Barycenters in Alexandrov spaces of curvature bounded below

Shin-ichi OHTA*

December 30, 2011

Abstract

We investigate barycenters of probability measures on proper Alexandrov spaces of curvature bounded below, and show that they enjoy several properties relevant to or different from those in metric spaces of curvature bounded above. We prove the reverse variance inequality, and show that the push forward of a measure to the tangent cone at its barycenter has the flat support.

Keywords: barycenter, Alexandrov space, variance inequality, Wasserstein space

Mathematics Subject Classification (2000): 53C21, 53C22

1 Introduction

In the Euclidean space $\mathbb{R}^n$, the barycenter of a probability measure μ (with finite second moment) is the point $z_\mu = \int_{\mathbb{R}^n} x \, d\mu(x)$. Among other ways, z_μ is determined as the unique minimizer of the function $w \longmapsto \int_{\mathbb{R}^n} |w-x|^2 \, d\mu(x)$ for $w \in \mathbb{R}^n$. This description makes sense in metric spaces (see Section 3 for the precise definition). Then the map $\mu \longmapsto z_\mu$ gives a canonical way of contracting a measure to a point, and there are various applications (see [Jo], [St2], [Oh3] and the references therein).

The behavior of barycenters is closely related to the curvature of X, and is well investigated for metric spaces of curvature bounded above (CAT-*spaces* for short). For instance, a barycenter z_μ of μ uniquely exists in a CAT(0)-space (nonpositively curved metric space), and then the map $\mu \longmapsto z_\mu$ is 1-Lipschitz with respect to the L^2-Wasserstein distance. In contrast to this, the behavior of barycenters in metric spaces of curvature bounded below (i.e., Alexandrov spaces) is less understood. Our aim of the present article is to verify that barycenters are interesting objects also in such spaces.

Our results can be divided into two types: quantitative estimates relevant to known results in CAT-spaces, and qualitative properties different from CAT-spaces. Our main result of the first kind is the reverse variance inequality (Theorems 4.8, 5.2) which is literally the reverse of the variance inequality known in CAT-spaces. As an application,

*Department of Mathematics, Kyoto University, Kyoto 606-8502, JAPAN (e-mail: sohta@math.kyoto-u.ac.jp). Supported in part by the Grant-in-Aid for Young Scientists (B) 20740036.

in the Wasserstein space over an Alexandrov space, any two geodesics emanating from the Dirac measure at their common barycenter have angle at most $\pi/2$ (Corollary 4.10). This is a very different phenomenon than CAT-spaces. Another main result (Theorem 4.11) asserts that the push forward of a measure to the tangent cone at its barycenter must have the flat support. In particular, the origin of a singular cone can not be a barycenter of a measure other than the Dirac measure at the origin (Corollary 4.12). This is also different from CAT-spaces, and seems to have further applications.

The organization of the article is as follows. After reviewing the basics of Alexandrov spaces and Wasserstein spaces in Section 2, we verify auxiliary lemmas on barycenters in general proper metric spaces in Section 3. Then Section 4 is devoted to the study of barycenters in Alexandrov spaces and our main results. Some estimates are improved in Section 5 in the particular case of nonnegative (or positive) curvature, and we compare them with nonpositively curved spaces.

Acknowledgements. This work stemmed from a discussion with Asuka Takatsu on her work [TY] with Takumi Yokota. I am grateful to them for valuable comments throughout the preparation of the article. Specifically, Yokota's recent work [Yo] was essential to improve the presentation of Theorem 4.11.

2 Preliminaries

We introduce some notations for later use. Let (X,d) be a metric space. The open ball of center $x \in X$ and radius $r>0$ will be denoted by $B(x,r)$. A rectifiable curve $\gamma : [0,l] \longrightarrow X$ is called a *geodesic* if it is locally minimizing and parametrized proportionally to the arc length. If γ is also globally minimizing, then it is said to be *minimal*. We call (X,d) a *geodesic space* if every pair of points is connected by a minimal geodesic. Denote by $\Gamma(X)$ the set of all minimal geodesics $\gamma : [0,1] \longrightarrow X$ equipped with the uniform topology induced from the distance $d_{\Gamma(X)}(\gamma,\eta) := \sup_{t\in[0,1]} d_X(\gamma(t),\eta(t))$. For each $t \in [0,1]$, define the evaluation map $e_t : \Gamma(X) \longrightarrow X$ by $e_t(\gamma) := \gamma(t)$. Observe that e_t is 1-Lipschitz.

Define $\mathcal{P}(X)$ as the set of all Borel probability measures on X, and define the subset $\mathcal{P}_2(X) \subset \mathcal{P}(X)$ as $\mu \in \mathcal{P}_2(X)$ if $\int_X d(w,x)^2\,d\mu(x) < \infty$ holds for some (and hence all) $w \in X$. We denote by $\mathcal{P}_c(X) \subset \mathcal{P}_2(X)$ the subset of compactly supported measures.

2.1 Alexandrov spaces

We review the basics of Alexandrov spaces of curvature bounded below. We refer to [ABN], [BGP], [OS] and [BBI] for further details.

For $k \in \mathbb{R}$, we denote by $\mathbb{M}^2(k)$ the two-dimensional simply-connected space form of constant sectional curvature k. Then a geodesic space (X,d) is called an *Alexandrov space of curvature* $\geq k$ if, given any three points $x,y,z \in X$ (with $d(x,y)+d(y,z)+d(z,x) \leq 2\pi/\sqrt{k}$ if $k>0$) and any minimal geodesic $\gamma : [0,1] \longrightarrow X$ from x to y, it holds that

$$d_X\big(z,\gamma(t)\big) \geq d_{\mathbb{M}^2(k)}\big(\tilde{z},\tilde{\gamma}(t)\big) \tag{2.1}$$

for all $t \in [0,1]$, where $\triangle\tilde{x}\tilde{y}\tilde{z} \subset \mathbb{M}^2(k)$ is a *comparison triangle* satisfying

$$d_X(x,y) = d_{\mathbb{M}^2(k)}(\tilde{x},\tilde{y}), \quad d_X(y,z) = d_{\mathbb{M}^2(k)}(\tilde{y},\tilde{z}), \quad d_X(z,x) = d_{\mathbb{M}^2(k)}(\tilde{z},\tilde{x}),$$

and $\tilde{\gamma} : [0, 1] \longrightarrow \mathbb{M}^2(k)$ is the unique minimal geodesic from $\tilde{x}$ to $\tilde{y}$. In the particular case of $k = 0$, (2.1) is written as

$$d\big(z, \gamma(t)\big)^2 \geq (1-t)d(z,x)^2 + td(z,y)^2 - (1-t)td(x,y)^2. \tag{2.2}$$

We present fundamental examples of Alexandrov spaces.

Example 2.1 (a) A complete Riemannian manifold is an Alexandrov space of curvature $\geq k$ if and only if its sectional curvature is not less than k everywhere.

(b) If (X, d) is an Alexandrov space of curvature $\geq k$, then the scaled metric space $(X, c \cdot d)$ with $c > 0$ is an Alexandrov space of curvature $\geq k/c^2$.

(c) Every Hilbert space is an Alexandrov space of nonnegative curvature.

(d) For a convex domain D in the Euclidean space $\mathbb{R}^n$, the boundary ∂D equipped with the length distance is an Alexandrov space of nonnegative curvature.

(e) The L^2-Wasserstein space over a compact Alexandrov space of nonnegative curvature is a compact (but infinite dimensional) Alexandrov space of nonnegative curvature. See the next subsection for more details.

We briefly discuss the infinitesimal structure of an Alexandrov space (X, d). Fix $z \in X$ and let $\hat{\Sigma}_z$ be the set of all (nontrivial) unit speed geodesics $\gamma : [0, l] \longrightarrow X$ with $\gamma(0) = z$. For $\gamma, \eta \in \hat{\Sigma}_z$, by virtue of the curvature bound (2.1), the joint limit

$$\angle_z(\gamma, \eta) := \arccos\left(\lim_{s,t \downarrow 0} \frac{s^2 + t^2 - d_X(\gamma(s), \eta(t))^2}{2st}\right) \in [0, \pi]$$

exists and is a pseudo-distance of $\hat{\Sigma}_z$. We define the *space of directions* $(\Sigma_z, \angle_z)$ at z as the completion of $\hat{\Sigma}_z/\sim$ with respect to $\angle_z$, where $\gamma \sim \eta$ if $\angle_z(\gamma, \eta) = 0$. The *tangent cone* (C_z, d_{C_z}) is the Euclidean cone over $(\Sigma_z, \angle_z)$, that is to say,

$$C_z := \Sigma_z \times [0, \infty)/\Sigma_z \times \{0\},$$
$$d_{C_z}\big((\gamma, s), (\eta, t)\big) := \sqrt{s^2 + t^2 - 2st \cos \angle_z(\gamma, \eta)}.$$

We also define the inner product of $\mathbf{u} = (\gamma, s), \mathbf{v} = (\eta, t) \in C_z$ by

$$\langle \mathbf{u}, \mathbf{v} \rangle_z := st \cos \angle_z(\gamma, \eta) = \frac{1}{2}\{s^2 + t^2 - d_{C_z}(\mathbf{u}, \mathbf{v})^2\}.$$

We will denote the origin of C_z by $\mathbf{o}_z$. In Riemannian manifolds, spaces of directions and tangent cones correspond to unit tangent spheres and tangent spaces, respectively.

Finite (Hausdorff) dimensional Alexandrov spaces are known to have remarkably nice local structure. For instance, spaces of directions and tangent cones become Alexandrov spaces of curvature ≥ 1 and ≥ 0, respectively, and (X, d) has a weak differentiable structure ([BGP], [OS]). However, infinite dimensional spaces can be much wilder: tangent cones may not be even geodesic ([Ha]).

Given $z \in X$, we take the subset $D_z \subset C_z$ consisting of elements $\mathbf{v} = (\gamma, t) \in C_z$ associated with some unit speed minimal geodesic $\gamma : [0, l] \longrightarrow X$ with $\gamma(0) = z$ and $l \geq t$. On D_z, we can define the *exponential map* $\exp_z : D_z \longrightarrow X$ by $\exp_z(\gamma, t) := \gamma(t)$. As a consequence of Lemmas 3.3, 4.2 below, there exists a measurable map $\log_z : X \longrightarrow D_z$ such that $\exp_z \circ \log_z = \mathrm{id}_X$. We call such a map $\log_z$ a *logarithmic map* at z.

2.2 Wasserstein spaces

We next explain (Kantorovich-Rubinstein-)Wasserstein spaces which play a key role in the geometric aspect of optimal transport theory. We refer to the recent comprehensive book of Villani [Vi] for further reading.

Let (X,d) be a proper metric space. For $\mu,\nu \in \mathcal{P}_2(X)$, we say that $\pi \in \mathcal{P}(X \times X)$ is a *coupling* of μ and ν if $\pi(A \times X) = \mu(A)$ and $\pi(X \times A) = \nu(A)$ hold for all Borel sets $A \subset X$. For instance, the product measure $\mu \times \nu$ is a coupling of μ and ν. Then we define the (L^2-) *Wasserstein distance* by

$$d_2^W(\mu,\nu) := \inf_{\pi} \left(\int_{X \times X} d(x,y)^2 \, d\pi(x,y) \right)^{1/2}, \tag{2.3}$$

where π runs over all couplings of μ and ν. Note that $d_2^W(\mu,\nu)$ is finite since $\mu,\nu \in \mathcal{P}_2(X)$. We call the metric space $(\mathcal{P}_2(X), d_2^W)$ the (L^2-) *Wasserstein space* over X.

The following lemma is concerned with the non-branching property. We say that a metric space (X,d) is *non-branching* if four points $x, y_0, y_1, y_2 \in X$ satisfy $d(x,y_0) = d(x,y_i) = d(y_0,y_i)/2$ for $i = 1,2$ only if $y_1 = y_2$. Observe that any Alexandrov space of curvature bounded below is non-branching (see also Remark 5.1).

Lemma 2.2 ([Vi, Corollary 7.32]) *Let (X,d) be a proper metric space. If (X,d) is non-branching, then so is $(\mathcal{P}_2(X), d_2^W)$.*

It is known by [LV, Theorem A.8] and [St4, Proposition 2.10] that the Wasserstein space over a compact geodesic space (X,d) is an Alexandrov space of nonnegative curvature if and only if so is (X,d) (recall Example 2.1(e)). However, over an Alexandrov space of curvature ≥ -1 but not of nonnegative curvature, the Wasserstein space is not an Alexandrov space of curvature $\geq k$ for any $k \in \mathbb{R}$ ([St4, Proposition 2.10]). Nonetheless, we see in [Oh2, Theorem 3.6] that the angle between two geodesics in the Wasserstein space makes sense. To be precise, for any minimal geodesics $\alpha, \beta : [0,\delta] \longrightarrow \mathcal{P}_c(X)$ with the common starting point $\alpha(0) = \beta(0) =: \mu$, the limit

$$\sigma_\mu(\alpha,\beta) := \lim_{t \downarrow 0} \frac{d_2^W(\alpha(t),\beta(t))}{t}$$

exists and, moreover, the angle

$$\angle_\mu\big(\dot\alpha(0), \dot\beta(0)\big) := \arccos\left(\frac{d_2^W(\mu,\alpha(\delta))^2 + d_2^W(\mu,\beta(\delta))^2 - \delta^2 \sigma_\mu(\alpha,\beta)^2}{2 d_2^W(\mu,\alpha(\delta)) d_2^W(\mu,\beta(\delta))} \right) \tag{2.4}$$

is independent of reparametrizations of α and β. This means that $(\mathcal{P}_2(X), d_2^W)$ carries a kind of Riemannian structure, and there are applications in gradient flow theory.

3 Barycenters in proper metric spaces

We verify some auxiliary lemmas on barycenters in general proper metric spaces.

Let (X, d) be a metric space. For $\mu \in \mathcal{P}_2(X)$, a *barycenter* (or a *center of mass*) of μ is a point in X which attains the infimum of the function

$$w \longmapsto \int_X d(w, x)^2 \, d\mu(x). \tag{3.1}$$

Note that the infimum is finite for $\mu \in \mathcal{P}_2(X)$. In the language of Wasserstein geometry, the Dirac measure δ_z at a barycenter z of μ is closest to μ among all Dirac measures. In the Euclidean space $\mathbb{R}^n$ with the standard distance structure, every $\mu \in \mathcal{P}_2(\mathbb{R}^n)$ admits the unique barycenter $\int_{\mathbb{R}^n} x \, d\mu(x)$. In general metric spaces, however, neither existence nor uniqueness can be expected:

Example 3.1 (a) Let X be the infinite dimensional ellipsoid of axes of lengths $c_n = (n+1)/2n$ with $n \in \mathbb{N}$, namely

$$X = \left\{ (x_1, x_2, \ldots) \in \mathbb{R}^\infty \,\middle|\, \sum_{n \in \mathbb{N}} \frac{x_n^2}{c_n^2} = 1 \right\}.$$

Then X is complete, but $\mu = (\delta_{(1,0,0,\ldots)} + \delta_{(-1,0,0,\ldots)})/2$ has no barycenter in X.

(b) Let X be the n-dimensional sphere $\mathbb{S}^n$ ($n \in \mathbb{N}$) and μ be the sum of one halves of Dirac measures on the north and south poles. Then every point on the equator is a barycenter of μ.

(c) Let X_ℓ be the Euclidean cone over a circle of length $\ell \in (0, 2\pi)$, and μ be the normalized uniform distribution on $B(\mathbf{o}, 1)$, where $\mathbf{o}$ is the origin of the cone. Cutting X_ℓ along a meridian and developing it in $\mathbb{R}^2$, we find that $\mathbf{o}$ is not a barycenter of μ. Then, by symmetry, there is $r_\ell \in (0, 2/3)$ such that every point on the circle $\partial B(\mathbf{o}, r_\ell)$ is a barycenter, and r_ℓ tends to 0 (resp. 2/3) as ℓ goes to 2π (resp. 0).

This is a typical example demonstrating the difference between nonnegatively and nonpositively curved spaces. On the one hand, the cone X_ℓ as above for $l \in (0, 2\pi)$ is an Alexandrov space of nonnegative curvature. On the other hand, for $\ell \geq 2\pi$, X_ℓ is a CAT(0)-space (see Subsection 5.1) and the origin is a unique barycenter of μ.

Nevertheless, it is easy to see existence in proper metric spaces.

Lemma 3.2 *If (X, d) is a proper metric space, then any $\mu \in \mathcal{P}_2(X)$ has a barycenter.*

Proof. Fix $z_0 \in X$ and take $r > 1$ large enough to satisfy $\mu(B(z_0, r)) \geq 1/2$ as well as $\int_{X \setminus B(z_0, r)} d(z_0, x)^2 \, d\mu(x) \leq 1$. Then we have

$$\int_X d(z_0, x)^2 \, d\mu(x) \leq r^2 \cdot \mu\big(B(z_0, r)\big) + 1 \leq r^2 + 1,$$

while for every $w \in X \setminus B(z_0, 3r)$

$$\int_X d(w, x)^2 \, d\mu(x) \geq \int_{B(z_0, r)} d(w, x)^2 \, d\mu(x) > (2r)^2 \cdot \mu\big(B(z_0, r)\big) \geq 2r^2$$

holds. Therefore it is sufficient to consider the infimum of (3.1) only for $w \in B(z_0, 3r)$, and it is achieved at some point due to the compactness of the closure of $B(z_0, 3r)$. $\square$

Next we consider the contraction of a measure to its barycenter. Although the following measurable selection property is rather standard, we give a sketch of proof for completeness.

Lemma 3.3 *Let (X, d) be a proper geodesic space. Then, for any $z \in X$, there exists a measurable map $\Phi : X \longrightarrow \Gamma(X)$ satisfying $e_0 \circ \Phi(x) = z$ and $e_1 \circ \Phi(x) = x$ for all $x \in X$.*

Proof. As (X, d) is proper, $(\Gamma(X), d_{\Gamma(X)})$ is also proper. We consider the map $F : X \longrightarrow 2^{\Gamma(X)}$ defined by $F(x) := e_0^{-1}(z) \cap e_1^{-1}(x)$ $(\neq \emptyset)$. We shall show that

$$\{x \in X \mid F(x) \cap G \neq \emptyset\} = e_1(G \cap \Gamma_z)$$

is a Borel set for every open set $G \subset \Gamma(X)$, where $\Gamma_z := e_0^{-1}(z)$. Then Kuratowski and Ryll-Nardzewski's classical selection theorem [KR] provides a measurable map $\Phi : X \longrightarrow \Gamma(X)$ with $\Phi(x) \in F(x)$ for all $x \in X$, as desired.

Fix a (nonempty) open set $G \subset \Gamma(X)$. For $\delta > 0$, let A_δ be the complement of the open δ-neighborhood of $\Gamma(X) \setminus (G \cap \Gamma_z)$. Note that $\bigcup_{\delta>0} A_\delta = G \cap \Gamma_z$. Given $\varepsilon > 0$, we consider the set U_ε of points $x \in X$ such that there is a rectifiable curve $\xi : [0, 1] \longrightarrow X$ with $\xi(0) = z$, $\xi(1) = x$ as well as $\inf_{\gamma \in A_\delta} \sup_{t \in [0,1]} d(\xi(t), \gamma(t)) < \varepsilon$. Observe that U_ε is an open set and that $\bigcap_{\varepsilon>0} U_\varepsilon = e_1(A_\delta)$. Hence $\bigcup_{\delta>0} e_1(A_\delta) = e_1(G \cap \Gamma_z)$ is a Borel set. $\square$

In particular, for any $\mu \in \mathcal{P}(X)$, we find that $\Pi = \Phi_\sharp \mu \in \mathcal{P}(\Gamma(X))$ satisfies $(e_0)_\sharp \Pi = \delta_z$ and $(e_1)_\sharp \Pi = \mu$.

Lemma 3.4 *Let (X, d) be a proper geodesic space. Given a barycenter z of $\mu \in \mathcal{P}_2(X)$ and $\Pi \in \mathcal{P}(\Gamma(X))$ so that $(e_0)_\sharp \Pi = \delta_z$ and $(e_1)_\sharp \Pi = \mu$, z is a barycenter of $(e_t)_\sharp \Pi$ for all $t \in [0, 1)$.*

Proof. Put $\mu_t := (e_t)_\sharp \Pi$ for $t \in [0, 1]$. Then $\int_X d(z, y)^2 \, d\mu_t(y) = t^2 \int_X d(z, x)^2 \, d\mu(x)$ clearly holds. Fix $w \in X$, $t \in (0, 1)$ and $\gamma \in \operatorname{supp} \Pi$. The triangle inequality verifies $d(w, \gamma(1)) \leq d(w, \gamma(t)) + d(\gamma(t), \gamma(1))$, and the convexity of the function $s \longmapsto s^2$ shows

$$\begin{aligned} d\big(w, \gamma(1)\big)^2 &\leq \frac{1}{t} d\big(w, \gamma(t)\big)^2 + \frac{1}{1-t} d\big(\gamma(t), \gamma(1)\big)^2 \\ &= \frac{1}{t} d\big(w, \gamma(t)\big)^2 + (1-t) d\big(z, \gamma(1)\big)^2. \end{aligned}$$

Hence we have

$$d\big(w, \gamma(t)\big)^2 \geq t d\big(w, \gamma(1)\big)^2 - (1-t) t d\big(z, \gamma(1)\big)^2. \tag{3.2}$$

Integrating (3.2) with respect to Π yields

$$\int_X d(w, y)^2 \, d\mu_t(y) \geq t \int_X d(w, x)^2 \, d\mu(x) - (1-t) t \int_X d(z, x)^2 \, d\mu(x).$$

As z is a barycenter of μ, this implies

$$\int_X d(w, y)^2 \, d\mu_t(y) \geq t^2 \int_X d(z, x)^2 \, d\mu(x) = \int_X d(z, y)^2 \, d\mu_t(y).$$

Therefore z is a barycenter of μ_t. The case of $t = 0$ is clear. $\square$

We remark that, in Lemma 3.4, z is not necessarily a unique barycenter of μ_t.

Example 3.5 Let $I_n := [-2^{-n+1}, 2^{-n+1}]$ for each $n \in \mathbb{N}$ and set

$$X := \left(\bigsqcup_{n\in\mathbb{N}} I_n \cup \{z\} \right) \Big/ \sim,$$

where $-2^{-n}, 2^{-n} \in I_n$ are identified with $-2^{-n}, 2^{-n} \in I_{n+1}$, respectively, and z is attached as the limit point of the sequence $\{2^{-n+1} \in I_n\}_{n\in\mathbb{N}}$ (or $\{-2^{-n+1} \in I_n\}_{n\in\mathbb{N}}$) as n goes to infinity. Observe that X is compact with respect to the length distance, but not locally simply connected at z. Now we consider unique minimal geodesics $\gamma_\pm : [0,1] \longrightarrow X$ from z to $\pm 1 \in I_1$, and put $\mu_t := (\delta_{\gamma_-(t)} + \delta_{\gamma_+(t)})/2$. Then z is a barycenter of μ_t for all $t \in [0,1]$, but $0 \in I_n$ is also a barycenter of μ_t for $t \in [2^{-n+1}, 1]$. Note that the point of this construction is branching geodesics in X, compare this with Lemma 4.3.

The persistence of barycenter along a geodesic in the Wasserstein space holds true only when contracting to the Dirac measure at the barycenter. That is to say, even if endpoints $\alpha(0)$, $\alpha(1)$ of a minimal geodesic $\alpha : [0,1] \longrightarrow \mathcal{P}_2(X)$ have a common berycenter z, it does not necessarily imply that z is a barycenter of $\alpha(t)$ for $t \in (0,1)$. In fact, we can show the following.

Proposition 3.6 *Let (M,g) be a Riemannian manifold satisfying the property:*

($*$) *For any minimal geodesic $\alpha : [0,1] \longrightarrow \mathcal{P}_2(M)$ such that a point z is a barycenter of both $\alpha(0)$ and $\alpha(1)$, z is also a barycenter of $\alpha(t)$ for all $t \in (0,1)$.*

Then (M,g) is flat.

Proof. Fix $z \in M$ and unit vectors $\mathbf{u}, \mathbf{v} \in T_zM$ with $\angle(\mathbf{u}, \mathbf{v}) = \pi/3$. Let γ, η be geodesics such that $\dot\gamma(0) = \mathbf{u}$ and $\dot\eta(0) = \mathbf{v}$. For $0 < \varepsilon \ll \tau \ll 1$, we put

$$\mu_0 := \frac{\tau}{\tau+\varepsilon}\delta_{\gamma(-2\varepsilon)} + \frac{\varepsilon}{\tau+\varepsilon}\delta_{\gamma(2\tau)}, \quad \mu_1 := \frac{\tau}{\tau+\varepsilon}\delta_{\eta(-\varepsilon)} + \frac{\varepsilon}{\tau+\varepsilon}\delta_{\eta(\tau)}.$$

Then $z = \gamma(0) = \eta(0)$ is the unique barycenter of both μ_0 and μ_1. Moreover, the optimal transport (minimal geodesic in the Wasserstein space) from μ_0 to μ_1 is done along geodesics $\xi : [0,1] \longrightarrow M$ from $\gamma(-2\varepsilon)$ to $\eta(-\varepsilon)$ as well as $\zeta : [0,1] \longrightarrow M$ from $\gamma(2\tau)$ to $\eta(\tau)$. Let us consider the midpoint of μ_0 and μ_1:

$$\mu_{1/2} = \frac{\tau}{\tau+\varepsilon}\delta_{\xi(1/2)} + \frac{\varepsilon}{\tau+\varepsilon}\delta_{\zeta(1/2)}.$$

Note that the angle $\angle\eta(\tau)z\zeta(1/2)$ coincides with $\arccos(2/\sqrt{7})$ if (M,g) is flat, and it is smaller (larger, resp.) than $\arccos(2/\sqrt{7})$ if the sectional curvature κ of the 2-plane spanned by $\mathbf{u}$ and $\mathbf{v}$ is positive (negative, resp.). However, the angle $\angle\eta(-\varepsilon)z\xi(1/2)$ can be arbitrarily close to $\arccos(2/\sqrt{7})$ for small $\varepsilon > 0$. Therefore $\angle\eta(\tau)z\zeta(1/2) < \angle\eta(-\varepsilon)z\xi(1/2)$ if $\kappa > 0$, and $\angle\eta(\tau)z\zeta(1/2) > \angle\eta(-\varepsilon)z\xi(1/2)$ if $\kappa < 0$. Thus the minimal geodesic between $\xi(1/2)$ and $\zeta(1/2)$ does not pass through z if $\kappa \neq 0$, so that z is not a barycenter of $\mu_{1/2}$. Hence ($*$) is false unless (M,g) is flat. $\square$

It is easy to see that ($*$) holds true in Hilbert spaces and, more generally, complete geodesic spaces satisfying equality in (2.2).

4 Barycenters in Alexandrov spaces

This section is the main part of the article. Throughout the section, (X,d) is a proper Alexandrov space of curvature ≥ -1. Due to the scaling property as in Example 2.1(b), choosing -1 as the lower bound does not lose any generality.

4.1 Preliminary lemmas

We start with preliminary lemmas for later convenience.

Lemma 4.1 *Fix $z \in X$ and take $\Pi, \Xi \in \mathcal{P}(\Gamma(X))$ with $(e_0)_\sharp \Pi = (e_0)_\sharp \Xi = \delta_z$ as well as $(e_1)_\sharp \Pi, (e_1)_\sharp \Xi \in \mathcal{P}_2(X)$. Then we have*

$$\lim_{t\downarrow 0} \frac{1}{t^2} \int_{\Gamma(X)\times\Gamma(X)} d\big(\gamma(t),\eta(t)\big)^2 \, d\Pi(\gamma) d\Xi(\eta) = \int_{\Gamma(X)\times\Gamma(X)} \lim_{t\downarrow 0} \frac{d(\gamma(t),\eta(t))^2}{t^2} \, d\Pi(\gamma) d\Xi(\eta).$$

Proof. Given $R>0$, we set $B_R := e_1^{-1}(B(z,R)) \subset \Gamma(X)$ and $B_R^c := \Gamma(X) \setminus B_R$. On the one hand, the dominated convergence theorem yields

$$\int_{B_R\times B_R} \lim_{t\downarrow 0} \frac{d(\gamma(t),\eta(t))^2}{t^2} \, d\Pi(\gamma) d\Xi(\eta) = \lim_{t\downarrow 0} \frac{1}{t^2} \int_{B_R\times B_R} d\big(\gamma(t),\eta(t)\big)^2 \, d\Pi(\gamma) d\Xi(\eta).$$

On the other hand, it follows from the triangle inequality that

$$\begin{aligned}
&\frac{1}{t^2} \int_{B_R^c\times\Gamma(X)} d\big(\gamma(t),\eta(t)\big)^2 \, d\Pi(\gamma) d\Xi(\eta) \\
&\leq \frac{2}{t^2} \int_{B_R^c} d\big(z,\gamma(t)\big)^2 \, d\Pi(\gamma) + \frac{2\Pi(B_R^c)}{t^2} \int_{\Gamma(X)} d\big(z,\eta(t)\big)^2 \, d\Xi(\eta) \\
&= 2 \int_{B_R^c} d\big(z,\gamma(1)\big)^2 \, d\Pi(\gamma) + 2\Pi(B_R^c) \int_{\Gamma(X)} d\big(z,\eta(1)\big)^2 \, d\Xi(\eta) \to 0
\end{aligned}$$

as R diverges to infinity. Combining these, we complete the proof. $\square$

Given $z \in X$, put $\Gamma_z := e_0^{-1}(z) \subset \Gamma(X)$. We define the one-to-one map $\Theta : \Gamma_z \longrightarrow D_z \subset C_z$ as the inverse of $(\gamma, s) \longmapsto \hat{\gamma}$, where $\hat{\gamma}(t) := \gamma(st)$.

Lemma 4.2 *The map $\Theta : \Gamma_z \longrightarrow C_z$ is measurable.*

Proof. It is sufficient to show that $\Theta^{-1}(B(\mathbf{v},r))$ is a Borel set for any $\mathbf{v} \in C_z$ and $r>0$. By approximation, we can assume that $\mathbf{v}$ is represented as $\mathbf{v} = (\gamma, s)$ with $\gamma \in \hat{\Sigma}_z$. Then we observe

$$\begin{aligned}
\Theta^{-1}\big(B(\mathbf{v},r)\big) &= \left\{ \eta \in \Gamma_z \,\middle|\, \lim_{t\downarrow 0} \frac{d(\gamma(st),\eta(t))}{t} < r \right\} \\
&= \bigcup_{N\in\mathbb{N}} \bigcap_{m\geq N} \big\{ \eta \in \Gamma_z \,\big|\, d\big(\gamma(s/m),\eta(1/m)\big) < r/m \big\}.
\end{aligned}$$

As every $\{\eta \in \Gamma_z \,|\, d(\gamma(s/m),\eta(1/m)) < r/m\}$ is clearly Borel, so is $\Theta^{-1}(B(\mathbf{v},r))$. $\square$

Composing Θ with the map $\Phi : X \longrightarrow \Gamma_z$ given by Lemma 3.3 ensures the existence of a measurable logarithmic map $\log_z : X \longrightarrow D_z$. Combination of Lemmas 2.2, 3.4 immediately shows the following.

Lemma 4.3 *Given a barycenter z of $\mu \in \mathcal{P}_2(X)$ and $\Pi \in \mathcal{P}(\Gamma(X))$ with $(e_0)_\sharp \Pi = \delta_z$ and $(e_1)_\sharp \Pi = \mu$, z is a unique barycenter of $(e_t)_\sharp \Pi$ for every $t \in [0,1)$.*

Proof. If μ_t admits a barycenter $z' \neq z$ for some $t \in (0,1)$, then z' is also a barycenter of μ since

$$d_2^W(\delta_{z'}, \mu) \le d_2^W(\delta_{z'}, \mu_t) + d_2^W(\mu_t, \mu) = d_2^W(\delta_z, \mu_t) + d_2^W(\mu_t, \mu) = d_2^W(\delta_z, \mu).$$

Then, however, the non-branching property (Lemma 2.2) yields $\delta_z = \delta_{z'}$, this is a contradiction. The case of $t = 0$ is clear. $\square$

The following lemma (to be improved in Lemma 4.6) is regarded as an infinitesimal (and quantitative) version of Lemma 4.3.

Lemma 4.4 *Let z be a barycenter of $\mu \in \mathcal{P}_2(X)$. Then, for any $\mathbf{v} \in \Sigma_z$, any logarithmic map $\log_z : X \longrightarrow C_z$ and $\Lambda := (\log_z)_\sharp \mu$, we have $\int_{C_z} \langle \mathbf{u}, \mathbf{v} \rangle_z \, d\Lambda(\mathbf{u}) \le 0$. In other words,*

$$\int_{C_z} d_{C_z}(\mathbf{v}, \mathbf{u})^2 d\Lambda(\mathbf{u}) \ge d_{C_z}(\mathbf{o}_z, \mathbf{v})^2 + \int_{C_z} d_{C_z}(\mathbf{o}_z, \mathbf{u})^2 \, d\Lambda(\mathbf{u}) \tag{4.1}$$

holds. In particular, $\mathbf{o}_z$ is a unique barycenter of Λ.

Proof. Let $\Phi : X \longrightarrow \Gamma(X)$ be the map associating $x \in X$ with the geodesic $\gamma \in \Gamma(X)$ so that $\gamma(t) = \bar{\gamma}(td(z,x))$ with $\log_z(x) = (\bar{\gamma}, d(z,x))$ (see also Lemma 3.3), and put $\Pi := \Phi_\sharp \mu$. Note that $[\Phi(x)](0) = z$ and $[\Phi(x)](1) = x$, thus $(e_0)_\sharp \Pi = \delta_z$ and $(e_1)_\sharp \Pi = \mu$.

As $\int_{C_z} \langle \mathbf{u}, \mathbf{v} \rangle_z \, d\Lambda(\mathbf{u})$ is continuous in $\mathbf{v}$, we can assume that $\mathbf{v} = (\eta, s)$ for some geodesic $\eta : [0, \varepsilon) \longrightarrow X$ with $\eta(0) = z$. Since z is a barycenter of μ, we have

$$0 \ge \frac{1}{t} \int_X \left\{ d(x,z)^2 - d\big(x, \eta(st)\big)^2 \right\} d\mu(x)$$

for $t \in (0, \varepsilon)$. For each x, by putting $\log_z(x) = (\gamma, d(z,x))$, it follows from the (directional) first variation formula ([OS, Fact (c-2)], [BBI, Proposition 4.5.2]) that

$$\lim_{t \downarrow 0} \frac{d(x,z)^2 - d(x, \eta(st))^2}{t} \ge 2d(x,z)s \cos\big(\dot{\gamma}(0), \dot{\eta}(0)\big) = 2\langle \log_z(x), \mathbf{v} \rangle_z.$$

Thus we obtain $\int_{C_z} \langle \mathbf{u}, \mathbf{v} \rangle_z \, d\Lambda(\mathbf{u}) \le 0$ by the dominated convegence theorem (as in Lemma 4.1). $\square$

4.2 Lang and Schroeder's inequality and key lemma

We introduce Lang and Schroeder's useful and important inequality. Their original version ([LS, Proposition 3.2]) is concerned with finitely supported measures, so that we slightly generalize it to arbitrary measures.

Lemma 4.5 *For any $z \in X$, $\mu \in \mathcal{P}_2(X)$, any logarithmic map $\log_z : X \longrightarrow C_z$ and $\Lambda := (\log_z)_\sharp \mu$, we have*

$$\int_{C_z \times C_z} \langle \mathbf{u}, \mathbf{v} \rangle_z \, d\Lambda(\mathbf{u}) d\Lambda(\mathbf{v}) \ge 0.$$

Proof. Similarly to Lemma 4.1, it is sufficient to consider μ satisfying $\operatorname{supp} \mu \subset B(z, R)$ for some $R > 0$. We approximate μ by finitely supported measures $\{\mu^i\}_{i \in \mathbb{N}}$ with respect to the weak convergence. Define the map $\Phi : X \longrightarrow \Gamma(X)$ as in Lemma 4.4 and put $\mu_t := (e_t \circ \Phi)_\sharp \mu$ and $\mu_t^i := (e_t \circ \Phi)_\sharp \mu^i$ for $t \in [0, 1]$. We also set $\Lambda^i := (\log_z)_\sharp \mu^i$ and deduce from [LS, Proposition 3.2] that $\int_{C_z \times C_z} \langle \mathbf{u}, \mathbf{v} \rangle_z \, d\Lambda^i(\mathbf{u}) d\Lambda^i(\mathbf{v}) \ge 0$, in other words,

$$2 \int_X d(z, x)^2 \, d\mu^i(x) \ge \int_{X \times X} \lim_{t \downarrow 0} \frac{d(x, y)^2}{t^2} \, d\mu_t^i(x) d\mu_t^i(y).$$

Note that the lower curvature bound of X implies

$$\int_{X \times X} \lim_{t \downarrow 0} \frac{d(x, y)^2}{t^2} \, d\mu_t^i(x) d\mu_t^i(y) \ge \big(1 + \theta_R(s)\big) \int_{X \times X} \frac{d(x, y)^2}{s^2} \, d\mu_s^i(x) d\mu_s^i(y)$$

for sufficiently small $s > 0$ independent of i, where $\lim_{s \downarrow 0} \theta_R(s) = 0$. As the closure of $B(x, R)$ is compact, letting $i \to \infty$ and then $s \downarrow 0$ yields (as in Lemma 4.1)

$$2 \int_{C_z} d_{C_z}(\mathbf{o}_z, \mathbf{u})^2 \, d\Lambda(\mathbf{u}) \ge \int_{C_z \times C_z} d_{C_z}(\mathbf{u}, \mathbf{v})^2 \, d\Lambda(\mathbf{u}) d\Lambda(\mathbf{v}).$$

This completes the proof. $\square$

The following lemma will be a key tool throughout the remainder of the article.

Lemma 4.6 *Let z be a barycenter of $\mu \in \mathcal{P}_2(X)$. Then, for any $\mathbf{v} \in \Sigma_z$, any logarithmic map $\log_z : X \longrightarrow C_z$ and $\Lambda := (\log_z)_\sharp \mu$, we have*

$$\int_{C_z} \langle \mathbf{u}, \mathbf{v} \rangle_z \, d\Lambda(\mathbf{u}) = 0. \tag{4.2}$$

Proof. Recall from Lemma 4.4 that $\int_{C_z} \langle \mathbf{u}, \mathbf{v} \rangle_z \, d\Lambda(\mathbf{u}) \le 0$ generally holds. Combining this with Lemma 4.5, we obtain $\int_{C_z \times C_z} \langle \mathbf{u}, \mathbf{w} \rangle_z \, d\Lambda(\mathbf{u}) d\Lambda(\mathbf{w}) = 0$. We next apply Lemma 4.5 to $(1 + \varepsilon)^{-1}(\Lambda + \varepsilon \delta_{\mathbf{v}})$ and find

$$\int_{C_z \times C_z} \langle \mathbf{u}, \mathbf{w} \rangle_z \, d\Lambda(\mathbf{u}) d\Lambda(\mathbf{w}) + 2\varepsilon \int_{C_z} \langle \mathbf{u}, \mathbf{v} \rangle_z \, d\Lambda(\mathbf{u}) + \varepsilon^2 \langle \mathbf{v}, \mathbf{v} \rangle_z \ge 0$$

for arbitrary $\varepsilon > 0$. As we saw that the first term vanishes, dividing both sides by ε and letting ε go to zero show $\int_{C_z} \langle \mathbf{u}, \mathbf{v} \rangle_z \, d\Lambda(\mathbf{u}) \ge 0$. $\square$

Remark 4.7 If every geodesic $\gamma : [0,\delta) \longrightarrow X$ can be extended to a slightly longer geodesic $\tilde{\gamma} : (-\varepsilon,\delta) \longrightarrow X$ (e.g., in Riemannian manifolds without boundary), then we can find a direction $-\mathbf{v} \in \Sigma_z$ with $\angle_z(\mathbf{v},-\mathbf{v}) = \pi$ for every $\mathbf{v} \in \Sigma_z$, and easily deduce (4.2) by comparing derivatives in the directions $\mathbf{v}$ and $-\mathbf{v}$. For instance, however, geodesics can not be extended beyond the origin of a singular cone. Lang and Schroeder's inequality is the key to overcome the difficulty arising from the absence of $-\mathbf{v}$.

The equation (4.2) is rewritten as

$$\int_{C_z} d_{C_z}(\mathbf{v},\mathbf{u})^2 d\Lambda(\mathbf{u}) = d_{C_z}(\mathbf{o}_z,\mathbf{v})^2 + \int_{C_z} d_{C_z}(\mathbf{o}_z,\mathbf{u})^2\, d\Lambda(\mathbf{u}). \tag{4.3}$$

It is essential in (4.3) that the barycenter is the origin of a cone. More generally, inequality (in the different directions) holds in (4.3) in nonnegatively or nonpositively curved spaces (see Theorem 5.2(i) and (5.4)).

4.3 Reverse variance inequality and applications

Lemma 4.6 enables us to extend Sturm's *reverse variance inequality* [St3, Lemma 8.4] to spaces in which geodesics may not be extended (see Remark 4.7).

Theorem 4.8 *Let z be a barycenter of $\mu \in \mathcal{P}_2(X)$. Then we have, for all $w \in X$,*

$$\int_X \cosh d(w,x) \frac{d(z,x)}{\sinh d(z,x)}\, d\mu(x) \le \cosh d(z,w) \int_X \cosh d(z,x) \frac{d(z,x)}{\sinh d(z,x)}\, d\mu(x).$$

Proof. Take a logarithmic map $\log_z : X \longrightarrow C_z$, put $\Lambda := (\log_z)_\sharp \mu$, and fix a minimal geodesic $\gamma : [0,1] \longrightarrow X$ from z to w. We deduce from (2.1) with $k = -1$ that

$$\begin{aligned}
&\int_X \{\cosh d(w,x) - \cosh d(z,w)\cosh d(z,x)\} \frac{d(z,x)}{\sinh d(z,x)}\, d\mu(x) \\
&\le -\int_{C_z} \sinh d(z,w) \sinh d(\mathbf{o}_z,\mathbf{u}) \cos\angle_z\big(\mathbf{u},\dot{\gamma}(0)\big) \frac{d(\mathbf{o}_z,\mathbf{u})}{\sinh d(\mathbf{o}_z,\mathbf{u})}\, d\Lambda(\mathbf{u}) = 0.
\end{aligned}$$

We used Lemma 4.6 in the last equality. $\square$

Applying Theorem 4.8 twice, we immediately obtain the following corollary.

Corollary 4.9 *Let $z,w \in X$ be barycenters of $\mu,\nu \in \mathcal{P}_2(X)$, respectively. Then we have*

$$\begin{aligned}
&\int_{X\times X} \cosh d(x,y) \frac{d(z,x)}{\sinh d(z,x)} \frac{d(w,y)}{\sinh d(w,y)}\, d\mu(x)d\nu(y) \\
&\le \cosh d(z,w) \int_X \cosh d(z,x) \frac{d(z,x)}{\sinh d(z,x)}\, d\mu(x) \int_X \cosh d(w,y) \frac{d(w,y)}{\sinh d(w,y)}\, d\nu(y).
\end{aligned}$$

Conversely, choosing $\nu = \delta_w$ in Corollary 4.9 recovers Theorem 4.8. See Theorem 5.2 below for the analogue in nonnegatively or positively curved spaces.

The next corollary, inspired by [TY, Remark 4.3] in connection with [CG, (3.10)], is concerned with an estimate in Wasserstein geometry. Recall (2.3) and (2.4) for the Wasserstein distance d_2^W and the angle between geodesics in $\mathcal{P}_c(X)$.

Corollary 4.10 *Suppose that $\mu, \nu \in \mathcal{P}_c(X) \setminus \{\delta_z\}$ have a common barycenter z, and let $\Pi, \Xi \in \mathcal{P}(\Gamma(X))$ satisfy $(e_0)_\sharp \Pi = (e_0)_\sharp \Xi = \delta_z$, $(e_1)_\sharp \Pi = \mu$ and $(e_1)_\sharp \Xi = \nu$. Then we have*

$$\angle_{\delta_z}\big(\dot\alpha(0), \dot\beta(0)\big) \le \frac{\pi}{2},$$

where we set $\alpha(t) := (e_t)_\sharp \Pi$ and $\beta(t) := (e_t)_\sharp \Xi$.

Proof. Note that, since $\alpha(t) \times \beta(t)$ is a coupling of $\alpha(t)$ and $\beta(t)$,

$$\lim_{t \downarrow 0} \frac{1}{t^2} d_2^W \big(\alpha(t), \beta(t)\big)^2 \le \liminf_{t \downarrow 0} \frac{1}{t^2} \int_{X \times X} d(x,y)^2 \, d[\alpha(t)](x) d[\beta(t)](y).$$

Take $R > 0$ such that $B(z,R) \supset \operatorname{supp} \mu \cup \operatorname{supp} \nu$ and observe, for $x, y \in B(z, tR)$,

$$\begin{aligned} &d(x,y)^2 = 2\{\cosh d(x,y) - 1\} + O(t^3), \\ &2\{\cosh d(z,x) \cosh d(z,y) - 1\} = d(z,x)^2 + d(z,y)^2 + O(t^3). \end{aligned}$$

Thus it follows from Corollary 4.9 with $z = w$ that

$$\begin{aligned} &\int_{X \times X} d(x,y)^2 \, d[\alpha(t)](x) d[\beta(t)](y) \\ &\le \int_{X \times X} \{d(z,x)^2 + d(z,y)^2\} \, d[\alpha(t)](x) d[\beta(t)](y) + O(t^3). \end{aligned}$$

Therefore we have $\lim_{t \downarrow 0} d_2^W(\alpha(t), \beta(t))^2/t^2 \le d_2^W(\delta_z, \mu)^2 + d_2^W(\delta_z, \nu)^2$, and hence

$$\cos \angle_{\delta_z}\big(\dot\alpha(0), \dot\beta(0)\big) = \frac{d_2^W(\delta_z, \mu)^2 + d_2^W(\delta_z, \nu)^2 - \lim_{t \downarrow 0} d_2^W(\alpha(t), \beta(t))^2/t^2}{2 d_2^W(\delta_z, \mu) d_2^W(\delta_z, \nu)} \ge 0.$$

□

Given $z \in X$, let $\mathcal{Q}_z \subset \mathcal{P}_c(X)$ be the set of measures adapting z as a barycenter. By virtue of Lemma 3.4, $\mathcal{Q}_z$ is starlike with the origin δ_z, however, Proposition 3.6 asserts that $\mathcal{Q}_z$ is not convex unless X is flat. In addition, Corollary 4.10 ensures that any pair of geodesics in $\mathcal{Q}_z$ emanating from δ_z has angle at most $\pi/2$. Lemma 4.3 shows that only points at the boundary of $\mathcal{Q}_z$ can also belong to some other stratum $\mathcal{Q}_w$.

4.4 Barycenters at the origins of tangent cones

Lemma 4.6 is also useful for deriving qualitative properties of barycenters. The following theorem (inspired by Example 3.1(c)) asserts that a barycenter can live only in an infinitesimally flat subset.

Theorem 4.11 *Let z be a barycenter of $\mu \in \mathcal{P}_2(X)$ and suppose that $(\log_z)_\sharp \mu$ has separable support for some logarithmic map $\log_z : X \longrightarrow C_z$. Then the support of $(\log_z)_\sharp \mu$ is contained in a subset $\mathcal{H} \subset C_z$ which is isometric to a Hilbert space.*

Proof. Put $\Lambda := (\log_z)_\sharp\mu$ and note that Lemma 4.6 yields $\int_{C_z \times C_z} \langle \mathbf{u}, \mathbf{v} \rangle_z \, d\Lambda(\mathbf{u}) d\Lambda(\mathbf{v}) = 0$. Then, as $\operatorname{supp}\Lambda$ is separable, Yokota's theorem [Yo, Theorems A, 27] can be applied and shows that $\operatorname{supp}\Lambda$ is contained in a subset which is isometric to a Hilbert space. $\square$

Corollary 4.12 *Suppose that, at a point $z \in X$, no pair of directions $\gamma, \eta \in \Sigma_z$ satisfies $\angle_z(p,q) = \pi$. Then, for $\mu \in \mathcal{P}_2(X)$ such that $(\log_z)_\sharp\mu$ has separable support for some logarithmic map $\log_z : X \longrightarrow C_z$, z is a barycenter of μ if and only if $\mu = \delta_z$.*

The assumption of separability in Theorem 4.11 holds true if μ has finite support or if C_z itself is separable (e.g., if (X,d) is finite dimensional). The author does not know if the separability of C_z generally follows from the properness of X.

In the finite dimensional case, the existence of a flat subset $\mathcal{H} \subset C_z$ as in Theorem 4.11 induces the isometric splitting $C_z = \mathcal{Y} \times \mathcal{H}$, where $\mathcal{Y}$ is an Alexandrov space of nonnegative curvature. The splitting theorem is also known in infinite dimensional Alexandrov spaces of nonnegative curvature ([Mi, Theorem 1]). Lytchak [Ly, Remark 5.6] claims the splitting of tangent cones of possibly infinite dimensional Alexandrov spaces, but then $\mathcal{Y}$ is not necessarily an Alexandrov space.

5 In nonnegatively or positively curved spaces

In this last section, we consider a proper Alexandrov space (X,d) of curvature ≥ 0 or ≥ 1 where we can simplify or improve some of our results in the previous sections.

We first observe that the uniqueness of a barycenter as in Lemma 4.3 can be derived in a more direct, quantitative way. To see this, in a proper Alexandrov space (X,d) of nonnegative curvature, take a barycenter z of $\mu \in \mathcal{P}_2(X)$ and $\Pi \in \mathcal{P}(\Gamma(X))$ with $(e_0)_\sharp\Pi = \delta_z$ and $(e_1)_\sharp\Pi = \mu$. We put $\mu_t := (e_t)_\sharp\Pi$ and observe that (2.2) improves (3.2) into

$$d\big(w, \gamma(t)\big)^2 \geq (1-t)d(w,z)^2 + td\big(w, \gamma(1)\big)^2 - (1-t)td\big(z, \gamma(1)\big)^2 \tag{5.1}$$

for any $w \in X$. As z is a barycenter of μ, the discussion as in Lemma 3.4 gives

$$\int_X d(w,y)^2 \, d\mu_t(y) \geq (1-t)d(z,w)^2 + \int_X d(z,y)^2 \, d\mu_t(y).$$

Hence z is a unique barycenter of μ_t.

Remark 5.1 The above proof also works when we weaken the inequality (5.1) to

$$d\big(w, \gamma(t)\big)^2 \geq \frac{1-t}{C^2} d(w,z)^2 + td\big(w, \gamma(1)\big)^2 - (1-t)td\big(z, \gamma(1)\big)^2, \tag{5.2}$$

where $C \geq 1$ is a fixed constant. This condition is regarded as a generalization of the 2-*uniform convexity* in Banach space theory, see [Oh1, Section 5], [Oh3] and the references therein for more discussion. The 2-uniform convexity (5.2) implies the non-branching property, so that the argument as in Lemma 4.3 is also applicable. To see the non-branching property, take two minimal geodesics $\gamma, \eta : [0,1] \longrightarrow X$ with $\gamma(1) = \eta(1)$ and

$\gamma(t) = \eta(t)$ for some $t \in (0,1)$. Then (5.2) implies

$$
\begin{aligned}
d\big(\eta(0), \gamma(t)\big)^2 &\geq \frac{1-t}{C^2} d\big(\eta(0), \gamma(0)\big)^2 + td\big(\eta(0), \gamma(1)\big)^2 - (1-t)td\big(\gamma(0), \gamma(1)\big)^2 \\
&= \frac{1-t}{C^2} d\big(\eta(0), \gamma(0)\big)^2 + t^2 d\big(\eta(0), \eta(1)\big)^2.
\end{aligned}
$$

As $d(\eta(0), \gamma(t)) = td(\eta(0), \eta(1))$, we have $\eta(0) = \gamma(0)$.

By a similar discussion to Theorem 4.8 and Corollary 4.9, we obtain the following.

Theorem 5.2 *Let (X, d) be a proper Alexandrov space, and let $z, w \in X$ be barycenters of $\mu, \nu \in \mathcal{P}_2(X)$, respectively.*

(i) *If (X, d) is of nonnegative curvature, then we have*

$$
\int_{X \times X} d(x,y)^2 \, d\mu(x) d\nu(y) \leq d(z,w)^2 + \int_X d(z,x)^2 \, d\mu(x) + \int_X d(w,y)^2 \, d\nu(y).
$$

(ii) *If (X, d) is of curvature ≥ 1, then we have*

$$
\begin{aligned}
&\int_{X \times X} \cos d(x,y) \frac{d(z,x)}{\sin d(z,x)} \frac{d(w,y)}{\sin d(w,y)} \, d\mu(x) d\nu(y) \\
&\geq \cos d(z,w) \int_X \cosh d(z,x) \frac{d(z,x)}{\sin d(z,x)} \, d\mu(x) \int_X \cos d(w,y) \frac{d(w,y)}{\sin d(w,y)} \, d\nu(y).
\end{aligned}
$$

The special case $\mu = \nu$ of Theorem 5.2(i) reduces to

$$
\int_{X \times X} d(x,y)^2 \, d\mu(x) d\mu(y) \leq 2 \int_X d(w,x)^2 \, d\mu(x) \tag{5.3}
$$

for all $w \in X$, without referring the barycenter. This is the global version of Lang and Schroeder's inequality (Lemma 4.5) used by Sturm [St1, Theorem 1.4, Proposition 1.7] to characterize Alexandrov spaces of nonnegative curvature among geodesic spaces. What is remarkable here is that (5.3) makes sense even in discrete spaces. See also [OP, Theorem 2.5] for another characterization by means of Ball's Markov type.

5.1 Barycenters in CAT(0)-spaces as a counterpoint

We close the article with a short review on rather well investigated barycenters in non-positively curved spaces which make an interesting contrast with our results. We refer to [Jo] and [St2] for more details.

A geodesic space (X, d) is called a CAT(0)-*space* if the reverse inequality of (2.2) holds, i.e., if

$$
d\big(z, \gamma(t)\big)^2 \leq (1-t)d(z,x)^2 + td(z,y)^2 - (1-t)td(x,y)^2
$$

holds for any three points $x, y, z \in X$ and any minimal geodesic $\gamma : [0,1] \longrightarrow X$ from x to y. In a complete CAT(0)-space, it is easy to see that every $\mu \in \mathcal{P}_2(X)$ admits a unique barycenter $z_\mu \in X$. Then the interesting fact is that the map $b : \mathcal{P}_2(X) \ni \mu \longmapsto z_\mu \in X$ is 1-Lipschitz with respect to the Wasserstein distance d_2^W.

The analogue of Lemma 4.3 is clear by Lemma 3.4 and the uniqueness of barycenters. It is well known that the *variance inequality*

$$\int_X d(w,x)^2 \, d\mu(x) \geq d(w,z_\mu)^2 + \int_X d(z_\mu,x)^2 \, d\mu(x) \tag{5.4}$$

holds for all $\mu \in \mathcal{P}_2(X)$ and $w \in X$. This implies the analogue of Lemma 4.4 (compare (5.4) with (4.1)) as well as the reverse inequality of Theorem 5.2(i)

$$\int_{X\times X} d(x,y)^2 \, d\mu(x)d\nu(y) \geq d(z_\mu,z_\nu)^2 + \int_X d(z_\mu,x)^2 \, d\mu(x) + \int_X d(z_\nu,y)^2 \, d\nu(y)$$

for $\mu,\nu \in \mathcal{P}_2(X)$.

Besides Proposition 3.6, we can construct a simpler example where the property $(*)$ is false. Consider two lines $L_0 = L_1 = \mathbb{R}$ and a segment $[0,1]$, and connect $0 \in [0,1]$ to $0 \in L_0$, and $1 \in [0,1]$ to $0 \in L_1$ as well. The resulting H-shaped space is a CAT(0)-space when we equip it with the length distance. Put $x = 1 \in L_0$, $y_0 = 1 \in L_1$, $y_1 = -1 \in L_1$, and set $\mu_0 := (\delta_x + \delta_{y_0})/2$ and $\mu_1 := (\delta_x + \delta_{y_1})/2$. Then the minimal geodesic α from μ_0 to μ_1 is given by $\alpha(t) = (\delta_x + \delta_{y_t})/2$ with $y_t = 1 - 2t \in L_1$. The unique barycenter of $\alpha(t)$ is $z_t := |1-2t|/2 \in [0,1]$, so that $z_0 = z_1 \neq z_t$ for $t \in (0,1)$.

As for Corollary 4.10, we see a completely different phenomenon in CAT(0)-spaces. It is easy to construct rays having angle π. For instance, let X be the four-pod, $X = \bigcup_{i=1}^4 \ell_i/\sim$ bound at $0 \in \ell_i = [0,\infty)$. Put $x_i = 1 \in \ell_i$ and consider $\mu = (\delta_{x_1} + \delta_{x_2})/2$ and $\nu = (\delta_{x_3} + \delta_{x_4})/2$. Then 0 is the common unique barycenter of μ and ν, and the pair of geodesics from δ_0 to μ and ν has angle π.

In the same four-pod X, let $\omega := \sum_{i=1}^4 \delta_{x_i}/4$. Then 0 is again the unique barycenter of ω, but the analogue of Theorem 4.11 is clearly false.

References

[ABN] A. D. Aleksandrov, V. N. Berestovskii and I. G. Nikolaev, Generalized Riemannian spaces (Russian), *Uspekhi Mat. Nauk* **41** (1986), 3–44, 240; English translation: *Russian Math. Surveys* **41** (1986), 1–54.

[BBI] D. Burago, Yu. Burago and S. Ivanov, *A course in metric geometry*, American Mathematical Society, Providence, RI, 2001.

[BGP] Yu. Burago, M. Gromov and G. Perel'man, A. D. Alexandrov spaces with curvatures bounded below (Russian), *Uspekhi Mat. Nauk* **47** (1992), 3–51, 222; English translation: *Russian Math. Surveys* **47** (1992), 1–58.

[CG] E. A. Carlen and W. Gangbo, Constrained steepest descent in the 2-Wasserstein metric, *Ann. of Math.* (2) **157** (2003), 807–846.

[Ha] S. Halbeisen, On tangent cones of Alexandrov spaces with curvature bounded below, *Manuscripta Math.* **103** (2000), 169–182.

[Jo] J. Jost, *Nonpositive curvature: geometric and analytic aspects*, Birkhäuser Verlag, Basel, 1997.

[KR] K. Kuratowski and C. Ryll-Nardzewski, A general theorem on selectors, *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **13** (1965), 397–403.

[LS] U. Lang and V. Schroeder, Kirszbrauns theorem and metric spaces of bounded curvature, *Geom. Funct. Anal.* **7** (1997), 535–560.

[LV] J. Lott and C. Villani, Ricci curvature for metric-measure spaces via optimal transport, *Ann. of Math.* **169** (2009), 903–991.

[Ly] A. Lytchak, Differentiation in metric spaces, *St. Petersburg Math. J.* **16** (2005), 1017–1041.

[Mi] A. Mitsuishi, A splitting theorem for infinite dimensional Alexandrov spaces with nonnegative curvature and its applications, *Geom. Dedicata* **144** (2010), 101–114.

[Oh1] S. Ohta, Markov type of Alexandrov spaces of nonnegative curvature, *Mathematika* **55** (2009), 177–189.

[Oh2] S. Ohta, Gradient flows on Wasserstein spaces over compact Alexandrov spaces, *Amer. J. Math.* **131** (2009), 475–516.

[Oh3] S. Ohta, Extending Lipschitz and Hölder maps between metric spaces, *Positivity* **13** (2009), 407–425.

[OP] S. Ohta and M. Pichot, A note on Markov type constants, *Arch. Math.* (*Basel*) **92** (2009), 80–88.

[OS] Y. Otsu and T. Shioya, The Riemannian structure of Alexandrov spaces, *J. Differential Geom.* **39** (1994), 629–658.

[St1] K.-T. Sturm, Metric spaces of lower bounded curvature, *Exposition. Math.* **17** (1999), 35–47.

[St2] K.-T. Sturm, Probability measures on metric spaces of nonpositive curvature, In: *Heat kernels and analysis on manifolds, graphs, and metric spaces* (Paris, 2002), 357–390, Contemp. Math. **338**, Amer. Math. Soc., Providence, RI, 2003.

[St3] K.-T. Sturm, A semigroup approach to harmonic maps, *Potential Anal.* **23** (2005), 225–277.

[St4] K.-T. Sturm, On the geometry of metric measure spaces. I, *Acta Math.* **196** (2006), 65–131.

[TY] A. Takatsu and T. Yokota, Cone structure of L^2-Wasserstein spaces, Preprint (2008). Available at http://arxiv.org/abs/0812.2752

[Vi] C. Villani, *Optimal transport, old and new*, Springer-Verlag, Berlin, 2009.

[Yo] T. Yokota, A rigidity theorem in Alexandrov spaces with lower curvature bound, to appear in *Math. Ann.*