as above. Consider $\mu_y = 0$ for y such that $\pi^{-1}(y) \notin \tilde{\mathcal{K}}$. Thus, there exist probability measures μ_y on X such that each μ_y is supported on $\pi^{-1}(y)$ and, for every $\omega \in C_c(X)$,

$$\int_Y \int_{\pi^{-1}(y)} \omega(x) \, d\mu_y \, d\nu = \int_X \omega(x) \, d\mu. \tag{4}$$

Since Y is a locally compact and separable metric space and $\pi : X \to Y$ is a Borel map, then $y \mapsto \mu_y$ is a Borel map for ν-a.e. $y \in Y$. Furthermore, note that equation (4) is equivalent to say that $\mu(A) = \int_Y \mu_y(A) \, d\nu(y)$ for every $A \in \mathcal{B}(X)$ and μ_y is concentrated on $\pi^{-1}(y)$ for ν-a.e. $y \in Y$. This concludes the proof. $\square$

Many interesting examples arise when we consider Theorem A for the case of product spaces with the Borel map π as the canonical projection on the first component, as follows.
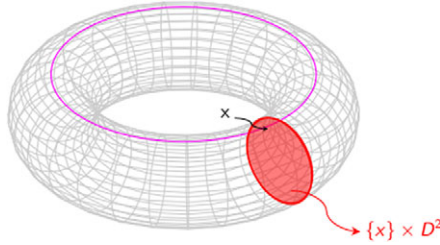
FIGURE 1. Representation of $\mathcal{F}^s = \{\{x\} \times D^2\}_{x \in S^1}$.

COROLLARY 3.1. *Let X and Y be locally compact and separable metric spaces. Let* $\mathrm{proj}_X : X \times Y \to X$ *be the canonical projection on the first component, take* $\gamma \in \mathcal{M}_+(X \times Y)$ *and set* $\mu = \mathrm{proj}_{X*}\gamma \in \mathcal{M}_+(X)$. *Then, there exist measures* $\gamma_x \in \mathcal{M}_+(X \times Y)$ *such that:*

(1) $x \mapsto \gamma_x$ *is a Borel map and* $\gamma_x \in \mathscr{P}(X \times Y)$ *for μ-a.e. $x \in X$;*
(2) $\gamma = \mu \otimes \gamma_x$, *i.e.* $\gamma(A) = \int_X \gamma_x(A)\, d\mu(x)$ *for every $A \in \mathcal{B}(X \times Y)$;*
(3) γ_x *is concentrated on* $\mathrm{proj}_X^{-1}(x)$ *for μ-a.e. $x \in X$.*

In fact, since γ_x is concentrated on $\mathrm{proj}_X^{-1}(x) = \{x\} \times Y$, we can consider each γ_x as a measure on Y, writing $\gamma(B) = \int_X \gamma_x(\{y : (x, y) \in B\})\, d\mu$ for every $B \in \mathcal{B}(X \times Y)$, adding a different point of view to disintegration. The following example illustrates this case of disintegration of measures.

Example 3.2. Consider a solid torus $S^1 \times D^2$. Let $\mathcal{F}^s = \{\{x\} \times D^2\}_{x \in S^1}$ be a foliation of $S^1 \times D^2$, as represented in Figure 1. Given a measure $\gamma \in \mathscr{P}(S^1 \times D^2)$, let $\mathrm{proj}_{S^1} : S^1 \times D^2 \to S^1$ be the canonical projection on the first component and set $\mu = \mathrm{proj}_{S^1*}\gamma$. Theorem A, with $\pi := \mathrm{proj}_{S^1}$, gives us a disintegration $\{\gamma_x : x \in S^1\}$ of γ along the leaves. Since the measures γ_x are concentrated on $\mathrm{proj}_{S^1}^{-1}(x) = \{x\} \times D^2$ for μ-a.e. $x \in S^1$, we can consider each γ_x as a measure on D^2. That is, we can define a probability on D^2 for each $x \in S^1$.

The point of view from Corollary 3.1 is, somehow, a generalization of cases of disintegration of a probability measure along leaves in a foliated compact Riemannian manifold, as in Example 3.2. In this sense, we remark that different versions of disintegration theorems can be related by taking suitable hypotheses, as we do in the following example.

Example 3.3. Let M_1 and M_2 be compact Riemannian manifolds and set the product space $\Sigma = M_1 \times M_2$. Let $\mathcal{F}^s = \{\{x\} \times M_2\}_{x \in M_1}$ be a foliation of Σ. Given $\gamma \in \mathscr{P}(M_1 \times M_2)$, let $\mathrm{proj}_{M_1} : \Sigma \to M_1$ be the canonical projection on M_1 and set $\mu = \mathrm{proj}_{M_1*}\gamma$. By Theorem A, there exists a family $\{\gamma_x : x \in M_1\} \subset \mathscr{P}(M_1 \times M_2)$ such that:

(1) $x \mapsto \gamma_x$ is a Borel map and $\gamma_x \in \mathscr{P}(M_1 \times M_2)$ for μ-a.e. $x \in M_1$;
(2) $\gamma(A) = \int_{M_1} \gamma_x(A)\, d\mu(x)$ for every $A \in \mathcal{B}(M_1 \times M_2)$;
(3) γ_x is concentrated on $\mathrm{proj}_{M_1}^{-1}(x)$ for μ-a.e. $x \in M_1$.

Furthermore, we can consider each γ_x as a probability on M_2, as stated above. Note that this result agrees with Rokhlin's disintegration theorem. In fact, let $\varrho : \Sigma \to \mathcal{F}^s$ be a map that associates each point $(x, y) \in \Sigma$ to the ζ element of $\mathcal{F}^s$ that contains (x, y). Consider $\hat{\gamma} = \varrho_* \gamma$. Note that $\mathcal{F}^s$ is a measurable partition of Σ and Σ is a complete separable metric space. So, Rokhlin's disintegration theorem describes a disintegration of γ relative to $\mathcal{F}^s$ by a family $\{\gamma_\zeta : \zeta \in \mathcal{F}^s\}$, such that, for $E \subset \Sigma$ measurable set:

(1) $\gamma_\zeta(\zeta) = 1$ for $\hat{\gamma}$-a.e. $\zeta \in \mathcal{F}^s$;

(2) $\zeta \mapsto \gamma_\zeta(E)$ is measurable;

(3) $\gamma(E) = \int \gamma_\zeta(E) \, d\hat{\gamma}(\zeta)$.

Rewrite $\mathcal{F}^s$ by $\{\zeta_x\}_{x \in M_1}$, where $\zeta_x = \{x\} \times M_2$ for each $x \in M_1$, and consider $\gamma'_x = \text{proj}_{M_1*} \gamma$. For each $x \in M_1$, let γ_{ζ_x} be the restriction of γ_ζ to $\zeta = \{x\} \times M_2$. Note that $\varrho^{-1}(\zeta) = \{x\} \times M_2 = \zeta_x = \text{proj}_{M_1}^{-1}(x)$, so

$$\int \gamma_\zeta(E) d\hat{\gamma}(\zeta) = \int \gamma_\zeta(E) \gamma(\varrho^{-1}(d\zeta))$$

$$= \int_{M1} \gamma_{\zeta_x}(E \cap \zeta_x) \gamma(\text{proj}_{M_1}^{-1}(dx))$$

$$= \int_{M1} \gamma_{\zeta_x}(E \cap \zeta_x) \, d\gamma'_x$$

and then

$$\gamma(E) = \int_{M1} \gamma_{\zeta_x}(E \cap \zeta_x) \, d\gamma'_x.$$

Moreover, note that γ_{ζ_x} is supported in $\text{proj}_{M_1}^{-1}(x) = \zeta_x$. Hence, we have a disintegration $\{\gamma_{\zeta_x} : \zeta_x \in \mathcal{F}^s\}$ along the leaves of Σ associated to γ'_x. In addition, it is possible to identify $\text{proj}_{M_1}^{-1}(x)$ with M_2, and write this disintegration by $\{\gamma_x : x \in M_1\} \subset \mathscr{P}(M_2)$, as desired.

Such an example is one of the possible roles of disintegration of measures. In dynamics, for instance, this kind of disintegration appears in several contexts. To name a few, in [**BM17**], the regularity of this kind of disintegration is investigated considering invariant measures for hyperbolic skew products. Specifically, for this purpose, a function that associates each x in X with a probability measure γ_x obtained via a disintegration of γ is analysed. In the next section, we will see that this type of application can be thought in a more general framework and it has important properties. We can also cite [**Gal17**, **GL20**], where the disintegration of Example 3.2 is used to study the behaviour of the transfer operator in a solenoidal map.

We can actually obtain uniqueness and absolute continuity of the disintegration in the context of Theorem A. If $\lambda \in \mathcal{M}_+(Y)$ is another measure, such that there exists a Borel map $y \mapsto \eta_y$ for which $\mu = \lambda \otimes \eta_y$, with η_y concentrated on $\pi^{-1}(y)$ for μ-a.e. $y \in Y$, then $\lambda|_C \ll \nu$, where $C = \{y \in Y : \eta_y(X) > 0\}$ and $\eta_y \ll \mu_y$ for ν-a.e. $y \in Y$. See the details in [**AP03**]. In the following proposition, we focus on the case of product spaces. Taking $\gamma \in \mathcal{M}_+(X \times Y)$ with $\mu = \text{proj}_{X*} \gamma$, we obtain uniqueness of γ_x and μ in $\gamma = \mu \otimes \gamma_x$.

PROPOSITION 3.4. *Let $X \times Y$, X be locally compact and separable metric spaces. Let $\mathrm{proj}_X : X \times Y \to X$ be the projection on the first component and consider $\gamma \in \mathcal{M}_+(X \times Y)$, $\nu \in \mathcal{M}_+(X)$ and $\mu = \mathrm{proj}_{X*}\gamma$. Let $x \mapsto \eta_x$ be a Borel $\mathcal{M}_+$-valued map defined on X such that:*

(1) $\gamma = \nu \otimes \eta_x$;

(2) η_x *is concentrated on* $\mathrm{proj}_X^{-1}(x)$ *for ν-a.e. $x \in X$.*

Then, the η_x are uniquely defined ν-a.a. by conditions (1) and (2). Moreover, for $C = \{x \in X : \eta_x(X \times Y) > 0\}$, $\nu|_C$ is absolutely continuous w.r.t. μ. In particular, $(\nu|_C/\mu)\eta_x = \gamma_x$ for μ-a.e. $x \in X$, where γ_x are the conditional probabilities as in Corollary 3.1.

Proof. Let η_x and η_x' be measures satisfying conditions (1) and (2). Let $(A_n)_n$ be a sequence of open sets such that the finite intersection is also an open set which generates $\mathcal{B}(X \times Y)$, the Borel σ-algebra of $X \times Y$. Consider $B \in \mathcal{B}(X \times Y)$ and $A = A_n \cap \mathrm{proj}_X^{-1}(B)$ for any $n \in \mathbb{N}$. By condition (1), we have that

$$\gamma(A) = \int_X \eta_x(A) \, d\nu = \int_B \eta_x(A_n) \, d\nu$$

and

$$\gamma(A) = \int_X \eta_x'(A) \, d\nu = \int_B \eta_x'(A_n) \, d\nu.$$

Therefore,

$$\int_B \eta_x(A_n) \, d\nu = \int_B \eta_x'(A_n) \, d\nu.$$

Given that B is arbitrary, then $\eta_x(A_n) = \eta_x'(A_n)$ for ν-a.e. $x \in X$ and for any $n \in \mathbb{N}$. So, there exists a set $N \subset X$, with $\nu(N) = 0$ such that $\eta_x(A_n) = \eta_x'(A_n)$ for any $n \in \mathbb{N}$, $x \in (X - N)$. Hence, for ν-a.e. $x \in X$, $\eta_x = \eta_x'$.

Let us denote $C = \{x \in X : \eta_x(X \times Y) > 0\}$. Let $\mathcal{G} \subset C$ be such that $\mu(\mathcal{G}) = 0$. So, $\mathrm{proj}_X^{-1}(\mathcal{G})$ is such that $\gamma(\mathrm{proj}_X^{-1}(\mathcal{G})) = 0$. Therefore, condition (2) implies

$$0 = \int_X \eta_x(\mathrm{proj}_X^{-1}(\mathcal{G})) \, d\nu = \int_{\mathcal{G}} \eta_x(X \times Y) \, d\nu.$$

Since $\eta_x(X \times Y) > 0$ in $C \supset \mathcal{G}$, we have that $\nu(\mathcal{G}) = 0$. That is, $\nu|_C \ll \mu$. Moreover, writing $\nu|_C = f\mu$ implies that $\gamma = f\mu \otimes \eta_x$. However, by Corollary 3.1, $\gamma = \mu \otimes \gamma_x$ and then $(\nu|_C/\mu)\eta_x = \gamma_x$. □

4. Disintegration maps

From the optimal transport perspective, Theorem A in fact deals with the disintegration of transport plans. In this sense, it is possible to define a function from Y to $\mathscr{P}(X)$ with certain properties which actually establishes the link between disintegration of measures and the geometric properties of the measure spaces. We will call such an object the *disintegration map*.

Definition 4.1. (Disintegration map) Let X and Y be locally compact and separable metric spaces. Consider a measure $\mu \in \mathcal{M}_+(X)$, a Borel map $\pi : X \to Y$ and a disintegration of μ given by Theorem A, so that $\mu = \nu \otimes \mu_y$. We define the disintegration map:

$$f : Y \to (\mathscr{P}(X), W_p)$$
$$y \mapsto \mu_y,$$

such that $\mu = \nu \otimes f(y)$, where W_p is the p-Wasserstein distance.

Remark 4.2. To clarify which measures are associated with the disintegration map, we will say that 'f is a disintegration map of μ w.r.t. ν'.

Although we may define such a map to a general W_p, our main interest will be when either $p = 1$ or $p = 2$. On the one hand, for $p = 1$, there is an explicit formula for the Wasserstein distance which we can apply to study the disintegration of measures in product spaces. On the other hand, when $p = 2$, the theory of optimal transport allows for a geometric characterization of $\mathscr{P}_2(X)$ in terms of the geometric properties of X. More precisely, the study of geodesics defined in $\mathscr{P}_2(X)$ and the convexity properties of certain functionals along these geodesics play a crucial role in the metric theory of gradient flows, which allows us to infer geometric properties of X itself. We shall address the case when $p = 2$ in the next section. Before that, we use the disintegration maps to further characterize the disintegration of measures in product spaces. In this section, consider the following definition of a disintegration map.

Definition 4.3. (Disintegration map—product spaces) Let X and Y be locally compact and separable metric spaces. Consider $\gamma \in \mathscr{P}(X \times Y)$ and a disintegration of γ given by Corollary 3.1, so that $\gamma = \mu \otimes \gamma_x$. In this case, the disintegration map reads as

$$f : X \to (\mathscr{P}(Y), W_1)$$
$$x \mapsto \gamma_x$$

such that $\gamma = \mu \otimes f(x)$.

In the lines of [**AP03**, **GM13**], we start by showing the following.

PROPOSITION 4.4. *A map* $f : X \to (\mathscr{P}(Y), W_1)$ *is a disintegration map if and only if it is Borel.*

Proof. Denote by $\text{Lip}_1(Y)$ the set of Lipschitz functions whose Lipschitz constants are less than or equal to 1. By the Arzelà–Ascoli theorem, the space $\text{Lip}_1(Y)$ is compact with respect to the uniform convergence [**AGS05**, Proposition 3.3.1]. Let $D \subset \text{Lip}_1(Y)$ be a countable dense subset and take $\varphi \in \text{Lip}_1(Y)$. If the measures $\nu_1, \nu_2 \in \mathscr{P}(Y)$ have bounded support, we can use the duality formula to obtain

$$W_1(\nu_1, \nu_2) = \sup_{\varphi \in Lip_1(Y)} \left\{ \int_Y \varphi \, d(\nu_1 - \nu_2) \right\} = \sup_{\varphi \in D} \int_Y \varphi \, d(\nu_1 - \nu_2);$$

see [AGS05, §7.1] for details. For all $\varphi \in \mathrm{Lip}_1(Y)$,

$$\psi_\varphi : X \to \mathbb{R}$$

$$x \mapsto \int_Y \varphi \, d(\nu - f(x))$$

is Borel. For f to be a Borel map, it suffices that

$$f^{-1}(B(\nu, r)) = \bigcap_{\varphi \in D} \psi_\varphi^{-1}((-r, r)) := A,$$

where $B(\nu, r)$ is an open ball of radius r centred at ν in $(\mathscr{P}(Y), W_1)$. In fact, if $x \in A$, then $|\psi_\varphi(x)| < r$ for all $\varphi \in D$ and $W_1(\nu, f(x)) < r$ (by the definition of ψ_φ), so that $f(x) \in B(\nu, r)$. In the same way, if $f(x) \in B(\nu, r)$, then $W_1(\nu, f(x)) < r$ and, by the duality formula,

$$|\psi_\varphi(x)| := \left| \int_Y \varphi \, d(\nu - f(x)) \right| < r$$

for every $\varphi \in D$, so that $x \in A$. Thus, one way is proven. Conversely, let $A \subset Y$ be an open subset. Note that

$$f(x)(A) = \int_Y \mathbb{1}_A(y) \, df(x).$$

Let I_φ be a function given by

$$I_\varphi : (\mathscr{P}(Y), W_1) \to \mathbb{R}$$

$$\lambda \mapsto \int_Y \varphi(y) \, d\lambda, \tag{5}$$

where φ is a lower semicontinuous function over Y. Since W_1 metrizes the weak* topology of $\mathscr{P}(Y)$ [AGS05, Ch. 7], the function I_φ is lower semicontinuous. For every $x \in X$, one obtains $\int_Y \varphi(y) \, df(x) = I_\varphi(f(x))$. By assumption, f is Borel and then $f(\cdot)(A) : X \to \mathbb{R}$ is a composition of a lower semicontinuous function and a Borel map, so it is a Borel map. Therefore, f is a disintegration map, which concludes the proof. □

The disintegration map in fact can be written in terms of the Monge problem. Given a transport map $T : X \to Y$ and the measures $\mu \in \mathscr{P}(X)$ and $\nu = T_*\mu \in \mathscr{P}(Y)$, the disintegration map is given by $x \mapsto \delta_{T(x)}$, where $\delta_{T(x)}$ is a Dirac measure at $T(x)$. It is possible to show that there exists a relationship among measures in $(\mathscr{P}(Y), W_1)$, via push forward of μ by disintegration maps, and the second marginal of transport plans induced by distinct transport maps.

LEMMA 4.5. *Let $T : X \to Y$ and $S : X \to Y$ be transport maps. Consider a measure $\mu \in \mathscr{P}(X)$, and the applications f and g given by*

$$f : X \to (\mathscr{P}(Y), W_1)$$
$$x \mapsto \delta_{T(x)},$$

$$g : X \to (\mathscr{P}(Y), W_1)$$
$$x \mapsto \delta_{S(x)}.$$

Then, $T_*\mu = S_*\mu$ if and only if $f_*\mu = g_*\mu$.

Proof. Define $\varphi(y) = \psi(\delta_y)$, where $\psi \in C((\mathscr{P}(Y), W_1))$ is arbitrarily chosen. Then,

$$\int_Y \psi \, d(f_*\mu) = \int_X \psi(f(x)) \, d\mu$$

$$= \int_X \psi(\delta_{T(x)}) \, d\mu$$

$$= \int_X \varphi(T(x)) \, d\mu$$

$$= \int_X \varphi(S(x)) \, d\mu$$

$$= \int_Y \psi \, d(g_*\mu).$$

Given the arbitrary choice of ψ, it follows that $f_*\mu = g_*\mu$. Conversely, consider $\varphi \in C(Y)$ and the application I_φ defined by equation (5). Note that

$$\int_{\mathscr{P}(Y)} I_\varphi(\lambda) \, d(f_*\mu) = \int_{\mathscr{P}(Y)} I_\varphi(\lambda) \, d(g_*\mu)$$

if and only if

$$\int_X I_\varphi(f(x)) \, d\mu = \int_X I_\varphi(g(x)) \, d\mu,$$

which in turn occurs if and only if

$$\int_X \left(\int_Y \varphi(y) \, df(x) \right) d\mu = \int_X \left(\int_Y \varphi(y) \, dg(x) \right) d\mu.$$

Since

$$\int_Y \varphi(y) \, df(x) = \varphi(T(x))$$

and

$$\int_Y \varphi(y) \, dg(x) = \varphi(S(x)),$$

the last equation can be written as

$$\int_X \varphi(T(x)) \, d\mu = \int_X \varphi(S(x)) \, d\mu.$$

Due to the arbitrary choice of φ, it follows that $T_*\mu = S_*\mu$. □

COROLLARY 4.6. *Let $T : X \to Y$ and $S : X \to Y$ be transport maps. Consider a measure $\mu \in \mathscr{P}(X)$, the disintegration maps f and g defined as in Lemma 4.5, and the measures $\gamma = \mu \otimes f(x)$ and $\eta = \mu \otimes g(x)$. Then, $f_*\mu = g_*\mu$ if and only if $\mathrm{proj}_{Y*}\gamma = \mathrm{proj}_{Y*}\eta$.*

Given $\gamma, \eta \in \Pi(\mu, \nu)$, as defined in §2, with $\gamma = \mu \otimes f(x)$ and $\eta = \mu \otimes g(x)$, we say that γ is *equivalent by disintegration* to η (and we denote $\gamma \approx \eta$) if $f_*\mu = g_*\mu$. With this equivalence in mind, it is possible to define an equivalent class among the transport plans as follows.

Definition 4.7. Given $\gamma \in \Pi(\mu, \nu)$ with $\gamma = \mu \otimes f(x)$, the transport class of γ is defined as the equivalence class $[\gamma] = \{\eta = \mu \otimes g(x) : g_*\mu = f_*\mu\}$.

Thus, we have that all transport plans induced by transport maps belong to the same transport class. Moreover, in the next proposition, we prove that it is possible to use equivalence by disintegration to assure the existence of a transport map.

PROPOSITION 4.8. *Consider a transport map $T : X \to Y$ such that $T_*\mu = \nu$ and $\gamma = \mu \otimes \delta_{T(x)}$ for a non-atomic measure $\mu \in \mathscr{P}(X)$. If $\eta \in [\gamma]$, then there exists a transport map $S : X \to Y$ such that $\eta = \mu \otimes \delta_{S(x)}$.*

Proof. By the approximation theorem [**Amb00**, Theorem 9.3] and by the definition of equivalent by disintegration, there exist a sequence of Borel functions $S_n : X \to Y$ such that $\eta = \lim\limits_{n \to \infty} \mu \otimes \delta_{S_n(x)}$ and $(S_n)_*\mu = \nu$ for every $n \in \mathbb{N}$. So $\mu \otimes \delta_{S_n(x)} \in [\gamma]$ for every $n \in \mathbb{N}$. Consider $\varphi(y) = |y|^2$ and observe that

$$\int_X |S_n(x)|^2 \, d\mu = \int_X \varphi(S_n(x)) \, d\mu = \int_Y \varphi(y) \, d\nu = \int_Y |y|^2 \, d\nu < \infty.$$

Let $\psi \in C((\mathscr{P}(Y), W_1))$ be a function given by $\psi(\delta_y) = |y|^2$ and let $\Delta \subset \mathscr{P}(Y)$ be the set of Dirac measures. Observe that ψ is Lipschitz over Δ w.r.t. W_1. So, take ψ any Lipschitz extension over $\mathscr{P}(Y)$. Since $(\delta_{S_n})_*\mu = (\delta_T)_*\mu$, we have that

$$\int_X |S_n(x)|^2 \, d\mu = \int_X \psi(\delta_{S_n(x)}) \, d\mu = \int_X \psi(\delta_{T(x)}) \, d\mu = \int_X |T(x)|^2 \, d\mu$$

for every $n \in \mathbb{N}$. Therefore, moving on to a subsequence, we can assume that S_n is weakly convergent to S. By [**Amb00**, Lemma 9.1], $\mu \otimes \delta_{S_n(x)}$ converges weakly to $\mu \otimes \delta_{S(x)}$. This concludes the proof. $\square$

In this context, the Monge problem can be interpreted as: minimize $\int_{X \times Y} c(x, y) \, d\gamma$ in a fixed transport class of $\Pi(\mu, \nu)$, that is, obtain

$$\min \left\{ \int_{X \times Y} c(x, y) \, d\gamma : \gamma \in [\mu \otimes \delta_T] \right\}$$

for a given transport map T. Regarding the Monge–Kantorovich problem, note that the second part of the proof of Lemma 4.5 applies to general transport plans, so that the second marginal can be fixed by the disintegration maps, as follows.

LEMMA 4.9. *Consider $\mu \in \mathscr{P}(X)$, the disintegration maps $f : X \to \mathscr{P}(Y)$ and $g : X \to \mathscr{P}(Y)$, and the transport plans $\gamma = \mu \otimes f(x)$ and $\eta = \mu \otimes g(x)$. Then, $f_*\mu = g_*\mu$ implies $\mathrm{proj}_{Y*}\gamma = \mathrm{proj}_{Y*}\eta$.*

The reciprocal of Lemma 4.9, however, is not true. Consider, for instance, the following transport problem. Three factories, x_1, x_2 and x_3, with the same production, supply 100% of their products to two stores, y_1 and y_2. Suppose store y_2 has a demand five times greater than store y_1. Let μ be the measure related to the production of the factories and ν be
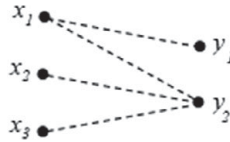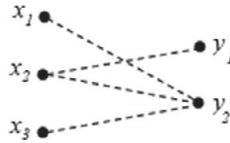
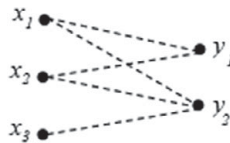FIGURE 2. Transport plan 1.

FIGURE 3. Transport plan 2.

FIGURE 4. Transport plan 3.

the measure related to the amount of delivered products to the stores. These measures are given by

$$\mu = \tfrac{1}{3}\delta_{x_1} + \tfrac{1}{3}\delta_{x_2} + \tfrac{1}{3}\delta_{x_3},$$

$$\nu = \tfrac{1}{6}\delta_{y_1} + \tfrac{5}{6}\delta_{y_2}.$$

Consider the transport plan illustrated in Figure 2, which divides the products which are produced in the factory x_1 between the two stores. The corresponding disintegration map f is given by $f(x_1) = \tfrac{1}{2}\delta_{y_1} + \tfrac{1}{2}\delta_{y_2}$, $f(x_2) = \delta_{y_2}$ and $f(x_3) = \delta_{y_2}$.

Note that since μ is of the type $\sum_i \alpha_i \delta_{x_i}$, we have that $f_*\mu = \sum_i \alpha_i \delta_{f(x_i)}$. Suppose, however, that there is a change in logistics and in the new transport plan, Figure 3, the factory x_1 delivers its production only to the store y_2 and the factory x_2 divides its production between the two stores. In this new situation, the disintegration map is given by $g(x_1) = \delta_{y_2}$, $g(x_2) = \tfrac{1}{2}\delta_{y_1} + \tfrac{1}{2}\delta_{y_2}$ and $g(x_3) = \delta_{y_2}$. Nevertheless, on the one hand, $f_*\mu = g_*\mu$, i.e., the transport class does not change. On the other hand, if in this new transport plan the factories x_1 and x_2 deliver to both stores, Figure 4, in such a way that x_1 sends 30% of its production to y_1 and 70% to y_2, and x_2 sends 20% of its production to y_1 and 80% to y_2, the disintegration map is given by $h(x_1) = 3/10\delta_{y_1} + 7/10\delta_{y_2}$, $h(x_2) = 2/10\delta_{y_1} + 8/10\delta_{y_2}$ and $h(x_3) = \delta_{y_2}$. In this case, $f_*\mu \neq h_*\mu$. So, the transport plans 1 and 3 do not belong to the same transport class.

Furthermore, if there is a small change in transport plan 3, so that x_1 sends 10% of its production to y_1 and 90% to y_2, and x_2 sends 40% of its production to y_1 and 60% to y_2, the new disintegration map is given by $k(x_1) = 1/10\delta_{y_1} + 9/10\delta_{y_2}$, $k(x_2) = 4/10\delta_{y_1} + 6/10\delta_{y_2}$ and $k(x_3) = \delta_{y_2}$. Therefore, $h_*\mu \neq k_*\mu$. Thus, with the previous definition for

transport class, the transport class changes when either the number of factories that deliver products to more than one store is changed or even if the fraction of delivered production is changed.

Therefore, to be compatible with the Monge–Kantorovich problem, we need another definition of transport class. Take $\mu \otimes f \in \Pi(\mu, \nu)$ and $\Lambda = f_* \mu$, and recall that $(\text{proj}_Y)_*(\mu \otimes f) = \nu$, then for every $\varphi \in C(Y)$,

$$
\begin{aligned}
\int_Y \varphi(y)\, d\nu &= \int_X \left(\int_Y \varphi(y)\, df(x) \right) d\mu \\
&= \int_X I_\varphi(f(x))\, d\mu \\
&= \int_{\mathscr{P}(Y)} I_\varphi(\lambda)\, d\Lambda(\lambda) \\
&= \int_{\mathscr{P}(Y)} \left(\int_Y \varphi(y)\, d\lambda \right) d\Lambda,
\end{aligned}
$$

where I_φ is given by equation (5) and $\lambda \in \mathscr{P}(Y)$. Hence, every probability Λ in $(\mathscr{P}(Y), W_1)$ satisfying

$$
\int_{\mathscr{P}(Y)} \lambda\, d\Lambda = \nu \tag{6}
$$

defines a transport class $[\gamma] = \{\mu \otimes f : f_* \mu = \Lambda\}$. As an example, a transport class $[\mu \times \nu]$ corresponds to the measure $\Lambda = \delta_\nu$. From this point of view, the Monge–Kantorovich problem in the class Λ can be thought as

$$
MK_\Lambda(c, \mu, \nu) = \inf_\gamma \left\{ \iint_{X \times Y} c(x, y)\, d\gamma : \gamma = \mu \otimes f,\, f_* \mu = \Lambda \right\}.
$$

This notion of transport class also allows us to see the Monge–Kantorovich problem as an abstract Monge problem between the spaces X and $\mathscr{P}(Y)$, considering the cost

$$
\tilde{c}(x, \lambda) = \int_Y c(x, y),\, d\lambda.
$$

In fact, for every disintegration map $f : X \to \mathscr{P}(Y)$, such that $f_* \mu = \Lambda$,

$$
\begin{aligned}
\int_X \tilde{c}(x, f(x))\, d\mu &= \int_X \left(\int_Y c(x, y)\, df(x) \right) d\mu \\
&= \int_{X \times Y} c(x, y)\, d(\mu \otimes f)
\end{aligned}
$$

and the problem $MK_\Lambda(c, \mu, \nu)$ is equivalent to the Monge problem with parameters $(\tilde{c}, \mu, \Lambda)$. Thus, the disintegration of transport plans introduces another perspective for the optimization problems proposed by Monge and Kantorovich.

5. *Absolute continuity from disintegration maps*

To connect what we just developed to the geometric properties of the measure spaces and the space of measures, we shall study how paths of measures are given by disintegration maps. Our interest hereafter is to use the 2-Wasserstein distance.

To construct such paths in 2-Wasserstein space, it is essential to analyse the weak continuity of the disintegration map. In the Wasserstein space, one can study a characterization of convergence. We say that $\{\mu_k\}$ *converges weakly* to μ when

$$\int \varphi \, \mu_k \longrightarrow \int \varphi \, d\mu$$

for any bounded continuous function φ. Moreover, Wasserstein distances metrize weak convergence, that is, if $\{\mu_k\}$ is a sequence in $\mathscr{P}_p(X)$ and $\mu \in \mathscr{P}_p(X)$, then μ_k converges weakly to μ is equivalent to $W_p(\mu_k, \mu) \longrightarrow 0$. When referring to continuity of the disintegration map, we use *weak continuity* meaning continuity with respect to weak convergence on $\mathscr{P}_p(X)$.

A map f on an metric space (Y, ν) is called *nearly continuous* if, for each $\varepsilon > 0$, there exist $\mathcal{K} \subset Y$ closed with $\nu(Y \backslash \mathcal{K}) < \varepsilon$ such that f restricted to $\mathcal{K}$ is continuous. Let f be a disintegration map of μ w.r.t. ν. If ν is absolutely continuous with respect to the volume measure of Y, we can apply Lusin's Theorem 2.4 to show that f is nearly weakly continuous. We also need the following lemma, which is actually a known fact within the optimal transport community. For completeness, we add it here in our context.

LEMMA 5.1. *Let (X, d) be a separable metric space. The 2-Wasserstein space $(\mathscr{P}(X), W_2)$ is a separable metric space.*

Proof. Let $\mathcal{D} \subset X$ be a countable and dense subset. Consider the space $\mathcal{M}$ defined by $\mathcal{M} := \{\nu \in \mathscr{P}(X) : \nu = \sum_j a_j \delta_{x_j}, \text{ with } a_j \in \mathbb{Q} \text{ and } x_j \in \mathcal{D}\}$. We want to show that $\mathcal{M}$ is dense in $(\mathscr{P}(X), W_2)$.

Given $\mu \in (\mathscr{P}(X), W_2)$, then for any $\varepsilon > 0$ and $x_0 \in \mathcal{D}$, there exists a compact set $K \subset X$ such that

$$\int_{X \backslash K} d(x_0, x)^2 \, d\mu \le \varepsilon^2.$$

Since K is compact, we may cover it with a family $\{B(x_k, \varepsilon/2) : 1 \le k \le N, x_k \in \mathcal{D}\}$ and define

$$B_k' := B(x_k, \varepsilon) \backslash \bigcup_{j < k} B(x_j, \varepsilon)$$

so that $\{B_k'\}$ are disjoint and still cover K. Define φ on X by

$$\begin{cases} \varphi(B_k' \cap K) = \{x_k\}, \\ \varphi(X - K) = \{x_0\}. \end{cases}$$

So $d(x, x_k) \le \varepsilon$ for every $x \in K$. Therefore,

$$\int_X d(x, \varphi(x))^2 \, d\mu \le \varepsilon^2 \int_K d\mu + \int_{X \backslash K} d(x, x_0)^2 \, d\mu \le 2\varepsilon^2$$

and then $W_2(\mu, \varphi_*\mu) \le 2\varepsilon^2$. Note that $\varphi_*\mu$ can be written as $\sum_{0 \le j \le N} \alpha_j \delta_{x_j}$, that is, μ might be approximated by a finite combination of Dirac masses. Moreover, the coefficients α_j might be replaced by rational coefficients (up to a small error in Wasserstein

distance): since Wasserstein distances are controlled by weighted total variation [**Vil09**, Theorem 6.15],

$$W_2\left(\sum_{0\le j\le N}\alpha_j\delta_{x_j},\sum_{0\le j\le N}\beta_j\delta_{x_j}\right)\le 2^{1/2}\left[\max_{k,l}\ d(x_k,x_l)\right]\sum_{0\le j\le N}|\alpha_j-\beta_j|^{1/2},$$

which can become small, taking suitable coefficients $\beta_j\in\mathbb{Q}$. Thus, it follows that $\mathcal{M}$ is dense in $\mathscr{P}(X)$. Consequently, $(\mathscr{P}(X),W_2)$ is separable. $\qquad\square$

PROPOSITION 5.2. *Let X, Y be locally compact, complete and separable metric spaces. Consider $\pi:X\to Y$ a Borel map, $\mu\in\mathcal{M}_+(X)$, vol_Y volume measure on Y and $\nu:=\pi_*\mu$. If $\nu\ll\mathrm{vol}_Y$, then the disintegration map of μ w.r.t. ν is nearly weakly continuous.*

Proof. Consider the 2-Wasserstein distance W_2 on $\mathscr{P}(X)$, with d a complete bounded metric for X. Here, W_2 metrizes the weak convergence of $\mathscr{P}(X)$ [**Vil09**, Theorem 6.9]. Furthermore, $(\mathscr{P}(X),W_2)$ is a separable space. Moreover, from Theorem A, the map $y\mapsto\mu_y\in\mathscr{P}(X)$ is measurable and ν is a Borel measure on Y, since $\nu\ll\mathrm{vol}_Y$. Then, by Lusin's theorem for Y and $\varepsilon>0$, there exists $\mathcal{K}\subset Y$ with $\nu(Y\backslash\mathcal{K})<\varepsilon$ such that the disintegration map restricted to $\mathcal{K}$ is weakly continuous. $\qquad\square$

Although Proposition 5.2 is relevant by itself, we would like to have conditions for which the disintegration map is weakly continuous at every point, so that, given any two points y, y' in Y, we can construct a path of conditional measures connecting μ_y and $\mu_{y'}$. Note that the 2-Wasserstein space is actually connected, and therefore we can always find a path connecting two measures. Nevertheless, we require this path to be specifically given by the disintegration map. This is not trivial indeed and we can easily construct examples in which this map is not weakly continuous.

Example 5.3. Consider $X=Y=[0,1]$. Let μ be the Lebesgue measure on X and take the map $\pi:X\to Y$ given by

$$\pi(x)=\begin{cases}2x & \text{if } x<\tfrac{1}{2},\\ 1 & \text{if } x\ge\tfrac{1}{2}.\end{cases}$$

Note that π is Borel measurable, since it is continuous. Define

$$\nu:=\pi_*\mu=\tfrac{1}{2}\lambda+\tfrac{1}{2}\delta_1,$$

where λ is the Lebesgue measure on Y and δ_1 is the Dirac measure at $y=1$. A disintegration $\{\mu_y\}_{y\in Y}$ of μ with respect to π is given by

$$\mu_y=\begin{cases}\delta_{\pi^{-1}(y)} & \text{if } y<1,\\ 2\lambda|_{[1/2,1]} & \text{if } y=1.\end{cases}$$

In this case, the disintegration map is not weakly continuous at $y=1$. In fact, consider a sequence $(y_n)_n$ in Y such that $y_n\longrightarrow y=1$. Note that

$$W_2^2(\mu_{y_n}, \mu_y) = \inf_{\gamma \in \Pi(\mu_{y_n}, \mu_y)} \int_{X \times X} d(x_1, x_2)^2 \, d\gamma$$

$$= \int_{X \times X} d(x_1, x_2)^2 \, d(\delta_{\pi^{-1}(y_n)} \times \mu_y)$$

$$= \int_X d(\pi^{-1}(y_n), x_2)^2 \, d\mu_y.$$

Then, at the limit $y_n \longrightarrow 1$, we have

$$W_2^2(\mu_{y_n}, \mu_y) \longrightarrow \int_X d\left(\frac{1}{2}, x_2\right)^2 d\mu_y \neq 0.$$

Therefore, f is not weakly continuous at $y = 1$. Furthermore, note that $\nu(\{1\}) = \frac{1}{2}$ and ν is not absolutely continuous with respect to λ. However, considering $Y = [0, 1)$ and guaranteeing the absolute continuity of ν, we can obtain a good approximation $\mathcal{K}$ in which f is weakly continuous.

We can actually ask for some additional hypotheses so that the disintegration map is weakly continuous. One possibility is to take π as a bijective and continuous map.

PROPOSITION 5.4. *Let X and Y be locally compact and separable metric spaces. Consider $\mu \in \mathcal{M}_+(X)$, $\pi : X \to Y$ a Borel map, $\nu := \pi_*\mu$, $\{\mu_y\}_{y \in Y}$ a disintegration of μ given by Theorem A, and f the disintegration map of μ w.r.t. ν. If π is bijective and continuous, then f is weakly continuous.*

Proof. On the one hand,

$$\mu(B) = \int_Y \mu_y(B) \, d\nu$$

$$= \int_Y \mu_y(B \cap \pi^{-1}(y)) \, d\nu.$$

On the other hand,

$$\mu(B) = \int_X \mathbb{1}_B \, d\mu$$

$$= \int_X \mathbb{1}_B \, d(\nu \circ \pi)$$

$$= \int_Y \mathbb{1}_B(\pi^{-1}(y)) \, d\nu$$

$$= \int_Y \mathbb{1}_B(\pi^{-1}(y)) \delta_{\pi^{-1}(y)}(\pi^{-1}(y)) \, d\nu$$

$$= \int_Y \delta_{\pi^{-1}(y)}(B \cap \pi^{-1}(y)) \, d\nu.$$

Then,

$$\int_Y \mu_y(B \cap \pi^{-1}(y)) \, d\nu = \int_Y \delta_{\pi^{-1}(y)}(B \cap \pi^{-1}(y)) \, d\nu.$$

Moreover, since μ_y is a probability and $B \cap \pi^{-1}(y)$ is either a singleton or an empty set, we have that

$$\delta_{\pi^{-1}(y)}(B \cap \pi^{-1}(y)) \geq \mu_y(B \cap \pi^{-1}(y)).$$

Then, $\mu_y(B \cap \pi^{-1}(y)) = \delta_{\pi^{-1}(y)}(B \cap \pi^{-1}(y))$ and it follows that $\mu_y = \delta_{\pi^{-1}(y)}$.

Suppose $y_n \longrightarrow y$. Since π is continuous, π^{-1} is continuous and then

$$\int g \, d\mu_{y_n} \longrightarrow \int g \, d\mu_y$$

for every bounded uniformly continuous function g. Therefore, f is weakly continuous. $\qquad\square$

Asking for the bijectivity of π is quite strong. We want to explore ways to ease this restriction and to obtain continuity at least for almost every point. In some examples for which π is a quotient map, we have the weak continuity of f.

Example 5.5. ℓ_q-product space from metric measures spaces.

Let (Y, d_Y, m_Y), (Z, d_Z, m_Z) denote the metric spaces (Y, d_Y) and (Z, d_Z) which are endowed with probability measures m_Y and m_Z, respectively. We shall call these metric measure spaces. For $q \in [1, \infty]$, define the ℓ_q-product space $X := Y \times_{\ell_q} Z$ as the product space $Y \times Z$ equipped with the measure $\mu = m_Y \times m_Z$ and the distance d_{ℓ_q} given by

$$d_{\ell_q}((y, z), (y'z')) = \begin{cases} [d_Y(y, y')^q + d_Z(z, z')^q]^{1/q} & \text{if } 1 \leq q < \infty, \\ \max\{d_Y(y, y'), d_Z(z, z')\} & \text{if } q = \infty. \end{cases}$$

Consider the projection $\pi : X \to Y$ and $\{\mu_y\}_{y \in Y}$ the disintegration of μ with respect to $\nu := \pi_* \mu$. Note that

$$d_{\ell_q}(\pi^{-1}(y), \pi^{-1}(y')) = d_Y(y, y')$$

for every $q \in [1, \infty]$ and, for every conditional measure $\mu_y, \mu_{y'}$,

$$\begin{aligned} W_2^2(\mu_y, \mu_{y'}) &= \inf_{\gamma \in \Pi(\mu_y, \mu_{y'})} \int d_{\ell_q}((y_1, z_1), (y_2, z_2))^2 \, d\gamma \\ &= \inf_{\gamma \in \Pi(\delta_y, \delta_{y'})} \int d_Y(y_1, y_2)^2 \, d\gamma \\ &= d_Y(y, y')^2. \end{aligned}$$

Then, $W_2(\mu_y, \mu_{y'}) = d_Y(y, y')$ and, therefore, the disintegration map is weakly continuous.

In fact, this is one example of a disintegration of measures associated with a foliation, called *metric measure foliation*, covered in [GKMS18]. For a precise definition, we need to introduce some concepts. Let (X, d) be a metric space. A *foliation* $\mathcal{F}$ of X is a partition of X into closed subsets. The elements of this partition are called *leaves*. Here, $\mathcal{F}$ is called a *metric foliation* if for every $F, F' \in \mathcal{F}$ and every $x \in F$,

$$d(F, F') = d(x, F'),$$

where $d(F, F') = \inf\{d(x, x') : x \in F, x' \in F'\}$ and $d(x, F') = d(\{x\}, F')$. In the case where each leaf is bounded, we say that $\mathcal{F}$ is bounded. Given a metric foliation $\mathcal{F}$ of X, define the equivalence relation:

$$x \sim x' \iff \text{there exists } F \in \mathcal{F} \text{ such that } x, x' \in F. \tag{7}$$

Consider $X^* := X/\sim$ the set of equivalence classes under equation (7) and the projection $p : X \to X^*$ onto X^*. We call X^* the *quotient space* and p the *quotient map*. Define a distance function d^* on X^* as

$$d^*(y, y') := d(p^{-1}(y), p^{-1}(y')) \tag{8}$$

for $y, y' \in X^*$. Note that p is a submetry: $p(B(x, r)) = B(p(x), r)$, where $B(x, r)$ is a ball centred at x with radius r. In fact,

$$
\begin{aligned}
B(p(x), r) &= \{y \in X^* : d^*(y, p(x)) < r\} \\
&= \{y \in X^* : d(p^{-1}(y), p^{-1}(p(x))) < r\} \\
&= p(B(x, r)).
\end{aligned}
$$

Therefore, p is 1-Lipschitz. In this notation, we define the following.

Definition 5.6. Let $\mathcal{F}$ be a metric foliation of (X, d, μ). Additionally, $\mathcal{F}$ is a *metric measure foliation* if $p_*\mu$ is locally finite Borel measure on X^*, and there exists a Borel subset $\Omega \subset X^*$ with $p_*\mu(X^*\backslash\Omega) = 0$ such that

$$W_2(\mu_y, \mu_{y'}) = d^*(y, y') \tag{9}$$

for any $y, y' \in \Omega$, where $\{\mu_y\}_{y \in Y}$ is a disintegration of μ with respect to $p_*\mu$.

Note that, in this case, the disintegration map is an isometry. This is the most important characteristic of a metric measure foliation for us. A very important example of a metric measure foliation is related to the action of isometry group.

Example 5.7. Let (X, d, μ) be a metric measure space and G a compact topological group. Let

$$G \times X \ni (g, x) \mapsto gx \in X$$

be an isometric action of G on X. Suppose this action is metric measure isomorphic, that is, for every $g \in G$, the map $X \ni x \mapsto gx \in X$ is an isometry preserving the measure μ. Consider $[x]$ the G-orbit of a point $x \in X$ and the quotient space X/G endowed with the distance

$$d_{X/G}([x], [x']) = \inf_{g, g' \in G} d(gx, g'x').$$

Consider $p : X \to X/G$ the projection map, that is, p is given by $x \mapsto [x]$. The family $\mathcal{F} := \{p^{-1}(y) : y \in X/G\}$ is a metric measure foliation on X.

Other interesting examples arise from Riemannian submersions of weighted Riemannian manifolds [**GKMS18**].
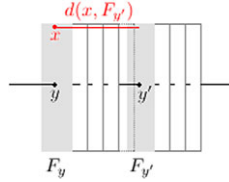
FIGURE 5. Idea of a metric measure foliation as a space fibration.

From Definition 5.6, we have a direct association of the 2-Wasserstein distance between conditional measures (given by Theorem A) and the distance between points to which these measures were indexed. It is clear that in this case, we have the weak continuity of the disintegration map, since it is a isometry. Note that we can relate the 2-Wasserstein distance between conditional measures to the distance between two points of the quotient space looking at a direction perpendicular to the leaves, through the leaves. Roughly speaking, we can think of a kind of space fibration, where one of the directions has been 'collapsed', and it becomes a parameter for the disintegration family, so that each conditional measure is supported on the underlying fibre (see Figure 5). That is, the measures μ_y are of the type $\delta_y \times \lambda$, where λ is a measure on the fibre, similar to Example 3.2.

With what we have seen so far, we have been able to classify some situation in which the disintegration map is weakly continuous.

PROPOSITION 5.8. *Let X be locally compact and separable metric space. Consider $\mu \in \mathcal{M}_+(X)$, a metric foliation $\mathcal{F}$ of X, the quotient space X^* and the quotient map $p : X \to X^*$. If there exists a metric measure foliation of X, with $\Omega = X^*$, then the disintegration map of μ w.r.t. $\nu = p_*\mu$ is weakly continuous.*

Proof. Given a metric measure foliation of X, we have for every $y, y' \in X^*$, $W_2(\mu_y, \mu_{y'}) = d^*(y, y')$, where $\mu_y = f(y)$ and $\mu_{y'} = f(y')$, and the disintegration map is denoted by f. Consider a sequence $(y_n)_n$ in X^* such that $y_n \longrightarrow y$. Note that $W_2(\mu_{y_n}, \mu_y) \longrightarrow 0$, since $d^*(y_n, y) \longrightarrow 0$. Therefore, we have the weak continuity of f.   □

Remark 5.9. If in Proposition 5.8 we do not ask for $\Omega = X^*$, the weak continuity of the disintegration map is given for $p_*\mu$-a.e. $y \in X^*$ due the definition of metric measure foliation. Although it seems to be a strong condition, in general cases of interest, we have it satisfied, as in Examples 5.5 and 5.7, for instance.

Remark 5.10. One can associate hypotheses about the map π and the type of disintegration obtained. In light of what we have seen, we may obtain the following.

(1)   Under the hypotheses of Proposition 5.4, that is, when the map π is bijective and continuous, the conditional measures given by Theorem A are Dirac deltas.

(2)   Under the hypotheses of Proposition 5.8, that is, in the metric measure foliation case, the conditional measures given by Theorem A are of the type $\delta_y \times \lambda$, where λ is a measure on the fibre.

The entire study carried out on the disintegration map makes clear a fundamental condition for it to be weakly continuous: the supports of the conditional measures $\{\mu_y\}$ must be disjoint. However, if we want some kind of absolute continuity of $\{\mu_y\}$ with respect to a reference measure, the supports must have μ-positive measure. In the cases of Propositions 5.4 and 5.8, we do not obtain supports with a μ-positive measure. Another way is to think about the absolute continuity of measures with respect to a reference measure on the fibres. This discussion is summarized in Theorem B. In the statement, we call f a *minimizing invariant*, if it maps a minimizing curve on Y to a minimizing curve on $(\mathscr{P}(X), W_2)$. Such a condition is fulfilled when π is a Riemannian submersion, for example.

THEOREM B. *Let X and Y be locally compact, complete, separable metric spaces. Consider $\pi : X \to Y$ a Borel map, μ in $\mathcal{M}_+(X)$ and $\nu := \pi_*\mu$. If the disintegration map of μ w.r.t. ν is weakly continuous and Y is path connected, then given two points $y, y' \in Y$:*

  (i)   *there exists a path on $(\mathscr{P}(X), W_2)$, given by the disintegration map, connecting μ_y and $\mu_{y'}$, the respective conditional measures given by Theorem A;*

 (ii)   *if X is a smooth compact Riemannian manifold equipped with a volume measure vol, $\mu \ll$ vol, π is such that $\pi^{-1}(y)$ has μ-positive measure for ν-almost every y, the disintegration map is minimizing invariant, and either μ_y or $\mu_{y'}$ is absolutely continuous w.r.t. vol, then all the measures μ_{y_t} on the path given by item (i) are absolutely continuous w.r.t. vol;*

(iii)   *if π is such that $\{\pi^{-1}(y)\}_{y \in Y}$ is a metric measure foliation of X, X a smooth compact Riemannian manifold, there exists a path given by the disintegration map connecting μ_y and $\mu_{y'}$, and if either μ_y or $\mu_{y'}$ is absolutely continuous with respect to the volume measure on the respective support fibre, then all the measures μ_{y_t} on this path are absolutely continuous with respect to the volume measure on the fibre.*

Proof. (i) This is a direct consequence of the weak continuity of the disintegration map, which will be denoted by f throughout the demonstration. Consider $y, y' \in Y$. Let ψ be a continuous curve in Y connecting y and y', that is, $\psi = \{y_t : t \in [0, 1], \psi(0) = y, \psi(1) = y'\}$. Taking $y_t \longrightarrow \bar{y} \in Y$, we have $f(y_t) \xrightarrow{w} f(\bar{y})$, that is, $W_2(\mu_{y_t}, \mu_{\bar{y}}) \longrightarrow 0$. Then, $\zeta = \{\mu_{y_t} : t \in [0, 1]\}$, where μ_{y_t} is the conditional measure associated with y_t via f, for every $t \in [0, 1]$, is a weakly continuous curve in $(\mathscr{P}(X), W_2)$ connecting μ_y and $\mu_{y'}$. This proves the first part of our theorem.

The proof of part (ii) will be done in a few steps. The idea is to use a sort of 'time-dependent' version of optimal transport. See [**Vil09**, Ch. 7], for example. In short, we will consider the curve ζ given by the disintegration map as an interpolation between probability measures, called *displacement interpolation*. To this end, we consider a transport problem, and we associate ζ with a random curve ξ in X.

Step 1: A minimizing curve in the space of measures. Consider the path $\zeta = \{\mu_{y_t} : t \in [0, 1]\}$ as constructed in part (i). Taking ψ as a minimizing curve in Y, ζ is a minimizing curve in $\mathscr{P}(X)$, since f is minimizing invariant. By abuse of notation, we use the weak continuous path $\zeta : [0, 1] \to (\mathscr{P}(X), W_2)$ while referring to this path in $(\mathscr{P}(X), W_2)$ given by the disintegration map f evaluated at the minimal curve ψ joining $[y, y']$ in Y. More precisely, we consider the composition $f \circ \psi$ to describe ζ.

Step 2: A random curve in X. We want to associate ζ with a random curve $\xi : [0, 1] \to X$. To set some notation, let e_t be the *evaluation map*, given by $e_t(\xi) = \xi(t) := \xi_t$, meaning the evaluation of ξ at t. We will also use some usual concepts related to geodesics in Riemannian manifolds throughout the demonstration. We suggest [**Jost11**] for a comprehensive reading.

Consider the curve $\zeta : [0, 1] \to \mathscr{P}_2(X)$ joining μ_0 and μ_1 (from Step 1), denoting $\mu_0 = \mu_y$ and $\mu_1 = \mu_{y'}$. We also denote $\mu_t = \mu_{y_t}$. Suppose that there is a transport problem associated with this curve, whose respective spatial distributions are modelled by these probability measures. Assume that the cost function for the transport between the initial point $x_0 \in X$ (at time 0) and the final point $x_1 \in X$ (at time 1), denoted by $c^{0,1}(x_0, x_1)$, is associated with a family of functionals parametrized by the initial and the final times. Denote by $\mathcal{A}^{0,1}$ the functional on the set of curves $[0, 1] \to X$, such that,

$$c^{0,1}(x_0, x_1) = \inf\{\mathcal{A}^{0,1}(\xi) : \xi_0 = x_0, \ \xi_1 = x_1, \ \xi \in \mathcal{C}([0, 1]; X)\}.$$

In other words, $c^{0,1}(x_0, x_1)$ is the minimal cost needed to go from point x_0 at initial time 0, to point x_1 at final time 1. Moreover, let $C^{0,1}(\mu_0, \mu_1)$ be the optimal transport cost between μ_0 and μ_1 for the cost $c^{0,1}(x_0, x_1)$. For $t_1, t_2 \in [0, 1]$, define

$$\mathcal{A}^{t_1,t_2}(\xi) = \frac{L(\xi)^2}{t_2 - t_1},$$

where $L(\xi)$ is the length of ξ, so that

$$c^{t_1,t_2}(x_0, x_1) = \frac{d(x_0, x_1)^2}{t_2 - t_1}$$

and

$$C^{t_1,t_2}(\mu_0, \mu_1) = \frac{W_2(\mu_0, \mu_1)^2}{t_2 - t_1}. \tag{10}$$

We want to show that there exists a random minimizer $\xi : [0, 1] \to X$, such that, $\text{law}(\xi_t) = \mu_t$ for every $t \in [0, 1]$. In other words, we want to show that ζ is a curve in the space of measures which interpolates all possible measures along the minimizing path joining μ_0 and μ_1. Such a curve is called displacement interpolation.

Step 3: ζ is a displacement interpolation. In this step, ξ will be constructed by dyadic approximation, according to couplings of measures in ζ associated with times $t = 1/2^k$. To achieve this, we will use the iterative construction in line with [**Vil09**, Theorem 7.21].

Let Γ be the set of minimizing curves in X. It will be necessary throughout the text to consider subsets of Γ in which the geodesics are defined for certain time intervals and endpoints (or endpoints regions). So, for $s, t \in [0, 1]$, $x_s, x_t \in X$, let $\Gamma^{s,t}_{x_s \to x_t}$ be the set of minimizing curves in X starting at x_s at time s and ending at x_t at time t. Similarly, for any two compact sets $K_s, K_t \subset X$, let $\Gamma^{s,t}_{K_s \to K_t}$ be the set of minimizing curves starting in K_s at time s and ending in K_t at time t.

Considering the measures along ζ, for $t_1, t_2, t_3 \in [0, 1]$, let $\gamma_{t_1 \to t_2}$ be an optimal transference plan between μ_{t_1} and μ_{t_2} for $c^{t_1,t_2}(x_{t_1}, x_{t_2})$, and let $\gamma_{t_2 \to t_3}$ be an optimal

transference plan between μ_{t_2} and μ_{t_3} for $c^{t_2,t_3}(x_{t_2}, x_{t_3})$. By Lemma 2.5, it is possible to take random variables $(\xi_{t_1}, \xi_{t_2}, \xi_{t_3})$ such that $\text{law}(\xi_{t_1}, \xi_{t_2}) = \gamma_{t_1 \to t_2}$, $\text{law}(\xi_{t_2}, \xi_{t_3}) = \gamma_{t_2 \to t_3}$ and $\text{law}(\xi_{t_i}) = \mu_{t_i}$ for $i = 1, 2, 3$. Since ζ is minimizing in $\mathscr{P}_2(X)$ (see Step 1) and $\mathscr{P}_2(X)$ is a geodesic space, it follows from equation (10) that

$$C^{t_1,t_2}(\mu_{t_1}, \mu_{t_2}) + C^{t_2,t_3}(\mu_{t_2}, \mu_{t_3}) = C^{t_1,t_3}(\mu_{t_1}, \mu_{t_3}).$$

This, in particular, implies:

(a) (ξ_{t_1}, ξ_{t_3}) is an optimal coupling of (μ_{t_1}, μ_{t_3}) for $c^{t_1,t_3}(\xi_{t_1}, \xi_{t_3})$;
(b) $c^{t_1,t_3}(\xi_{t_1}, \xi_{t_3}) = c^{t_1,t_2}(\xi_{t_1}, \xi_{t_2}) + c^{t_2,t_3}(\xi_{t_2}, \xi_{t_3})$ almost surely.

Let (ξ_0, ξ_1) be an optimal coupling of (μ_0, μ_1). Consider optimal transference plans $\gamma_{0 \to 1/2}$, $\gamma_{1/2 \to 1}$, as above, and construct random variables $(\xi_0^{(1)}, \xi_{1/2}^{(1)}, \xi_1^{(1)})$ such that $(\xi_0^{(1)}, \xi_{1/2}^{(1)})$ is an optimal coupling of $(\mu_0, \mu_{1/2})$ for $c^{0,1/2}(\xi_0^{(1)}, \xi_{1/2}^{(1)})$; $(\xi_{1/2}^{(1)}, \xi_1^{(1)})$ is an optimal coupling of $(\mu_{1/2}, \mu_1)$ for $c^{1/2,1}(\xi_{1/2}^{(1)}, \xi_1^{(1)})$, and $\text{law}(\xi_i^{(1)}) = \mu_i$ for $i = 0, \frac{1}{2}, 1$. Moreover, item (a) implies that $(\xi_0^{(1)}, \xi_1^{(1)})$ is an optimal coupling of (μ_0, μ_1), and item (b) implies

$$c^{0,1}(\xi_0^{(1)}, \xi_1^{(1)}) = c^{0,1/2}(\xi_0^{(1)}, \xi_{1/2}^{(1)}) + c^{1/2,1}(\xi_{1/2}^{(1)}, \xi_1^{(1)})$$

almost surely. Iterating this process, at the step k, we have random variables $(\xi_0^{(k)}, \xi_{1/2^k}^{(k)}, \xi_{2/2^k}^{(k)}, \ldots, \xi_1^{(k)})$, so that for any two $i, j \leq 2^k$, $(\xi_{i/2^k}^{(k)}, \xi_{j/2^k}^{(k)})$ is an optimal coupling of $(\mu_{i/2^k}, \mu_{j/2^k})$. Furthermore, for $i_1, i_2, i_3 \leq 2^k$,

$$c^{i_1/2^k, i_3/2^k}(\xi_{i_1/2^k}^{(k)}, \xi_{i_3/2^k}^{(k)}) = c^{i_1/2^k, i_2/2^k}(\xi_{i_1/2^k}^{(k)}, \xi_{i_2/2^k}^{(k)}) + c^{i_2/2^k, i_3/2^k}(\xi_{i_2/2^k}^{(k)}, \xi_{i_3/2^k}^{(k)})$$

almost always.

We want to extend the random variables $\xi^{(k)}$, defined for times $i/2^k$, $i \leq 2^k$, to continuous curves $(\xi^{(k)})_{0 \leq t \leq 1}$. For this, note that for all times $s, t \in [0, 1]$, $s < t$, there exists a Borel map $S_{s \to t} : X \times X \to \mathcal{C}([s, t]; X)$ such that for all $x, z \in X$, $S(x, z)$ belongs to $\Gamma_{x \to z}^{s,t}$ [**Vil09**, Proposition 7.16]. Indeed, let $E_{s,t}$ be the function given by $E_{s,t}(\xi) := (\xi_s, \xi_t)$ for $\xi : [s, t] \to X$ minimizing curve. Since X is a geodesic space, any two points of X can be joined by at least one minimizing curve, so $E_{s,t}$ is onto $X \times X$. Moreover, $E_{s,t}$ is a continuous map between complete separable metric spaces, and $E_{s,t}^{-1}(x, z)$ is compact for every x, z. Therefore, $E_{s,t}$ admits a measurable right-inverse $S_{s \to t}$ [**Del75**], that is, $E_{s,t} \circ S_{s \to t} = \text{Id}$. Thus, $S_{s \to t}$ is a measurable recipe to join two points x, z by a minimizing curve, which was to be proved.

For $t \in (i/2^k, (i + 1)/2^k)$, define $\xi_t^{(k)}$ by $e_t(S_{i/2^k \to (i+1)/2^k}(\xi_{i/2^k}, \xi_{(i+1)/2^k}))$. Then, the law of $(\xi_t^{(k)})_{0 \leq t \leq 1}$ is a probability on $C(X)$. Let us denote it by $\Theta^{(k)}$. Note that $(e_t)_* \Theta^{(k)} = \mu_t$ for every $t = i/2^k$, $i \leq 2^k$, and $\Theta^{(k)}$ is concentrated on Γ.

In short, from ζ, we iteratively constructed probability measures $\Theta^{(k)}$ on Γ up to a step k. We want to pass to the limit as $k \longrightarrow \infty$. Given $\varepsilon > 0$, since μ_0 and μ_1 are Radon measures, there exist compact sets K_0 and K_1 such that $\mu_0(X \setminus K_0) \leq \varepsilon$, $\mu_1(X \setminus K_1) \leq \varepsilon$. Also, the set $\Gamma_{K_0 \to K_1}^{0,1}$ is compact [**Vil09**, Definition 7.13 and Example 7.15] and

$$\Theta^{(k)}(\Gamma \setminus \Gamma^{0,1}_{K_0 \to K_1}) = \mathbb{P}((\xi_0, \xi_1) \notin K_0 \times K_1)$$
$$\leq \mathbb{P}(\xi_0 \notin K_0) + \mathbb{P}(\xi_1 \notin K_1)$$
$$= \mu_0(X \setminus K_0) + \mu_1(X \setminus K_1)$$
$$\leq 2\varepsilon.$$

Then, we can take a subsequence of $(\Theta^{(k)})_k$ that converges weakly to Θ. Since Γ is closed in the topology of uniform convergence [**Vil09**, Theorem 7.16(v)], Θ is supported in Γ. Moreover, given a constant a, for every $t = 1/2^a \in [0, 1]$, if $k > a$, we have $(e_t)_* \Theta^{(k)} = \mu_t$ and, passing to the limit $k \longrightarrow \infty$, $(e_t)_* \Theta = \mu_t$. Finally, since μ_t depends continuously on t, to show that $(e_t)_* \Theta = \mu_t$ for every $t \in [0, 1]$, it suffices to show that $(e_t)_* \Theta$ is continuous as a function of t. In other words, we need to show that, given u a bounded continuous function on X, $U(t) = \mathbb{E} u(\xi_t)$ is a continuous function of t if ξ is random geodesic with law Θ. In fact, since $t \mapsto \xi_t$ is continuous and the composition of continuous functions is also continuous, $t \mapsto u(\xi_t)$ is continuous. Moreover, let $\{u_n\}$ be Lebesgue integrable functions such that $u_n \longrightarrow u$. Since u is bounded, $|u_n| \leq g$ for some integrable function g; by Lebesgue's dominated convergence theorem, $\mathbb{E} u_n \longrightarrow \mathbb{E} u$. From these results, the continuity of $U(t)$ follows, as we wanted.

Accordingly, we constructed ξ such that for each $t \in [0, 1]$, μ_t is the law of ξ_t, where $(\xi_t)_{0 \leq t \leq 1}$ is a dynamical optimal coupling of (μ_0, μ_1). In other terms, we say that ζ is displacement interpolation.

Step 4: An important observation about displacement interpolation. By [**Vil09**, Theorem 8.5], if $\{\mu_t\}$ is a displacement interpolation between two compactly supported probability measures on X, and $t_0 \in (0, 1)$ is given, then, for every $t \in [0, 1]$, the transport map $T_{t_0 \to t}$ between the points $\xi(t_0)$ and $\xi(t)$ is well defined μ_{t_0}-almost everywhere and it is Lipschitz continuous. In other words, $T_{t_0 \to t}$ is a solution of the Monge problem between μ_{t_0} and μ_t. For the completeness of the text, we will comment briefly on the proof.

Note that $(e_0, e_1, e_0, e_1)_* (\Theta \otimes \Theta) = \gamma_{0 \to 1} \otimes \gamma_{0 \to 1}$. So, if one property holds true $\gamma_{0 \to 1} \otimes \gamma_{0 \to 1}$-a.a. for quadruples, this property, for the endpoints of pairs of curves, holds true $\Theta \otimes \Theta$-a.a. Since $\gamma_{0 \to 1}$ is optimal, it has a property named c-cyclical monotonicity [**Vil09**, Theorem 5.10], so that $c(x, y) + c(\tilde{x}, \tilde{y}) \leq c(x, \tilde{y}) + c(\tilde{x}, y)$, $\Theta \otimes \Theta(dx, dy, d\tilde{x}, d\tilde{y})$-a.a. Thus, $c(\xi(0), \xi(1)) + c(\tilde{\xi}(0), \tilde{\xi}(1)) \leq c(\xi(0), \tilde{\xi}(1)) + c(\tilde{\xi}(0), \xi(1))$, $\Theta \otimes \Theta(d\xi, d\tilde{\xi})$-a.a.

Moreover, by Mather's shortening lemma [**Vil09**, Theorem 8.1 and Corollary 8.2],

$$\sup_{0 \leq t \leq 1} d(\xi_t, \tilde{\xi}_t) \leq C_K d(\xi_{t_0}, \tilde{\xi}_{t_0}), \tag{11}$$

where C_K is a constant. Suppose that Θ is supported on a compact set S. Equation (11) defines a closed set for all pairs of curves $\xi, \tilde{\xi} \in S \otimes S$. Let $e_{t_0}(S)$ be the union of all $\xi(t_0)$, when ξ varies over S, and $T_{t_0 \to t}$ by $T_{t_0 \to t}(\xi(t_0)) = \xi(t)$. Note that if $\xi, \tilde{\xi}$ in S are such that $\xi(t_0) = \tilde{\xi}(t_0)$, then equation (11) implies $\xi = \tilde{\xi}$. Moreover, $T_{t_0 \to t}$ is Lipschitz-continuous. So, $(\xi(t_0), T_{t_0 \to t}(\xi(t_0)))$ is a Monge coupling of (μ_{t_0}, μ_t).

Step 5: Absolute continuity of measures μ_t on ζ with respect to the volume measure on X, when μ_0 and μ_1 are compactly supported. Without loss of generality, let us suppose that μ_1 is absolutely continuous with respect to the volume measure on X, vol. If μ_0 and

μ_1 are compactly supported, then ζ has a compact support. Indeed, let A_0, $A_1 \subset X$ be the compact supports of μ_0, μ_1 and $\gamma_{0 \to 1}$ be the transference plan with marginals μ_0 and μ_1. Consider the canonical projections (proj_1), (proj_2) on the first and second components, respectively. Since $(\text{proj}_1)_* \gamma_{0 \to 1} = \mu_0$ and $(\text{proj}_2)_* \gamma_{0 \to 1} = \mu_1$,

$$(\text{proj}_1)_* \gamma_{0 \to 1}(X \times X) = \mu_0(X) = \mu_0(A_0) = \gamma_{0 \to 1}(A_0 \times X),$$

$$(\text{proj}_2)_* \gamma_{0 \to 1}(X \times X) = \mu_1(X) = \mu_1(A_1) = \gamma_{0 \to 1}(X \times A_1).$$

Therefore, $\gamma_{0 \to 1}$ is concentrated in a compact set $A_0 \times A_1$. Moreover, since $\gamma_{0 \to 1} = (e_0, e_1)_* \Theta$ and the evaluation map is continuous, Θ is concentrated in a compact set. The compactness of the ζ support follows from $(e_t)_* \Theta = \mu_t$.

We can use Step 4, and there is a Lipschitz map T solving the Monge problem between μ_t and μ_1, $t \in (0, 1)$. Let N be a set such that the volume measure is zero and consider $T(N)$. If $T(N)$ is not Borel measurable, consider a negligible Borel set that contains $T(N)$ (which by abuse of notation, we will continue denoting $T(N)$). Note that $N \subset T^{-1}(T(N))$, so

$$\mu_t(N) \le \mu_t(T^{-1}(T(N))) = (T_* \mu_t)(T(N)) = \mu_1(T(N))$$

and then $\mu_t(N) = 0$, since $\text{vol}(T(N)) \le \|T\|_{\text{Lip}} \text{vol}(N) = 0$ (this inequality occurs since T is Lipschitz and $\|T\|_{\text{Lip}}$ stands for the Lipschitz constant) and $\mu_1 \ll \text{vol}$ by hypothesis. So, $\mu_t(N) = 0$ for every Borel set N such that $\text{vol}(N) = 0$, that is, $\mu_t \ll \text{vol}$.

Step 6: Absolute continuity of measures μ_t of ζ with respect to the volume measure on X: the general case. Without loss of generality, suppose $\mu_1 \ll \text{vol}$. Let us assume that there is some case in which neither μ_0 nor μ_1 is compactly supported. We will prove our statement (ii) by contradiction. Suppose that μ_τ, for $\tau \in (0, 1)$, is not absolutely continuous with respect to the volume measure on X. Then, there exists a set $Z_\tau \subset X$, such that $\text{vol}(Z_\tau) = 0$ and $\mu_\tau(Z_\tau) > 0$. Consider $\mathcal{Z} := \{\xi \in \Gamma : \xi_\tau \in Z_\tau\}$ so that $\Theta(\mathcal{Z}) = \mathbb{P}(\xi_{s\tau} \in Z_\tau) = \mu_\tau(Z_\tau) > 0$. Since Θ is a regular measure, there exists $\mathcal{K} \subset \mathcal{Z}$ compact such that $\Theta(\mathcal{K}) > 0$. So, if we set

$$\Theta' := \frac{\Theta \, \mathbb{1}_{\mathcal{K}}}{\Theta(\mathcal{K})}$$

and consider $\gamma'_{0 \to 1} := (e_0, e_1)_* \Theta'$ and $\mu'_t = (e_t)_* \Theta'$, we have

$$\mu'_t \le \frac{(e_t)_* \Theta}{\Theta(\mathcal{K})} = \frac{\mu_t}{\Theta(\mathcal{K})}.$$

In this way, (μ'_t) is a displacement interpolation and, considering the previous equation for $t = 1$, $\mu'_1 \ll \mu_1 \ll \text{vol}$. Note that now, μ'_τ is concentrated on $e_\tau(\mathcal{K}) \subset e_\tau(\mathcal{Z}) \subset Z_\tau$ and then μ'_τ is singular. However, μ'_0 is supported in $e_0(\mathcal{K})$ and μ'_1 is supported in $e_1(\mathcal{K})$, which are compact. This is the case of Step 5. Then, $\mu'_\tau \ll \text{vol}$, which is a contradiction.

(iii) In this item, we denote $Y := \Omega$, where Ω is the subset of the quotient space X^* on which the metric measure foliation is defined (see Definition 5.6), and $\pi := p$, where p is the quotient map. The weak continuity of f in this case was proved in Proposition 5.8. Let ψ be a minimizing curve on Y connecting y and y'. Consider the path $\zeta = \{\mu_{y_t} : t \in [0, 1]\}$

as constructed in item (i). Since ψ was taken as a minimizing curve and f is an isometry, ζ is minimizing.

Observe that every optimal transference plan between μ_y and $\mu_{y'}$ is supported on $\{(x, x') \in \pi^{-1}(y) \times \pi^{-1}(y') : d^*(y, y') = d(x, x')\}$. In fact, let $\gamma_{y \to y'}$ be an optimal transference plan for $\mu_y, \mu_{y'}$. Since $\mathrm{supp}(\mu_y) \subset \pi^{-1}(y)$, $\mathrm{supp}(\mu_{y'}) \subset \pi^{-1}(y')$ and π is 1-Lipschitz, we have $\gamma_{y \to y'}$ supported on $\{(x, x') \in \pi^{-1}(y) \times \pi^{-1}(y') : d^*(y, y') \le d(x, x')\}$. Consider the set $\Upsilon := \{(x, x') \in \pi^{-1}(y) \times \pi^{-1}(y') : d^*(y, y') < d(x, x')\}$. If $\gamma_{y \to y'}(\Upsilon) > 0$, then

$$
\begin{aligned}
\gamma_{y \to y'}(\Upsilon)\, d^*(y, y')^2 &\le \gamma_{y \to y'}(\Upsilon)\, d(x, x')^2 \\
&< \int_\Upsilon d(x, x')^2 \, d\gamma_{y \to y'}.
\end{aligned}
$$

So,

$$
\begin{aligned}
d^*(y, y')^2 &< \int_\Upsilon d(x, x')^2 \, d\gamma_{y \to y'} + \gamma_{y \to y'}(X \times X \setminus \Upsilon)\, d^*(y, y')^2 \\
&= \int_\Upsilon d(x, x')^2 \, d\gamma_{y \to y'} + \int_{X \times X \setminus \Upsilon} d^*(y, y')^2 \, d\gamma_{y \to y'} \\
&\le \int_\Upsilon d(x, x')^2 \, d\gamma_{y \to y'} + \int_{X \times X \setminus \Upsilon} d(x, x')^2 \, d\gamma_{y \to y'} \\
&= \int_{X \times X} d(x, x')^2 \, d\gamma_{y \to y'}.
\end{aligned}
$$

Then, $d^*(y, y')^2 < W_2(\mu_y, \mu_{y'})^2$, which is a contradiction.

Therefore, we have the transport between μ_y and $\mu_{y'}$ orthogonal to the leaves. Furthermore, since ζ is minimizing, the existence of the optimal transport plan is guaranteed.

Suppose without loss of generality that μ_y is absolutely continuous with respect to the volume measure of the leaf y. Then, the support of μ_y contains more than one point. Considering the transport problem described above, for each x in the support of μ_y such that (x, x') is in the support of $\gamma_{y \to y'}$ for some x', each one of the intermediate leaves y_t must contain a corresponding point x_{y_t}. Thus, each leaf will have the distribution $\mu_{y(t)}$ absolutely continuous with respect to the volume measure of the respective leaf. $\square$

Remark 5.11. Since ζ is a displacement interpolation, it is a constant speed geodesic, by [**AG13**, Theorem 2.10]. That is, the path in $(\mathscr{P}(X), W_2)$ given by the disintegration map f evaluated at the minimal curve ψ in Y is a constant speed geodesic in $\mathscr{P}_2(X)$.

Remark 5.12. One could, for instance, take the disintegration map in the domain of weak continuity given by Proposition 5.2.

Remark 5.13. Theorem B(iii) holds, for instance, in the case of the disintegration of the volume measure in the solid torus (Example 3.2), or in the context of Examples 3.3, 5.5 and 5.7.

REFERENCES

[AG13] L. Ambrosio and N. Gigli. *A User's Guide to Optimal Transport (Lecture Notes in Mathematics, 2062)*. Springer-Verlag, Berlin, 2013.

[AGS05] L. Ambrosio, N. Gigli and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Birkhäuser Verlag, Basel, 2005.

[Amb00] L. Ambrosio. Lecture notes on optimal transport problems. *Mathematical Aspects of Evolving Interfaces (Lecture Notes in Mathematics, 1812)*. Springer, Berlin, 2000.

[AP03] L. Ambrosio and A. Pratelli. *Existence and Stability Results in the L1 Theory of Optimal Transportation (Lecture Notes in Mathematics, 1813)*. Springer, Berlin, 2003.

[BM17] O. Butterley and I. Melbourne. Disintegration of invariant measures for hyperbolic skew products. *Israel J. Math.* **219** (2017), 171–188.

[CP97] J. T. Chang and D. Pollard. Conditioning as disintegration. *Stat. Neerl.* **51**(3) (1997), 287–317.

[Del75] C. Dellacherie. Ensembles analytiques. Théorèmes de séparation et applications. *Séminaire de Probabilités (Lecture Notes in Mathematics, 465)*. Ed. P. A. Meyer. Springer, Berlin, 1975.

[DM78] C. Dellacherie and P. A. Meyer. *Probabilities and Potential (North-Holland Mathematics Studies, 29)*. North-Holland Publishing Company, Amsterdam, 1978.

[Fed69] H. Federer. *Geometric Measure Theory (Grundlehren der Mathematischen Wissenschaften, 153)*. Springer, New York, 1969.

[Gal17] S. Galatolo. Statistical properties of dynamics: introduction to the functional analytic approach. *Preprint*, 2017, arXiv:1510.02615v2.

[GKMS18] F. Galaz-García, M. Kell, A. Mondino and G. Sosa. On quotients of spaces with Ricci curvature bounded below. *J. Funct. Anal.* **275** (2018), 1368–1446.

[GL20] S. Galatolo and R. Lucena. Spectral gap and quantitative statistical stability for systems with contracting fibers and Lorenz-like maps. *Discrete Contin. Dyn. Syst.* **40** (2020), 1309–1360.

[GM13] L. Granieri and F. Maddalena. Transport problems and disintegration maps. *ESAIM Control Optim. Calc. Var.* **19**(3) (2013), 888–905.

[Jost11] J. Jost. *Riemannian Geometry and Geometric Analysis*. Springer-Verlag, Berlin, 2011.

[OV14] K. Oliveira and M. Viana. *Fundamentos da Teoria Ergódica*. Sociedade Brasileira de Matemática, Rio de Janeiro, 2014.

[Par67] K. R. Parthasarathy. *Probability Measures on Metric Spaces*. Academic Press, New York, 1967.

[Rok52] V. A. Rokhlin. *On the Fundamental Ideas of Measure Theory (American Mathematical Society Translation, 71)*. American Mathematical Society, Providence, RI, 1952.

[Sim12] D. Simmons. Conditional measures and conditional expectation; Rohlin's disintegration theorem. *Discrete Contin. Dyn. Syst.* **32**(7) (2012), 2565–2582.

[Stu06] K. T. Sturm. On the geometry of metric measure spaces. *Acta Math.* **196** (2006), 65–131.

[Tue75] T. Tjur. *A Constructive Definition of Conditional Distributions (Institute of Mathematical Statistics, 13)*. University of Copenhagen, Copenhagen, 1975.

[Var16] R. Varão. Center foliation: absolute continuity, disintegration and rigidity. *Ergod. Th. & Dynam. Sys.* **36** (2016), 256–275.

[Vil03]   C. Villani. *Topics in Optimal Transportation (Graduate Studies in Mathematics, 58)*. American
              Mathematical Society, Providence, RI, 2003.
[Vil09]   C. Villani. *Optimal Transport: Old and New (Grundlehren der Mathematischen Wissenschaften,
              338)*. Springer-Verlag, Berlin, 2009.
[Von32]   J. Von Neumann. Zur Operatorenmethode in der klassischen Mechanik. *Ann. of Math. (2)* **33** (1932),
              587–642.