J. Math. Soc. Japan
Vol. 68, No. 3 (2016) pp. 1297–1323
doi: 10.2969/jmsj/06831297

# Convex functions and barycenter on CAT(1)-spaces of small radii

By Takumi YOKOTA

(Received Jan. 1, 2015)

**Abstract.** We use the convexity of a certain function discovered by W. Kendall on small metric balls in CAT(1)-spaces to show that any probability measure on a complete CAT(1)-space of small radius admits a unique barycenter. We also present various properties of barycenter on those spaces. This extends the results previously known for CAT(0)-spaces and CAT(1)-spaces of small diameter.

## 1. Introduction.

In this paper, we are concerned with local geometry of CAT($\kappa$)-spaces particularly with a positive number $\kappa > 0$. They are metric spaces with $\kappa \in \mathbb{R}$ as an upper bound for their curvature in the sense of Alexandrov which is defined in terms of geodesic triangles. The precise definition is given in Definition 2 below.

Let $(X, d)$ be a metric space. By a *geodesic* in it, we mean a curve $\gamma : I \to X$ of constant speed defined on an interval $I \subset \mathbb{R}$ which realizes the distance between points on its image, i.e., there is a constant $|\gamma'| \geq 0$ with $d(\gamma(s), \gamma(t)) = |\gamma'| \cdot |s - t|$ for any $s, t \in I$. We say that a function $f : X \to \mathbb{R} \cup \{\infty\}$ is *convex* when the function $f(\gamma(\,\cdot\,))$ is convex on $I$ for any geodesic $\gamma : I \to X$. When $X$ is a product of two metric spaces $Y_1$ and $Y_2$ equipped with a natural product metric, this amounts to that $f(\gamma_1(\,\cdot\,), \gamma_2(\,\cdot\,))$ is convex on $I$ for any pair of geodesics $\gamma_i : I \to Y_i$, $i = 1, 2$.

It is well-known that the distance function $d : Y \times Y \to [0, \infty)$ of a CAT(0)-space $(Y, d)$ is convex. The following theorem is the main tool that we avail ourselves of in our approach, which states that any small ball in a CAT($\kappa$)-space with $\kappa > 0$ also admits such a convex function. Here and hereafter, $B(o, \cdot\,)$ and $\overline{B}(o, \cdot\,)$ denote open and closed metric balls centered at $o \in Y$ respectively. We also use $R_\kappa := \pi/\sqrt{\kappa}$ and $\cos_\kappa r := \cos(\sqrt{\kappa} \cdot r)$ for $\kappa > 0$ and $r > 0$.

THEOREM A. *Let $(Y, d)$ be a CAT($\kappa$)-space with $\kappa > 0$ and $r < R_\kappa/2$. For any $h > \tilde{h} > 0$ with $h \leq \cos_\kappa r$, $\nu \in \mathbb{R}$ and $o \in Y$, the function $\Phi^{(\kappa)}_{\nu,\tilde{h}} : B(o, r) \times B(o, r) \to [0, \infty)$ given by*

$$(x, y) \longmapsto \left( \frac{1}{\kappa} \cdot \frac{1 - \cos_\kappa d(x, y)}{\cos_\kappa d(x, o) \cos_\kappa d(y, o) - \tilde{h}^2} \right)^{\nu+1}$$

*is convex, provided that $2(2\nu + 1)\tilde{h}^2(h^2 - \tilde{h}^2) \geq 1$.*

---

2010 *Mathematics Subject Classification.* Primary 53C23.
*Key Words and Phrases.* CAT(1)-space, convex function, barycenter, Banach–Saks property.

In the terminology of Kendall [**23**], Theorem A says that the ball $B(o,r)$ in its statement has *convex geometry* and it extends his result in [**23**] to general CAT($\kappa$)-spaces. The function in Theorem A appeared in [**23**] after a similar function was used by Jäger–Kaul [**14**] in their proof of a uniqueness theorem for harmonic maps. Kendall [**23**] proved Theorem A for the unit sphere of the Euclidean space; see Theorem 34 below. He also remarked that Theorem A holds for a *regular geodesic ball* [**22**, Definition 1.6] with a positive upper bound for the sectional curvature in a complete Riemannian manifold. A regular geodesic ball is characterized as a ball which itself becomes a CAT($\kappa$)-space with some $\kappa > 0$, e.g. Kuwae [**25**]. Since a proof of Theorem A is not found in the literature, we decided to give a detailed proof in the appendix.

We also add that convex functions have been playing key roles in the theory of harmonic maps. Among others, Ishihara [**13**] gave a characterization of harmonic maps between Riemannian manifolds by means of convex functions and Jost–Xin–Yang [**18**] proved a Liouville type theorem for harmonic maps into a certain subset of the Euclidean sphere by constructing convex functions. Convex functions are also behind the Liouville type theorem for harmonic maps into singular spaces formulated by Kuwae–Sturm [**27**].

Next, we proceed to the second topic in the title of the paper.

Definition 1 (Barycenter). For a metric space $(X,d)$ and $p \in [1,\infty)$, we let $\mathcal{P}_p(X)$ be the set of all Borel probability measures $\mu$ on $X$ with $\int_X d^p(\,\cdot\,,x_0)\,d\mu < \infty$ for some (hence all) $x_0 \in X$. For a probability measure $\mu \in \mathcal{P}_2(X)$, we consider the function $F_\mu : X \to [0,\infty)$ given by $F_\mu(x) := (1/2)\int_X d^2(\,\cdot\,,x)\,d\mu$. We call a point of $X$ where $F_\mu$ attains its global minimum (resp. local minimum) a *barycenter* (resp. a *Karcher mean*) of $\mu$.

Barycenter of probability measures is also referred to as *center of mass* or *Fréchet mean* in the literature. If $\mu$ is in $\mathcal{P}_1(X)$, we consider the function $x \longmapsto \int_X d^2(\,\cdot\,,x) - d^2(\,\cdot\,,x_0)\,d\mu$ with $x_0 \in X$ being fixed instead to define a barycenter of $\mu$. The theory of barycenter of probability measures on CAT(0)-spaces has been developed by many authors; see e.g. Sturm [**34**].

We now state a main theorem of this paper. We say that a measure $\mu$ on a space $X$ is *concentrated* on a subset $S \subset X$ if $\mu(X \setminus S) = 0$. The *radius* of a metric space $(X,d)$ is defined as $\mathrm{rad}(X) := \inf_{x\in X}\sup_{y\in X} d(x,y)$.

Theorem B. *Let $(Y,d)$ be a complete* CAT($\kappa$)*-space with $\kappa > 0$. Suppose that $\mu \in \mathcal{P}_2(Y)$ is concentrated on a ball $B(o,r)$ with $o \in Y$ and $r < R_\kappa/2$. Then $\mu$ admits a barycenter $b(\mu) \in B(o,r)$ and it is unique in $Y$. In particular, if the radius of $(Y,d)$ is less than $R_\kappa/2$, any $\mu \in \mathcal{P}_2(Y)$ admits a unique barycenter in $Y$.*

The condition on the radius that $r < R_\kappa/2$ is sharp. We prove Theorem B by combining Theorem A and the Ekeland principle (Lemma 11). We expect further applications of our Theorem A in the geometry of CAT($\kappa$)-spaces, cf. Jost [**17**].

In addition to the theorems stated above, we will also obtain Banach–Saks–Kakutani type theorems for CAT($\kappa$)-spaces in Theorems C and D and prove existence of minimizers of some convex functions on CAT($\kappa$)-spaces with $\kappa > 0$ in Theorem E in the subsequent sections. They extend the theorems of Jost [**15**], [**16**] proved for CAT(0)-spaces to

CAT($\kappa$)-spaces.

The organization of this paper is as follows: After some preparation in the next section, we prove Theorem B in Section 3. Then Sections 4 and 5 are devoted to a collection of some properties of barycenter of probability measures on CAT($\kappa$)-spaces. Among them is Jensen's inequality for convex functions (Theorem 25), cf. Kuwae [**25**]. The proof of Theorem A is given in the appendix.

## 2. Preliminaries.

In this section, we recall some rudimentary definitions and facts on the geometry of CAT($\kappa$)-spaces. The textbook [**6**] by Burago–Burago–Ivanov is one of the standard references of the Alexandrov geometry.

We begin with the definition of CAT($\kappa$)-spaces. For any real number $\kappa \in \mathbb{R}$, let $(M_\kappa, d_\kappa)$ be the model surface, i.e., the simply-connected surface with the distance induced by the complete Riemannian metric of constant curvature $\kappa$. We will also use $(S^2, \overline{d})$ instead of $(M_1, d_1)$. We let $R_\kappa := \pi/\sqrt{\kappa}$ for $\kappa > 0$ and $R_\kappa := +\infty$ for $\kappa \leq 0$.

Definition 2 (CAT($\kappa$)-space). We call a metric space $(Y, d)$ a CAT($\kappa$)-*space* if it is an $R_\kappa$-geodesic space, i.e., any two points $x, y \in Y$ with $d(x,y) < R_\kappa$ are connected by a geodesic, and it satisfies the following: For any three points $x, y, z \in Y$ with $d(x,y) + d(y,z) + d(z,x) < 2R_\kappa$ and a geodesic $\gamma : [0,1] \to Y$ with $\gamma(0) = y$ and $\gamma(1) = z$,

$$d(\gamma(t), x) \leq d_\kappa(\bar{\gamma}(t), \bar{x}) \text{ for any } t \in [0,1].$$

Here, $\{\bar{x}, \bar{y}, \bar{z}\} \subset (M_\kappa, d_\kappa)$ is an isometric copy of the three-point subset $\{x, y, z\} \subset (Y, d)$ and $\bar{\gamma} : [0,1] \to M_\kappa$ is the geodesic with $\bar{\gamma}(0) = \bar{y}$ and $\bar{\gamma}(1) = \bar{z}$.

In this paper, we persist in using the letter $Y$ to denote a CAT($\kappa$)-space. Complete Riemannian manifolds with sectional curvature at most $\kappa$ and injectivity radius not less than $R_\kappa$ are typical examples of CAT($\kappa$)-spaces. The upper curvature bound $\kappa \in \mathbb{R}$ of a CAT($\kappa$)-space changes accordingly as its distance is rescaled by a positive number and a CAT($\kappa$)-space is also a CAT($\kappa'$)-space for $\kappa' > \kappa$.

Our main interest is in CAT($\kappa$)-spaces with radius $< R_\kappa/2$ for some $\kappa > 0$. A metric ball of radius $< R_\kappa/2$ in a CAT($\kappa$)-space is a simple example. One of the theorems of Fujiwara–Nagano–Shioya [**11**, Theorem 1.7], cf. Balser–Lytchak [**4**, Proposition 1.2], also provides such spaces.

Here we collect some notations used throughout this paper without giving the precise definitions:

- For $\kappa > 0$ and $r \in \mathbb{R}$, $\cos_\kappa r := \cos(\sqrt{\kappa} \cdot r)$ and $\sin_\kappa r := \sin(\sqrt{\kappa} \cdot r)/\sqrt{\kappa}$.
- In a CAT($\kappa$)-space $(Y, d)$, a geodesic connecting two points $x, y \in Y$ with $d(x,y) < R_\kappa$ is unique up to parameterization. We denote by $\gamma_{xy} : [0,1] \to Y$ the geodesic with $\gamma_{xy}(0) = x$ and $\gamma_{xy}(1) = y$.
- $\tilde{\angle}_\kappa(x; y, z) \in [0, \pi]$ denotes the *comparison angle* for three points $x, y, z$ in $(Y, d)$. For example, it is defined for $\kappa > 0$ by

$$\cos \tilde{\angle}_\kappa(x;y,z) := \frac{\cos_\kappa d(y,z) - \cos_\kappa d(x,y)\cos_\kappa d(x,z)}{\kappa \cdot \sin_\kappa d(x,y)\sin_\kappa d(x,z)}$$

if $y,z \neq x$ and $d(x,y)+d(y,z)+d(z,x) < 2R_\kappa$.

- $(\Sigma_x, \angle_x)$ and $(C_x, |\cdot|)$ denote the *space of directions* and the *tangent cone* at a point $x \in Y$ respectively with $o_x \in C_x := \Sigma_x \times [0,\infty)/\Sigma_x \times \{0\}$ being the *vertex*.
- $\uparrow_x^y \in \Sigma_x$ with $x \neq y \in Y$ denotes the equivalence class of any geodesic from $x$ to $y$, and $\angle_x(y,z) := \angle_x(\uparrow_x^y, \uparrow_x^z) \in [0,\pi]$ denotes the *angle*.
- For $x,y \in Y$ with $x \neq y$, we set $\log_x y := d(x,y) \cdot \uparrow_x^y \in C_x$ and $\log_x x := o_x \in C_x$. We also use $\angle_x(u,y) := \angle_x(u, \uparrow_x^y)$ for $u \in C_x \setminus \{o_x\}$.
- $|u| := |u - o_x|$ and $\langle u,v\rangle := (|u|^2 + |v|^2 - |u-v|^2)/2$ for two vectors $u,v \in C_x$ at $x \in Y$. If $u,v \neq o_x$, $\langle u,v\rangle = |u||v|\cos\angle_x(u,v)$.
- For a geodesic $\gamma : I \to Y$ and $t_0 \in I \subset \mathbb{R}$, $\gamma'(t_0+)$ and $\gamma'(t_0-)$ denote the equivalence classes in the tangent cone $C_{\gamma(t_0)}$ at the point $\gamma(t_0) \in Y$ represented by the geodesics $\gamma_\pm : [0,\varepsilon) \to Y$ given by $\gamma_\pm(t) := \gamma(t_0 \pm t)$ respectively with small $\varepsilon > 0$.
- For a function $f : I \to \mathbb{R}$ and $t_0 \in I \subset \mathbb{R}$ with $t_0 \neq \sup I$,

$$f'(t_0+) := \lim_{h \to 0+} \frac{f(t_0+h) - f(t_0)}{h}.$$

We list some basic facts on CAT($\kappa$)-spaces which we will make use of.

FACT 3 (Angle monotonicity/comparison). *For any three points $x,y,z$ in a CAT($\kappa$)-space with $y,z \neq x$ and $d(x,y)+d(y,z)+d(z,x) < 2R_\kappa$ and a point $y' := \gamma_{xy}(t)$ for some $t \in (0,1)$,*

$$\tilde{\angle}_\kappa(x;y,z) \geq \tilde{\angle}_\kappa(x;y',z) \geq \angle_x(y,z).$$

FACT 4 (Local uniform convexity). *For any $\kappa, r, \varepsilon > 0$ with $r < R_\kappa/2$, there is $\delta_\kappa(\varepsilon; r) > 0$ such that*

$$d(m(x,y),o) \leq r - \delta_\kappa(\varepsilon;r)$$

*for any $x,y \in \overline{B}(o,r)$ with $d(x,y) \geq \varepsilon r$ in a CAT($\kappa$)-space $(Y,d)$ with $o \in Y$. Here $m(x,y) := \gamma_{xy}(1/2) \in Y$ is the midpoint of $x$ and $y$.*

The function

$$\delta_\kappa(\varepsilon;r) := \inf\left\{r - d(m(x,y),o) : x,y \in \overline{B}(o,r) \subset Y \text{ with } d(x,y) \geq \varepsilon r\right\}$$

resembles what is called the *modulus of convexity* in the theory of Banach spaces, cf. Gelander–Karlsson–Margulis [**12**].

Although it will not be required later, Ohta's lemma [**30**, Lemma 3.1] gives an explicit estimate for $\delta_\kappa(\varepsilon;r) > 0$ in Fact 4. He states his lemma for CAT(1)-spaces of diameter $< \pi/2$, but his proof actually shows the following.

LEMMA 5 ($k$-convexity of CAT($\kappa$)-spaces, Ohta [**30**]). *Any geodesic $\gamma_{xy} : [0,1] \to Y$ connecting points $x, y \in \overline{B}(o, (R_\kappa - \varepsilon)/2)$ in a* CAT($\kappa$)*-space $(Y, d)$ with $\kappa, \varepsilon > 0$ and $o \in Y$ satisfies*

$$d^2(\gamma_{xy}(t), o) \le (1-t)d^2(x, o) + td^2(y, o) - \frac{k}{2}t(1-t)d^2(x, y) \tag{6}$$

*for any $t \in [0,1]$ with $k := (R_\kappa - \varepsilon)\tan(\varepsilon/2) > 0$.*

The following fact is needed in our proof of Theorem A.

FACT 7 (First variation formula, cf. [**6**, Exercise 4.5.10]). *Suppose that $\lambda, \mu : I \to Y$ are two geodesics in a* CAT($\kappa$)*-space $(Y, d)$ with $x := \lambda(t_0)$, $y := \mu(t_0)$ and $d(x, y) < R_\kappa$ for some $t_0 \in I$ with $t_0 \ne \sup I$. Then the function $d(t) := d(\lambda(t), \mu(t))$ satisfies*

$$d(t_0+) = -\langle \lambda'(t_0+), \uparrow_x^y \rangle - \langle \mu'(t_0+), \uparrow_y^x \rangle.$$

We close this section with a simple consequence of the combination of some of the above facts. For $\kappa \in \mathbb{R}$, we say that a subset $C \subset X$ of a metric space $(X, d)$ is *$R_\kappa$-convex* if any geodesic connecting points $x, y \in C$ with $d(x, y) < R_\kappa$ does not leave $C \subset X$.

FACT 8 (Chebyshev property of convex subsets). *For a closed $R_\kappa$-convex subset $C \subset Y$ and a point $p \in Y$ of a complete* CAT($\kappa$)*-space $(Y, d)$ with $d(p, C) < R_\kappa/2$, there exists a unique point $\pi_C(p) \in C$ with $d(p, \pi_C(p)) = d(p, C)$. It also holds that $\tilde{\angle}_\kappa(\pi_C(p); p, c) \ge \angle_{\pi_C(p)}(p, c) \ge \pi/2$ for any $c \in C$ if they are defined.*

## 3. Proof of Theorem B.

In this section, we present a proof of Theorem B stated in the introduction after some comment and preparation.

Theorem B is well-known for CAT(0)-spaces, e.g. Sturm [**34**], and is also known for CAT(1)-spaces of diameter $< \pi/2$, e.g. Kuwae [**25**]. The proofs for these spaces rely on the *$k$-convexity* of them as in Inequality (6) with $k = 2$ for CAT(0)-spaces and $k = (\pi - 2\varepsilon)\tan\varepsilon > 0$ for CAT(1)-spaces of diameter $\le \pi/2 - \varepsilon$, cf. Karcher [**20**, Theorem 1.2]. The notion of $k$-convexity for metric spaces was introduced and studied by Ohta [**30**].

We here recall some definitions. For a function $\varphi$ defined on a neighborhood of a point $x \in Y$ of a CAT($\kappa$)-space $(Y, d)$, we define its *directional derivative* by $D\varphi[\log_x y] := \varphi \circ \gamma_{xy}(0+) \in \mathbb{R} \cup \{\pm\infty\}$, if exists, for $y \in Y$ with $0 < d(x, y) < R_\kappa$. If $\varphi$ is locally Lipschitz at $x$, we extend it to a Lipschitz function $D\varphi$ on the tangent cone $C_x$ at $x$.

The following lemma is well-known, e.g. Lytchak [**28**, Lemma 7.3].

LEMMA 9. *Let $\varphi$ be a convex function defined on a neighborhood of a point $x$ of a* CAT($\kappa$)*-space $(Y, d)$ with $\kappa \in \mathbb{R}$. Suppose that $D\varphi[\xi] < 0$ for some $\xi \in \Sigma_x$ and $\varphi$ is locally Lipschitz at $x$. Then there exists a unique vector $\xi_x \in \Sigma_x$ such that*

$$D\varphi[\eta] \ge D\varphi[\xi_x]\langle \eta, \xi_x \rangle \text{ for any } \eta \in \Sigma_x.$$

In the above lemma, $\xi_x \in \Sigma_x$ is the point where the function $D\varphi$ restricted on $\Sigma_x$ attains its minimum. We put $\nabla_x^- \varphi := (-D\varphi[\xi_x]) \cdot \xi_x \in C_x$ in the situation of Lemma 9 or $\nabla_x^- \varphi := o_x \in C_x$ when $D\varphi[\xi] \geq 0$ for any $\xi \in C_x$, and we call the vector $\nabla_x^- \varphi$ the (*negative*) *gradient* of $\varphi$ at $x$.

Jensen's inequality is one of the properties that we expect for barycenter or Karcher mean of probability measures. We here give a quick proof of it, because it has something in common with our proof of Theorem B. Another version of Jensen's inequality is proved in Theorem 25 in the subsequent section.

Proposition 10 (Jensen's inequality, cf. Kendall [**22**, Lemma 7.2]). *Let $(Y, d)$ be a complete $\mathrm{CAT}(\kappa)$-space with $\kappa \in \mathbb{R}$ and $\varphi : Y \to \mathbb{R} \cup \{\infty\}$ be a lower-semicontinuous convex function. Suppose that $\mu \in \mathcal{P}_2(Y)$ has its Karcher mean $b(\mu) \in Y$ and $\mu$ is concentrated on $B(b(\mu), R_\kappa) \subset Y$. Then*

$$\varphi(b(\mu)) \leq \int_Y \varphi \, d\mu$$

*if $\varphi$ is locally Lipschitz at $b(\mu)$.*

Proof. Since our assumption implies that $\varphi \equiv \infty$ and nothing remains to prove if $\varphi(b(\mu)) = \infty$, we may assume that $\varphi(b(\mu)) < \infty$ in the proof.

For any $y \in B(b(\mu), R_\kappa)$, the function $\varphi \circ \gamma_{b(\mu)y}$ is convex on $[0, 1]$, and we have

$$\begin{aligned} \varphi(y) - \varphi(b(\mu)) &\geq D\varphi\big[\log_{b(\mu)} y\big] \\ &\geq -\big\langle \nabla_{b(\mu)}^- \varphi, \log_{b(\mu)} y \big\rangle. \end{aligned}$$

We finish the proof by integrating this inequality. Indeed, it follows from Fact 7 that the directional derivative $DF$ of the function $F := F_\mu$ at a point $x \in Y$ is given by

$$DF[\xi] = -\int_Y \big\langle \xi, \log_x y \big\rangle d\mu(y) \text{ for } \xi \in C_x,$$

and we know that $DF[\xi] \geq 0$ for any $\xi \in C_{b(\mu)}$ at the Karcher mean $b(\mu) \in Y$. $\square$

We also invoke the following lemma.

Lemma 11 (Ekeland principle, e.g. Ekeland [**8**]). *Let $f : X \to \mathbb{R}$ be a lower-semicontinuous function on a complete metric space $(X, d)$ with $\inf_X f > -\infty$. For any point $x_0 \in X$ and $\varepsilon > 0$, we can find a point $x_\varepsilon \in X$ for which $d(x_\varepsilon, x_0) \leq (f(x_0) - \inf_X f)/\varepsilon$ and*

$$f(y) \geq f(x_\varepsilon) - \varepsilon \cdot d(y, x_\varepsilon) \textit{ for any } y \in X.$$

Now we are in a position to begin our proof of Theorem B. The following proof was inspired by that of Kendall [**22**, Theorem 7.3].

PROOF OF THEOREM B. Recall that $\mu \in \mathcal{P}_2(Y)$ is concentrated on $B(o,r) \subset Y$ for some $o \in Y$ and $r < R_\kappa/2$ and we would like to find a point where the function $F := F_\mu$ attain the minimum. We start with the following observations.

CLAIM 12. *For any $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that*

$$F(x) > \inf_{B(o,r)} F + \delta \text{ for any } x \in Y \setminus B(o, r+\varepsilon).$$

PROOF. For a point $x \in Y$ with $d(x,o) \geq 2r$, we have

$$F(x) > r^2/2 > F(o) \geq \inf_{B(o,r)} F.$$

For a point $x \in Y$ with $r + \varepsilon \leq d(x,o) < 2r$, we let $\gamma : [0, d(x,o)] \to Y$ be the unit speed geodesic from $o$ to $x$ and put $x' := \gamma(2r + \varepsilon - d(x,o)) \in B(o,r)$. We shall show that $d(x,y) > d(x',y) + \sqrt{2}\delta$ for any $y \in B(o,r)$, which implies that $F(x) > F(x') + \delta$, cf. Afsari [**1**]. To show this, we fix $y \in B(o,r)$ and set $x_0 := \gamma(r + (\varepsilon/2)) \in Y$. Note that $d(y,x_0), d(x,x_0) = d(x',x_0) > \varepsilon/2$.

If $d(x_0,x) + d(x,y) + d(y,x_0) \geq 2R_\kappa$, we note $4r > d(x_0,x') + d(x',y) + d(y,x_0)$ and get

$$d(x,y) \geq 2R_\kappa - d(x_0,x') - d(y,x_0) > d(x',y) + 2(R_\kappa - 2r).$$

If $d(x_0,x) + d(x,y) + d(y,x_0) < 2R_\kappa$, we have

$$\begin{aligned}
\tilde{\angle}_\kappa(x_0; x', y) &\leq \tilde{\angle}_\kappa(x_0; o, y) < (\pi/2) - \delta'; \\
\tilde{\angle}_\kappa(x_0; x, y) &\geq \angle_{x_0}(x,y) \\
&\geq \pi - \angle_{x_0}(x', y) \\
&\geq \pi - \tilde{\angle}_\kappa(x_0; o, y) > (\pi/2) + \delta'
\end{aligned}$$

for some $\delta' = \delta'(\varepsilon) > 0$. These inequalities confirm the claim. □

CLAIM 13. *There exist $r' \in (0,r)$ and $\delta > 0$ such that*

$$DF\left[\uparrow_x^o\right] < -\delta \text{ for any } x \in B(o, R_\kappa/2) \setminus B(o, r').$$

PROOF. We choose $r' \in (0,r)$ so that $r - r' > 0$ is small enough with

$$\begin{aligned}
\delta &:= -R_\kappa \mu(B(o,r) \setminus B(o,r')) \\
&\quad + \int_{B(o,r')} (r' - d(y,o))(\cos d(y,o) - \cos r')\, d\mu(y) > 0.
\end{aligned}$$

Then for any $x \in B(o, R_\kappa/2) \setminus B(o, r')$ we have

$$\begin{aligned}
-DF[\uparrow_x^o] &= \int_Y \langle \uparrow_x^o, \log_x y\rangle d\mu(y)\\
&\geq \int_{B(o,r)} d(x,y)\cos\tilde{\angle}_\kappa(x;y,o)\,d\mu(y) > \delta.
\end{aligned}$$

This confirms the claim. □

We continue our proof of Theorem B. For the function $F := F_\mu$ and each $\varepsilon > 0$, we appeal to Lemma 11 to find a point $x(\varepsilon) \in Y$ such that $\lim_{\varepsilon\to 0+} F(x(\varepsilon)) \to \inf_Y F$ and

$$F(y) \geq F(x(\varepsilon)) - \varepsilon \cdot d(y, x(\varepsilon)) \text{ for any } y \in Y.$$

By the choice of $x(\varepsilon)$,

$$DF[\xi] \geq -\varepsilon|\xi| \text{ for any } \xi \in C_{x(\varepsilon)}. \tag{14}$$

Combined with this, Claims 12 and 13 imply that $\limsup_{\varepsilon\to 0+} d(x(\varepsilon), o) < r' < r$.

According to Theorem A, the function $\Phi := \Phi^{(\kappa)}_{\nu,\tilde{h}} : B(o,r)\times B(o,r) \to [0,\infty)$ with appropriate $\tilde{h} < h := \cos_\kappa r$ and $\nu \in \mathbb{R}$ is convex. We then use Fact 7 to derive for any $y \in B(o,r)$ and $\varepsilon, \varepsilon' > 0$ that

$$\begin{aligned}
&\Phi(y,y) - \Phi(x(\varepsilon), x(\varepsilon'))\\
&\quad\geq D\Phi\big[\log_{(x(\varepsilon),x(\varepsilon'))}(y,y)\big]\\
&\quad= D\Phi(\,\cdot\,, x(\varepsilon'))\big[\log_{x(\varepsilon)} y\big] + D\Phi(x(\varepsilon),\,\cdot\,)\big[\log_{x(\varepsilon')} y\big]\\
&\quad\geq -\big\langle \nabla^-_{x(\varepsilon)}\Phi(\,\cdot\,, x(\varepsilon')), \log_{x(\varepsilon)} y\big\rangle - \big\langle \nabla^-_{x(\varepsilon')}\Phi(x(\varepsilon),\,\cdot\,), \log_{x(\varepsilon')} y\big\rangle.
\end{aligned}$$

Since $\Phi(y,y) = 0$, we integrate this inequality and use (14) to obtain that

$$\begin{aligned}
-\Phi(x(\varepsilon), x(\varepsilon')) &\geq DF\big[\nabla^-_{x(\varepsilon)}\Phi(\,\cdot\,, x(\varepsilon'))\big] + DF\big[\nabla^-_{x(\varepsilon')}\Phi(x(\varepsilon),\,\cdot\,)\big]\\
&\geq -\varepsilon\big|\nabla^-_{x(\varepsilon)}\Phi(\,\cdot\,, x(\varepsilon'))\big| - \varepsilon'\big|\nabla^-_{x(\varepsilon')}\Phi(x(\varepsilon),\,\cdot\,)\big|\\
&\geq -C(\varepsilon + \varepsilon')
\end{aligned}$$

with some constant $C < \infty$ depending only on $h$, $\tilde{h}$ and $\nu$. This says that $\Phi(x(\varepsilon), x(\varepsilon')) \to 0$ or equivalently $d(x(\varepsilon), x(\varepsilon')) \to 0$ as $\varepsilon, \varepsilon' \to 0+$. Therefore, a sequence $(x(\varepsilon_i))_{i\in\mathbb{N}}$ with $\varepsilon_i \to 0+$ as $i \to \infty$ is a Cauchy sequence in $Y$ with $\limsup_{i\to\infty} d(x(\varepsilon_i), o) < r$ and it converges to some point $b(\mu) \in B(o,r)$, which turns out to be a barycenter of $\mu$.

We notice that the above argument also establishes the uniqueness of barycenter of $\mu$ in $Y$. Now the proof of Theorem B is complete. □

We give a few easy corollaries of Theorem B. First of all, by inspecting our proof of Theorem B, we have the following characterisation of the barycenter.

Corollary 15. *In the situation of Theorem B, suppose a point $z \in B(o,r)$ satisfies that*

$$DF_\mu[\xi] \geq 0 \text{ for any } \xi \in C_z.$$

*Then $z$ is the unique barycenter $b(\mu)$ in $B(o,r)$.*

It is known that the barycenter of a probability measure $\mu \in \mathcal{P}_1(Y)$ on a complete CAT(0)-space $Y$ lies in the closed convex hull of a subset on which $\mu$ is concentrated, e.g. [**34**, Proposition 6.1]. For CAT($\kappa$)-spaces, we can prove something similar as well. In the statement below, $\overline{\mathrm{conv}}(S) \subset Y$ denotes the *closed convex hull* of a subset $S \subset Y$, i.e., the smallest closed $R_\kappa$-convex subset of $Y$ containing $S$.

Corollary 16. *In the situation of Theorem B, suppose that $\mu \in \mathcal{P}_2(Y)$ is concentrated on a subset $S$ of a ball $B(o,r) \subset Y$. Then $b(\mu) \in \overline{\mathrm{conv}}(S)$.*

Proof. We slightly modify our proof of Theorem B to know that there exists a point $\bar{b}(\mu) \in C := \overline{\mathrm{conv}}(S) \subset B := B(o,r)$ such that $F_\mu(x) \geq F_\mu(\bar{b}(\mu))$ for any $x \in C$.

If we suppose that $\bar{b}(\mu) \neq b(\mu)$, then it follows from Corollary 15 that there exist a geodesic $\gamma : [0,1] \to Y$ with $\gamma(0) = \bar{b}(\mu)$ and $\varepsilon > 0$ such that

$$F_\mu(\gamma(t)) < F_\mu(\bar{b}(\mu)) - 2\varepsilon t \text{ for any } t > 0.$$

By assumption, we have $\gamma(t) \notin C$ and

$$\tilde{\angle}_\kappa(\pi_C(\gamma(t)); \gamma(t), x) \geq \angle_{\pi_C(\gamma(t))}(\gamma(t), x) \geq \pi/2$$

for any $t > 0$ and $x \in C \setminus \{\pi_C(\gamma(t))\}$.

Since $\mathrm{diam}(S) < R_\kappa$, we notice that there exists $\delta > 0$ such that

$$d^2(\gamma(t), x) > d^2(\pi_C(\gamma(t)), x) - 2\varepsilon t$$

for any $x \in C$ and any $t \in (0, \delta)$. This implies that

$$F_\mu(\gamma(t)) > F_\mu(\pi_C(\gamma(t))) - \varepsilon t \geq F_\mu(\bar{b}(\mu)) - \varepsilon t$$

for any $t \in (0,\delta)$, which is a contradiction. We therefore conclude that $b(\mu) = \bar{b}(\mu) \in C$. □

Remark 17. In [**27**] and [**26**], a minimizer of the function $F_\mu$ restricted on the closed convex hull of the support of $\mu \in \mathcal{P}_1(X)$ is called a *pure barycenter* of $\mu$. The *support* of a measure $\mu$ on a metric space $X$ is defined as

$$\mathrm{supp}[\mu] := \{x \in X : \mu(B(x,r)) > 0 \text{ for any } r > 0\}.$$

On a complete separable metric space, $\mathrm{supp}[\mu]$ is the minimal closed subset on which $\mu$

is concentrated. Corollary 16 says that the barycenter and the pure barycenter coincide for $\mu \in \mathcal{P}_1(Y)$ as in the corollary on a complete separable CAT($\kappa$)-space $Y$ with $\kappa > 0$.

## 4. Properties of barycenters.

In this section, we collect some properties of Karcher mean or barycenter of probability measures on CAT($\kappa$)-spaces with $\kappa > 0$, which we proved to exist in Theorem B. We will utilize Theorem A along with the fact that the product of two CAT($\kappa$)-spaces equipped with the direct product metric is again a CAT($\kappa$)-space.

Some properties of barycenter of probability measures on CAT(0)-spaces are well-known, e.g. Sturm [**34**]. Our results in this section extend some of them to the context of CAT($\kappa$)-spaces. We do not attempt to exhaust such possible extensions. We here also add that Ohta [**31**] investigated properties of barycenter of probability measures on proper Alexandrov spaces of curvature $\geq \kappa$.

Throughout this section, we always assume that

- $(Y, d)$ stands for a complete CAT($\kappa$)-space with $\kappa > 0$;
- $\mu \in \mathcal{P}_2(Y)$ is concentrated on $B := B(o, r)$ with $o \in Y$ and $r < R_\kappa/2$ and has its barycenter $b(\mu) \in B$;
- $\Phi := \Phi^{(\kappa)}_{\nu,\tilde{h}} : B \times B \to [0, \infty)$ is the convex function in Theorem A with suitable parameters $\nu > -1/2$ and $\tilde{h} > 0$ with $\tilde{h} < h := \cos_\kappa r$.

Before we commence, we remark that a simple estimate says

$$\left(\frac{4d^2(x,y)}{\pi^2(1-\tilde{h}^2)}\right)^{\nu+1} \leq \Phi(x,y) \leq \left(\frac{d^2(x,y)}{2(h^2-\tilde{h}^2)}\right)^{\nu+1} \tag{18}$$

for any $x, y \in B(o, r)$.

### 4.1. Variance inequality.

Proposition 19 (Variance inequality, cf. [**34**, Proposition 4.4]). *For any $x \in \overline{B} := \overline{B}(o, r)$,*

$$\int_Y d^2(\,\cdot\,, x) - d^2(\,\cdot\,, b(\mu))\, d\mu \geq c \cdot d^\alpha(x, b(\mu))$$

*with some constants $c > 0$ and $\alpha > 2$ depending only on $\kappa$ and $r$.*

Proof. For any $x \in \overline{B}$ with $x \neq b(\mu)$, we apply Lemma 11 with $x_0 := x$ and $\varepsilon := 2(F(x) - F(b(\mu)))/d(x, b(\mu)) > 0$ to obtain a point $x_\varepsilon \in \overline{B}$ such that $d(x_\varepsilon, x) \leq d(x, b(\mu))/2$ and the function $F := F_\mu$ satisfies

$$F(y) \geq F(x_\varepsilon) - \varepsilon \cdot d(y, x_\varepsilon) \text{ for any } y \in \overline{B}.$$

Then we use the argument in the proof of Theorem B, with $\Phi$ being extended to $\overline{B} \times \overline{B}$ if necessary, to see that $\Phi(x_\varepsilon, b(\mu)) \leq C\varepsilon$ for some constant $C < \infty$ and hence by (18)

$$F(x) - F(b(\mu)) \geq \frac{d(x, b(\mu))}{2C} \cdot d^{2(\nu+1)}(x_\varepsilon, b(\mu)) \geq \frac{1}{C}\left(\frac{d(x, b(\mu))}{2}\right)^{2\nu+3}.$$

This proves the proposition with $\alpha := 2\nu + 3 > 2$. □

Next we seek a variance inequality which fits better to CAT($\kappa$)-spaces with $\kappa > 0$. We begin with a simple definition. For three points $x, y, z$ of a metric space $(X, d)$ and $\kappa \in \mathbb{R}$, we put

$$\langle \overrightarrow{xy}, \overrightarrow{xz} \rangle_\kappa := d(x, y)d(x, z) \cdot \cos \tilde{\angle}_\kappa(x; y, z).$$

To complement, we define $\langle \overrightarrow{xy}, \overrightarrow{xz} \rangle_\kappa := 0$ if $d(x, y)d(x, z) = 0$, and we declare that $\langle \overrightarrow{xy}, \overrightarrow{xz} \rangle_\kappa := -\infty$ if $d(x, y)d(x, z) > 0$ but $\kappa > 0$ and $\tilde{\angle}_\kappa(x; y, z)$ is not well-defined. We refer to $\langle \overrightarrow{xy}, \overrightarrow{xz} \rangle_\kappa$ as the *inner product* of $(X, d)$, cf. [**35**]. This is a minor modification of the notation introduced by Berg–Nikolaev [**5**], who gave a new characterization of CAT(0)-spaces, cf. Sato [**32**].

Proposition 20. *We have*

$$\int_Y \int_Y \left\langle \overrightarrow{b(\mu)x}, \overrightarrow{b(\mu)y} \right\rangle_\kappa d\mu(x) d\mu(y) \leq 0. \tag{21}$$

This inequality is obtained by integrating the following one, cf. [**35**, Inequality (33)]: For any $x \in Y$ with $d(x, b(\mu)) < R_\kappa$,

$$\begin{aligned} 0 \geq -DF\big[\log_{b(\mu)} x\big] &= \int_Y \big\langle \log_{b(\mu)} x, \log_{b(\mu)} y \big\rangle d\mu(y) \\ &\geq \int_Y \left\langle \overrightarrow{b(\mu)x}, \overrightarrow{b(\mu)y} \right\rangle_\kappa d\mu(y). \end{aligned} \tag{22}$$

Inequality (21) is similar but opposite to the one which appears and is called the *Lang–Schroeder–Sturm inequality* in [**35**, Proposition 22]. That inequality holds in Alexandrov spaces of curvature $\geq \kappa$ with $\kappa \in \mathbb{R}$.

Inequality (22) with $u$ instead of $\log_{b(\mu)} x$ yields that

$$\int_Y \big|\log_{b(\mu)} y - u\big|^2 - \big|\log_{b(\mu)} y - o\big|^2 d\mu(y) \geq \big|u - o\big|^2$$

for any $u \in C_{b(\mu)}$ with $o := o_{b(\mu)} \in C_{b(\mu)}$. This is the variance inequality for the push-forward measure $(\log_{b(\mu)})_* \mu \in \mathcal{P}_2(C_{b(\mu)})$ by the map $\log_{b(\mu)} : B(b(\mu), R_\kappa) \to C_{b(\mu)}$, which means that it has the vertex as the unique barycenter, cf. Ohta [**31**].

We can restate Inequality (22) as follows.

Proposition 23 (Curved variance inequality, cf. Ohta [**31**]). *For any point $x \in Y$ with $d(x, b(\mu)) < R_\kappa$,*

$$\int_Y \cos_\kappa d(\,\cdot, x) \frac{d(\,\cdot, b(\mu))}{\sin_\kappa d(\,\cdot, b(\mu))}\, d\mu$$
$$\leq \cos_\kappa d(b(\mu), x) \int_Y \cos_\kappa d(\,\cdot, b(\mu)) \frac{d(\,\cdot, b(\mu))}{\sin_\kappa d(\,\cdot, b(\mu))}\, d\mu.$$

Ohta [**31**] proved the curved reverse variance inequality in proper Alexandrov spaces of curvature $\geq \kappa$ with $\kappa \in \mathbb{R}$.

### 4.2. Contraction property.

Proposition 24 (Contraction property, cf. [**34**, Theorem 6.3]). *For any another $\nu \in \mathcal{P}_2(Y)$ concentrated on $B \subset Y$ with its barycenter $b(\nu) \in B$,*

$$\Phi(b(\mu), b(\nu)) \leq \inf_\pi \iint_{Y\times Y} \Phi(x, y)\, d\pi(x, y)$$

*and*

$$d(b(\mu), b(\nu)) \leq C \cdot d_p^W(\mu, \nu)$$

*with some constants $C < \infty$ and $p > 1$ depending only on $\kappa$ and $r$.*

In the above statement,

$$d_p^W(\mu, \nu) := \inf_\pi \left( \iint_{Y\times Y} d^p(x, y)\, d\pi(x, y) \right)^{1/p}$$

denotes the so-called $L^p$-*Wasserstein distance* between $\mu$ and $\nu$ usually defined for $p \geq 1$, and the infimum is taken over all *couplings* $\pi \in \mathcal{P}_2(Y \times Y)$ of $\mu$ and $\nu$, i.e., the push-forward measures of $\pi$ by the projections $\mathrm{pr}_i : Y \times Y \to Y$, $i = 1, 2$, onto the factors satisfy that $(\mathrm{pr}_1)_*\pi = \mu$ and $(\mathrm{pr}_2)_*\pi = \nu$.

Proof. It is easy to see that $(b(\mu), b(\nu)) \in Y \times Y$ is a barycenter of any coupling $\pi \in \mathcal{P}_2(Y \times Y)$ of $\mu$ and $\nu$ if $Y \times Y$ is equipped with the direct product metric $d_{Y\times Y}$. Indeed, for any $(x, y) \in B \times B$ in a neighborhood of $(b(\mu), b(\nu))$,

$$F_\pi(x, y) = \frac{1}{2} \int_{Y\times Y} d^2_{Y\times Y}((\cdot, \cdot), (x, y))\, d\pi = \frac{1}{2} \int_Y d^2(\cdot, x)\, d\mu + \frac{1}{2} \int_Y d^2(\cdot, y)\, d\nu$$
$$\geq F_\mu(b(\mu)) + F_\nu(b(\nu)) = F_\pi(b(\mu), b(\nu)).$$

Like in the proof of Theorem B, we obtain

$$\Phi(x, y) - \Phi(b(\mu), b(\nu)) \geq D\Phi\big[\log_{(b(\mu), b(\nu))}(x, y)\big]$$
$$\geq -\big\langle \nabla^-_{(b(\mu), b(\nu))}\Phi, \log_{(b(\mu), b(\nu))}(x, y)\big\rangle$$

for any $(x, y) \in B \times B$. Integrating this inequality, we obtain

$$\iint_{Y\times Y} \Phi(x,y)\,d\pi(x,y) - \Phi(b(\mu),b(\nu)) \geq DF_\pi\left[\nabla^-_{(b(\mu),b(\nu))}\Phi\right] \geq 0$$

for any coupling $\pi \in \mathcal{P}_2(Y\times Y)$ of $\mu$ and $\nu$, which proves the first inequality.

The second inequality with $p := 2(\nu+1) > 1$ follows from the first one and (18). □

### 4.3. Jensen's inequality.

As promised above, we present another version of Proposition 10 for CAT($\kappa$)-spaces. Because of its potential applications, we give a full statement.

THEOREM 25 (Jensen's inequality, cf. Kuwae [**25**]). *Let $(Y,d)$ be a complete CAT($\kappa$)-space with $\kappa > 0$. Suppose that $\mu \in \mathcal{P}_2(Y)$ is concentrated on a ball $B \subset Y$ of radius $r < R_\kappa/2$ and $b(\mu) \in B$ is the barycenter of $\mu$. Then for any lower-semicontinuous convex function $\varphi : Y \to \mathbb{R}\cup\{\infty\}$*

$$\varphi(b(\mu)) \leq \int_Y \varphi\,d\mu.$$

PROOF. Since our proof is identical to that of the main theorem of Kuwae [**25**], cf. [**34**, First Proof of Theorem 6.2], we try to keep the description short.

First of all, we remark that $\varphi$ is bounded below on any ball of radius $< R_\kappa$ and hence the integral $\int_Y \varphi\,d\mu \in (-\infty,\infty]$ is well-defined.

Moreover, as was remarked in [**25**], we may assume that $\varphi$ is bounded above on $Y$ by replacing $\varphi$ with $\varphi_n$ given by

$$\varphi_n(x) := \inf_{y\in B}[\varphi(y) + n\Phi(x,y)] \text{ for } n\in\mathbb{N} \text{ and } x \in B := B(b(\mu),r).$$

It is easy to see by using the convexity of $\Phi(\cdot,\cdot)$ that $\varphi_n$ is a lower-semicontinuous convex function bounded above for each $n\in\mathbb{N}$ and $\varphi_n(x) \nearrow \varphi(x)$ as $n\nearrow\infty$ for $x\in B$.

We define a subset $Y_\varphi := \{(x,t) \in Y\times[\inf_Y\varphi, \sup_Y\varphi] : \varphi(x) \leq t\}$. Then $Y_\varphi$ is a closed $R_\kappa$-convex subset of a complete CAT($\kappa$)-space $Y\times\mathbb{R}$ equipped with the direct product metric and we may further assume that $Y_\varphi$ is contained in a ball of radius $< R_\kappa/2$ in $Y\times\mathbb{R}$. We consider the map $\hat\varphi : Y \to Y_\varphi$ assigning $(x,\varphi(x)) \in Y_\varphi$ to each $x\in Y$. Then the push-forward measure $\hat\mu := \hat\varphi_*\mu \in \mathcal{P}_2(Y\times\mathbb{R})$ is concentrated on $Y_\varphi$ and has a barycenter

$$b(\hat\mu) = \left(b(\mu), \int_Y \varphi\,d\mu\right) \in B\times[\inf_Y\varphi, \sup_Y\varphi].$$

Now Corollary 16 induces that $b(\hat\mu)$ lies in $Y_\varphi$. This finishes the proof. □

We close this section with a few comments:

Firstly, the theory of CAT(0)-space valued martingales has been explored by e.g. Sturm [**33**] and Christiansen–Sturm [**7**]. They defined such martingales by using barycenter of probability measures. In this paper, we observe that some of the facts on barycenter in CAT(0)-spaces also hold in CAT($\kappa$)-spaces with $\kappa > 0$ as well. These circumstances would suggest that the analogue of their theory in CAT($\kappa$)-spaces of radius $< R_\kappa/2$ with $\kappa > 0$ is possible. Although we do not pursue this issue further in this paper, the author wishes to come back in a future work.

Secondly, Navas [**29**], cf. Es-Sahib–Heinich [**10**], presented a construction of another barycenter map $\mathrm{bar}^\star : \mathcal{P}_1(X) \to X$ enjoying that $\mathrm{bar}^\star(\delta_x) = x$ with $\delta_x \in \mathcal{P}_1(X)$ being the Dirac measure for all $x \in X$ and

$$d(\mathrm{bar}^\star(\mu), \mathrm{bar}^\star(\nu)) \leq d_1^W(\mu, \nu) \text{ for all } \mu, \nu \in \mathcal{P}_1(X)$$

on any complete separable metric space $(X, d)$ of nonpositive curvature in the sense of Busemann. Here $d_1^W$ denotes the $L^1$-*Wasserstein distance*, cf. Proposition 24.

A metric space of *nonpositive curvature in the sense of Busemann* is a geodesic space $(X, d)$ for which $d : X \times X \to [0, \infty)$ is convex, e.g. [**12**], and CAT(0)-spaces are special examples of them.

It seems that, with the use of the convex function $\Phi^{(\kappa)}_{\nu,\tilde{h}} : Y \times Y \to [0, \infty)$ in Theorem A instead of the distance function, Navas's construction carries over into any complete separable CAT($\kappa$)-space $Y$ of radius $< R_\kappa/2$ with $\kappa > 0$. Moreover, the ergodic theorem of the form of Austin [**2**] and Navas [**29**] can also be extended to such CAT($\kappa$)-spaces.

## 5. Banach–Saks Property of CAT($\kappa$)-spaces.

In this section, we state and prove Theorems C and D for CAT($\kappa$)-spaces, one of which generalizes a theorem of Jost [**15**] stated for CAT(0)-spaces.

Kakutani [**19**] proved the *Banach–Saks property* of uniformly convex Banach spaces: any bounded sequence $(x_n)_{n\in\mathbb{N}}$ of points of an uniformly convex Banach space $B$ has a subsequence, still denoted $(x_n)_{n\in\mathbb{N}}$, for which the sequence $(m_n)_{n\in\mathbb{N}}$ of the arithmetic means $m_n := (1/n)\sum_{i=1}^n x_i \in B$ converges to a point of $B$. As the main result of this section, we formulate this property for CAT($\kappa$)-spaces.

We start with the following definitions.

Definition 26. For a subset $A \subset X$ of a metric space $(X, d)$, we define its *circumradius* as $\mathrm{rad}_X(A) := \inf_{x\in X} \mathrm{rad}_x(A)$, where $\mathrm{rad}_x(A) := \sup_{a\in A} d(a, x)$ for $x \in X$. A point $x \in X$ giving $\mathrm{rad}_x(A) = \mathrm{rad}_X(A)$ is called a *circumcenter* of $A \subset X$. The *radius* of $(X, d)$ is defined as $\mathrm{rad}(X) := \mathrm{rad}_X(X)$.

It is easy to see by using the local convexity that any subset $A \subset Y$ of a complete CAT($\kappa$)-space $Y$ with $\kappa \in \mathbb{R}$ and $\mathrm{rad}_Y(A) < R_\kappa/2$ has a unique circumcenter contained in the closed convex hull $\overline{\mathrm{conv}}(A) \subset Y$ of $A$, cf. Balser–Lytchak [**4**, Lemma 3.3]. The famous Bruhat–Tits fixed point theorem states that any group acting isometrically on a CAT(0)-space fixes the circumcenter of a bounded orbit. This can be easily extended to any isomeric action on a CAT(1)-space with an orbit of circumradius $< \pi/2$, cf. [**4**, Proposition 1.4].

Definition 27 (Weak convergence [**15**]). Let $(p_n)_{n\in\mathbb{N}}$ be a sequence of points in a CAT($\kappa$)-space $(Y, d)$ with $\mathrm{rad}_o(\{p_n\}) < R_\kappa/2$ for some point $o \in Y$. We say that $(p_n)_{n\in\mathbb{N}}$ *converges weakly* to $o$ if $\pi_\gamma(p_n) \to o$ as $n \to \infty$ for any geodesic $\gamma : [0, 1] \to Y$

with $\gamma(0) = o$. Here, $\pi_\gamma(p_n) \in \gamma([0,1]) \subset Y$ denotes the closest point to $p_n$ on the image of $\gamma$, cf. Fact 8.

The following lemma is a Banach–Alaoglu type result for CAT($\kappa$)-spaces.

LEMMA 28 (cf. Jost [**15**, Theorem 2.1]). *Let $(Y, d)$ be a complete* CAT($\kappa$)*-space with $\kappa > 0$. Any sequence $(p_n)_{n\in\mathbb{N}}$ of points in $Y$ with $\mathrm{rad}_Y(\{p_n\}) < R_\kappa/2$ has a subsequence which converges weakly to a point in $Y$.*

Although a proof of this lemma seems to be scattered in the literature, we present a short proof of it here for convenience of the reader, cf. Bačák [**3**, Section 2.2], Espínola–Fernández-León [**9**, Corollary 4.4].

PROOF. We let $\Lambda_0 := \mathbb{N}$ and take a decreasing sequence $\{\Lambda_n\}_{n\in\mathbb{N}}$ of infinite subsets of $\mathbb{N}$ as follows: Suppose that we have chosen $\Lambda_{n-1} \subset \mathbb{N}$. We put

$$r_n := \inf_{\Lambda} \mathrm{rad}_Y(\{p_\lambda : \lambda \in \Lambda\}),$$

where $\Lambda$ runs over all infinite subsets of $\Lambda_{n-1} \setminus \{\min \Lambda_{n-1}\}$, and choose an infinite subset $\Lambda_n \subset \Lambda_{n-1} \setminus \{\min \Lambda_{n-1}\}$ such that

$$r'_n := \mathrm{rad}_Y(\{p_\lambda : \lambda \in \Lambda_n\})$$

satisfies that $r'_n - r_n \to 0+$ as $n \to \infty$. Then $r_n$ is nondecreasing in $n \in \mathbb{N}$ and hence the limit value $\lim_{n\to\infty} r_n = \lim_{n\to\infty} r'_n < R_\kappa/2$ exists.

Now the local convexity yields that the sequence $(o_n)_{n\in\mathbb{N}}$ with $o_n \in Y$ being the circumcenter of $\{p_\lambda : \lambda \in \Lambda_n\}$ is a Cauchy sequence in $Y$. We denote the limit of $(o_n)_{n\in\mathbb{N}}$ as $o \in Y$. Then the local convexity again yields that the sequence $(q_n)_{n\in\mathbb{N}}$ with $q_n := p_{\min \Lambda_n} \in Y$ converges weakly to $o \in Y$ as $n \to \infty$. This finishes the proof. $\square$

The following fact follows from the definition of weak convergence and Fact 8.

FACT 29. *Suppose that a sequence $(p_n)_{n\in\mathbb{N}}$ in a* CAT($\kappa$)*-space $(Y, d)$ converges weakly to $o \in Y$ with $\limsup_{n\to\infty} d(p_n, o) < R_\kappa/2$ and $\liminf_{n\to\infty} d(p_n, o) \geq \rho \geq 0$. For any point $x \in B(o, R_\kappa/2) \setminus \{o\}$, we have $\liminf_{n\to\infty} d(p_n, x) > \rho$.*

Now we are in a position to state and prove the main theorem of this section. The following theorem generalizes a theorem of Jost [**15**], cf. Remark in [**15**, Section 5].

THEOREM C (cf. Jost [**15**, Theorem 2.2]). *Let $(Y, d)$ be a complete* CAT($\kappa$)*-space with $\kappa \in \mathbb{R}$ and $(p_n)_{n\in\mathbb{N}}$ be a sequence of points in $Y$ with $\mathrm{rad}_Y(\{p_n\}) < R_\kappa/2$. Then it has a subsequence, still denoted as $(p_n)_{n\in\mathbb{N}}$, for which the sequence $(m_n)_{n\in\mathbb{N}}$ of barycenters of finitely and uniformly supported probability measures $(1/n)\sum_{i=1}^n \delta_{p_i} \in \mathcal{P}_2(Y)$ converges to a point in $Y$.*

Our proof of Theorem C uses only a few properties of CAT($\kappa$)-spaces and it also works for more general convex spaces, cf. Kell [**21**].

Proof. We may assume that $\kappa > 0$ because the proof of this theorem for nonpositive $\kappa \le 0$ is reduced to that for positive $\kappa > 0$.

Lemma 28 states that $(p_n)_{n\in\mathbb{N}}$ has a subsequence, still denoted as $(p_n)_{n\in\mathbb{N}}$, which converges weakly to a point $o \in Y$. It follows from its proof that $\sup_{n\in\mathbb{N}} d(p_n, o) < R_\kappa/2$ and we may further assume that the limit $\rho := \lim_{n\to\infty} d(p_n, o) \in [0, R_\kappa/2)$ exists. Then it follows that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} d^2(p_i, o) = \rho^2.$$

We put $B := B(o, R_\kappa/2)$ and

$$\Lambda(I) := \inf_{x\in B} \left[ \frac{1}{\#I} \sum_{i\in I} d^2(p_i, x) \right]$$

for a finite subset $I \subset \mathbb{N}$ of cardinality $\#I < \infty$.

We start our proof by making the following observation.

Claim 30. *For each $k, N \in \mathbb{N}$, we put $I_k^N := \{(k-1)2^N + 1, \dots, k2^N\} \subset \mathbb{N}$. If $(p_n)_{n\in\mathbb{N}}$ satisfies*

$$\sup\left\{ \liminf_{k\to\infty} \Lambda\big(I_k^N\big) : N \in \mathbb{N} \right\} = \rho^2,$$

*then the sequence $(m_n)_{n\in\mathbb{N}}$ of the barycenters $m_n \in B$ of probability measures $(1/n)\sum_{i=1}^{n} \delta_{p_i} \in \mathcal{P}_2(Y)$ obtained in Theorem B converges to $o$.*

Proof. By assumption, for any $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$\liminf_{k\to\infty} \Lambda\big(I_k^N\big) > \rho^2 - \varepsilon.$$

This implies that

$$\rho^2 \ge \liminf_{n\to\infty} \left[ \frac{1}{n} \sum_{i=1}^{n} d^2(p_i, m_n) \right] \ge \liminf_{k\to\infty} \frac{1}{k} \sum_{l=1}^{k} \Lambda\big(I_l^N\big) > \rho^2 - \varepsilon.$$

Since $\varepsilon > 0$ is taken arbitrarily,

$$\lim_{n\to\infty} \left[ \frac{1}{n} \sum_{i=1}^{n} \big( d^2(p_i, o) - d^2(p_i, m_n) \big) \right] = 0.$$

Combining this and the variance inequality (Proposition 19), we infer that $d(m_n, o) \to 0$ as $n \to \infty$. $\square$

We thus may assume that $\rho > 0$ and start a process of extracting a subsequence from $(p_n)_{n\in\mathbb{N}}$. We set $J_k^0 := \{k\}$ for each $k \in \mathbb{N}$ and we construct inductively a sequence $(J_k^N)_{k\in\mathbb{N}}$ of subsets of $\mathbb{N}$ of cardinality $2^N$ for each $N \in \mathbb{N}$ such that $J_k^N = J_l^{N-1} \cup J_m^{N-1}$ for some $l \neq m$, $\max J_k^N < \min J_{k+1}^N$, $J_1^N \subset J_1^{N+1}$ and

$$\lim_{k\to\infty} \Lambda\big(J_k^N\big) = \overline{\Lambda}^N := \limsup_{l,m\to\infty} \Lambda\big(J_l^{N-1} \cup J_m^{N-1}\big)$$

It is clear that $\overline{\Lambda}^N \leq \rho^2$ and $\overline{\Lambda}^{N+1} \geq \overline{\Lambda}^N$ for each $N \in \mathbb{N}$. We intend to show that $\lim_{N\to\infty} \overline{\Lambda}^N = \rho^2$ by proving the following

Claim 31. *If $\overline{\Lambda}^N < (\rho-\varepsilon)^2$ for some $N \in \mathbb{N}$ and small $\varepsilon > 0$, then $\overline{\Lambda}^{N+1} > \overline{\Lambda}^N + \delta(\varepsilon)$ for some $\delta(\varepsilon) > 0$.*

Proof. For each $l \in \mathbb{N}$, Theorem B states that there exists a unique point $m_l^N \in B$ which is a barycenter of $(1/2^N)\sum_{j\in J_l^N} \delta_{p_j} \in \mathcal{P}_2(Y)$ and satisfies that

$$\Lambda\big(J_l^N\big) = \frac{1}{2^N} \sum_{j\in J_l^N} d^2\big(p_j, m_l^N\big).$$

Fix large $l \in \mathbb{N}$ with $\Lambda\big(J_l^N\big) < (\rho-\varepsilon)^2$. Since $(p_n)_{n\in\mathbb{N}}$ converges weakly to $o$, we find $m \gg l$ such that $d(p_j, m_l^N) > \rho$ for each $j \in J_m^N$ by Fact 29. Then

$$\frac{1}{2^N} \sum_{j\in J_m^N} d^2\big(p_j, m_l^N\big) > \rho^2 > (\rho-\varepsilon)^2 > \Lambda\big(J_m^N\big) = \frac{1}{2^N} \sum_{j\in J_m^N} d^2\big(p_j, m_m^N\big)$$

and hence by the triangle inequality

$$2\max\big\{d(m_{l\cup m}^{N+1}, m_l^N), d(m_{l\cup m}^{N+1}, m_m^N)\big\} \geq d(m_l^N, m_m^N) > \varepsilon,$$

where $m_{l\cup m}^{N+1} \in B$ is the barycenter of $(1/2^{N+1})\sum_{j\in J_l^N \cup J_m^N} \delta_{p_j} \in \mathcal{P}_2(Y)$ in $B$.

By the variance inequality (Proposition 19), we acquire that

$$\begin{aligned}
\Lambda\big(J_l^N \cup J_m^N\big) &= \frac{1}{2^{N+1}}\Bigg(\sum_{j\in J_l^N} d^2\big(p_j, m_{l\cup m}^{N+1}\big) + \sum_{j\in J_m^N} d^2\big(p_j, m_{l\cup m}^{N+1}\big)\Bigg) \\
&\geq \frac{1}{2}\Bigg(\Lambda\big(J_l^N\big) + \Lambda\big(J_m^N\big) + c\cdot\left(\frac{\varepsilon}{2}\right)^{\alpha}\Bigg)
\end{aligned}$$

with constants $c > 0$ and $\alpha > 2$ from Proposition 19 and hence $\overline{\Lambda}^{N+1} > \overline{\Lambda}^N + c'\cdot\varepsilon^{\alpha}$ for some $c' > 0$. □

We let $s : \mathbb{N} \to \cap_{N\in\mathbb{N}} \cup_{k\in\mathbb{N}} J_k^N \subset \mathbb{N}$ be the order-preserving bijection. Then the sequence $(q_n)_{n\in\mathbb{N}}$ with $q_n := p_{s(n)}$ satisfies the assumption of Claim 30. This completes the proof of Theorem C. □

During this work, we also came up with another formulation of Banach–Saks property of CAT($\kappa$)-spaces. We need the following definition to state it.

Definition 32 (Inductive mean value [**34**], cf. [**33**]). Given a sequence $(p_i)_{i=1}^n$ of points in a convex subset of a metric space $(X, d)$ on which any two points are connected by a unique geodesic, we define a point $s_n \in X$ inductively by choosing $s_1 := p_1$ and $s_n$ for $n \geq 2$ as the unique point satisfying

$$n \cdot d(s_n, s_{n-1}) = d(s_{n-1}, p_n) = \frac{n}{n-1} \cdot d(s_n, p_n).$$

It would be illustrative to express this as

$$s_n \text{ “}=\text{” } \left(1 - \frac{1}{n}\right) s_{n-1} + \frac{1}{n} p_n.$$

Following Sturm [**34**, Definition 4.6], we write $s_n := (1/n) \overrightarrow{\sum}_{i=1}^n p_i$ and call it the *inductive mean value* of $(p_i)_{i=1}^n$.

Theorem D. *Let $(Y, d)$ be a complete* CAT($\kappa$)*-space with $\kappa \in \mathbb{R}$ and $(p_n)_{n\in\mathbb{N}}$ be a sequence of points in $Y$ with* $\mathrm{rad}_Y(\{p_n\}) < R_\kappa/2$. *Then it has a subsequence, still denoted as $(p_n)_{n\in\mathbb{N}}$, for which the sequence $(s_n)_{n\in\mathbb{N}}$ of inductive mean values $s_n := (1/n) \overrightarrow{\sum}_{i=1}^n p_i \in Y$ converges to a point in $Y$.*

Proof. We may assume that $\kappa > 0$ because the proof of this theorem for nonpositive $\kappa \leq 0$ is reduced to that for positive $\kappa > 0$.

By Lemma 28, $(p_n)_{n\in\mathbb{N}}$ has a subsequence which converges weakly to a point $o \in Y$. We do not change the notation in taking subsequences. It follows from its proof that $\sup_{n\in\mathbb{N}} d(p_n, o) < R_\kappa/2$ and we may further assume that the limit $\rho := \lim_{n\to\infty} d(p_n, o) \in [0, R_\kappa/2)$ exists.

We start the proof for the case of $\rho = 0$. In this case, we further take a subsequence such that $\lim_{n\to\infty} n \cdot d(p_n, o) = 0$. Then we claim that

$$\liminf_{n\to\infty} d(s_n, o) = 0. \tag{33}$$

Since $d(s_n, o) \leq \max\{d(s_{n-1}, o), d(p_n, o)\}$, this means that $d(s_n, o)$ is arbitrarily small for any large $n \gg 1$ which proves the theorem if $\rho = 0$. If we assume that $\inf_{n\geq n_0} d(s_n, o) > \varepsilon_0 > 0$ for some fixed $\varepsilon_0 > 0$ and $n_0 \gg 1$, we have

$$\begin{aligned} d(s_n, o) - d(s_{n-1}, o) &\leq d(s_n, p_n) + d(p_n, o) - d(s_{n-1}, o) \\ &\leq \frac{1}{n}(2n \cdot d(p_n, o) - d(s_{n-1}, o)), \end{aligned}$$

and hence

$$d(s_n, o) \le d(s_{n_0}, o) - \varepsilon_0 \sum_{k=n_0+1}^{n} \frac{1}{k} \to -\infty \text{ as } n \to \infty.$$

This is a contradiction.

We assume that $\rho > 0$ in the rest of the proof. We once again take a subsequence from $(p_n)_{n\in\mathbb{N}}$ as follows: Suppose that we have chosen $p_{n-1}$. If $s_{n-1} = o$, we leave $p_n$ unchanged; if $s_{n-1} \neq o$, we use Fact 29 and replace $p_n$ by its successor with larger $n$ so that

$$d(p_n, s_{n-1}) > d(p_n, o).$$

Then for any $n \ge 2$ with $s_{n-1} \neq o$, by the angle monotonicity, we have

$$\begin{aligned} \cos_\kappa \tilde{\angle}_\kappa(s_{n-1}; s_n, o) &\ge \cos_\kappa \tilde{\angle}_\kappa(s_{n-1}; p_n, o) \\ &> \cos_\kappa d(p_n, o) - \cos_\kappa d(s_{n-1}, o) \cos_\kappa d(s_{n-1}, p_n) \\ &> \cos_\kappa d(p_n, o)(1 - \cos_\kappa d(s_{n-1}, o)) > 0, \end{aligned}$$

We now verify for any $\varepsilon > 0$ that there exist numbers $N = N(\varepsilon) \in \mathbb{N}$ and $\delta = \delta(\varepsilon) > 0$ such that $d(s_n, o) < d(s_{n-1}, o) - (\delta/n)$ if $n > N$ and $d(s_{n-1}, o) > \varepsilon$. Indeed, for any large $n \gg 1$ with $d(s_{n-1}, o) > \varepsilon$, we use that $d(s_n, s_{n-1}) < R_\kappa/n$ to see

$$\begin{aligned} d^2(s_n, o) &= d^2(s_{n-1}, o) + d^2(s_n, s_{n-1}) - 2d(s_{n-1}, o)d(s_n, s_{n-1}) \cos \tilde{\angle}_0(s_{n-1}; s_n, o) \\ &< d^2(s_{n-1}, o) + d^2(s_n, s_{n-1}) - 2d(s_{n-1}, o)d(s_n, s_{n-1}) \cos \tilde{\angle}_\kappa(s_{n-1}; s_n, o) \\ &< d^2(s_{n-1}, o) - (2\delta/n). \end{aligned}$$

This implies (33) and that $d(s_n, o)$ is arbitrarily small for any large $n \gg 1$. Now the proof of Theorem D is complete. ☐

We conclude this section with the following theorem for CAT($\kappa$)-spaces, which is an analogue of the theorem of Jost [**15**], [**16**] proved for CAT(0)-spaces, cf. Jost [**17**].

THEOREM E (cf. Jost [**15**], [**16**], [**17**]). *Let $(Y, d)$ be a complete* CAT($\kappa$)*-space with $\kappa > 0$. Suppose that a lower-semicontinuous convex function $f : Y \to \mathbb{R} \cup \{\infty\}$ satisfies that*

$$\operatorname{rad}_Y(f^{-1}(-\infty, c_0]) < R_\kappa/2 \textit{ for some } c_0 > \inf\nolimits_Y f.$$

*Then $f$ attains its minimum in $Y$.*

Although it seems Theorem E may have many alternative proofs, we follow Jost's original argument in his proof of [**15**, Theorem 2.3]. It is a nice application of what we have proved.

PROOF. If $(p_n)_{n\in\mathbb{N}}$ is a minimizing sequence of a convex function $f$, i.e., $f(p_n) \to$

$\inf_Y f$ as $n \to \infty$, our proof of Theorem C says that it has a subsequence, still denoted as $(p_n)_{n\in\mathbb{N}}$, for which the sequence $(m_n)_{n\in\mathbb{N}}$ of barycenters converges to a point $o \in Y$ with $\mathrm{rad}_o(\{p_n\}) < R_\kappa/2$. Then Jensen's inequality (Theorem 25) yields that

$$f(o) \leq \liminf_{n\to\infty} f(m_n) \leq \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} f(p_i) = \inf_Y f.$$

This proves that $f(o) = \inf_Y f$. □

## A. Proof of Theorem A.

In this appendix, we describe a proof of Theorem A stated in the introduction. As mentioned there, it generalizes Kendall's result in [**23**]. Before working on Theorem A, we recall his result and its proof.

For $h > 0$, we consider a small upper hemisphere

$$S^{n-1}_{+,h} := \{x = (x_1, \dots, x_n) \in \mathbb{R}^n : |x| = 1, x_1 > h\},$$

and equip it with the spherical distance $\overline{d}$. Here and in what follows, $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^n$.

THEOREM 34 (Kendall [**23**]). *For any $h > \tilde{h} > 0$ and $\nu \in \mathbb{R}$, the function $\Phi_{\nu,\tilde{h}} : S^{n-1}_{+,h} \times S^{n-1}_{+,h} \to [0,\infty)$ given by*

$$(x,y) \longmapsto \left(\frac{|x-y|^2}{2(x_1 y_1 - \tilde{h}^2)}\right)^{\nu+1}$$

*is convex, provided that $2(2\nu+1)\tilde{h}^2(h^2 - \tilde{h}^2) \geq 1$.*

Theorem 34 says that a small upper hemisphere $S^{n-1}_{+,h} \subset \mathbb{R}^n$ has *convex geometry.* Besides this, Kendal [**23**] also showed that an open hemisphere $S^{n-1}_{+} := S^{n-1}_{+,0} \subset \mathbb{R}^n$ does not have convex geometry. Notice that the function $\Phi_{\nu,\tilde{h}}$ in Theorem 34 is nothing but $\Phi^{(1)}_{\nu,\tilde{h}}$ in Theorem A with $o := (1, 0, \dots, 0) \in S^{n-1}_{+}$ up to a constant multiple. In the above statement, we made a slight improvement for the condition on the parameters. We notice that $\Phi^{(\kappa)}_{\nu,\tilde{h}}(x,y)$ approaches to a constant multiple of $d(x,y)$ as $\kappa, \tilde{h} \to 0$ and $\nu \to -1/2$.

Since our proof of Theorem A relies on some calculation carried out in [**23**], we first recall Kendall's argument.

KENDALL'S PROOF OF THEOREM 34. We fix two geodesics $\lambda, \mu : I \to (S^{n-1}_{+,h}, \overline{d})$ and consider the function $\Phi(t) := \Psi^{\nu+1}(t) := \Phi_{\nu,\tilde{h}}(\lambda(t), \mu(t))$ with $t \in I$. Roughly speaking, we need to prove that $\Phi'' \geq 0$ on $I$; see Proof of Theorem A for the precise meaning of this. We do this at $t = 0 \in I$.

To begin with, we put

$$x := \lambda(0),\ y := \mu(0) \in S^{n-1} \subset \mathbb{R}^n; \quad u := \lambda'(0),\ v := \mu'(0) \in \mathbb{R}^n.$$

Then $\lambda''(0) = -|u|^2 x$ and $\mu''(0) = -|v|^2 y$. We also put

$$p := \frac{1}{2}|x-y|^2 \text{ and } q := x_1 y_1 - \tilde{h}^2$$

and note that $q > h^2 - \tilde{h}^2 > 0$.

Then, with $\psi := \Psi(0) = p/q$ and $\psi' := \Psi'(0)$, we know that $\Psi'(0) = (p' - \psi q')/q$ and

$$\Psi''(0) = (p'' - \psi q'' - 2\psi' q')/q,$$

where

$$\begin{aligned} p' &= \langle u - v, x - y \rangle = \psi' q + \psi q'; \\ p'' &= |u-v|^2 - (|u|^2 + |v|^2)p \geq (p')^2/2p - (|u|^2+|v|^2)p; \\ q' &= u_1 y_1 + x_1 v_2; \\ q'' &= 2u_1 v_1 - (|u|^2 + |v|^2)(q + \tilde{h}^2). \end{aligned}$$

Together with that $(\psi q')^2/2p - 2\psi u_1 v_1 \geq 0$ [**23**] and $(q')^2 \leq |u|^2 + |v|^2$, which follows from $u_1 = \sqrt{1 - x_1^2}|u|$ and $v_1 = \sqrt{1-y_1^2}|v|$, cf. Proof of Theorem A below, the above inequalities yield that

$$\begin{aligned} \Psi''(0) &\geq \frac{1}{q}\left[\tilde{h}^2 \psi (|u|^2 + |v|^2) - \psi' q'\right] + \frac{(\psi')^2}{2\psi} \\ &\geq \frac{1}{q}\left[\tilde{h}^2 \psi (q')^2 - \psi' q'\right] + \frac{(\psi')^2}{2\psi}. \end{aligned}$$

Recalling that $\Phi = \Psi^{\nu+1}$, we acquire

$$\Phi''(0) \geq (\nu+1)\frac{\tilde{h}^2}{q}(|u|^2 + |v|^2)\Phi(0) \geq 0$$

if $\psi' = 0$, and

$$\begin{aligned} \Phi''(0) &= (\nu+1)\psi^{\nu-1}\left[\psi \Psi''(0) + \nu(\psi')^2\right] \\ &\geq (\nu+1)\psi^{\nu-1}\left[\frac{1}{q}\left(\tilde{h}\psi q' - \frac{\psi'}{2\tilde{h}}\right)^2 + \left(\nu + \frac{1}{2} - \frac{1}{4q\tilde{h}^2}\right)(\psi')^2\right] \\ &\geq C(\nu, h, \tilde{h})(\Phi'(0))^2 \end{aligned}$$

in general, where

$$C(\nu,h,\tilde{h}) := \frac{1}{\nu+1}\left(\nu+\frac{1}{2}-\frac{1}{4\tilde{h}^2(h^2-\tilde{h}^2)}\right)\left(\frac{h^2-\tilde{h}^2}{2(1-h^2)}\right)^{\nu+1}. \tag{35}$$

This constant is nonnegative provided that the parameters satisfy the condition in the statement. This ends a quick review of the proof of Theorem 34. $\square$

Now we turn to the proof of Theorem A. To prove this theorem, we mimic Kendall's proof of Theorem 34. However, we have to be careful in dealing with general CAT($\kappa$)-spaces in which geodesics may branch.

PROOF OF THEOREM A. In proving the convexity of $\Phi^{(\kappa)}_{\nu,\tilde{h}}$, we may assume that $(Y,d)$ is a CAT(1)-space and $\kappa = 1$ by rescaling the metric. We fix two geodesics $\lambda,\mu : I \to B(o,r) \subset Y$ with $r < \pi/2$ and consider the function $\Phi(t) := \Psi^{\nu+1}(t) := \Phi^{(\kappa)}_{\nu,\tilde{h}}(\lambda(t),\mu(t))$ with $t \in I$.

We plan to prove that $\Phi'' \geq 0$ on $I$ in the barrier sense, i.e., for every $t_0 \in I \setminus \partial I$, there exists a $C^2$-function $\overline{\Phi}$ defined on a neighborhood of $t_0 \in I$ such that $\Phi(\,\cdot\,) \geq \overline{\Phi}(\,\cdot\,)$, $\overline{\Phi}(t_0) = \Phi(t_0)$ and $\overline{\Phi}''(t_0) \geq 0$. This suffices to prove the convexity of the function $\Phi$ on $I$. We do this at $t_0 = 0 \in I$ and may assume that $x := \lambda(0)$ and $y := \mu(0)$ are distinct, because the constant function $\overline{\Phi} \equiv 0$ does the job if $x = y$.

We choose $\overline{x} = \underline{x}$ and $\overline{y} = \underline{y} \in S^2$ such that, with $\underline{o} := (1,0,0) \in S^2$, $\{\underline{x},\underline{y},\underline{o}\} \subset S^2$ is an isometric copy of $\{x,y,o\} \subset B(o,r)$. As $d(x,o)+d(y,o)+d(x,y) < 2\pi$, this is possible. We also take geodesics $\overline{\lambda},\underline{\lambda},\overline{\mu},\underline{\mu} : I \to (S^2,\overline{d})$ such that

$$\overline{\lambda}(0) = \underline{\lambda}(0) = \overline{x};\quad \overline{\mu}(0) = \underline{\mu}(0) = \overline{y};$$
$$|\overline{\lambda}'| = |\underline{\lambda}'| = |\lambda'|;\quad |\overline{\mu}'| = |\underline{\mu}'| = |\mu'|;$$
$$\langle \overline{\lambda}'(0+), \log_{\overline{x}} \overline{y}\rangle = \langle \lambda'(0+), \log_x y\rangle;\quad \langle \overline{\mu}'(0+), \log_{\overline{y}} \overline{x}\rangle = \langle \mu'(0+), \log_y x\rangle;$$
$$\langle \underline{\lambda}'(0+), \log_{\underline{x}} \underline{o}\rangle = \langle \lambda'(0+), \log_x o\rangle;\quad \langle \underline{\mu}'(0+), \log_{\underline{y}} \underline{o}\rangle = \langle \mu'(0+), \log_y o\rangle;$$

and the points $\overline{\lambda}(t)$ and $\overline{\mu}(t)$ with some $t > 0$ live in the same closed hemisphere with the great circle running through $\overline{x}$ and $\overline{y}$ as the equator.

By construction, if $\gamma \in \{\lambda,\mu\}$ is nontrivial, i.e., $|\gamma'| > 0$, then

$$\tilde{\angle}_1\,(\gamma(0);\gamma(t),o) \geq \angle_{\gamma(0)}(\gamma'(0+),o) = \tilde{\angle}_1\,(\underline{\gamma}(0);\underline{\gamma}(t),\underline{o})$$

for any $t > 0$, and by the triangle inequality for the angle $\angle_{\gamma(0)}$ on $\Sigma_{\gamma(0)}$,

$$\begin{aligned}\tilde{\angle}_1\,(\gamma(0);\gamma(t),o) &\geq \angle_{\gamma(0)}(\gamma'(0-),o)\\ &\geq \pi - \angle_{\gamma(0)}(\gamma'(0+),o) = \tilde{\angle}_1\,(\underline{\gamma}(0);\underline{\gamma}(t),\underline{o}),\end{aligned} \tag{36}$$

for any $t < 0$. This yields that $d(\gamma(t),o) \geq \overline{d}\big(\underline{\gamma}(t),\underline{o}\big)$ for $t \in I$ and hence two geodesics $\underline{\lambda}$ and $\underline{\mu}$ are contained in $S^2_{+,h}$ with $0 < h \leq \cos r$.

For $t \in I$, we put

$$P(t) := 1 - \cos d(\lambda(t), \mu(t)) \text{ and } Q(t) := \cos d(\lambda(t), o) \cos d(\mu(t), o) - \tilde{h}^2,$$

and define

$$\overline{P}(t) := 1 - \cos \overline{d}\big(\overline{\lambda}(t), \overline{\mu}(t)\big) \text{ and } \underline{Q}(t) := \cos \overline{d}(\underline{\lambda}(t), \underline{o}) \cos \overline{d}\big(\underline{\mu}(t), \underline{o}\big) - \tilde{h}^2,$$

If $\lambda$ and $\mu$ are nontrivial and either of $\theta_x := \angle_x(\lambda'(0+), y)$ or $\theta_y := \angle_y(\mu'(0+), x)$ is zero, we alter the definition as

$$\overline{P}(t) := 1 - \cos\left[\overline{d}\big(\overline{\lambda}(t), \overline{y}\big) + \overline{d}(\overline{x}, \overline{\mu}(t)) - \overline{d}(\overline{x}, \overline{y})\right]$$

for $t \geq 0$. This applies to $\overline{P}(t)$ for $t \leq 0$ as well if either of $\angle_x(\lambda'(0-), y)$ or $\angle_y(\mu'(0-), x)$ is zero. It is clear that $\overline{P}(0) = P(0)$ and $\underline{Q}(0) = Q(0)$. It also follows from Fact 7 that $P'(0+) = \overline{P}'(0)$ and $Q'(0+) = \underline{Q}'(0)$.

Claim 37. *$P \geq \overline{P}$ and $Q \leq \underline{Q}$ on a neighborhood of $0 \in I$.*

Proof. Since actually we have already seen this claim for $Q$, it only remains to check that $P(t) \geq \overline{P}(t)$ for all $t \in I$ near 0. We do this only for positive $t > 0$, since the proof for negative $t < 0$ is the same except for the additional use of an inequality similar to Inequality (36).

If one of $\lambda$ and $\mu$ is trivial or one of $\theta_x$ and $\theta_y$ is $\pi$, the image of a geodesic extending the geodesic connecting $x$ and $y$ contains one of $\lambda(t)$ and $\mu(t)$ for any $t > 0$. Then the angle comparison implies that

$$d(\lambda(t), \mu(t)) \geq \overline{d}\big(\overline{\lambda}(t), \overline{\mu}(t)\big)$$

and this proves the claim in this case. We thus assume that both of $\lambda$ and $\mu$ are nontrivial in the rest of the proof.

If both of $\theta_x$ and $\theta_y$ are between 0 and $\pi$, we fix small $t > 0$ and take $\overline{\lambda_t}, \overline{\mu_t} \in S^2$ such that $\{x, y, \lambda(t)\}$ and $\{\mu(t), y, \lambda(t)\} \subset Y$ are isometric to $\big\{\overline{x}, \overline{y}, \overline{\lambda_t}\big\}$ and $\big\{\overline{\mu_t}, \overline{y}, \overline{\lambda_t}\big\} \subset S^2$ respectively and the open regions spanned by $\big\{\overline{x}, \overline{y}, \overline{\lambda_t}\big\}$ and $\{\overline{x}, \overline{y}, \overline{\mu_t}\}$ have nonempty intersection. Then we have

$$\begin{aligned}
\overline{\theta}_x &:= \tilde{\angle}_1\left(\overline{x}; \overline{\lambda_t}, \overline{y}\right) = \tilde{\angle}_1\left(x; \lambda(t), y\right) \geq \theta_x; \\
\overline{\theta}_y &:= \tilde{\angle}_1\left(\overline{y}; \overline{\mu_t}, \overline{x}\right) = \tilde{\angle}_1\left(y; \mu(t), \lambda(t)\right) + \tilde{\angle}_1\left(y; \lambda(t), x\right) \geq \theta_y.
\end{aligned}$$

The distance $\overline{d}\big(\overline{\lambda_t}, \overline{\mu_t}\big)$ for small $t \ll \min\{\theta_x, \theta_y, \pi - \theta_y\}$ is expressed in two ways as

$$\begin{aligned}
\overline{d}\big(\overline{\lambda_t}, \overline{\mu_t}\big) &= \operatorname{Leng}\big(\overline{\theta}_x - \tilde{\angle}_1\left(\overline{x}; \overline{\mu_t}, \overline{y}\right); d(\lambda(t), x), \overline{d}(\overline{\mu_t}, \overline{x})\big) \\
&= \operatorname{Leng}\big(\overline{\theta}_y - \tilde{\angle}_1\left(\overline{y}; \overline{\lambda_t}, \overline{x}\right); d(\mu(t), y), d(\lambda(t), y)\big),
\end{aligned}$$

which depend only on $d(x,y)$, $d(\lambda(t),x)$, $d(\mu(t),y)$, $\overline{\theta}_x$ and $\overline{\theta}_y$ as every input in the above expression is determined by them. Here, $\mathrm{Leng}(\theta;a,b)\geq 0$ denotes the length of the third edge of a geodesic triangle in $(S^2,\overline{d})$ with two edges of length $a,b>0$ and the angle $\theta\in[0,\pi]$ between them.

Since $\mathrm{Leng}(\theta;a,b)$ with fixed $a,b>0$ is monotone increasing in $\theta\in[0,\pi]$ and hence $\overline{d}\big(\overline{\lambda_t},\overline{\mu_t}\big)$ is monotone increasing in $\overline{\theta}_x$ and $\overline{\theta}_y$, we deduce that

$$d(\lambda(t),\mu(t))=\overline{d}\big(\overline{\lambda_t},\overline{\mu_t}\big)\geq\overline{d}\big(\overline{\lambda}(t),\overline{\mu}(t)\big)\,.$$

If one of $\theta_x$ or $\theta_y$ is zero, say $\theta_y=0$, then

$$\begin{aligned}d(\lambda(t),\mu(t))&\geq d(\lambda(t),y)-d(\mu(t),y)\\&\geq\overline{d}\big(\overline{\lambda}(t),\overline{y}\big)+\overline{d}(\overline{\mu}(t),\overline{x})-\overline{d}(\overline{x},\overline{y})\,.\end{aligned}$$

Therefore we gather that $P(t)\geq\overline{P}(t)$ for any $t>0$ in a neighborhood of $0\in I$ and finish the proof of the claim. □

It is easy to see that $\overline{P}$ and $\underline{Q}$ are $C^2$-functions. This is also the case even if $\overline{P}(t)$ is defined differently for $t\geq 0$ and $t\leq 0$. For example, if $\theta_x>0$ and $\theta_y=0$, a simple calculation using

$$\begin{aligned}d/dt|_{t=0}\cos\overline{d}\big(\overline{\lambda}(t),\overline{y}\big)&=-|\lambda'|\sin d(x,y)\cos\theta_x;\\ d^2/dt^2\big|_{t=0}\cos\overline{d}\big(\overline{\lambda}(t),\overline{y}\big)&=-|\lambda'|^2\cos d(x,y)\end{aligned}$$

yields that two functions

$$\begin{aligned}t\longmapsto{}&1-\cos\overline{d}\big(\overline{\lambda}(t),\overline{\mu}(t)\big)\\&=1-\cos(|\lambda'|t)\cos\big[d(x,y)-|\mu'|t\big]-\sin(|\lambda'|t)\sin\big[d(x,y)-|\mu'|t\big]\cos\theta_x;\\ t\longmapsto{}&1-\cos\big[\overline{d}\big(\overline{\lambda}(t),\overline{y}\big)+\overline{d}(\overline{\mu}(t),\overline{x})-\overline{d}(\overline{x},\overline{y})\,\big]\\&=1-\cos\overline{d}\big(\overline{\lambda}(t),\overline{y}\big)\cos(|\mu'|t)-\sin\overline{d}\big(\overline{\lambda}(t),\overline{y}\big)\sin(|\mu'|t)\end{aligned}$$

have the same first and second derivatives at $t=0$.

Now we see that the function $\overline{\Phi}$ given by

$$\overline{\Phi}(t):=\overline{\Psi}^{\nu+1}(t),\ \text{ where }\overline{\Psi}(t):=\overline{P}(t)/\underline{Q}(t)\text{ for }t\in I$$

is a barrier function of $\Phi$ on a neighborhood of $0\in I$ with the required properties.

We already know that $\overline{\Phi}$ is a $C^2$-function and $\Phi\geq\overline{\Phi}$ on a neighborhood of $0\in I$ with $\Phi(0)=\overline{\Phi}(0)$. Fact 7 induces that $\Phi(0+)=\overline{\Phi}'(0)$. It also follows from Kendall's computation recalled in the proof of Theorem 34 above that $\overline{\Phi}''(0)\geq 0$. We now check this. All of the following equations and inequalities are evaluated at $t=0$.

We have

$$\overline{P}'' = \left|\overline{\lambda}' - \overline{\mu}'\right|^2 - (|\lambda'|^2 + |\mu'|^2)P \geq (P')^2/2P - (|\lambda'|^2 + |\mu'|^2)P;$$
$$\underline{Q}'' = 2\left(\underline{\lambda}'\right)_1\left(\underline{\mu}'\right)_1 - (|\lambda'|^2 + |\mu'|^2)(Q + \tilde{h}^2);$$

and Fact 7 induces that

$$\begin{aligned} Q'(0+) = &- \sin d(x,o)\cos d(y,o)\big\langle \lambda'(0+), \uparrow_o^x \big\rangle \\ &- \cos d(x,o)\sin d(y,o)\big\langle \mu'(0+), \uparrow_o^y \big\rangle, \end{aligned}$$

which implies that $(Q'(0+))^2 \leq |\lambda'|^2 + |\mu'|^2$. We use them to obtain

$$\overline{\Psi}'' = \frac{1}{Q}\Big[\overline{P}'' - \Psi\underline{Q}'' - 2\overline{\Psi}'\underline{Q}'\Big] \geq \frac{1}{Q}\Big[\tilde{h}\Psi(\underline{Q}')^2 - \overline{\Psi}'\underline{Q}'\Big] + \frac{(\overline{\Psi}')^2}{2\Psi}$$

Therefore, we acquire that

$$\overline{\Phi}'' = (\nu+1)\Psi^{\nu-1}\big[\Psi\overline{\Psi}'' + \nu(\overline{\Psi}')^2\big] \geq C(\nu, h, \tilde{h})(\overline{\Phi}')^2,$$

with $C(\nu, h, \tilde{h})$ given in (35). Now the proof of Theorem A is complete. $\square$

We close this appendix with supplementary remarks on Theorem A.

Remark 38 (cf. [**23**, Corollaries A, B]). The above proof says that the function $\Phi$ there satisfies

$$\Phi''(t) \geq C(\nu, h, \tilde{h})\left(\Phi'(t+)\right)^2$$

for any $t \in I \setminus \partial I$ in the barrier sense.

If $\Phi'(t_0+) = 0$ at some $t_0 \in I \setminus \partial I$, the poof of Theorem 34 says that

$$\overline{\Phi}'' \geq (\nu+1)\frac{\tilde{h}^2}{Q}\left(|\lambda'|^2 + |\mu'|^2\right)\Phi \text{ at } t = t_0.$$

Therefore, we infer that $\Phi'' > 0$ on $I$ in the barrier sense as long as all of $\Phi(\,\cdot\,)$, $|\lambda'|$, $|\mu'|$ and $C(\nu, h, \tilde{h})$ are positive.

Remark 39. For a ball $B(o,r) \subset Y$ in a CAT($\kappa$)-space $(Y, d)$ with $\kappa > 0$ as in Theorem A, the function $\Phi^{(\kappa)}_{\nu,\tilde{h}}(\,\cdot\,, x_0)$, with $x_0 \in B(o,r)$ fixed, is convex on $B(o,r)$.

Furthermore, simpler functions $\varphi : B(o, R_\kappa/2) \to [0,\infty)$ given by

$$x \longmapsto \frac{1}{\kappa}\left(1 - \cos_\kappa d(x,o)\right)$$

and $\varphi_{x_0} : B(o, R_\kappa/2) \to [0,\infty)$, with $x_0 \in B(o, R_\kappa/2)$ fixed, given by

$$x \longmapsto \frac{1}{\kappa} \cdot \frac{1-\cos_\kappa d(x,x_0)}{\cos_\kappa d(x,o)}$$

are also proved to be convex on $B(o,R_\kappa/2) \subset Y$, cf. Kendall [**24**]. In fact, it is easy to see that they are convex on the open hemisphere $(S^{n-1}_+, \overline{d})$ and thus we can prove their convexity in the barrier sense along any geodesic in $B(o,R_\kappa/2) \subset Y$ as was done in our proof of Theorem A.

ACKNOWLEDGEMENTS. The author is indebted to Prof. Kazuhiro Kuwae for suggesting this work and for various discussions while preparing this manuscript. He also thanks Prof. Jürgen Jost for the email conversation concerning our Theorem C, Prof. Koichi Nagano and Prof. Burkhard Wilking for their comments and the anonymous referee for carefully reading our manuscript. This work was done during the author's stay in the University of Münster under JSPS Postdoctoral Fellowship for Research Abroad.

## References

[1] B. Afsari, Riemannian $L^p$ center of mass: existence, uniqueness, and convexity, Proc. Amer. Math. Soc., **139** (2011), 655–673.
[2] T. Austin, A CAT(0)-valued pointwise ergodic theorem, J. Topol. Anal., **3** (2011), 145–152.
[3] M. Bačák, The proximal point algorithm in metric spaces, Israel J. Math., **194** (2013), 689–701.
[4] A. Balser and A. Lytchak, Centers of convex subsets of buildings, Ann. Global Anal. Geom., **28** (2005), 201–209.
[5] I.-D. Berg and I.-G. Nikolaev, Quasilinearization and curvature of Aleksandrov spaces, Geom. Dedicata, **133** (2008), 195–218.
[6] D. Burago, Y. Burago and S. Ivanov, A course in metric geometry, Graduate Studies in Math., **33**, Amer. Math. Soc., Providence, RI, 2001.
[7] T. Christiansen and K. T. Sturm, Expectations and martingales in metric spaces, Stochastics, **80** (2008), 1–17.
[8] I. Ekeland, Nonconvex minimization problems, Bull. Amer. Math. Soc., **1** (1979), 443–474.
[9] R. Espínola and A. Fernández-León, CAT(k)-spaces, weak convergence and fixed points, J. Math. Anal. Appl., **353** (2009), 410–427.
[10] A. Es-Sahib and H. Heinich, Barycentre canonique pour un espace métrique ȩurbure négative, (French) Séminaire de Probabilités, XXXIII, 355–370, Lecture Notes in Math., **1709**, Springer, Berlin, 1999.
[11] K. Fujiwara, K. Nagano and T. Shioya, Fixed point sets of parabolic isometries of CAT(0)-spaces, Comment. Math. Helv., **81** (2006), 305–335.
[12] T. Gelander, A. Karlsson and G. A. Margulis, Superrigidity, generalized harmonic maps and uniformly convex spaces, Geom. Funct. Anal., **17** (2008), 1524–1550.
[13] T. Ishihara, A mapping of Riemannian manifolds which preserves harmonic functions, J. Math. Kyoto Univ., **19** (1979), 215–229.
[14] W. Jäger and H. Kaul, Uniqueness and stability of harmonic maps and their Jacobi fields, Manuscripta Math., **28** (1979), 269–291.
[15] J. Jost, Equilibrium maps between metric spaces, Calc. Var. Partial Differential Equations, **2** (1994), 173–204.
[16] J. Jost, Convex functionals and generalized harmonic maps into spaces of nonpositive curvature, Comment. Math. Helv., **70** (1995), 659–673.
[17] J. Jost, Generalized harmonic maps between metric spaces, Geometric analysis and the calculus of variations, Int. Press, Cambridge, MA, 1996, 143–174.
[18] J. Jost, Y. Xin and L. Yang, The regularity of harmonic maps into spheres and applications to Bernstein problems, J. Differential Geom., **90** (2012), 131–176.
[19] S. Kakutani, Weak convergence in uniformly convex spaces, Tôhoku Math. J., **45** (1938), 188–193.

[20] H. Karcher, Riemannian center of mass and mollifier smoothing, Comm. Pure Appl. Math., **30** (1977), 509–541.
[21] M. Kell, Uniformly Convex Metric Spaces, Anal. Geom. Metr. Spaces, **2** (2014), 359–380.
[22] W. Kendall, Probability, convexity, and harmonic maps with small image, I, Uniqueness and fine existence, Proc. London Math. Soc. (3), **61** (1990), 371–406.
[23] W. Kendall, Convexity and the hemisphere, J. London Math. Soc. (2), **43** (1991), 567–576.
[24] W. Kendall, From stochastic parallel transport to harmonic maps, New directions in Dirichlet forms, 49–115, AMS/IP Stud. Adv. Math., **8**, Amer. Math. Soc., Providence, RI, 1998.
[25] K. Kuwae, Jensen's inequality over CAT($\kappa$)-space with small diameter, Potential theory and stochastics in Albac, 173–182, Theta Ser. Adv. Math., **11**, Theta, Bucharest, 2009.
[26] K. Kuwae, Jensen's inequality on convex spaces, Calc. Var. Partial Differential Equations, **49** (2014), 1359–1378.
[27] K. Kuwae and K.-T. Sturm, On a Liouville type theorem for harmonic maps to convex spaces via Markov chains, Proceedings of RIMS Workshop on Stochastic Analysis and Applications, RIMS Kôkyûroku Bessatsu, 2008, 177–191.
[28] A. Lytchak, Open map theorem for metric spaces, Algebra i Analiz, **17** (2005), 139–159; translation in St. Petersburg Math. J., **17** (2006), 477–491.
[29] A. Navas, An $L^1$ ergodic theorem with values in a non-positively curved space via a canonical barycenter map, Ergodic Theory Dynam. Systems, **33** (2013), 609–623.
[30] S.-i. Ohta, Convexities of metric spaces, Geom. Dedicata, **125** (2007), 225–250.
[31] S.-i. Ohta, Barycenters in Alexandrov spaces of curvature bounded below, Adv. Geom., **12** (2012), 571–587.
[32] T. Sato, An alternative proof of Berg and Nikolaev's characterization of CAT(0)-spaces via quadrilateral inequality, Arch. Math. (Basel), **93** (2009), 487–490.
[33] K.-T. Sturm, Nonlinear martingale theory for processes with values in metric spaces of nonpositive curvature, Ann. Probab., **30** (2002), 1195–1222.
[34] K.-T. Sturm, Probability measures on metric spaces of nonpositive curvature, Heat kernels and analysis on manifolds, graphs, and metric spaces, (Paris, 2002), 357–390, Contemp. Math., **338**, AMS, 2003.
[35] T. Yokota, A rigidity theorem in Alexandrov spaces with lower curvature bound, Math. Annalen, **353** (2012), 305–331.

Takumi Yokota
Research Institute for Mathematical Sciences
Kyoto University
Kyoto 606-8502, Japan
E-mail: takumiy@kurims.kyoto-u.ac.jp