ABSOLUTELY CONTINUOUS CURVES IN EXTENDED WASSERSTEIN-ORLICZ SPACES

STEFANO LISINI

Abstract. In this paper we extend a previous result of the author [Lis07] of characterization of absolutely continuous curves in Wasserstein spaces to a more general class of spaces: the spaces of probability measures endowed with the Wasserstein-Orlicz distance constructed on extended Polish spaces (in general non separable), recently considered in [AGS14]. An application to the geodesics of this Wasserstein-Orlicz space is also given.

1. Introduction

In this paper we extend a previous result of the author [Lis07] to a more general class of spaces. The result in [Lis07] concerns the representation of absolutely continuous curves with finite energy in the Wasserstein space $(\mathscr{P}(X, \mathsf{d}), W_p)$ (the space of Borel probability measures on a Polish metric space (X, d), endowed with the p-Wasserstein distance induced by d) by means of superposition of curves of the same kind on the space (X, d). The superposition is described by a probability measure on the space of continuous curves in (X, d) representing the curve in $(\mathscr{P}(X, \mathsf{d}), W_p)$ and satisfying a suitable property.

Here we extend the previous representation result in two directions: in the first one we consider a so-called extended Polish space (X, τ, d) instead of a Polish space (X, d); in the second one we consider the ψ-Orlicz-Wasserstein distance induced by an increasing convex function $\psi : [0, +\infty) \to [0, +\infty]$ instead of the p-Wasserstein distance modelled on the particular case of $\psi(r) = r^p$ for $p > 1$.

The class of extended Polish spaces was introduced in the recent paper [AGS14]. The authors consider a Polish space (X, τ), i.e. τ is a separable topology on X induced by a distance δ on X such that (X, δ) is complete. The Wasserstein distance is defined between Borel probability measures on (X, τ) and constructed by means of an extended distance d on X that can assume the value $+\infty$. The minimization problem defining the extended Wasserstein distance makes sense between Borel probability measures on (X, τ), assuming that the extended distance d is lower semi continuous with respect to τ.

A typical example of extended Polish space is the abstract Wiener space (X, τ, γ) where (X, τ) is a separale Banach space and τ is the topology induced by the norm, γ is a gaussian reference measure on X with zero mean and supported on all the space. The extended distance is given by $\mathsf{d}(x, y) = |x - y|_H$ if $x - y \in H$, where H is the Cameron-Martin space associated to γ in X and $|\cdot|_H$ is the Hilbertian norm of H, and $\mathsf{d}(x, y) = +\infty$ if $x - y \notin H$ (see for instance [Str11]).

2010 *Mathematics Subject Classification.* 49J27, 49J52.

Key words and phrases. Spaces of probability measures, Wasserstein-Orlicz distance, absolutely continuous curves, superposition principle, geodesic in spaces of probability measures.

The Wasserstein-Orlicz distance is still unexplored. At the author's knowledge, only the papers [Stu11] and, more recently, [Kuw13] deal with this kind of spaces. In the paper [FGY11, Remark 3.19], the authors discuss the possibility to use this kind of Wasserstein-Orlicz distance to extend their results for equation of the form $\partial_t u - \operatorname{div}(u\nabla H(u^{-1}\nabla u) = 0$ to the case of a convex function H with non power growth.

Only the particular case of the Wasserstein-Orlicz distance W_∞, corresponding to the function $\psi(s) = 0$ if $s \in [0,1]$ and $\psi(s) = +\infty$ if $s \in (1,+\infty)$ has been deeply investigated. The extension of the representation Theorem of [Lis07] to the W_∞ case has been proved in [ADM14]. Another refinement of the representation Theorem of [Lis07] is contained in [BS11, Sec. 5]. The problem of the validity of the representation Theorem of [Lis07] in the case of a general Wasserstein-Orlicz space is raised in the last section of [AGS13].

For the precise statement of the result we address to Theorem 3.1. The strategy of the proof is similar to the one used to prove Theorem 5 of [Lis07], but there are several additional difficulties because (X, d), in general, is non separable and the function ψ that induces the Wasserstein-Orlicz distance is not homogeneous.

The paper is structured as follows: in Section 2 we introduce the framework of our study and some preliminary results, in Section 3 we state and prove the main theorem of the paper, and finally in Section 4 we apply the main theorem in order to characterize the geodesics of the Wasserstein-Orlicz space.

2. Notation and preliminary results

2.1. **Extended Polish spaces and probability measures.** Given a set X, we say that $\mathsf{d} : X \times X \to [0,+\infty]$ is an extended distance if

- $\mathsf{d}(x,y) = \mathsf{d}(y,x)$ for every $x, y \in X$,
- $\mathsf{d}(x,y) = 0$ if and only if $x = y$,
- $\mathsf{d}(x,y) \le \mathsf{d}(x,z) + \mathsf{d}(z,y)$ for every $x, y, z \in X$.

$(X, \mathsf{d}))$ is called extended metric space. We observe that the only difference between a distance and an extended distance is that $\mathsf{d}(x,y)$ could be equal to $+\infty$.

We say that (X, τ, d) is a Polish extended space if:

(i) τ is a topology on X and (X, τ) is Polish, i.e. τ is induced by a distance δ such that the metric space (X, δ) is separable and complete;
(ii) d is an extended distance on X and (X, d) is a complete extended metric space;
(iii) For every sequence $\{x_n\} \subset X$ such that $\mathsf{d}(x_n, x) \to 0$ with $x \in X$, we have that $x_n \to x$ with respect to the topology τ;
(iv) d is lower semicontinuous in $X \times X$, with respect to the $\tau \times \tau$ topology; i.e.,

$$\liminf_{n\to+\infty} \mathsf{d}(x_n, y_n) \ge \mathsf{d}(x,y), \qquad \forall (x,y) \in X \times X, \quad \forall (x_n, y_n) \to (x,y) \text{ w.r.t. } \tau \times \tau. \tag{1}$$

In the sequel, the class of compact sets, the class of Borel sets $\mathscr{B}(X)$, the class $C_b(X)$ of bounded continuous functions and the class $\mathscr{P}(X)$ of Borel probability measures, are always referred to the topology τ, even when d is a distance.

We say that a sequence $\mu_n \in \mathscr{P}(X)$ *narrowly converges* to $\mu \in \mathscr{P}(X)$ if

$$\lim_{n\to+\infty} \int_X \varphi(x)\, d\mu_n(x) = \int_X \varphi(x)\, d\mu(x) \qquad \forall \varphi \in C_b(X). \tag{2}$$

It is well known that the narrow convergence is induced by a distance on $\mathscr{P}(X)$ (see for instance [AGS05, Remark 5.1.1]) and we call *narrow topology* the topology induced by this distance. In particular the compact subsets of $\mathscr{P}(X)$ coincides with sequentially compact subsets of $\mathscr{P}(X)$.
We also recall that if $\mu_n \in \mathscr{P}(X)$ narrowly converges to $\mu \in \mathscr{P}(X)$ and $\varphi : X \to (-\infty, +\infty]$ is a lower semi continuous (with respect to τ) function bounded from below, then

$$\liminf_{n\to+\infty} \int_X \varphi(x)\, d\mu_n(x) \geq \int_X \varphi(x)\, d\mu(x). \tag{3}$$

A subset $\mathscr{T} \subset \mathscr{P}(X)$ is said to be tight if

$$\forall \varepsilon > 0 \quad \exists K_\varepsilon \subset X \text{ compact} : \mu(X \setminus K_\varepsilon) < \varepsilon \quad \forall \mu \in \mathscr{T}, \tag{4}$$

or, equivalently, if there exists a function $\varphi : X \to [0, +\infty]$ with compact sublevels $\lambda_c(\varphi) := \{x \in X : \varphi(x) \leq c\}$, such that

$$\sup_{\mu\in\mathscr{T}} \int_X \varphi(x)\, d\mu(x) < +\infty. \tag{5}$$

By Prokhorov Theorem, a set $\mathscr{T} \subset \mathscr{P}(X)$ is tight if and only if $\mathscr{T}$ is relatively compact in $\mathscr{P}(X)$. In particular, the Polish condition on τ guarantees that all Borel probability measures $\mu \in \mathscr{P}(X)$ are tight.

2.2. **Orlicz spaces.** Given

$$\begin{gathered} \psi : [0, +\infty) \to [0, +\infty] \text{ convex, lower semicontinuous, non-decreasing, } \psi(0) = 0, \\ \lim_{x\to+\infty} \psi(x) = +\infty, \end{gathered} \tag{6}$$

a measure space (Ω, ν) and a ν-measurable function $u : \Omega \to \mathbb{R}$, the $L^\psi_\nu(\Omega)$ Orlicz norm of u is defined by

$$\|u\|_{L^\psi_\nu(\Omega)} := \inf \left\{ \lambda > 0 : \int_\Omega \psi\Big(\frac{|u|}{\lambda}\Big)\, d\nu \leq 1 \right\}.$$

The Orlicz space $L^\psi_\nu(\Omega) := \{u : \Omega \to \mathbb{R},\ \text{measurable} : \|u\|_{L^\psi_\nu(\Omega)} < +\infty\}$ is a Banach space. For the theory of the Orlicz spaces we refer to the complete monography [RR91].

Given a bounded sequence $\{w_n\} \subset L^\psi_\nu(\Omega)$, the following property of lower semi continuity of the norm holds:

$$\liminf_{n\to\infty} w_n(x) \geq w(x) \quad \text{for } \nu\text{-a.e. } x \in \Omega \quad \Longrightarrow \quad \liminf_{n\to\infty} \|w_n\|_{L^\psi_\nu(\Omega)} \geq \|w\|_{L^\psi_\nu(\Omega)}. \tag{7}$$

Indeed, denoting by $\lambda_n := \|w_n\|_{L^\psi_\nu(\Omega)}$ and $\lambda := \liminf_n \lambda_n$, up to extract a subsequence we can assume that $\lambda = \lim_n \lambda_n$. By the lower semicontinuity and the monotonicity of ψ we have

$$\liminf_{n\to\infty} \psi\Big(\frac{w_n(x)}{\lambda_n}\Big) \geq \psi\Big(\frac{w(x)}{\lambda}\Big) \qquad \text{for } \nu\text{-a.e. } x \in \Omega.$$

Finally, by Fatou's lemma

$$1 \geq \liminf_{n\to\infty} \int_\Omega \psi\Big(\frac{w_n(x)}{\lambda_n}\Big)\, d\nu(x) \geq \int_\Omega \psi\Big(\frac{w(x)}{\lambda}\Big)\, d\nu(x)$$

which shows that $\lambda \geq \|w\|_{L^\psi_\nu(\Omega)}$.

We denote by $\psi^* := [0,+\infty) \to [0,+\infty]$ the conjugate of ψ defined by $\psi^*(y) = \sup_{x\geq 0}\{xy - \psi(x)\}$. The following generalized Hölder's inequality holds

$$\int_\Omega u(x)v(x)\,d\nu(x) \leq 2\|u\|_{L^\psi_\nu(\Omega)}\|v\|_{L^{\psi^*}_\nu(\Omega)}, \tag{8}$$

and the following equivalence between the Orlicz norm in $L^\psi_\nu(\Omega)$ and the dual norm of $L^{\psi^*}_\nu(\Omega)$ holds

$$\|u\|_{L^\psi_\nu(\Omega)} \leq \sup\left\{\int_\Omega |u(x)v(x)|\,d\nu(x) : v \in L^{\psi^*}_\nu(\Omega), \|v\|_{L^{\psi^*}_\nu(\Omega)} \leq 1\right\} \leq 2\|u\|_{L^\psi_\nu(\Omega)}. \tag{9}$$

In the statement of our main theorem we will assume, in addition to (6), that ψ is superlinear at $+\infty$, i.e.

$$\lim_{x\to+\infty}\frac{\psi(x)}{x} = +\infty, \tag{10}$$

and it has null right derivative at 0, i.e.

$$\lim_{x\to 0}\frac{\psi(x)}{x} = 0. \tag{11}$$

It is easy to check that conditions (10) and (11) are equivalent to assume that $\psi^*(y) > 0$ and $\psi^*(y) < +\infty$ for every $y > 0$.

Typical examples of admissible ψ satisfying (6), (10) and (11) are:

- $\psi(x) = x^p$ for $p \in (1,+\infty)$ and the corresponding Orlicz norm is the standard L^p norm;
- $\psi(x) = 0$ if $x \in [0,1]$ and $\psi(x) = +\infty$ if $x \in (1,+\infty)$ and the corresponding Orlicz norm is the L^∞ norm;
- $\psi(x) = e^x - x - 1$, exponential growth;
- $\psi(x) = e^{x^p} - 1$ for $p \in (1,+\infty)$, power exponential growth;
- $\psi(x) = (1+x)\ln(1+x) - x$, $L\log L$-growth.

2.3. **Continuous curves.** Given (X,τ,d) an extended Polish space, $I := [0,T]$, $T > 0$, we denote by $C(I;X)$ the space of continuous curves in X with respect to the topology τ. $C(I;X)$ is a Polish space with the metric

$$\delta_\infty(u,\tilde u) = \sup_{t\in I}\delta(u(t),\tilde u(t)), \tag{12}$$

where δ is a complete and separable metric on X inducing τ.

Given ψ satisfying (6), we say that a curve $u : I \to X$ belongs to $AC^\psi(I;(X,\mathsf{d}))$, if there exists $m \in L^\psi(I)$ such that

$$\mathsf{d}(u(s),u(t)) \leq \int_s^t m(r)\,dr \qquad \forall s,t \in I, \quad s \leq t. \tag{13}$$

We also denote by $AC(I;(X,\mathsf{d}))$ the set $AC^\psi(I;(X,\mathsf{d}))$ for $\psi(r) = r$. We call a curve $u \in AC^\psi(I;(X,\mathsf{d}))$ an absolutely continuous curve with finite L^ψ-energy.

It can be proved that (see [AGS05, Theorem 1.1.2]) for every $u \in AC^{\psi}(I;(X,\mathsf{d}))$, there exists the following limit, called metric derivative,

$$|u'|(t) := \lim_{h\to 0} \frac{\mathsf{d}(u(t+h),u(t))}{|h|} \qquad \text{for } \mathscr{L}^1\text{-a.e. } t\in I. \tag{14}$$

The function $t\mapsto |u'|(t)$ belongs to $L^{\psi}(I)$ and it is the minimal one that satisfies (13).

The following Lemma will be useful in the proof of our main theorem.

Lemma 2.1. *Let ψ be satisfying (6), (10) and (11). If $u : I \to (X,\mathsf{d})$ is right continuous at every point and continuous except at most a countable set, and*

$$\limsup_{h\to 0^+} \left\| \frac{\mathsf{d}(u(\cdot+h),u(\cdot))}{h} \right\|_{L^{\psi}(I)} < +\infty, \tag{15}$$

where u is extended for $t > T$ as $u(t) = u(T)$, then $u \in AC^{\psi}(I;(X,\mathsf{d}))$.

Proof. Since I is bounded, by the assumptions on u we have that the d-closure of $u(I)$ is compact in (X,d). Consequently $u(I)$ is d-separable. We consider a sequence $\{y_n\}_{n\in\mathbb{N}}$ dense in $(u(I),\mathsf{d})$. We fix $n\in\mathbb{N}$. Defining $u_n : I \to \mathbb{R}$ by $u_n(t) := \mathsf{d}(u(t),y_n)$, the triangular inequality implies

$$|u_n(t+h) - u_n(t)| \le \mathsf{d}(u(t+h),u(t)), \qquad \forall t\in I, h>0. \tag{16}$$

Given a test function $\eta\in C_c^{\infty}(I)$ and $h>0$, recalling Hölder inequality (8) we obtain

$$\begin{aligned}\left|\int_I u_n(t)\frac{\eta(t-h)-\eta(t)}{h}\,dt\right| &= \left|\int_I \eta(t)\frac{u_n(t+h)-u_n(t)}{h}\,dt\right| \\ &\le 2\left\|\frac{u_n(\cdot+h)-u_n(\cdot)}{h}\right\|_{L^{\psi}(I)} \|\eta\|_{L^{\psi^*}(I)}.\end{aligned}$$

By the last inequality, (15) and (16), passing to the limit for $h\to 0$ we have that

$$\left|\int_I u_n(t)\eta'(t)\,dt\right| \le C\,\|\eta\|_{L^{\psi^*}(I)}. \tag{17}$$

The linear functional $\mathscr{L}_n : (C_c^{\infty}(I),\|\cdot\|_{L^{\psi^*}(I)}) \to \mathbb{R}$ defined by $\mathscr{L}_n(\eta) = \int_I u_n(t)\eta'(t)\,dt$, by (17), is bounded and we still denote by $\mathscr{L}_n$ its extension to $E^{\psi^*}(I)$, the closure of $C_c^{\infty}(I)$ with respect to the norm $\|\cdot\|_{L^{\psi^*}(I)}$. Since, by (10) and (11), ψ^* is continuous and strictly positive on $(0,+\infty)$, $\mathscr{L}_n$ is uniquely represented by an element $v_n\in L^{\psi^{**}}(I)$ (see Theorem 6, pag. 105 of [RR91]). The element v_n coincides with the distributional derivative of u_n and then $u_n\in AC^{\psi}(I;\mathbb{R})$ (we observe that $\psi^{**}=\psi$ because ψ is convex and lower semi continuous). We denote by $u_n'(t)$ the pointwise derivative of u_n which exists for a.e. $t\in I$.

We introduce the negligible set

$$N = \bigcup_{n\in\mathbb{N}} \{t\in I : u_n'(t) \text{ does not exists}\},$$

and we define $m(t) := \sup_{n\in\mathbb{N}} |u_n'(t)|$ for all $t\in I\setminus N$. By the density of $\{y_n\}_{n\in\mathbb{N}}$ in $u(I)$, we have that for all $t,s\in I$, with $s<t$,

$$\mathsf{d}(u(t),u(s)) = \sup_{n\in\mathbb{N}} |u_n(t)-u_n(s)| \le \sup_{n\in\mathbb{N}} \int_s^t |u_n'(r)|\,dr \le \int_s^t m(r)\,dr. \tag{18}$$

We show that $m \in L^\psi(I)$. Actually, by (16), if $t \in I \setminus N$ then

$$|u_n'(t)| = \lim_{h\to 0^+} \frac{|u_n(t+h) - u_n(t)|}{h} \leq \liminf_{h\to 0^+} \frac{\mathsf{d}(u(t+h), u(t))}{h},$$

which implies $m(t) \leq \liminf_{h\to 0^+} \frac{\mathsf{d}(u(t+h), u(t))}{h}$. By (15) and (7) we conclude. $\square$

2.4. **The $\mathcal{M}(I;X)$ space.** We denote by $\mathscr{M}(I;X)$ the space of curves $u : I \to X$ which are Lebesgue measurable as functions with values in (X,τ). We denote by $\mathcal{M}(I;X)$ the quotient space of $\mathscr{M}(I;X)$ with respect to the equality $\mathscr{L}^1$-a.e. in I. The space $\mathcal{M}(I;X)$ is a Polish space endowed with the metric

$$\delta_1(u,v) := \int_0^T \tilde{\delta}(u(t), v(t))\, dt,$$

where $\tilde{\delta}(x,y) := \min\{\delta(x,y), 1\}$ is a bounded distance still inducing τ and δ is a distance inducing τ.

The space $\mathcal{M}(I;X)$ coincides with $L^1(I;(X,\tilde{\delta}))$. It is well known that $\delta_1(u_n,u) \to 0$ as $n \to +\infty$ if and only if $u_n \to u$ in measure as $n \to +\infty$; i.e.

$$\lim_{n\to+\infty} \mathscr{L}^1(\{t \in I : \delta(u_n(t), u(t)) > \sigma\}) = 0, \qquad \forall\, \sigma > 0.$$

We recall a useful compactness criterion in $\mathcal{M}(I;X)$, [RS03, Theorem 2].

Theorem 2.2. *A family $\mathscr{A} \subset \mathcal{M}(I;X)$ is precompact if there exists a function $\Psi : X \to [0,+\infty]$ whose sublevels $\lambda_c(\Psi) := \{x \in X : \Psi(x) \leq c\}$ are compact for every $c \geq 0$, such that*

$$\sup_{u\in\mathscr{A}} \int_0^T \Psi(u(t))\, dt < +\infty, \tag{19}$$

and there exists a map $g : X \times X \to [0,\infty]$ lower semi continuous with respect to $\tau \times \tau$ such that

$$g(x,y) = 0 \quad \Longrightarrow \quad x = y$$

and

$$\lim_{h\to 0^+} \sup_{u\in\mathscr{A}} \int_0^{T-h} g(u(t+h), u(t))\, dt = 0.$$

2.5. **Push forward of probability measures.** If Y, Z are topological spaces, $\mu \in \mathscr{P}(Y)$ and $F : Y \to Z$ is a Borel map (or a μ-measurable map), the *push forward of μ through F*, denoted by $F_\#\mu \in \mathscr{P}(Z)$, is defined as follows:

$$F_\#\mu(B) := \mu(F^{-1}(B)) \qquad \forall B \in \mathscr{B}(Z). \tag{20}$$

It is not difficult to check that this definition is equivalent to

$$\int_Z \varphi(z)\, d(F_\#\mu)(z) = \int_Y \varphi(F(y))\, d\mu(y) \tag{21}$$

for every bounded Borel function $\varphi : Z \to \mathbb{R}$. More generally (21) holds for every $F_\#\mu$-integrable function $\varphi : Z \to \mathbb{R}$.

We recall the following composition rule: for every $\mu \in \mathscr{P}(Y)$ and for all Borel maps $F : Y \to Z$ and $G : Z \to W$, we have

$$(G \circ F)_{\#}\mu = G_{\#}(F_{\#}\mu).$$

The following continuity property holds:

$$F : Y \to Z \quad \text{continuous} \quad \Longrightarrow \quad F_{\#} : \mathscr{P}(Y) \to \mathscr{P}(Z) \quad \text{narrowly continuous.}$$

We say that $\mu \in \mathscr{P}(Y)$ is concentrated on the set A if $\mu(X \setminus A) = 0$. It follows from the definition that $F_{\#}\mu$ is concentrated on $F(A)$ if μ is concentrated on A.

The support of a Borel probability measure $\mu \in \mathscr{P}(Y)$ is the closed set defined by $\operatorname{supp}\mu = \{y \in Y : \mu(U) > 0, \forall U \text{ neighborhood of } y\}$. μ is concentrated on $\operatorname{supp}\mu$ and it is the smallest closed set on which μ is concentrated.

In general we have $F(\operatorname{supp}\mu) \subset \operatorname{supp} F_{\#}\mu \subset \overline{F(\operatorname{supp}\mu)}$ for $F : Y \to Z$ continuous.

It follows that $F_{\#}\mu(\operatorname{supp} F_{\#}\mu \setminus F(\operatorname{supp}\mu)) = 0$.

The following Lemma is fundamental in our proof of Theorem 3.1. It allows to recover a pointwise bound assuming an integral bound.

Lemma 2.3. *Let Y be a Polish space and $\{\mu_n\}_{n\in\mathbb{N}} \subset \mathscr{P}(Y)$ be a sequence narrowly convergent to $\mu \in \mathscr{P}(Y)$ as $n \to +\infty$. Let $F_n : Y \to [0, +\infty)$ be a sequence of μ_n-measurable functions such that*

$$\sup_{n\in\mathbb{N}} \int_Y F_n(y)\, d\mu_n(y) < +\infty. \tag{22}$$

Then there exists a subsequence μ_{n_k} such that

(23)

$$\text{for } \mu\text{-a.e. } \bar{y} \in \operatorname{supp}\mu \quad \exists y_{n_k} \in \operatorname{supp}\mu_{n_k} : \lim_{k\to+\infty} y_{n_k} = \bar{y} \qquad \text{and} \qquad \sup_{k\in\mathbb{N}} F_{n_k}(y_{n_k}) < +\infty.$$

Proof. Let us define the sequence $\nu_n := (\mathsf{i} \times F_n)_{\#}\mu_n \in \mathscr{P}(Y \times \mathbb{R})$, where i denotes the identity map in Y. We denote by $\pi^1 : Y \times \mathbb{R} \to Y$ and $\pi^2 : Y \times \mathbb{R} \to \mathbb{R}$ the projections defined by $\pi^1(y, z) = y$ and $\pi^2(y, z) = z$. The set $\{\nu_n\}_{n\in\mathbb{N}}$ is tight because $\{\pi^1_{\#}\nu_n\}_{n\in\mathbb{N}}$ and $\{\pi^2_{\#}\nu_n\}_{n\in\mathbb{N}}$ are tight. Indeed $\pi^1_{\#}\nu_n = \mu_n$ is narrowly convergent, and $\pi^2_{\#}\nu_n = (F_n)_{\#}\mu_n$ has first moments uniformly bounded because

$$\int_{\mathbb{R}} |z|\, d\pi^2_{\#}\nu_n(z) = \int_Y |F_n(y)|\, d\mu_n(y),$$

$F_n \geq 0$ and (22) holds. By Prokhorov's Theorem there exists $\nu \in \mathscr{P}(Y \times \mathbb{R})$ and a subsequence $\{\nu_{n_k}\}_{k\in\mathbb{N}} \subset \mathscr{P}(Y \times \mathbb{R})$ narrowly convergent to ν. Since $\pi^1_{\#}\nu_n = \mu_n$ and $\pi^1_{\#}\nu_{n_k} \to \pi^1_{\#}\nu$ as $k \to +\infty$ we have that $\pi^1_{\#}\nu = \mu$.

Let $\bar{y} \in \pi^1(\operatorname{supp}\nu)$, and we observe that $\mu(\operatorname{supp}\mu \setminus \pi^1(\operatorname{supp}\nu)) = 0$. By definition of $\bar{y}$ there exists $z \in \mathbb{R}$ such that $(\bar{y}, z) \in \operatorname{supp}\nu$. Let $h \in \mathbb{N}$ and $D_{1/h}(\bar{y}, z) := B_{1/h}(\bar{y}) \times (z - 1/h, z + 1/h)$ where $B_r(\bar{y})$ denotes the open ball of radius r and center $\bar{y}$, when a distance in Y is fixed. By (3), with φ the characteristic function of $D_{1/h}(\bar{y}, z)$, we obtain

$$\liminf_{k\to+\infty} \nu_{n_k}(D_{1/h}(\bar{y}, z)) \geq \nu(D_{1/h}(\bar{y}, z)) > 0.$$

Then there exists $k(h) \in \mathbb{N}$ such that

$$\nu_{n_k}(D_{1/h}(\bar{y}, z)) > 0 \qquad \forall k \geq k(h). \tag{24}$$

By definition of ν_n

$$\begin{aligned}\nu_{n_k}(D_{1/h}(\bar y,z)) &= \mu_{n_k}(\{y\in Y:(\mathsf{i}\times F_{n_k})(y)\in D_{1/h}(\bar y,z)\})\\ &= \mu_{n_k}(\{y\in Y:(y,F_{n_k}(y))\in B_{1/h}(\bar y)\times(z-1/h,z+1/h)\}).\end{aligned}\tag{25}$$

By (24) and (25) we have that

$$\operatorname{supp}\mu_{n_k}\cap\{y\in Y:(y,F_{n_k}(y))\in B_{1/h}(\bar y)\times(z-1/h,z+1/h)\}\neq\emptyset\qquad\forall k\geq k(h).\tag{26}$$

Since we can choose the application $h\mapsto k(h)$ strictly increasing, by (26) we can select a sequence $y_{n_k}\in\operatorname{supp}\mu_{n_k}\cap\{y\in Y:(y,F_{n_k}(y))\in B_{1/h}(\bar y)\times(z-1/h,z+1/h)\}$. By definition $y_{n_k}\to\bar y$ and $F_{n_k}(y_{n_k})\to z$ as $k\to+\infty$. Since $F_{n_k}(y_{n_k})$ converges in $\mathbb{R}$ we obtain the bound in (23). □

2.6. **The extended Wasserstein-Orlicz space** $(\mathscr{P}(X),W_\psi)$. Given $\mu,\nu\in\mathscr{P}(X)$ we define the set of admissible plans $\Gamma(\mu,\nu)$ as follows:

$$\Gamma(\mu,\nu):=\{\gamma\in\mathscr{P}(X\times X):\pi^1_\#\gamma=\mu,\ \pi^2_\#\gamma=\nu\},$$

where $\pi^i:X\times X\to X$, for $i=1,2$, are the projections on the first and the second component, defined by $\pi^1(x,y)=x$ and $\pi^2(x,y)=y$.

Given ψ satisfying (6), the ψ-Wasserstein-Orlicz extended distance between $\mu,\nu\in\mathscr{P}(X)$ is defined by

$$\begin{aligned}W_\psi(\mu,\nu)&:=\inf_{\gamma\in\Gamma(\mu,\nu)}\inf\left\{\lambda>0:\int_{X\times X}\psi\Big(\frac{\mathsf{d}(x,y)}{\lambda}\Big)\,d\gamma(x,y)\leq 1\right\}\\ &=\inf_{\gamma\in\Gamma(\mu,\nu)}\|\mathsf{d}(\cdot,\cdot)\|_{L^\psi_\gamma(X\times X)}.\end{aligned}\tag{27}$$

It is easy to check that

$$W_\psi(\mu,\nu)=\inf\left\{\lambda>0:\inf_{\gamma\in\Gamma(\mu,\nu)}\int_{X\times X}\psi\Big(\frac{\mathsf{d}(x,y)}{\lambda}\Big)\,d\gamma(x,y)\leq 1\right\}$$

which is the definition given in [Stu11] (see also [Kuw13]).

When the set of $\gamma\in\Gamma(\mu,\nu)$ such that $\|\mathsf{d}(\cdot,\cdot)\|_{L^\psi_\gamma(X\times X)}<+\infty$ is empty, then $W_\psi(\mu,\nu)=+\infty$. Otherwise it is not difficult to show that a minimizer $\gamma\in\Gamma(\mu,\nu)$ in (27) exists. We denote by $\Gamma^\psi_o(\mu,\nu)$ the set of minimizers in (27). We observe that

$$\gamma\in\Gamma^\psi_o(\mu,\nu)\qquad\Longleftrightarrow\qquad\int_{X\times X}\psi\Big(\frac{\mathsf{d}(x,y)}{W_\psi(\mu,\nu)}\Big)\,d\gamma(x,y)\leq 1.\tag{28}$$

Since ψ satisfies (6) it is well defined $\psi^{-1}(s)$ for every $s>0$, with the convention that in the case that $\psi(r)=+\infty$ for $r>r_0$ and $\psi(r_0)<+\infty$ we define $\psi^{-1}(s)=r_0$ for every $s>\psi(r_0)$.

Moreover if $\gamma\in\Gamma^\psi_o(\mu,\nu)$ then

$$\int_{X\times X}\mathsf{d}(x,y)\,d\gamma(x,y)\leq\psi^{-1}(1)W_\psi(\mu,\nu).\tag{29}$$

Indeed, for $\mu\neq\nu$ (the other case is trivial) using Jensen's inequality and (28)

$$\psi\Big(\int_{X\times X}\frac{\mathsf{d}(x,y)}{W_\psi(\mu,\nu)}\,d\gamma(x,y)\Big)\leq\int_{X\times X}\psi\Big(\frac{\mathsf{d}(x,y)}{W_\psi(\mu,\nu)}\Big)\,d\gamma(x,y)\leq 1$$

and (29) follows.

Being (X, d) complete, $(\mathscr{P}(X), W_\psi)$, is complete too (the proof of [AGS05, Proposition 7.1.5] works also in the case of the extended distance d and the Orlicz-Wasserstein distance).

We observe that (X, d) is embedded in $(\mathscr{P}(X), W_\psi)$ via the map $x \mapsto \delta_x$ and it holds

$$W_\psi(\delta_x, \delta_y) = \frac{1}{\psi^{-1}(1)} d(x, y). \tag{30}$$

Thanks to the compatibility condition (iii) in the definition of extended Polish space we also have the following fundamental property:

$$W_\psi(\mu_n, \mu) \to 0 \quad \Longrightarrow \quad \mu_n \to \mu \text{ narrowly in } \mathscr{P}(X). \tag{31}$$

The space $(\mathscr{P}(X), W_\psi)$ is an extended Polish space, when in $\mathscr{P}(X)$ we consider the narrow topology.

3. Main theorem

In this section we state and prove our main result: a characterization of absolutely continuous curves with finite L^ψ-energy in the extended ψ-Wasserstein-Orlicz space $(\mathscr{P}(X), W_\psi)$.

Before to state the result, we define, for every $t \in I$, the *evaluation map* $e_t : C(I; X) \to X$ in this way

$$e_t(u) = u(t) \tag{32}$$

and we observe that e_t is continuous.

Theorem 3.1. *Let ψ be satisfying* (6), (10) *and* (11). *Let (X, τ, d) be an extended Polish space and $I := [0, T]$, $T > 0$. If $\mu \in AC^\psi(I; (\mathscr{P}(X), W_\psi))$, then there exists $\eta \in \mathscr{P}(C(I; X))$ such that*

(i) *η is concentrated on $AC^\psi(I; (X, \mathsf{d}))$,*
(ii) $(e_t)_\# \eta = \mu_t \qquad \forall t \in I$,
(iii) *for a.e. $t \in I$, the metric derivative $|u'|(t)$ exists for $\eta-$a.e. $u \in C(I; X)$ and it holds the equality*
$$|\mu'|(t) = \| |u'|(t) \|_{L^\psi_\eta(C(I;X))} \qquad \text{for a.e. } t \in I.$$

Proof. We preliminary assume that

$$|\mu'| = 1 \qquad \text{for a.e. } t \in I, \tag{33}$$

and we will remove this assumption in Step 6 of this proof. We also assume for simplicity that $I = [0, 1]$.

For any integer $N \geq 1$, we divide the unitary interval I in 2^N equal parts, and we denote by t^i the points

$$t^i := \frac{i}{2^N} \qquad i = 0, 1, \ldots, 2^N.$$

We also denote by $\boldsymbol{X}_N$ the product space

$$\boldsymbol{X}_N := X_0 \times X_1 \times \ldots \times X_{2^N}$$

where X_i, with $i = 0, 1, \ldots, 2^N$, are $2^N + 1$ copies of the same space X.

Choosing optimal plans

$$\gamma_N^i \in \Gamma_o^\psi(\mu_{t^i}, \mu_{t^{i+1}}) \qquad i = 0, 1, \ldots, 2^N - 1,$$

there exists (see for instance [AGS05, Lemma 5.3.2 and Remark 5.3.3]) a measure $\gamma_N \in \mathscr{P}(\boldsymbol{X}_N)$ such that

$$\pi^i_\# \gamma_N = \mu_{t^i} \qquad \text{and} \qquad \pi^{i,i+1}_\# \gamma_N = \gamma^i_N,$$

where we denoted by $\pi^i : \boldsymbol{X}_N \to X_i$ the projection on the i-th component and by $\pi^{i,j} : \boldsymbol{X}_N \to X_i \times X_j$ the projection on the (i,j)-th component.
We define $\sigma : \boldsymbol{X}_N \to \mathscr{M}(I;X)$, and we use the notation $\boldsymbol{x} = (x_0, \dots, x_{2^N}) \mapsto \sigma_{\boldsymbol{x}}$, by

$$\sigma_{\boldsymbol{x}}(t) := x_i \qquad \text{if} \quad t \in [t^i, t^{i+1}), \qquad i = 0, 1, \dots, 2^N - 1.$$

Finally, we define the sequence of probability measures

$$\eta_N := \sigma_\# \gamma_N \in \mathscr{P}(\mathcal{M}(I;X)).$$

Step 1. (Tightness of $\{\eta_N\}_{N\in\mathbb{N}}$ in $\mathscr{P}(\mathcal{M}(I;X))$) In order to prove the tightness of $\{\eta_N\}_{N\in\mathbb{N}}$ in $\mathscr{P}(\mathcal{M}(I;X))$ (we recall that $\mathcal{M}(I;X)$ is a Polish space with the metric δ_1) it is sufficient to show the existence of a function $\Phi : \mathcal{M}(I;X) \to [0,+\infty]$ whose sublevels $\lambda_c(\Phi) := \{u \in \mathcal{M}(I;X) : \Phi(u) \le c\}$ are compact in $\mathcal{M}(I;X)$ for any $c \in \mathbb{R}_+$, and

$$\sup_{N\in\mathbb{N}} \int_{\mathcal{M}(I;X)} \Phi(u)\, d\eta_N(u) < +\infty. \tag{34}$$

First of all we observe that $\mathscr{A} := \{\mu_t : t \in I\}$ is compact in $(\mathscr{P}(X), W_\psi)$ (because it is a continuous image of a compact) and consequently in $\mathscr{P}(X)$. Since, by Prokhorov's Theorem, $\mathscr{A}$ is tight in $\mathscr{P}(X)$ there exists a function $\Psi : X \to [0,+\infty]$ whose sublevels $\lambda_c(\Psi) := \{x \in X : \Psi(x) \le c\}$ are compact in X for any $c \in \mathbb{R}_+$, such that

$$\sup_{t\in I} \int_X \Psi(x)\, d\mu_t(x) < +\infty. \tag{35}$$

We define $\Phi : \mathcal{M}(I;X) \to [0,+\infty]$ by

$$\Phi(u) := \int_0^1 \Psi(u(t))\, dt + \sup_{h\in(0,1)} \int_0^{1-h} \frac{\mathsf{d}(u(t+h), u(t))}{h}\, dt.$$

The compactness of the sublevels $\lambda_c(\Phi)$ in $\mathcal{M}(I;X)$ follows by Theorem 2.2 with the choice $g(x,y) = \mathsf{d}(x,y)$. In order to prove (34) we begin to show that

$$\sup_{N\in\mathbb{N}} \int_{\mathcal{M}(I;X)} \int_0^1 \Psi(u(t)))\, dt\, d\eta_N(u) < +\infty. \tag{36}$$

By the definition of η_N we have

$$\begin{aligned}\int_{\mathcal{M}(I;X)}\int_0^1 \Psi(u(t))\,dt\,d\eta_N(u) &= \int_{\boldsymbol{X}_N}\int_0^1 \Psi(\sigma_{\boldsymbol{x}}(t))\,dt\,d\gamma_N(\boldsymbol{x})\\ &= \int_{\boldsymbol{X}_N}\sum_{i=0}^{2^N-1}\int_{t^i}^{t^{i+1}} \Psi(x_i)\,dt\,d\gamma_N(\boldsymbol{x})\\ &= \int_{\boldsymbol{X}_N}\frac{1}{2^N}\sum_{i=0}^{2^N-1}\Psi(x_i)\,d\gamma_N(\boldsymbol{x})\\ &= \frac{1}{2^N}\sum_{i=0}^{2^N-1}\int_X \Psi(x)\,d\mu_{t^i}(x)\\ &\le \frac{1}{2^N}\sum_{i=0}^{2^N-1}\sup_{t\in I}\int_X \Psi(x)\,d\mu_t(x) = \sup_{t\in I}\int_X \Psi(x)\,d\mu_t(x)\end{aligned}$$

and (36) follows by (35). The second bound that we have to show is

$$\sup_{N\in\mathbb{N}}\int_{\mathcal{M}(I;X)}\sup_{h\in(0,1)}\int_0^{1-h}\frac{\mathsf{d}(u(t+h),u(t))}{h}\,dt\,d\eta_N(u) < +\infty. \tag{37}$$

First of all we prove that for $\boldsymbol{x}\in\boldsymbol{X}_N$ we have

$$\sup_{h\in(0,1)}\int_0^{1-h}\frac{\mathsf{d}(\sigma_{\boldsymbol{x}}(t+h),\sigma_{\boldsymbol{x}}(t))}{h}\,dt \le 2\sum_{i=0}^{2^N-1}\mathsf{d}(x_i,x_{i+1}). \tag{38}$$

We fix $h\in(0,1)$. When $h<2^{-N}$ we have that $\sigma_{\boldsymbol{x}}(t+h)=\sigma_{\boldsymbol{x}}(t)$ for every $t\in[t^i,t^{i+1}-h]$ and $i=0,\dots,2^N-1$. Then

$$\int_0^{1-h}\mathsf{d}(\sigma_{\boldsymbol{x}}(t+h),\sigma_{\boldsymbol{x}}(t))\,dt = \sum_{i=0}^{2^N-1}\int_{t^i}^{t^{i+1}}\mathsf{d}(\sigma_{\boldsymbol{x}}(t+h),\sigma_{\boldsymbol{x}}(t))\,dt = h\sum_{i=0}^{2^N-2}\mathsf{d}(x_i,x_{i+1}). \tag{39}$$

Now we assume that $h\ge 2^{-N}$ and we take the integer $k(h)=[h2^N]$, where $[a]:=\max\{n\in\mathbb{Z}: n\le a\}$ is the integer part of the real number a. Since the triangular inequality yields

$$\mathsf{d}(\sigma_{\boldsymbol{x}}(t+h),\sigma_{\boldsymbol{x}}(t)) \le \sum_{i=0}^{k(h)}\mathsf{d}(\sigma_{\boldsymbol{x}}(t+t^{i+1}),\sigma_{\boldsymbol{x}}(t+t^i)),$$

we have that

$$
\begin{aligned}
\int_0^{1-h} \mathsf{d}(\sigma_{\boldsymbol{x}}(t+h), \sigma_{\boldsymbol{x}}(t))\, dt &\leq \int_0^{1-t^{k(h)}} \mathsf{d}(\sigma_{\boldsymbol{x}}(t+h), \sigma_{\boldsymbol{x}}(t))\, dt \\
&\leq \int_0^{1-t^{k(h)}} \sum_{i=0}^{k(h)} \mathsf{d}(\sigma_{\boldsymbol{x}}(t+t^{i+1}), \sigma_{\boldsymbol{x}}(t+t^i))\, dt \\
&= \sum_{i=0}^{k(h)} \frac{1}{2^N} \sum_{j=0}^{2^N-k(h)-1} \mathsf{d}(x_{i+j+1}, x_{i+j}).
\end{aligned} \tag{40}
$$

Observing that in (40) the term $\mathsf{d}(x_{k+1}, x_k)$, for every $k = 0, 1, \ldots, 2^N - 1$, is counted at most $k(h)+1$ times, we obtain that

$$
\int_0^{1-h} \mathsf{d}(\sigma_{\boldsymbol{x}}(t+h), \sigma_{\boldsymbol{x}}(t))\, dt \leq \frac{k(h)+1}{2^N h} h \sum_{j=0}^{2^N-1} \mathsf{d}(x_{j+1}, x_j) \leq 2h \sum_{j=0}^{2^N-1} \mathsf{d}(x_{j+1}, x_j), \tag{41}
$$

because

$$
\frac{k(h)+1}{h2^N} \leq \frac{k(h)+1}{k(h)} \leq 2.
$$

The inequality (38) follows from (41) and (39). Finally, by (38), (29) taking into account the optimality of the plans $\pi^{i,i+1}_{\#}\gamma_N$, and (33) we have

$$
\begin{aligned}
\int_{\mathcal{M}(I;X)} \sup_{h\in(0,1)} \int_0^{1-h} \frac{\mathsf{d}(u(t+h), u(t))}{h}\, dt\, d\eta_N(u) &\leq 2 \int_{\boldsymbol{X}_N} \sum_{i=0}^{2^N-1} \mathsf{d}(x_i, x_{i+1})\, d\gamma_N(\boldsymbol{x}) \\
&\leq 2\psi^{-1}(1) \sum_{i=0}^{2^N-1} W_\psi(\mu_{t^i}, \mu_{t^{i+1}}) \\
&\leq 2\psi^{-1}(1) \sum_{i=0}^{2^N-1} \frac{1}{2^N} = 2\psi^{-1}(1)
\end{aligned} \tag{42}
$$

and (37) follows.

Then, by Prokhorov's Theorem, there exist $\eta \in \mathscr{P}(\mathcal{M}(I;X))$ and a subsequence N_n such that $\eta_{N_n} \to \eta$ narrowly in $\mathscr{P}(\mathcal{M}(I;X))$ as $n \to +\infty$.

Step 2. (η is concentrated on BV right continuous curves) We apply Lemma 2.3 in order to show that η-a.e. $u \in \operatorname{supp}\eta$ has a right continuous BV representative.

Given a curve $u : [a,b] \to X$, we denote by $\mathsf{pV}(u,[a,b]) = \sup\{\sum_{i=1}^n \mathsf{d}(u(t_i), u(t_{i+1})) : a = t_1 < t_2 < \ldots < t_n < t_{n+1} = b\}$ its pointwise variation and by $\mathsf{eV}(u,[a,b]) = \inf\{\mathsf{pV}(w,[a,b]) : w(t) = u(t) \text{ for a.e. } t \in (a,b)\}$ its essential variation.

We define $F_N : \mathcal{M}(I;X) \to [0,+\infty)$ by

$$
F_N(u) = \begin{cases} \mathsf{eV}(u, I) & \text{if } u \in \operatorname{supp}\eta_N, \\ 0 & \text{if } u \notin \operatorname{supp}\eta_N. \end{cases} \tag{43}
$$

If u is a.e. equal to $\sigma_{\boldsymbol{x}}$ then $\mathsf{eV}(u,I) = \mathsf{pV}(\sigma_{\boldsymbol{x}}, I)$. Taking into account this equality, the proof of bound (36) shows that

$$\sup_{N\in\mathbb{N}} \int_{\mathcal{M}(I;X)} F_N(u)\, d\eta_N(u) < +\infty. \tag{44}$$

Since $F_N \geq 0$ by definition, we apply Lemma 2.3 with the choice $Y = \mathcal{M}(I;X)$ and $\mu_n = \eta_{N_n}$. We still denote by η_{N_n} the subsequence of η_{N_n} given by Lemma 2.3. Let $u \in \mathrm{supp}(\eta)$ be such that (23) holds and we denote by $u_{N_n} \in \mathrm{supp}(\eta_{N_n})$ such that $u_{N_n} \to u$ in $\mathcal{M}(I;X)$ and C a constant independent of n such that

$$F_{N_n}(u_{N_n}) \leq C. \tag{45}$$

Moreover, up to extract a further subsequence, we can also assume that $u_{N_n}(t) \to u(t)$ with respect to the distance δ for a.e. $t \in I$. Since $u_{N_n} \in \mathrm{supp}(\eta_{N_n})$ we can choose the piecewise constant right continuous representative of u_{N_n}, still denoted by u_{N_n}. From (45) we obtain that

$$\mathsf{eV}(u_{N_n}) = \mathsf{pV}(u_{N_n}) \leq C. \tag{46}$$

Defining the increasing functions $v_n : I \to \mathbb{R}$ by $v_n(t) = \mathsf{pV}(u_{N_n}, [0,t])$, from the Helly theorem, up to extract a further subsequence still denoted by v_n, there exists an increasing function $v : I \to \mathbb{R}$ such that $v_n(t)$ converges to $v(t)$ for every $t \in I$ (we observe that for (46) $v \leq C$). Since the set of discontinuity points of v is at most countable we can redefine a right continuous function $\bar{v}$ by $\bar{v}(t) = \lim_{s\to t^+} v(t)$. Since

$$\mathsf{d}(u_{N_n}(t), u_{N_n}(s)) \leq v_n(s) - v_n(t) \qquad \forall\, t, s \in I, \quad t \leq s, \tag{47}$$

from the property (1) it follows that

$$\mathsf{d}(u(t), u(s)) \leq \bar{v}(s) - \bar{v}(t) \qquad \text{for a.e. } t, s \in I, \quad t \leq s. \tag{48}$$

Since (X, d) is complete, by (48) we can choose the representative of u, $\bar{u} : I \to X$ defined by $\bar{u}(t) = \lim_{s\to t^+} u(t)$, which is right continuous by (48).

We have just proved that η-a.e. $u \in \mathrm{supp}\,\eta$ is equivalent (with respect to the a.e. equality) to a d-right continuous function with pointwise d-bounded variation, continuous at every points except at most a countable set.

Step 3. (Proof of (i))

Since we want to apply Lemma 2.1, we prove that

$$\sup_{h\in(0,1)} \left\| \frac{\mathsf{d}(u(\cdot + h), u(\cdot))}{h} \right\|_{L^\psi(0,1-h)} < +\infty, \qquad \text{for } \eta-\text{a.e. } u \in \mathcal{M}(I;X). \tag{49}$$

Let us define the sequence of lower semi continuous functions $f_N : \mathcal{M}(I;X) \to [0, +\infty]$ by

$$f_N(u) := \sup_{1/2^N \leq h < 1} \int_0^{1-h} \psi\Big(\frac{\mathsf{d}(u(t+h), u(t))}{2h}\Big)\, dt,$$

that satisfies the monotonicity property

$$f_N(u) \leq f_{N+1}(u) \qquad \forall\, u \in \mathcal{M}(I;X). \tag{50}$$

For $h \in [2^{-N}, 1)$, and $u \in \operatorname{supp}(\eta_N)$, by the monotonicity of ψ, the discrete Jensen's inequality and taking into account that $(k(h)+1)/(2h) \le 2^N$, we have that

$$
\begin{aligned}
&\int_0^{1-h} \psi\Big(\frac{\mathsf{d}(u(t+h), u(t))}{2h}\Big)\, dt \\
&= \int_0^{1-h} \psi\Big(\frac{\mathsf{d}(x_{k(t+h)}, x_{k(t)})}{2h}\Big)\, dt \\
&= \int_0^{1-t^{k(h)}} \psi\Big(\frac{\mathsf{d}(x_{k(t+h)}, x_{k(t)})}{2h}\Big)\, dt \\
&\le \int_0^{1-t^{k(h)}} \psi\Big(\frac{1}{k(h)+1} \sum_{i=0}^{k(h)} \frac{k(h)+1}{2h} \mathsf{d}(x_{k(t)+i+1}, x_{k(t)+i})\Big)\, dt \\
&\le \int_0^{1-t^{k(h)}} \frac{1}{k(h)+1} \sum_{i=0}^{k(h)} \psi\Big(\frac{k(h)+1}{2h} \mathsf{d}(x_{k(t)+i+1}, x_{k(t)+i})\Big)\, dt \\
&= \sum_{j=0}^{2^N-k(h)-1} 2^{-N} \frac{1}{k(h)+1} \sum_{i=0}^{k(h)} \psi\Big(2^N \mathsf{d}(x_{j+i+1}, x_{j+i})\Big) \\
&\le \sum_{j=0}^{2^N-1} 2^{-N} \psi\Big(2^N \mathsf{d}(x_{j+1}, x_j)\Big).
\end{aligned}
$$

It follows that

$$
f_N(u) \le \sum_{j=0}^{2^N-1} 2^{-N} \psi\Big(2^N \mathsf{d}(x_{j+1}, x_j)\Big)
$$

for every $u \in \operatorname{supp}(\eta_N)$. Integrating the last inequality, taking into account that $W_\psi(\mu_{t^j}, \mu_{t^{j+1}}) \le 2^{-N}$ and

$$
\int_{\boldsymbol{X}_N} \psi\Big(\frac{\mathsf{d}(x_{j+1}, x_j)}{W_\psi(\mu_{t^{j+1}}, \mu_{t^j})}\Big)\, d\gamma_N(\boldsymbol{x}) \le 1,
$$

we obtain that

$$
\begin{aligned}
\int_{\mathcal{M}(I;X)} f_N(u)\, d\eta_N(u) &\le \sum_{j=0}^{2^N-1} 2^{-N} \int_{\boldsymbol{X}_N} \psi\Big(2^N \mathsf{d}(x_{j+1}, x_j)\Big)\, d\gamma_N(\boldsymbol{x}) \\
&\le \sum_{j=0}^{2^N-1} 2^{-N} \int_{\boldsymbol{X}_N} \psi\Big(\frac{\mathsf{d}(x_{j+1}, x_j)}{W_\psi(\mu_{t^{j+1}}, \mu_{t^j})}\Big)\, d\gamma_N(\boldsymbol{x}) \le 1.
\end{aligned}
$$

The lower semi continuity of f_N, the monotonicity (50) of f_N and the last inequality yield

$$
\int_{\mathcal{M}(I;X)} f_N(u)\, d\eta(u) \le 1 \qquad \forall\, N \in \mathbb{N},
$$

and consequently, by monotone convergence Theorem, we have that

$$
\int_{\mathcal{M}(I;X)} \sup_{N\in\mathbb{N}} f_N(u)\, d\eta(u) \le 1,
$$

and

$$\sup_{N\in\mathbb{N}} f_N(u) < +\infty \quad \text{ for } \eta - \text{a.e. } u \in \mathcal{M}(I;X). \tag{51}$$

Since

$$\sup_{N\in\mathbb{N}} f_N(u) = \sup_{0<h<1} \int_0^{1-h} \psi\Big(\frac{\mathsf{d}(u(t+h),u(t))}{2h}\Big)\,dt,$$

and $\int_0^{1-h}\psi\Big(\frac{\mathsf{d}(u(t+h),u(t))}{2h}\Big)\,dt \le C$ implies $\Big\|\frac{\mathsf{d}(u(\cdot+h),u(\cdot))}{h}\Big\|_{L^\psi(0,1-h)} \le \max\{C,1\}$ we obtain (49).

Finally, taking into account Step 2, we can associate to η-a.e. $u \in \operatorname{supp}\eta$ a right continuous representative $\bar u$, with at most a countable points of discontinuity satisfying (15). By Lemma 2.1 this representative belongs to $AC^\psi(I;(X,\mathsf{d}))$.

Defining the canonical immersion $T : C(I;X) \to \mathcal{M}(I;X)$ and observing that it is continuous, we define the new Borel probability measure $\tilde\eta \in \mathscr{P}(C(I;X))$ by $\tilde\eta(B) = \eta(T(B))$. For the previous steps $\tilde\eta$ is concentrated on $AC^\psi(I;(X,\mathsf{d}))$.

Step 4. (Proof of (ii)) In order to show (ii) we prove that for every $t \in I$,

$$\int_{C(I;X)} \varphi(u(t))\,d\tilde\eta(u) = \int_X \varphi(x)\,d\mu_t(x) \qquad \forall \varphi \in C_b(X). \tag{52}$$

Let $\varphi \in C_b(X)$. Since $g : I \to \mathbb{R}$ defined by

$$g(t) := \int_X \varphi(x)\,d\mu_t(x)$$

is uniformly continuous in I, we have that the sequence of piecewise constant functions $g_N : I \to \mathbb{R}$ defined by

$$g_N(t) := g(t^i) = \int_X \varphi(x)\,d\mu_{t^i}(x) \qquad \text{if } t \in [t^i, t^{i+1}),$$

converges uniformly to g in I when $N \to +\infty$. Then, for every test function $\zeta \in C_b(I)$, we have that

$$\lim_{N\to+\infty} \int_0^1 \zeta(t) g_N(t)\,dt = \int_0^1 \zeta(t) g(t)\,dt. \tag{53}$$

On the other hand

$$\begin{aligned}\int_0^1 \zeta(t) g_N(t)\,dt &= \int_0^1 \zeta(t) \int_{\mathcal{M}(I;X)} \varphi(u(t))\,d\eta_N(u)\,dt \\ &= \int_{\mathcal{M}(I;X)} \int_0^1 \zeta(t)\varphi(u(t))\,dt\,d\eta_N(u).\end{aligned}$$

Since the map

$$u \mapsto \int_0^1 \zeta(t)\varphi(u(t))\,dt$$

is continuous and bounded from $\mathcal{M}(I;X)$ to $\mathbb{R}$, then by the narrow convergence of η_{N_n} we have

$$\lim_{n\to+\infty} \int_{\mathcal{M}(I;X)} \int_0^1 \zeta(t)\varphi(u(t))\,dt\,d\eta_{N_n}(u) = \int_{\mathcal{M}(I;X)} \int_0^1 \zeta(t)\varphi(u(t))\,dt\,d\eta(u).$$

By Fubini's Theorem and the definition of $\tilde\eta$

$$\begin{aligned}\int_{\mathcal{M}(I;X)}\int_0^1 \zeta(t)\varphi(u(t))\,dt\,d\eta(u) &= \int_{C(I;X)}\int_0^1 \zeta(t)\varphi(u(t))\,dt\,d\tilde\eta(u)\\ &= \int_0^1 \zeta(t)\int_{C(I;X)}\varphi(u(t))\,d\tilde\eta(u)\,dt.\end{aligned}$$

By the uniqueness of the limit then

$$\int_0^1 \zeta(t)\int_{C(I;X)}\varphi(u(t))\,d\tilde\eta(u)\,dt = \int_0^1 \zeta(t)\int_X \varphi(x)\,d\mu_t(x)\,dt \qquad \forall \zeta\in C_b(I),$$

from which

$$\int_{C(I;X)}\varphi(u(t))\,d\tilde\eta(u) = \int_X \varphi(x)\,d\mu_t(x) \qquad \text{for a.e. } t\in I. \tag{54}$$

Since the applications $t\mapsto \int_X \varphi(x)\,d\mu_t(x)$ and $t\mapsto \int_{C(I;X)}\varphi(u(t))\,d\tilde\eta(u)$ are continuous, (54) is true for every $t\in I$ and (52) is proved.

Step 5. (Proof of (iii))

First of all we check that for a.e. $t\in I$, $|u'|(t)$ exists for $\tilde\eta-$a.e. $u\in C(I;X)$. We set $\Lambda := \{(t,u)\in I\times C(I;X) : |u'|(t) \text{ does not exist}\}$. Λ is a Borel subset of $I\times C(I;X)$ since the maps $G_h : I\times C(I;X)\to\mathbb{R}$ defined by $G_h(t,u) := \frac{d(u(t+h),u(t))}{|h|}$ are lower semi continuous for every $h\neq 0$, and $\Lambda = \{(t,u)\in I\times C(I;X) : \liminf_{h\to 0} G_h(t,u) < \limsup_{h\to 0} G_h(t,u)\}$. Since $\tilde\eta$ is concentrated on $AC(I;(X,\mathsf{d}))$ curves, we have that for $\tilde\eta$-a.e. $u\in C(I;X)$, $\mathscr{L}^1(\{t\in I : (t,u)\in\Lambda\}) = 0$ and then Fubini's Theorem implies that for a.e. $t\in I$, $\tilde\eta(\{u\in C(I;X) : (t,u)\in\Lambda\}) = 0$.

Let $a,b\in I$ such that $a<b$ and let $h>0$ such that $b+h\in I$. Recalling that $k(h) = [2^N h]$, for every $N\in\mathbb{N}$ such that $2^{-N}\le h$, by the monotonicity of ψ and the discrete Jensen's inequality we have

$$\begin{aligned}&\int_{\mathcal{M}(I;X)}\int_a^b \psi\Big(\frac{k(h)}{k(h)+1}\frac{\mathsf{d}(u(t+h),u(t))}{h}\Big)\,dt\,d\eta_N(u)\\ &\le \int_{\boldsymbol{X}_N}\int_a^b \psi\Big(\frac{k(h)}{k(h)+1}\frac{\mathsf{d}(x_{k(t+h)},x_{k(t)})}{h}\Big)\,dt\,d\gamma_N(\boldsymbol{x})\\ &\le \int_{\boldsymbol{X}_N}\int_a^b \psi\Big(\frac{1}{k(h)+1}\sum_{i=0}^{k(h)}\frac{k(h)}{h}\mathsf{d}(x_{k(t)+i+1},x_{k(t)+i})\Big)\,dt\,d\gamma_N(\boldsymbol{x})\\ &\le \int_{\boldsymbol{X}_N}\int_a^b \sum_{i=0}^{k(h)}\frac{1}{k(h)+1}\psi\Big(\frac{k(h)}{h}\mathsf{d}(x_{k(t)+i+1},x_{k(t)+i})\Big)\,dt\,d\gamma_N(\boldsymbol{x}).\end{aligned}$$

Since $k(h)/h \le 2^N$ and, by (33), $W_\psi(\mu_{t^k}, \mu_{t^{k+1}}) \le 2^{-N}$, we have that

$$\begin{aligned}
&\int_{\boldsymbol{X}_N} \int_a^b \sum_{i=0}^{k(h)} \frac{1}{k(h)+1} \psi\Big(\frac{k(h)}{h} \mathsf{d}(x_{k(t)+i+1}, x_{k(t)+i})\Big)\, dt\, d\gamma_N(\boldsymbol{x}) \\
&\le \int_{\boldsymbol{X}_N} \int_a^b \sum_{i=0}^{k(h)} \frac{1}{k(h)+1} \psi\Big(2^N \mathsf{d}(x_{k(t)+i+1}, x_{k(t)+i})\Big)\, dt\, d\gamma_N(\boldsymbol{x}) \\
&\le \int_a^b \sum_{i=0}^{k(h)} \frac{1}{k(h)+1} \int_{\boldsymbol{X}_N} \psi\Big(\frac{\mathsf{d}(x_{k(t)+i+1}, x_{k(t)+i})}{W_\psi(\mu_{t^{k(t)+i+1}}, \mu_{t^{k(t)+i}})}\Big)\, d\gamma_N(\boldsymbol{x})\, dt \le b-a,
\end{aligned}$$

where we used the inequality

$$\int_{\boldsymbol{X}_N} \psi\Big(\frac{\mathsf{d}(x_{k(t)+i+1}, x_{k(t)+i})}{W_\psi(\mu_{t^{k(t)+i+1}}, \mu_{t^{k(t)+i}})}\Big)\, d\gamma_N(\boldsymbol{x}) \le 1.$$

It follows that

$$\int_{\mathcal{M}(I;X)} \frac{1}{b-a} \int_a^b \psi\Big(\frac{k(h)}{k(h)+1} \frac{\mathsf{d}(u(t+h), u(t))}{h}\Big)\, dt\, d\eta_N(u) \le 1$$

and then, passing to the limit along the sequence η_{N_n},

$$\int_{C(I;X)} \frac{1}{b-a} \int_a^b \psi\Big(\frac{\mathsf{d}(u(t+h), u(t))}{h}\Big)\, dt\, d\tilde{\eta}(u) \le 1.$$

Taking into account (i), Fubini's Theorem and Lebesgue differentiation Theorem we obtain

$$\int_{C(I;X)} \psi\Big(|u'|(t)\Big)\, d\tilde{\eta}(u) \le 1 \qquad \text{for a.e. } t \in I \tag{55}$$

and this shows that

$$\||u'|(t)\|_{L^\psi_{\tilde{\eta}}(C(I;X))} \le 1 = |\mu'|(t) \qquad \text{for a.e. } t \in I.$$

Step 6. (Conclusion) Finally we have to remove the assumption (33). Let $\mu \in AC^\psi(I; (\mathscr{P}(X), W_\psi))$ with length $L := \int_0^T |\mu'|(t)\, dt$.

If $L = 0$, then $\mu_t = \mu_0$ for every $t \in I$ and μ is represented by $\eta := \sigma_\# \mu_0$, where $\sigma : X \to C(I;X)$ denotes the function $\sigma(x) = c_x$, $c_x(t) := x$ for every $t \in I$.

When $L > 0$ we can reparametrize μ by its arc-length (see Lemma 1.1.4(b) of [AGS05] for the details). We define the increasing function $\boldsymbol{s} : I \to [0, L]$ by $\boldsymbol{s}(t) := \int_0^t |\mu'|(r)\, dr$ observing that $\boldsymbol{s}$ is absolutely continuous with pointwise derivative

$$\boldsymbol{s}'(t) = |\mu'|(t) \qquad \text{for a.e. } t \in I. \tag{56}$$

Defining $\boldsymbol{s}^{-1} : I \to [0, L]$ by $\boldsymbol{s}^{-1}(s) = \min\{t \in I : \boldsymbol{s}(t) = s\}$ it is easy to check that the new curve $\hat{\mu} : [0, L] \to \mathscr{P}(X)$ defined by $\hat{\mu}_s = \mu_{\boldsymbol{s}^{-1}(s)}$ satisfies $|\hat{\mu}'|(s) = 1$ for a.e. $s \in [0, L]$ and $\mu_t = \hat{\mu}_{\boldsymbol{s}(t)}$. By the previous steps, we represent $\hat{\mu}$ by a measure $\hat{\eta}$ concentrated on $AC^\psi([0, L]; (X, \mathsf{d}))$. Denoting by $F : C([0, L]; X) \to C(I; X)$ the map defined by $F(\hat{u}) = \hat{u} \circ \boldsymbol{s}$, we represent μ by $\eta := F_\# \hat{\eta}$. Clearly $(e_t)_\# \eta = (e_t \circ F)_\# \hat{\eta} = \hat{\mu}_{\boldsymbol{s}(t)} = \mu_t$. Moreover, η is concentrated on curves u of the form $u(t) = \hat{u}(\boldsymbol{s}(t))$ with $\hat{u} \in AC^\psi([0, L]; (X, \mathsf{d}))$. Since $\boldsymbol{s}$

is monotone and $AC(I;\mathbb{R})$ and $\hat{u}$ is $AC([0,L];(X,\mathsf{d}))$ then $\hat{u}\circ \boldsymbol{s}$ is $AC(I;(X,\mathsf{d}))$, and the metric derivative satisfies

$$|u'|(t) \le |\hat{u}'|(\boldsymbol{s}(t))\boldsymbol{s}'(t) \qquad \text{for a.e. } t \in I. \tag{57}$$

Let $t\in I$ such that $\boldsymbol{s}'(t)$ and $|\mu'|(t)$ exist and $\boldsymbol{s}'(t) = |\mu'|(t) > 0$. Taking into account (55) and Jensen's inequality we have for $h>0$

$$\begin{aligned}\int_{C(I;X)} \psi\Big(\frac{\mathsf{d}(u(t+h),u(t))}{\boldsymbol{s}(t+h)-\boldsymbol{s}(t)}\Big)\,d\eta(u) &= \int_{C([0,L];X)} \psi\Big(\frac{\mathsf{d}(\hat{u}(\boldsymbol{s}(t+h)),u(\boldsymbol{s}(t)))}{\boldsymbol{s}(t+h)-\boldsymbol{s}(t)}\Big)\,d\hat{\eta}(\hat{u})\\ &\le \int_{C([0,L];X)} \psi\Big(\frac{1}{\boldsymbol{s}(t+h)-\boldsymbol{s}(t)}\int_{\boldsymbol{s}(t)}^{\boldsymbol{s}(t+h)} |\hat{u}'|(r)\,dr\Big)\,d\hat{\eta}(\hat{u})\\ &\le \frac{1}{\boldsymbol{s}(t+h)-\boldsymbol{s}(t)}\int_{\boldsymbol{s}(t)}^{\boldsymbol{s}(t+h)}\int_{C([0,L];X)} \psi\Big(|\hat{u}'|(r)\Big)\,d\hat{\eta}(\hat{u})\,dr \le 1.\end{aligned}$$

By Fatou's lemma, taking into account that η is concentrated on $AC(I;(X,\mathsf{d}))$ curves, we obtain the inequality

$$\int_{C(I;X)} \psi\Big(\frac{|u'|(t)}{|\mu'|(t)}\Big)\,d\eta(u) \le 1. \tag{58}$$

On the other hand, if $|\mu'|(t) = 0$ on a set $J\subset I$ of positive measure, then for η-a.e. u we have $|u'|(t) = 0$ for a.e. $t\in J$ because of the inequality (57). Taking into account this observation and (58) we obtain the inequality

$$\||u'|(t)\|_{L^\psi_\eta(C(I;X))} \le |\mu'|(t), \qquad \text{for a.e. } t\in I. \tag{59}$$

We prove that η is concentrated on $AC^\psi(I;(X,\mathsf{d}))$. For every $v\in L^{\psi^*}(I)$, $v\ge 0$, $\|v\|_{L^{\psi^*}(I)}\le 1$, from (59) we have that

$$\int_I \||u'|(t)\|_{L^\psi_\eta(C(I;X))} v(t)\,dt \le \int_I |\mu'|(t)v(t)\,dt.$$

By the inequality (9) it follows that, for every $w\in L^{\psi^*}_\eta(C(I;X))$, $w\ge 0$, $\|w\|_{L^{\psi^*}_\eta(C(I;X))}\le 1$,

$$\int_I\int_{C(I;X)} |u'|(t)w(u)\,d\eta(u)v(t)\,dt \le 4\||\mu'|\|_{L^\psi(I)}.$$

By Fubini Theorem and (9) we obtain that

$$\|\||u'|(t)\|_{L^\psi(I)}\|_{L^\psi_\eta(C(I;X))} \le 4\||\mu'|\|_{L^\psi(I)}$$

and (i) holds.

In order to show the opposite inequality of (59), we assume that $t\in I$ is such that $|u'|(t)$ exists for $\eta-$a.e. $u\in C(I;X)$ and $\lambda_t := \||u'|(t)\|_{L^\psi_\eta(C(I;X))} > 0$. We fix $\varepsilon>0$. Since $\int_{C(I;X)}\psi\Big(\frac{|u'|(t)}{\lambda_t}\Big)\,d\eta(u)\le 1$ and ψ is strictly increasing on an interval of the form (r_0,r_1) where $r_0\ge 0$, $r_1\le+\infty$ and $\psi(r)=0$ for $r<r_0$, $\psi(r)=+\infty$ for $r>r_1$, we have that

$$\int_{C(I;X)}\psi\Big(\frac{|u'|(t)}{\lambda_t+\varepsilon}\Big)\,d\eta(u) < 1.$$

For $h > 0$, let $\gamma_{t,t+h} := (e_t, e_{t+h})_\# \eta$. Taking into account that η is concentrated on $AC^\psi(I; (X, \mathsf{d}))$, we have

$$\begin{aligned}
\limsup_{h\to 0^+} \int_{X\times X} \psi\Big(\frac{\mathsf{d}(x,y)}{h(\lambda_t+\varepsilon)}\Big)\, d\gamma_{t,t+h}(x,y) &= \limsup_{h\to 0^+} \int_{C(I;X)} \psi\Big(\frac{\mathsf{d}(u(t),u(t+h))}{h(\lambda_t+\varepsilon)}\Big)\, d\eta(u) \\
&\le \int_{C(I;X)} \limsup_{h\to 0^+} \psi\Big(\frac{\mathsf{d}(u(t),u(t+h))}{h(\lambda_t+\varepsilon)}\Big)\, d\eta(u) \qquad (60)\\
&= \int_{C(I;X)} \psi\Big(\frac{|u'|(r)}{\lambda_t+\varepsilon}\Big)\, d\eta(u) < 1.
\end{aligned}$$

Consequently there exists $\bar{h}$ (depending on ε and t) such that

$$\int_{X\times X} \psi\Big(\frac{\mathsf{d}(x,y)}{h(\lambda_t+\varepsilon)}\Big)\, d\gamma_{t,t+h}(x,y) \le 1 \qquad \forall h \in (0,\bar{h}).$$

Since $\gamma_{t,t+h} \in \Gamma(\mu_t, \mu_{t+h})$, the last inequality shows that

$$W_\psi(\mu_t, \mu_{t+h}) \le h(\lambda_t + \varepsilon) \qquad \forall h \in (0, \bar{h}).$$

Finally, dividing by h and passing to the limit for $h \to 0^+$ we obtain

$$|\mu'|(t) \le \| |u'|(t) \|_{L^\psi_\eta(C(I;X))} \qquad \text{for a.e. } t \in I.$$

□

Remark 3.2. The following example shows that the assumptions on ψ are necessary for the validity of Theorem 3.1.

Since ψ is convex, if (10) and (11) are not satisfied there exist $a, b \in \mathbb{R}$ such that $0 < a \le b < +\infty$ and $at \le \psi(t) \le bt$ for every $t \ge 0$. Then it holds $aW_1(\mu,\nu) \le W_\psi(\mu,\nu) \le bW_1(\mu,\nu)$, where W_1 denotes the distance W_ϕ for $\phi(t) = t$. Given two distinct points $x_0, x_1 \in X$, consider the curve $\mu : [0,1] \to \mathscr{P}(X)$ defined by $\mu_t = (1-t)\delta_{x_0} + t\delta_{x_1}$. We observe that $\mathrm{supp}(\mu_t) = \{x_0, x_1\}$ for $t \in (0,1)$ and $\mathrm{supp}(\mu_i) = \{x_i\}$ for $i = 0, 1$. Clearly μ is Lipschitz with respect to the distance W_ψ and in particular $\mu \in AC^\psi(I; X)$. If there is a measure η satisfying properties (i) and (ii) of Theorem 3.1, then for η-a.e. u there holds $u(i) = x_i$ for $i = 0, 1$ and $u(t) \in \{x_0, x_1\}$ for every $t \in (0,1)$ and u cannot be continuous.

4. Geodesics in $(\mathscr{P}((X,\mathsf{d})), W_\psi)$

We apply Theorem 3.1 in order to characterize the geodesics of the metric space $(\mathscr{P}(X), W_\psi)$ in terms of the geodesics of the space (X, d).

In this section I denotes the unitary interval $[0,1]$.

We say that $u : I \to X$ is a constant speed geodesic in (X, d) if

$$\mathsf{d}(u(t), u(s)) = |t-s|\mathsf{d}(u(0), u(1)) \qquad \forall s,t \in I. \tag{61}$$

We define the set

$$G(X,\mathsf{d}) := \{u : I \to X : u \text{ is a constant speed geodesic of } (X,\mathsf{d})\}.$$

Proposition 4.1. *Let (X, τ, d) be an extended Polish space and ψ be satisfying* (6). *If $\eta \in \mathscr{P}(C(I;X))$ is concentrated on $G(X,\mathsf{d})$ and $\gamma_{0,1} := (e_0, e_1)_\#\eta \in \Gamma^\psi_o((e_0)_\#\eta, (e_1)_\#\eta)$, then the curve $\mu : I \to \mathscr{P}(X)$ defined by $\mu_t = (e_t)_\#\eta$ is a constant speed geodesic in $(\mathscr{P}(X), W_\psi)$.*

Proof. Since $\gamma_{0,1} := (e_0, e_1)_\# \eta \in \Gamma_o^\psi(\mu_0, \mu_1)$, the following inequality holds

$$\int_{X \times X} \psi\Big(\frac{\mathsf{d}(x,y)}{W_\psi(\mu_0,\mu_1)}\Big)\, d\gamma_{0,1}(x,y) \le 1. \tag{62}$$

Since η is concentrated on constant speed geodesics and $\gamma_{s,t} := (e_s, e_t)_\# \eta \in \Gamma(\mu_s, \mu_t)$ we have, for every $t, s \in I$, $t \neq s$.

$$\begin{aligned}
\int_{X \times X} \psi\Big(\frac{\mathsf{d}(x,y)}{W_\psi(\mu_0,\mu_1)}\Big)\, d\gamma_{0,1}(x,y) &= \int_{C(I;X)} \psi\Big(\frac{\mathsf{d}(u(0),u(1))}{W_\psi(\mu_0,\mu_1)}\Big)\, d\eta(u) \\
&= \int_{C(I;X)} \psi\Big(\frac{\mathsf{d}(u(t),u(s))}{|t-s| W_\psi(\mu_0,\mu_1)}\Big)\, d\eta(u) \\
&= \int_{X \times X} \psi\Big(\frac{\mathsf{d}(x,y)}{|t-s| W_\psi(\mu_0,\mu_1)}\Big)\, d\gamma_{t,s}(x,y).
\end{aligned} \tag{63}$$

From (62) and (63) it follows that

$$W_\psi(\mu_t, \mu_s) \le |t-s| W_\psi(\mu_0, \mu_1) \qquad \forall s, t \in I. \tag{64}$$

By the triangular inequality we conclude that equality holds in (64). $\square$

Theorem 4.2. *Let (X, τ, d) be an extended Polish space and ψ be satisfying* (6), (10) *and* (11). *Let $\mu : I \to \mathscr{P}(X)$ be a constant speed geodesic in $(\mathscr{P}(X), W_\psi)$ and $\eta \in \mathscr{P}(C(I;X))$ a measure representing μ in the sense that (i), (ii) and (iii) of Theorem 3.1 hold. Then $\gamma_{s,t} := (e_s, e_t)_\# \eta$ belongs to $\Gamma_o^\psi(\mu_s, \mu_t)$ for every $s, t \in I$. If, in addition, ψ is strictly convex and*

$$\int_{X \times X} \psi\Big(\frac{\mathsf{d}(x,y)}{W_\psi(\mu_0,\mu_1)}\Big)\, d\gamma_{0,1}(x,y) = 1, \tag{65}$$

then η is concentrated on $G(X, \mathsf{d})$.

Proof. Let $L = W_\psi(\mu_0, \mu_1)$. Since μ is a constant speed geodesic and (iii) of Theorem 3.1 holds

$$L = |\mu'|(r) = \| |u'|(r) \|_{L^\psi_\eta(C(I;X))} \qquad \text{for a.e. } r \in I. \tag{66}$$

Let $t, s \in I$, $t \neq s$. Since, by (66), it holds

$$\frac{1}{t-s} \int_s^t \int_{C(I;X)} \psi\Big(\frac{|u'|(r)}{L}\Big)\, d\eta(u)\, dr \le 1,$$

Fubini's theorem and Jensen's inequality yield

$$\int_{C(I;X)} \psi\Big(\frac{1}{t-s} \int_s^t \frac{|u'|(r)}{L}\, dr\Big)\, d\eta(u) \le 1. \tag{67}$$

By the monotonicity of ψ and (67) we obtain

$$\int_{C(I;X)} \psi\Big(\frac{\mathsf{d}(u(s),u(t))}{|t-s| L}\Big)\, d\eta(u) \le 1.$$

Since $|t-s|L = W_\psi(\mu_s,\mu_t)$ we have

$$\int_{C(I;X)} \psi\Big(\frac{\mathsf{d}(u(s),u(t))}{W_\psi(\mu_s,\mu_t)}\Big)\,d\eta(u) \le 1 \tag{68}$$

and, recalling (28), this shows that $\gamma_{s,t}$ is optimal.

Assuming (65) and using Jensen's inequality we have

$$\begin{aligned} 1 &= \int_{C(I;X)} \psi\Big(\frac{\mathsf{d}(u(0),u(1))}{L}\Big)\,d\eta(u) \le \int_{C(I;X)} \psi\Big(\int_0^1 \frac{|u'|(t)}{L}\,dt\Big)\,d\eta(u) \\ &\le \int_{C(I;X)} \int_0^1 \psi\Big(\frac{|u'|(t)}{L}\Big)\,dt\,d\eta(u) = \int_0^1 \int_{C(I;X)} \psi\Big(\frac{|u'|(t)}{L}\Big)\,d\eta(u)\,dt \le 1. \end{aligned} \tag{69}$$

It follows that equality holds in (69) and, still by Jensen's inequality, we have

$$\psi\Big(\int_0^1 \frac{|u'|(t)}{L}\,dt\Big) = \int_0^1 \psi\Big(\frac{|u'|(t)}{L}\Big)\,dt, \qquad \text{for } \eta\text{-a.e. } u \in C(I;X). \tag{70}$$

The strict convexity of ψ implies that, if u satisfies the equality in (70), then $|u'|$ is constant, say $|u'|(t) = L_u$ for a.e. $t \in I$. Analogously equality in (69) shows that $\psi\Big(\frac{\mathsf{d}(u(0),u(1))}{L}\Big) = \psi\Big(\frac{L_u}{L}\Big)$ for η-a.e. $u \in C(I;X)$. The strict monotonicity of ψ implies that $\mathsf{d}(u(0),u(1)) = L_u$ and we conclude that $u \in G(X,\mathsf{d})$ for η-a.e. $u \in C(I;X)$. □

Acknowledgements

The author would like to thank Luigi Ambrosio, Nicola Gigli and Giuseppe Savaré for helpful conversations on the topic of this paper. The author is supported by a MIUR-PRIN 2010-2011 grant for the project Calculus of Variations. The author is member of the GNAMPA group of the Istituto Nazionale di Alta Matematica (INdAM).

References

[ADM14] Luigi Ambrosio and Simone Di Marino. Equivalent definitions of BV space and of total variation on metric measure spaces. To appear on J. Funct. Anal. http://dx.doi.org/10.1016/j.jfa.2014.02.002, 2014.

[AGS05] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savarè. *Gradient flows in metric spaces and in the Wasserstein spaces of probability measures*. Birkhäuser, 2005.

[AGS13] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. Density of lipschitz functions and equivalence of weak gradients in metric measure spaces. *Rev. Mat. Iberoamericana*, 29:969–986, 2013.

[AGS14] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below. *Inventiones mathematicae*, 195(2):289–391, 2014.

[BS11] Lorenzo Brasco and Filippo Santambrogio. An equivalent path functional formulation of branched transportation problems. *Discrete Contin. Dyn. Syst.*, 29(3):845–871, 2011.

[FGY11] Alessio Figalli, Wilfrid Gangbo, and Türkay Yolcu. A variational method for a class of parabolic PDEs. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 10(1):207–252, 2011.

[Kuw13] Kazumasa Kuwada. Gradient estimate for Markov kernels, Wasserstein control and Hopf-Lax formula. To appear in RIMS Kôkyûroku Bessatsu. Preprint available at http://www.math.ocha.ac.jp/kuwada/papers.html, 2013.

[Lis07] Stefano Lisini. Characterization of absolutely continuous curves in Wasserstein spaces. *Calc. Var. Partial Differential Equations*, 28(1):85–120, 2007.

[RR91] M. M. Rao and Z. D. Ren. *Theory of Orlicz spaces*, volume 146 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker Inc., New York, 1991.

[RS03] Riccarda Rossi and Giuseppe Savaré. Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 2(2):395–431, 2003.

[Str11] Daniel W. Stroock. *Probability theory*. Cambridge University Press, Cambridge, second edition, 2011. An analytic view.

[Stu11] Karl-Theodor Sturm. Generalized Orlicz spaces and Wasserstein distances for convex-concave scale functions. *Bull. Sci. Math.*, 135(6-7):795–802, 2011.

Dipartimento di Matematica "F.Casorati", Università degli Studi di Pavia, 27100 Pavia, Italy

E-mail address: `stefano.lisini@unipv.it`