TSUKUBA J. MATH.
Vol. 41 No. 1 (2017), 43–80

CONVEX FUNCTIONS AND p-BARYCENTER ON CAT(1)-SPACES OF SMALL RADII

By

Takumi YOKOTA

Abstract. We establish unique existence of p-barycenter of any probability measure for $p \geq 2$ on CAT(1)-spaces of small radii. In our proof, we employ Kendall's convex function on a ball of CAT(1)-spaces instead of the convexity of distance function. Various properties of p-barycenter on those spaces are also presented. They extend the author's previous work [Yo].

1. Introduction

In this paper, we extend our previous work [Yo] on barycenter of probability measures on CAT(1)-spaces and study p-barycenter of them for some real number $p \geq 1$. CAT(κ)-spaces are metric spaces with $\kappa \in \mathbf{R}$ as an upper bound for the curvature in the sense of Alexandrov which is defined in terms of the convexity of distance function. The precise definition is given in Definition 3 below.

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

In [Yo] we studied 2-barycenter, usually called *barycenter*, *center of mass* or *Fréchet mean* in the literature, of probability measures on CAT(1)-spaces. We

2010 *Mathematics Subject Classification*: 53C23.
Key words and phrases: CAT(1)-space, Convex function, p-barycenter, Banach–Saks property.
Received August 1, 2016.
Revised March 22, 2017.

remark that 1-barycenter, also called *median*, e.g. Yang [Ya], is a generalization of *Fermat*(*–Torricelli*) *points* of plane triangles and *Steiner points* in Sakai [Sa]. For example, p-barycenter appears in the works of Afsari [Af], Naor–Silberman [NS] and Kuwae [Ku2, Ku3].

The theory of barycenter of probability measures on CAT(0)-spaces has been developed by many authors; See e.g. Sturm [St]. It is well-known that the distance function $d : Y \times Y \to [0, \infty)$ of a CAT(0)-space (Y, d) is convex in the sense of Definition 2 below. The following theorem is the main tool that we use in our approach, which states that any small ball in a CAT(κ)-space with $\kappa > 0$ also admits such a convex function. Here and hereafter, $B(o, \cdot)$ and $\bar{B}(o, \cdot)$ 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 (Kendall [Ke2], Jost [Jo2] and [Yo]). *Let (Y, d) be a CAT(κ)-space with $\kappa > 0$ and $r < R_\kappa/2$. For any $h > \tilde{h} > 0$ with $h \le \cos_\kappa r$, $\nu \in \mathbf{R}$ and $o \in Y$, the function $\Phi^{(\kappa)}_{\nu, \tilde{h}} : B(o, r) \times B(o, r) \to [0, \infty)$ given by*

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

is convex provided $2(2\nu + 1)\tilde{h}^2(h^2 - \tilde{h}^2) \ge 1$.

Kendall [Ke2] proved Theorem A for the unit sphere of the Euclidean space and remarked that it also holds for Riemannian manifolds. Jost [Jo2] gave an application of Theorem A. A detailed proof of Theorem A can be found in the appendix of [Yo].

We now state the main theorem of this paper. We say that a measure μ on a space X is *concentrated* on a subset $S \subset X$ if $\mu(X \setminus S) = 0$. We notice that $\mu \in \mathscr{P}_p(X)$ for any $p \in [1, \infty)$ if $\mu \in \mathscr{P}(X)$ is concentrated on a bounded subset of a metric space X. 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(κ)-space with $\kappa > 0$. Suppose $\mu \in \mathscr{P}(Y)$ is concentrated on a ball $B(o, r)$ with $o \in Y$ and $r < R_\kappa/2$. Then μ admits a p-barycenter for any $p \ge 1$, which is the unique p-barycenter in Y and the unique p-Karcher mean in $B(o, r)$ if $p \ge 2$. In particular, if $\mathrm{rad}(Y) < R_\kappa/2$ and $p \ge 2$, any $\mu \in \mathscr{P}(Y)$ admits a unique p-barycenter $b^p(\mu)$ in Y.*

This generalizes the main result of [Yo]. The upper bound $R_\kappa/2$ for the radius is almost sharp, cf. Remark 66 below. The combination of our result, i.e., Theorem B, Corollary 42 and Theorem 57 below, extends the result [Af, Theorem 2.1] of Afsari to general CAT(κ)-spaces.

In addition to Theorem B above, we also establish an analogue of the Banach–Saks–Kakutani type theorem for p-barycenter on CAT(κ)-spaces as Theorems C and D below. They extend the theorems of Jost [Jo, Theorem 2.2] and the author [Yo, Theorem C].

The structure of this paper is as follows: Section 2 consists of several definitions and properties of CAT-spaces. In Section 3, we prove propositions pertaining to the local convexity of CAT(1)-spaces, which might be of independent interest. We prove Theorem B in Section 4. Then Sections 5 and 6 are devoted to a collection of several properties of p-barycenter of probability measures on CAT(κ)-spaces, some of which might also be new on CAT(0)-spaces.

In this paper, we reuse almost all of the materials from our previous work [Yo]. For this reason, there must be substantial text overlap between them.

2. Preliminaries

In this section, we recall some rudimentary definitions and facts on the geometry of CAT-spaces. The textbook [BBI] by Burago–Burago–Ivanov is one of the standard references of the Alexandrov geometry. A reader who is familiar with them can safely skip this section.

Definition 2 (Convex function). Let (X, d) be a metric space. A *geodesic* is a curve $\gamma : I \to X$ defined on an interval $I \subset \mathbf{R}$ for which 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 $f : X \to \mathbf{R} \cup \{\infty\}$ is *convex* if 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 $\gamma_i : I \to Y_i$, $i = 1, 2$.

For a real number $\kappa \in \mathbf{R}$, we let (M_κ, d_κ) be the model surface, i.e., the simply-connected surface with the distance induced by the complete Riemannian metric of constant curvature κ. We will also use $(\mathbf{S}^2, d_{\mathbf{S}^2})$ instead of (M_1, d_1) later. We let $R_\kappa := \pi/\sqrt{\kappa}$ for $\kappa > 0$ and $R_\kappa := +\infty$ for $\kappa \leq 0$.

DEFINITION 3 (CAT(κ)-space). We call a metric space (Y,d) a *CAT(κ)-space* if it is an R_κ-geodesic space, i.e., any two points $x, y \in Y$ with $d(x,y) < R_\kappa$ are connected by a geodesic, and

$$d(x,\gamma(t)) \le d_\kappa(\bar{x},\bar{\gamma}(t))$$

holds for any three points $x,y,z \in Y$ with $d(x,y)+d(y,z)+d(z,x) < 2R_\kappa$, a geodesic $\gamma:[0,1] \to Y$ with $\gamma(0)=y$ and $\gamma(1)=z$ and $t \in [0,1]$. 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 $\bar{\gamma}(1)=\bar{z}$.

We persist in using the letter Y to denote a CAT-space. Unit spheres of Hilbert spaces and complete Riemannian manifolds with sectional curvature at most κ and injectivity radius at least R_κ are typical examples of CAT(κ)-spaces. CAT(κ)-spaces are also CAT(κ')-spaces for $\kappa' > \kappa$ and the upper curvature bound $\kappa \in \mathbf{R}$ of a CAT(κ)-space changes accordingly as its distance is rescaled by a positive number.

In this paper, we stick to the same notations as in [Yo], which we here recollect without giving precise definitions. In the rest of this section, (X,d) and (Y,d) denote a metric space and a CAT(κ)-space for some $\kappa \in \mathbf{R}$ respectively.

- $[x,y] := \{z \in X : d(x,z)+d(z,y) = d(x,y)\} \subset X$ for $x,y \in X$.
- $\gamma_{xy}:[0,1] \to Y$ denotes the unique geodesic with $\gamma_{xy}(0)=x$ and $\gamma_{xy}(1)=y$ for two points $x,y \in Y$ with $d(x,y) < R_\kappa$.
- $\tilde{\angle}_\kappa(x;y,z) \in [0,\pi]$ denotes the *comparison angle* for three points $x,y,z \in Y$. 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 $x \notin \{y,z\}$ and $d(x,y)+d(y,z)+d(z,x) < 2R_\kappa$, where $\cos_\kappa r := \cos(\sqrt{\kappa}\cdot r)$ and $\sin_\kappa r := \sin(\sqrt{\kappa}\cdot r)/\sqrt{\kappa}$ for $r \in \mathbf{R}$.
- $(\Sigma_x, \angle_x)$ 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*.
- $\uparrow_x^y \in \Sigma_x$ denotes the equivalence class of a 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 $x,y,z \in Y$ with $x \notin \{y,z\}$.
- $\log_x y := d(x,y)\cdot \uparrow_x^y \in C_x$ and $\log_x x := o_x \in C_x$ for $x,y \in Y$ with $x \ne y$.
- $|u| := |u - o_x|$ and $\langle u,v\rangle := (|u|^2+|v|^2-|u-v|^2)/2$ for vectors $u,v \in C_x$ at $x \in Y$.

- For a function φ defined on a neighborhood of $x \in Y$, $D\varphi[\log_x y] := (d/dt)^+|_{t=0}\varphi \circ \gamma_{xy}(t) \in \mathbf{R} \cup \{\pm\infty\}$ for $y \in Y$ with $0 < d(x, y) < R_\kappa$, if exists, denotes the *directional derivative*. If φ is locally Lipschitz at x, $D\varphi$ is extended to a Lipschitz function on $(C_x, |\cdot|)$.

We list some basic facts on CAT(κ)-spaces which we will make use of later.

Fact 4 (Angle monotonicity/comparison). *For any three points $x, y, z \in Y$ with $x \notin \{y, z\}$ and $d(x, y) + d(y, z) + d(z, x) < 2R_\kappa$ and a point $y' \in [x, y]\backslash\{x\}$,*

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

Fact 5 (Local uniform convexity). *For any $\kappa, r, \varepsilon > 0$ with $r < R_\kappa/2$, there is $\delta_\kappa(\varepsilon; r) > 0$ with*

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

for any $x \in Y$ and $y, z \in \bar{B}(x, r)$ with $d(y, z) \geq \varepsilon r$. Here $m(y, z) := \gamma_{yz}(1/2) \in Y$ is the midpoint of y and z.

It is known that $\delta_1(\varepsilon; r) = r - \arccos(\cos r/\cos(\varepsilon r/2))$ for any $\varepsilon > 0$ and $r < \pi/2$, e.g. Espínola–Fernández-León [EF]. Propositions 9 and 21 below also give estimates for $\delta_\kappa(\varepsilon; r)$.

The following fact is used along with Theorem A in our argument.

Fact 6 (First variation formula, cf. [BBI, Exercise 4.5.10]). *For any two geodesics $\lambda, \mu : [0, 1] \to Y$ representing $\lambda'(0+) \in C_x$ and $\mu'(0+) \in C_y$ with $x := \lambda(0)$, $y := \mu(0)$ and $d(x, y) < R_\kappa$ in (Y, d), we have*

$$\left.\frac{d^+}{dt}\right|_{t=0} d(\lambda(t), \mu(t)) = -\langle \lambda'(0+), \uparrow_x^y \rangle - \langle \mu'(0+), \uparrow_y^x \rangle.$$

For $\kappa \in \mathbf{R}$, we say that a subset $C \subset X$ of a metric space (X, d) is *R_κ-convex* if any geodesic connecting points $x, y \in C$ with $d(x, y) < R_\kappa$ does not leave C. For a subset $S \subset Y$ of a CAT(κ)-space Y, $\overline{\mathrm{conv}}(S) \subset Y$ denotes the *closed convex hull* of S, i.e., the smallest closed R_κ-convex subset containing S.

Fact 7 (Chebyshev property of convex subsets). *Suppose (Y, d) is complete. For any closed R_κ-convex subset $C \subset Y$ and a point $x \in Y$ of Y with $d(x, C) < R_\kappa/2$, there exists a unique point $\pi_C(x) \in C$ with $d(x, \pi_C(x)) = d(x, C)$. It also holds that $\tilde{\angle}_\kappa(\pi_C(x); x, c) \geq \angle_{\pi_C(x)}(x, c) \geq \pi/2$ for any $c \in C$ if they are defined.*

FACT 8 (e.g. Lytchak [Ly, Lemma 7.3]). *For a Lipschitz convex function* φ *defined on a neighborhood of a point* $x \in Y$, *there exists a vector* $\nabla_x^- \varphi \in C_x$ *with*

$$D\varphi[\eta] \geq -\langle \nabla_x^- \varphi, \eta \rangle \quad \text{for any } \eta \in C_x.$$

We call $\nabla_x^- \varphi$ *the* (negative) gradient *of* φ *at* x.

3. Local Convexity of CAT(1)-Spaces

In this section, we make a detour and discuss local p-uniform convexity of the distance function of CAT(1)-spaces. Propositions 9 and 21 below are the main result of this section. They are not used in our proof of Theorem B but might be of independent interest. A reader in a hurry can safely skip this section.

PROPOSITION 9 (p-uniform convexity of CAT(κ)-spaces, cf. Ohta [Oh]). *For any* $\kappa > 0$, $r < R_\kappa/2$ *and* $p \in (1, \infty)$, *there exists a constant* $k_p > 0$ *with the following property: Let* (Y, d) *be a* CAT(κ)*-space with* $\kappa > 0$. *Then*

$$d^p(x, \gamma_{yz}(t)) \leq (1-t)d^p(x, y) + td^p(x, z) - \frac{k_p}{2} t(1-t) d^{\max\{p,2\}}(y, z) \tag{10}$$

holds for any geodesic $\gamma_{yz} : [0, 1] \to Y$ *connecting* $y, z \in \bar{B}(x, r)$ *with* $x \in Y$ *and* $t \in [0, 1]$.

DEFINITION 11 (p-uniformly convex space, [NS], [Ku3, Ku2]). A geodesic space, i.e., an ∞-geodesic space, (X, d) is called a *p-uniformly convex space* for $p \geq 2$ if there exists a constant $c_p > 0$ for which

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

holds for any $x \in X$, a geodesic $\gamma : [0, 1] \to X$ with $y := \gamma(0)$ and $z := \gamma(1)$ and $t \in [0, 1]$.

COROLLARY 12. *Any* CAT(κ)*-space* (Y, d) *with* $\kappa > 0$ *and* diam $Y < R_\kappa/2$ *is a p-uniformly convex space for all* $p \in [2, \infty)$.

Ohta [Oh] proved Inequality (10) with $p = 2$ and the sharp constant $k_2 = 2r/\tan r$. We refer to Naor–Silberman [NS] and Kuwae [Ku2, Ku3] for p-uniformly convex spaces. It is not possible to improve the power $\max\{p, 2\}$ to p in Inequality (10), e.g. [NS], [Ku3]. Inequality (10) might be a candidate for a definition of p-uniformly convex spaces when $p < 2$, but it forces the space to have finite diameter.

Our proof of Proposition 9 is naturally divided into two cases. We only deal with the case $p \le 2$ here. The other case $p > 2$ follows from an argument in the proof of a more general result (Proposition 67), which we defer to the appendix.

We start with the following observation.

Lemma 13. *For any $p \in [1,2]$ and $r < \pi/2$, we have*

$$c_p^{\mathbf{S}} := 8 \inf_{\{x,y,z\}} \frac{d_{\mathbf{S}^2}^p(x,y) - d_{\mathbf{S}^2}^p(x,m(y,z))}{d_{\mathbf{S}^2}^2(y,z)} > 0, \tag{14}$$

where the infimum is taken over all $\{x,y,z\} \subset (\mathbf{S}^2, d_{\mathbf{S}^2})$ with $d_{\mathbf{S}^2}(x,y) = d_{\mathbf{S}^2}(x,z) \le r$ and $y \ne z$.

Proof. We mimic the argument of Ohta [Oh]. For $\{x,y,z\} \subset (\mathbf{S}^2, d_{\mathbf{S}^2})$ with $y \ne z$, we put

$$a := d_{\mathbf{S}^2}(x,y), \quad b := d_{\mathbf{S}^2}(x,z), \quad c := d_{\mathbf{S}^2}(y,z)/2, \quad d := d_{\mathbf{S}^2}(x,\gamma_{yz}(1/2))$$

and

$$f(a,b,c) := \frac{2}{c^2}\left(\frac{1}{2}a^p + \frac{1}{2}b^p - d^p\right) \ge 0.$$

The equality holds only if $p = 1$ and $\{x,y,z\}$ lies on a great circle.

If $a = b$, we know $d < a = b$ and $\cos a = \cos c \cos d$. As the function $a \mapsto a^{p-1}/\tan a$ is nonincreasing in a on $(0,\pi/2)$ if $p \le 2$, we have

$$\frac{\partial}{\partial a} f(a,a,c) = \frac{2p}{c^2} \tan a \left(\frac{a^{p-1}}{\tan a} - \frac{d^{p-1}}{\tan d}\right) < 0,$$

which implies $f(a,a,c) \ge f(r,r,c) > 0$ for any $a \le r$ and $c > 0$. Since

$$\lim_{c \to 0} f(r,r,c) = \frac{pr^{p-1}}{\tan r} \quad \text{and} \quad \lim_{c \to r} f(r,r,c) = \frac{2}{r^{2-p}},$$

we know that the infimum in (14) is positive. □

Proof of Proposition 9 for $p \le 2$. It suffices to prove Inequality (10) when $t = 1/2$, $\kappa = 1$ and (Y,d) is isometric to $(\mathbf{S}^2, d_{\mathbf{S}^2})$. We fix $x,y,z \in (\mathbf{S}^2, d)$ with $y,z \in \bar{B}(x,r)$ and put $w := m(y,z) \in \mathbf{S}^2$. The argument is divided into several cases.

If $d(x,w) < (1/2)(d(x,y) + d(x,z)) - (1/8)d(y,z)$, we have

$$d^p(x,w) + \frac{d^2(y,z)}{8^p(R_\kappa)^{2-p}} \le d^p(x,w) + \left(\frac{d(y,z)}{8}\right)^p$$
$$\le \left(d(x,w) + \frac{d(y,z)}{8}\right)^p$$
$$< \left(\frac{1}{2}(d(x,y)+d(x,z))\right)^p \le \frac{1}{2}(d^p(x,y)+d^p(x,z)).$$

If $(1/2)d(y,z) < |d(x,y) - d(x,z)|$, we use the following p-uniform convexity:

$$\left(\frac{a+b}{2}\right)^p + \frac{c_p^{\mathbf{R}}}{8}(a-b)^2 \le \frac{1}{2}(a^p+b^p) \quad \text{for any } 0 \le a,b \le \frac{R_\kappa}{2}$$

with $c_p^{\mathbf{R}} := p(p-1)(R_\kappa/2)^{p-2} > 0$. This yields

$$d^p(x,w) + \frac{c_p^{\mathbf{R}}}{32}d^2(y,z) < \left(\frac{1}{2}(d(x,y)+d(x,z))\right)^p + \frac{c_p^{\mathbf{R}}}{8}|d(x,y)-d(x,z)|^2$$
$$\le \frac{1}{2}(d^p(x,y)+d^p(x,z)).$$

We now deal with the remaining case. We may assume $d(x,y) \ge d(x,z)$. Let $E \subset \mathbf{S}^2$ be the great circle passing through w and perpendicular to $[x,w]$. We also let y' be the point in $E \cap [x,y]$ and $z' \in \mathbf{S}^2 \backslash \{z\}$ be the point for which $\{x,z',w\}$ is isometric to $\{x,z,w\}$. Then $\{w,y,y'\}$ is isometric to $\{w,z',y'\}$.

With the triangle inequality, the assumptions yields

$$2d(w,y') \ge d(w,y) + d(x,w) - d(x,y)$$
$$\ge \frac{d(y,z)}{2} + \frac{1}{2}(d(x,z)-d(x,y)) - \frac{d(y,z)}{8} \ge \frac{d(y,z)}{8},$$

while the choice of y' and z' yields

$$2d(x,y') \le d(x,y') + d(y',z') + d(x,z') = d(x,y) + d(x,z). \tag{15}$$

We combine them with Lemma 13 to conclude

$$d^p(x,w) + \frac{c_p^{\mathbf{S}}}{8}\left(\frac{d(y,z)}{8}\right)^2 \le d^p(x,w) + \frac{c_p^{\mathbf{S}}}{8}(2d(w,y'))^2$$
$$\le d^p(x,y')$$
$$\le \frac{1}{2}(d^p(x,y)+d^p(x,z)).$$

This completes the proof of Proposition 9 for $p \le 2$. □

Corollary 16 (p-variance inequality, cf. [NS, Ku2]). *Suppose* $\mu \in \mathscr{P}(Y)$ *is concentrated on* $S \subset Y$ *and its p-barycenter* $b^p(\mu)$ *lies in* $C := \bigcap_{s \in S} \bar{B}(s, r) \subset Y$ *for some* $r < R_\kappa/2$ *and* $p \in (1, \infty)$. *Then, with the constant* $k_p > 0$ *in Inequality* (10),

$$F_\mu^p(y) - F_\mu^p(b^p(\mu)) \geq \frac{k_p}{2p} d^{\max\{p,2\}}(y, b^p(\mu))$$

holds for any $y \in C$.

Proof. We choose $z := b^p(\mu)$ in Inequality (10). Then we divide it by $1 - t$ and let $t \to 1$ to obtain the desired inequality. □

Corollary 17 (cf. Kuwae [Ku2]). *Suppose* $C \subset Y$ *is a closed* R_κ*-convex subset and* $p \in (1, \infty)$. *Then, with the constant* $k_p > 0$ *in Inequality* (10) *for* $r < R_\kappa/2$,

$$d^p(x, y) - d^p(x, \pi_C(x)) \geq \frac{k_p}{2} d^{\max\{p,2\}}(y, \pi_C(x))$$

holds for any $x \in Y$ *and* $y \in C$ *with* $d(x, y) < r$.

Proof. The proof is essentially the same as that of Corollary 16. □

There is another notion of convexity of metric spaces.

Definition 18 (Uniform p-convex spaces, Foertsch [Fo], Kell [Kel]). Let (X, d) be a geodesic space. For $a, b \geq 0$ and $p \in [1, \infty)$, we put $\mathscr{M}_p(a, b) := ((a^p + b^p)/2)^{1/p}$ and $\mathscr{M}_\infty(a, b) := \max\{a, b\}$.

(1) We call (X, d) a *uniformly p-convex space* for $p \in (1, \infty]$ if there exists $\rho_p(\varepsilon) > 0$ for any $\varepsilon > 0$ with

$$d(x, m(y, z)) \leq (1 - \rho_p(\varepsilon)) \mathscr{M}_p(d(x, y), d(x, z)) \tag{19}$$

for any $x, y, z \in X$ with $d(y, z) > \varepsilon \mathscr{M}_p(d(x, y), d(x, z))$.

(2) We call (X, d) a *uniformly* 1*-convex space* if there exists $\rho_1(\varepsilon) > 0$ for any $\varepsilon > 0$ with Inequality (19) with $p = 1$ holds for any $x, y, z \in X$ with

$$d(y, z) > |d(x, y) - d(x, z)| + \varepsilon \mathscr{M}_1(d(x, y), d(x, z)). \tag{20}$$

Foertsch [Fo] investigated the above uniform 1- and ∞-convexity under the names *uniform distance* and *ball convexity*. Subsequently Kell [Kel] introduced the above uniform p-convexity for $p \in (1, \infty)$. He proved that uniformly p-convex spaces for some $p \geq 1$ are uniformly q-convex for all $q \in [p, \infty]$ and that CAT(0)-spaces are uniformly p-convex for all $p \in [1, \infty]$. He also remarked that p-uniformly convex spaces in the sense of Definition 11 are uniformly p-convex spaces in the sense of Definition 18 for any $p \in [2, \infty)$.

As for CAT(κ)-spaces, we can prove

PROPOSITION 21. *On any CAT(κ)-space (Y, d) with $\kappa > 0$ and for any $r < R_\kappa/2$, Inequality* (19) *holds with $p \in (1, \infty)$ for any $x, y, z \in Y$ with $y, z \in \bar{B}(x, r)$ and with $p = 1$ for any $x, y, z \in Y$ with $y, z \in \bar{B}(x, r)$ satisfying Inequality* (20). *In particular, any CAT(κ)-space Y with* diam $Y < R_\kappa/2$ *is a uniformly p-convex space in the sense of Definition* 18 *for all $p \in [1, \infty]$.*

PROOF. Our proof is similar to that of Proposition 9 for $p \leq 2$ presented above. It suffices to prove in the case (Y, d) is isometric to the unit sphere $(\mathbf{S}^2, d_{\mathbf{S}^2})$ and $p = 1$. For any three points $x, y, z \in (\mathbf{S}^2, d)$ satisfying Inequality (20), we suppose $d(x, y) \geq d(x, z)$ and put $w := m(y, z) \in \mathbf{S}^2$. We reuse the notations $y', z' \in \mathbf{S}^2$ used in our proof of Proposition 9 for $p \leq 2$.

We may assume

$$M := \mathscr{M}_1(d(x, y), d(x, z)) := \frac{1}{2}(d(x, y) + d(x, z)) \geq r/4. \tag{22}$$

If $M < r/4$, we have $\max\{d(x, y), d(x, z)\} < r/2$ and choose $\hat{y}, \hat{z} \in \mathbf{S}^2$ with

$$d(x, \hat{\star}) = 2d(x, \star) < r \quad \text{for } \star \in \{y, z\} \text{ and } d(\hat{y}, \hat{z}) = 2d(y, z).$$

Then the CAT(1)-inequality for $(\mathbf{S}^2, 2d)$ implies $2d(x, w) \leq d(x, \hat{w})$ with $\hat{w} := m(\hat{y}, \hat{z})$ and Inequality (19) for x, y, z follows from that for x, $\hat{y}$, $\hat{z}$.

We may also assume $d(x, w) > (1 - (\varepsilon/4))M$, because otherwise we have nothing to prove. Inequality (20) yields

$$2(d(w, y) - d(x, y)) = d(y, z) - d(x, y) + d(x, z) - 2M > (\varepsilon - 2)M$$

and hence by the triangle inequality we obtain

$$2d(w, y') \geq d(x, w) + d(w, y) - d(x, y) > \frac{\varepsilon}{4}M.$$

Inequality (15) means $d(x, y') \leq M$. Combining with Lemma 13 and Inequality (22), we conclude

$$d(x,w) \leq d(x,y') - \frac{c_1^{\mathbf{S}}}{8}(2d(w,y'))^2 \leq \left(1 - \frac{c_1^{\mathbf{S}} r}{512}\varepsilon^2\right)M.$$

The last statement of the proposition follows from [Kel, Lemma 1.4] or Proposition 9 and Fact 5. This completes the proof. □

4. Proof of Theorem B

In this section, we present a proof of Theorem B stated in Introduction after making some comment.

Theorem B is known for CAT(0)-spaces and other spaces, cf. Sturm [St], Naor–Silberman [NS], Kuwae [Ku2, Ku3]. In those cases, the proof relies on the convexity of the distance function of those spaces. We instead exploit Theorem A to prove Theorem B for CAT(κ)-spaces. Theorem B with $p = 2$ was proved in [Yo].

The following examples explain the subtlety of the uniqueness of p-barycenter when p is equal or close to 1.

Example 23. Let $x \neq y \in X$ be two points of a metric space (X, d). Suppose a probability measure $\mu \in \mathscr{P}_1(X)$ is concentrated on

$$\{z \in X : x \in [y,z] \text{ or } y \in [x,z]\}.$$

If x and y are 1-barycenters of μ, then so is any point $w \in [x, y] \subset X$. This happens for example when $\mu = (1/2)(\delta_x + \delta_y) \in \mathscr{P}(X)$.

Example 24 (e.g. Afsari [Af, Remark 2.4]). For four points $x_0, \ldots, x_3 \in (\mathbf{S}^2, d_{\mathbf{S}^2})$ with

$$r := d_{\mathbf{S}^2}(x_0, x_i) \quad \text{and} \quad D := d_{\mathbf{S}^2}(x_i, x_j)$$

for each $1 \leq i \neq j \leq 3$, we consider $\mu := (1/3)\sum_{i=1}^{3}\delta_{x_i} \in \mathscr{P}(\mathbf{S}^2)$. If p and r are close to 1 and $\pi/2$ respectively, we have $F_\mu^p(x_i) < F_\mu^p(x_0)$ for $i \neq 0$ and μ has at least three p-barycenters.

Now we begin our proof of Theorem B. Our proof is naturally divided into two parts.

4.1. Existence. We start with the existence of p-barycenter. For this, we prove the following more general theorem. Our proof was inspired by that of Kendall [Ke, Theorem 7.3] and is similar to that of [Yo, Theorem B].

THEOREM 25. *Suppose Y, $r < R_\kappa/2$ and $\mu \in \mathscr{P}(Y)$ are as in Theorem* B *in Introduction and $p \geq 1$. Then any sequence $(x_n)_{n\in\mathbf{N}}$ in Y with $F_\mu^p(x_n) \to \inf_Y F_\mu^p$ as $n \to \infty$ has a subsequence which converges to a p-barycenter of μ.*

We first prove the following lemma. Inequality (27) is similar to the definition of the weak convergence of Jost [Jo], cf. Lemma 29 below.

LEMMA 26. *Let (Y, d) be a complete $CAT(\kappa)$-space with $\kappa \in \mathbf{R}$. Suppose $\Phi(x_n, \cdot) : C \to [0, \infty)$ is a convex function on a closed R_κ-convex subset $C \subset B(o, r)$ with $o \in Y$ and $r < R_\kappa/2$ for all $n \in \mathbf{N}$ with*

$$\sup_{n\in\mathbf{N}, y\in C} \Phi(x_n, y) < \infty.$$

Then there exist an infinite subset $\mathscr{N} \subset \mathbf{N}$ and a point $x_\infty \in C$ with

$$\liminf_{\mathscr{N}\ni n\to\infty} \Phi(x_n, y) - \Phi(x_n, x_\infty) \geq 0 \tag{27}$$

for any $y \in C$.

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

$$\varphi_n := \inf_{\Lambda} \inf_{y\in C} \sup_{i\in\Lambda} \Phi(x_i, y),$$

where Λ runs over all infinite subsets of $\Lambda_{n-1} \setminus \{\min \Lambda_{n-1}\}$, and choose an infinite subset $\Lambda_n \subset \Lambda_{n-1} \setminus \{\min \Lambda_{n-1}\}$ for which

$$\varphi_n' := \inf_{y\in C} \sup_{i\in\Lambda_n} \Phi(x_i, y) \geq \varphi_n$$

satisfies $\varphi_n' - \varphi_n \to 0$ as $n \to \infty$. Then φ_n is nondecreasing in $n \in \mathbf{N}$ and hence the limit value

$$\varphi_\infty := \lim_{n\to\infty} \varphi_n = \lim_{n\to\infty} \varphi_n' \leq \sup_{n\in\mathbf{N}, y\in C} \Phi(x_n, y) < \infty$$

exists. We put

$$r_\infty := \inf_{(y_n)} \left\{ \liminf_{n\to\infty} d(o, y_n) : \sup_{i\in\Lambda_n} \Phi(x_i, y_n) \to \varphi_\infty \text{ as } n \to \infty \right\} \leq r,$$

where the infimum is taken over all such sequences $(y_n)_{n\in\mathbf{N}}$ in C.

Then there exists a sequence $(y_n)_{n\in\mathbf{N}}$ with

$$d(o, y_n) \to r_\infty \quad \text{and} \quad \sup_{i\in\Lambda_n} \Phi(x_i, y_n) \to \varphi_\infty \quad \text{as } n \to \infty.$$

It follows from Fact 5 that $(y_n)_{n\in\mathbf{N}}$ is a Cauchy sequence in $C \subset Y$ and hence converges in Y. The infinite subset $\mathcal{N} := \{\min \Lambda_n : n \in \mathbf{N}\}$ and the limit point $x_\infty := \lim_{n\to\infty} y_n \in C$ fulfill Inequality (27). This finishes the proof. □

Definition 28 (Weak convergence [Jo]). Let $(x_n)_{n\in\mathbf{N}}$ be a sequence of points in a CAT(κ)-space (Y, d) with $\limsup_{n\to\infty} d(x_n, x_\infty) < R_\kappa/2$ for some point $x_\infty \in Y$. We say that $(x_n)_{n\in\mathbf{N}}$ *converges weakly* to x_∞ if $\pi_\gamma(x_n) \to x_\infty$ as $n \to \infty$ for any geodesic $\gamma : [0, 1] \to Y$ with $\gamma(0) = x_\infty$. Here, $\pi_\gamma(x_n) \in \gamma([0, 1]) \subset Y$ denotes the closest point to x_n on the image of γ, cf. Fact 7.

The following is a Banach–Alaoglu type result for CAT(κ)-spaces.

Lemma 29 (cf. Jost [Jo, Theorem 2.1]). *Let (Y, d) be a complete CAT(κ)-space with $\kappa > 0$. Any sequence $(x_n)_{n\in\mathbf{N}}$ of points in $B(o, r)$ with $o \in Y$ and $r < R_\kappa/2$ has a subsequence which converges weakly to a point in Y.*

A proof of this lemma can be found in e.g. [Yo]. As hinted above, Lemma 29 follows from Lemma 26. For reader's convenience, we give a proof here.

Proof of Lemma 29. We apply Lemma 26 with

$$\Phi(x_n, \cdot) := d(x_n, \cdot) \quad \text{and} \quad C := \bigcap_{n\in\mathbf{N}} \bar{B}(x_n, R_\kappa/2) \cap \bar{B}(o, r)$$

to obtain a subsequence, still denoted $(x_n)_{n\in\mathbf{N}}$, and a point $x_\infty \in C$ for which

$$\liminf_{n\to\infty} d(x_n, y) - d(x_n, x_\infty) \geq 0 \tag{30}$$

holds for any $y \in C$. This yields $\limsup_{n\to\infty} d(x_n, x_\infty) < R_\kappa/2$.

We now suppose that there is a geodesic $\gamma : [0,1] \to Y$ with $\gamma(0) = x_\infty$ and

$$\limsup_{n\to\infty} d(x_\infty, \pi_\gamma(x_n)) > 0.$$

Then by Inequality (30) and Fact 5 the midpoint $w_n := m(x_\infty, \pi_\gamma(x_n)) \in \gamma((0,1))$ of x_∞ and $\pi_\gamma(x_n)$ satisfies

$$d(x_n, w_n) < d(x_n, \pi_\gamma(x_n))$$

for some large $n \gg 1$ and this is a contradiction. □

We will later use the following fact, which follows from Fact 7.

FACT 31. *For any sequence $(x_n)_{n\in\mathbf{N}}$ which converges weakly to $x_\infty \in Y$ in a $CAT(\kappa)$-space (Y,d) with $\limsup_{n\to\infty} d(x_n, x_\infty) < R_\kappa/2$,*

$$\liminf_{n\to\infty} d(x_n, y) > \liminf_{n\to\infty} d(x_n, x_\infty) \quad \text{and} \quad \liminf_{n\to\infty} d(x_n, y) \geq d(x_\infty, y)$$

hold for any point $y \in B(x_\infty, R_\kappa/2)\backslash\{x_\infty\}$.

We also invoke the following lemma.

LEMMA 32 (Ekeland principle, e.g. Ekeland [Ek]). *Let $f : X \to \mathbf{R}$ 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*

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

PROOF OF THEOREM 25. We recall that $\mu \in \mathscr{P}(Y)$ 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 $F := F_\mu^p$ attains its minimum for $p \geq 1$. According to Theorem A, the function $\Phi := \Phi_{\nu,\tilde{h}}^{(\kappa)} : B(o,r) \times B(o,r) \to [0,\infty)$ with appropriate $\tilde{h} < h := \cos_\kappa r$ and $\nu \in \mathbf{R}$ is convex.

We start with the following observations. Similar claims are verified in [Yo] when $p = 2$ and their proofs can be easily adapted to our case $p \geq 1$.

Claim 33 ([Yo, Claim 12], cf. Afsari [Af], Claim 59 below). *For any* $\varepsilon > 0$, *there exists* $\delta = \delta(\varepsilon) > 0$ *with*

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

Claim 34 ([Yo, Claim 13]). *There exist* $r' \in (0, r)$ *and* $\delta' > 0$ *with*

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

We appeal to Lemma 32 to find a sequence $(z_n)_{n \in \mathbf{N}} \subset Y$ for which $d(x_n, z_n) \to 0$ as $n \to \infty$ and

$$F(y) \geq F(z_n) - \frac{1}{n} \cdot d(y, z_n) \quad \text{for any } y \in Y \text{ and } n \in \mathbf{N}.$$

By the choice of z_n, we have $F(z_n) \to \inf_Y F$ as $n \to \infty$ and

$$DF[\xi] = -\int_Y \langle \xi, \log_{z_n} y \rangle d^{p-2}(y, z_n)\, d\mu(y) \geq -\frac{1}{n} |\xi| \tag{35}$$

for any $\xi \in C_{z_n}$. Then Claims 33 and 34 imply $\limsup_{n \to \infty} d(o, z_n) \leq r' < r$.

Lemma 26 states that there is a subsequence, still denoted $(z_n)_{n \in \mathbf{N}}$, and a point $z_\infty \in \bar{B}(o, r')$ for which Inequality (27) holds. We intend to prove that a subsequence of $(z_n)_{n \in \mathbf{N}}$ converges to z_∞ and thus assume that this is not the case. Inequality (27) allows us to take a further subsequence with

$$\inf_{m \neq n \in \mathbf{N}} \Phi(z_m, z_n) > \frac{1}{2} \limsup_{n \to \infty} \Phi(z_n, z_\infty) > 0$$

and hence $\inf_{m \neq n \in \mathbf{N}} d(z_m, z_n) > 2\delta$ for some small $\delta > 0$. Then the collection $\{B(z_n, \delta)\}_{n \in \mathbf{N}}$ of the balls is mutually disjoint and $\mu(B(z_n, \delta)) \to 0$ as $n \to \infty$. We put $M := \max\{\delta^{p-2}, (R_\kappa)^{p-2}\} < \infty$.

We fix $\varepsilon > 0$ and put $y_\varepsilon \in [z_\infty, y]$ as the point with $d(z_\infty, y_\varepsilon) = \varepsilon d(z_\infty, y)$ for each $y \in B(o, r)$. The map $y \mapsto y_\varepsilon$ is continuous on $B(o, r)$.

We then use the convexity of Φ and Fact 6 to derive for any $y \in B(o, r)$

$$\begin{aligned} \Phi(y, y) - \Phi(z_n, y_\varepsilon) &\geq D\Phi[\log_{(z_n, y_\varepsilon)}(y, y)] \\ &= D\Phi(z_n, \cdot)[\log_{y_\varepsilon} y] + D\Phi(\cdot, y_\varepsilon)[\log_{z_n} y]. \end{aligned}$$

We put $d_\delta(\cdot, \cdot) := \chi_{[\delta, \infty)}(d(\cdot, \cdot)) d(\cdot, \cdot)$, where $\chi_{[\delta, \infty)}(s) := \delta_s([\delta, \infty))$ with $\delta_s \in \mathscr{P}(\mathbf{R})$ being the Dirac measure centered at $s \in \mathbf{R}$. We shall estimate the integrals of the above two terms multiplied by $d_\delta^{p-2}(z_n, y)$.

First we have

$$\begin{aligned}
&\int_Y D\Phi(z_n,\cdot)[\log_{y_\varepsilon} y] d_\delta^{p-2}(z_n,y)\,d\mu(y)\\
&\quad\geq \frac{1-\varepsilon}{\varepsilon}\int_Y (\Phi(z_n,y_\varepsilon)-\Phi(z_n,z_\infty))d(z_\infty,y)d_\delta^{p-2}(z_n,y)\,d\mu(y)\\
&\quad\geq \frac{1-\varepsilon}{\varepsilon}\int_Y \min\{\Phi(z_n,y_\varepsilon)-\Phi(z_n,z_\infty),0\}d(z_\infty,y)d_\delta^{p-2}(z_n,y)\,d\mu(y),
\end{aligned}$$

with which the dominated convergence theorem yields

$$\liminf_{n\to\infty}\int_Y D\Phi(z_n,\cdot)[\log_{y_\varepsilon} y] d_\delta^{p-2}(z_n,y)\,d\mu(y)\geq 0.$$

In the following, $C<\infty$ denotes a fixed large constant depending only on κ, r and p. For example, we have

$$|D\Phi(\cdot,y_\varepsilon)[\log_{z_n} y]-D\Phi(\cdot,z_\infty)[\log_{z_n} y]|\leq C\varepsilon$$

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

$$\int_{B(z_n,\delta)} D\Phi(\cdot,z_\infty)[\log_{z_n} y]d^{p-2}(z_n,y)\,d\mu(y)\leq C\mu(B(z_n,\delta))\delta^{p-1}.$$

Second we have

$$\begin{aligned}
&\int_Y D\Phi(\cdot,y_\varepsilon)[\log_{z_n} y]d_\delta^{p-2}(z_n,y)\,d\mu(y)+C(\mu(B(z_n,\delta))+M\varepsilon)\\
&\quad\geq \int_Y D\Phi(\cdot,z_\infty)[\log_{z_n} y]d^{p-2}(z_n,y)\,d\mu(y)\\
&\quad\geq -\int_Y \langle\nabla_{z_n}^-\Phi(\cdot,z_\infty),\log_{z_n} y\rangle d^{p-2}(z_n,y)\,d\mu(y)\\
&\quad\geq -\frac{1}{n}|\nabla_{z_n}^-\Phi(\cdot,z_\infty)|\\
&\quad\geq -\frac{C}{n},
\end{aligned}$$

which yields

$$\liminf_{n\to\infty}\int_Y D\Phi(\cdot,y_\varepsilon)[\log_{z_n} y]d_\delta^{p-2}(z_n,y)\,d\mu(y)\geq -CM\varepsilon.$$

Therefore we conclude

$$\limsup_{n\to\infty} \Phi(z_n, z_\infty) \int_Y d_\delta^{p-2}(z_n, y)\, d\mu(y)$$

$$\leq \limsup_{n\to\infty} \int_Y \Phi(z_n, y_\varepsilon) d_\delta^{p-2}(z_n, y)\, d\mu(y) + CM\varepsilon \leq 2CM\varepsilon.$$

Since $\varepsilon > 0$ is arbitrarily and

$$\int_Y d_\delta^{p-2}(z_n, \cdot)\, d\mu \geq \min\{\delta^{p-2}, (R_\kappa)^{p-2}\}(1 - \mu(B(z_n, \delta))) > 0,$$

we conclude that $(z_n)_{n\in\mathbf{N}}$ and hence $(x_n)_{n\in\mathbf{N}}$ converge to $z_\infty \in B(o, r)$ and thus

$$F(z_\infty) = \lim_{n\to\infty} F(x_n) = \inf_Y F,$$

which means that z_∞ is a p-barycenter of μ.

Now the proof of Theorem 25 is complete. □

4.2. Uniqueness. We now proceed to the uniqueness part of Theorem B. For this, we prove the following more general theorem.

THEOREM 36. *Suppose Y, $r < R_\kappa/2$ and $\mu \in \mathscr{P}(Y)$ are as in Theorem* B *in Introduction and $p \geq 2$. Then a point $z \in B(o, r)$ with*

$$DF_\mu^p[\xi] \geq 0 \quad \text{for any } \xi \in C_z \tag{37}$$

is the unique p-barycenter of μ. In particular, the p-barycenter $b^p(\mu)$ of μ is unique if $p \geq 2$.

To prove this, we need the following result from [Yo] for barycenter of probability measures on CAT(κ)-spaces.

PROPOSITION 38 (Variance inequality [Yo, Proposition 19]). *Suppose (Y, d) and $\mu \in \mathscr{P}(Y)$ are as in Theorem 36. Let $b(\mu) := b^2(\mu) \in B(o, r)$ be the barycenter of μ. For any $x \in B(o, r)$, we have*

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

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

PROPOSITION 39. *Suppose* (Y, d) *and* $\mu \in \mathscr{P}(Y)$ *are as in Theorem* 36. *If a point* $z \in B(o, r)$ *satisfies Inequality* (37) *and* $E_\mu^{p-2}(z) := \int_Y d^{p-2}(z, \cdot)\, d\mu \in (0, \infty)$, *then* z *is the barycenter of the weighted probability measure* $\tilde{\mu} := (E_\mu^{p-2}(z))^{-1} d^{p-2}(z, \cdot)\mu \in \mathscr{P}(Y)$.

PROOF. By assumption, we have

$$DF_{\tilde{\mu}}^2[\xi] = -\int_Y \langle \xi, \log_z y \rangle \, d\tilde{\mu}(y) = (E_\mu^{p-2}(z))^{-1} DF_\mu^p[\xi] \geq 0$$

for any $\xi \in C_z$. It follows from the characterization of the barycenter established in [Yo, Corollary 15] that z is the barycenter of $\tilde{\mu}$. □

PROOF OF THEOREM 36. We may assume that μ is not a Dirac measure. Hölder's inequality yields

$$\left(\int_Y d^p(x, \cdot)\, d\mu\right)^{2/p} \left(\int_Y d^p(z, \cdot)\, d\mu\right)^{(p-2)/p} - \int_Y d^p(z, \cdot)\, d\mu$$
$$\geq E_\mu^{p-2}(z) \int_Y d^2(x, \cdot) - d^2(z, \cdot)\, d\tilde{\mu}$$

for any $x \in B(o, r)$, where $\tilde{\mu}$ is the probability measure defined in Proposition 39. Then, Propositions 38 and 39 yield

$$\left(\int_Y d^p(x, \cdot)\, d\mu\right)^{2/p} - \left(\int_Y d^p(z, \cdot)\, d\mu\right)^{2/p} \geq c E_\mu^{p-2}(z) \left(\int_Y d^p(z, \cdot)\, d\mu\right)^{(2-p)/p} d^\alpha(x, z) \tag{40}$$

for any $x \in B(o, r)$. Combined with Claims 33 and 34, this implies that $z \in B(o, r)$ is the unique p-barycenter of μ. □

4.3. The Other Cases. As for p-barycenter of probability measures on CAT(1)-spaces with $p \in [1, 2)$, we can prove the following, cf. Afsari [Af].

THEOREM 41. *Let* (Y, d) *be a complete* $CAT(\kappa)$*-space with* $\kappa > 0$. *Suppose* $\mu \in \mathscr{P}(Y)$ *is concentrated on a subset* $S \subset B(o, r)$ *of* $\operatorname{diam}(S) \leq R_\kappa/2$ *with* $o \in Y$ *and* $r < R_\kappa/2$. *For an increasing convex function* $U : [0, \infty) \to [0, \infty)$, *consider the function* $F(x) := \int_Y U(d(x, \cdot))\, d\mu$ *for* $x \in Y$. *If* U *is not strictly convex, assume* *also that μ is not concentrated on the union of images of geodesics passing through two points* (*cf. Example* 23). *Then F admits a unique minimizer in Y, which is also a unique local minimizer of F in $B(o,r)$.*

Corollary 42. *Let (Y,d) be as in Theorem* 41 *and $p \in [1,2)$. Suppose $\mu \in \mathscr{P}(Y)$ is concentrated on $B(o,r)$ with $o \in Y$ and $r < R_\kappa/4$ and also assume that μ is not concentrated on the union of geodesics passing through two points if $p = 1$. Then μ admits a unique p-barycenter $b^p(\mu)$ in Y, which is also a unique p-Karcher mean of μ in $B(o,r)$.*

Proof of Theorem 41. We first notice that $C := \overline{\mathrm{conv}}(S \cup \{o\}) \subset Y$ is a closed R_κ-convex subset with

$$S \subset C \subset \bigcap_{x \in C} \bar{B}(x, R_\kappa/2) \cap \bar{B}(o,r).$$

Then it follows that $F|_C : C \to [0,\infty)$ is a convex function. Indeed

$$U(d(x,w)) \le U\left(\frac{1}{2}(d(x,y) + d(x,z))\right) \le \frac{1}{2}(U(d(x,y)) + U(d(x,z))) \tag{43}$$

for any $x \in Y$ and $y \ne z \in \bar{B}(x, R_\kappa/2)$ with $w := m(y,z) \in \bar{B}(x, R_\kappa/2)$ being a midpoint of y, z with equalities only if either $d(x,y) = d(x,z) \in \{0, R_\kappa/2\}$ or U is not strictly convex and $\{x,y,z\}$ is on a geodesic. This yields $F(w) < (1/2)(F(y) + F(z))$ for any $y \ne z \in C$ by assumption and hence the uniqueness of a minimizer of $F|_C$.

It is easy to check that $F(x) > \inf_C F$ for any $x \in Y \backslash C$. Indeed, we have $F(x) \ge U(R_\kappa/2) > F(o)$ if $d(x,C) \ge R_\kappa/2$ and $F(x) > F(\pi_C(x))$ by Fact 7 if $0 < d(x,C) < R_\kappa/2$. Now the existence of a minimizer of $F|_C$ and hence of F follows from e.g. [Yo, Theorem E].

If $x \in B(o,r) \backslash C$ and $x' \in [x, \pi_C(x)] \backslash \{x\}$, then by Facts 4 and 7 we have $d(x',y) < d(x,y)$ for any $y \in C$ and hence $F(x') < F(x)$, which means that x is not a local minimizer of F and a local minimizer of F in $B(o,r)$ is a minimizer of F.

Now the proof of Theorem 41 is complete. □

The following proposition characterizes 1-barycenter.

PROPOSITION 44 (cf. Yang [Ya, Theorem 2.2]). *Let* (Y,d) *be a* $CAT(\kappa)$*-space with* $\kappa \in \mathbf{R}$. *Suppose* $\mu \in \mathscr{P}(Y)$ *is concentrated on a subset* $S \subset Y$. *Define*

$$H(z) := \sup_{\xi \in \Sigma_z} \int_{Y \setminus \{z\}} \langle \xi, \uparrow_z^y \rangle \, d\mu(y) = -\inf_{\xi \in \Sigma_z} DF_\mu^1[\xi]$$

for $z \in Y$ *with* $S \subset B(z, R_\kappa)$. *Then* z *satisfies* $DF_\mu^1[\xi] \geq 0$ *for any* $\xi \in C_z$ *if and only if* $H(z) \leq \mu(\{z\})$.

In particular, if (Y,d) *and* $\mu \in \mathscr{P}(Y)$ *are as in Theorem* 41, *then* $z \in B(o,r)$ *is a* 1*-barycenter of* μ *if and only if* $H(z) \leq \mu(\{z\})$.

PROOF. We set $F := F_\mu^1$. If $DF[\xi] \geq 0$ for any $\xi \in C_z$, then we have $H(z) \leq 0 \leq \mu(\{z\})$. For a fixed $x \in Y$ in a neighborhood of z and any $x' \in [x,z]$ with $\varepsilon := d(x',z) > 0$, Fact 6 and the dominated convergence theorem yield

$$\begin{aligned} F(x') - \varepsilon\mu(\{z\}) &= \int_{Y \setminus \{z\}} d(x', \cdot) \, d\mu \\ &= F(z) + \varepsilon DF[\uparrow_z^x] + o(\varepsilon) \\ &\geq F(z) - \varepsilon H(z) + o(\varepsilon), \end{aligned}$$

where $o(\varepsilon)/\varepsilon \to 0$ as $\varepsilon \to 0$. This proves the proposition. □

DEFINITION 45. We define an ∞*-barycenter* of a probability measure $\mu \in \mathscr{P}(X)$ on a metric space (X,d) as a point where the function

$$x \mapsto \operatorname*{ess\,sup}_X d(x,\cdot) := \inf\left\{ \sup_{X \setminus N} d(x,\cdot) : N \subset X \text{ with } \mu(N) = 0 \right\}$$

attains its minimum.

The definition and proof of the unique existence of ∞-barycenter is essentially the same as those of circumcenter of subsets of $\mathrm{CAT}(\kappa)$-spaces.

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 Fact 5 that any subset $A \subset Y$ of a complete $\mathrm{CAT}(\kappa)$-space (Y,d) with $\kappa \in \mathbf{R}$ 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 [BL].

PROPOSITION 46. *Let (Y, d) be a complete CAT(κ)-space with $\kappa \in \mathbf{R}$. Suppose $\mu \in \mathscr{P}(Y)$ is concentrated on a subset $S \subset Y$ with $\mathrm{rad}_Y(S) < R_\kappa/2$. Then μ admits a unique ∞-barycenter $b^\infty(\mu)$ in Y and $b^\infty(\mu)$ is contained in the closed convex hull $\overline{\mathrm{conv}}(S) \subset Y$ of S.*

We omit the proof of this proposition.

5. Properties of p-Barycenter

In this section, we establish several properties of p-barycenter of probability measures on CAT(κ)-spaces with $\kappa > 0$, which we proved to exist in Theorem B. We exploit Theorem A in our argument here as well.

A number of properties of barycenter of probability measures on CAT(0)-spaces are known, e.g. Sturm [St]. We also add that Ohta [Oh2] investigated barycenter of probability measures on proper Alexandrov spaces of curvature $\geq \kappa$. A couple of properties of barycenter on CAT(κ)-spaces are established in [Yo]. Our results in this section extend some of them to the context of p-barycenter on CAT(κ)-spaces. We do not attempt to exhaust such possible extensions. Some of them might be new on CAT(0)-space as well.

Throughout this section, we usually assume the following unless otherwise stated.

ASSUMPTION 47. • *(Y, d) stands for a complete CAT(κ)-space with $\kappa > 0$.*
• *$\mu \in \mathscr{P}(Y)$ is a probability measure concentrated on $B(o, r)$ with $o \in Y$ and $r < R_\kappa/2$ and hence it admits a p-barycenter $b^p(\mu) \in B(o, r)$ for $p \in [1, \infty]$.*
• *$\Phi := \Phi^{(\kappa)}_{\nu, \tilde{h}} : \bar{B}(o, r) \times \bar{B}(o, r) \to [0, \infty)$ is the convex function in Theorem A extended to the closure of the domain with suitable parameters $\nu > -1/2$ and $\tilde{h} > 0$ with $\tilde{h} < h := \cos_\kappa r$.*

We remark that a simple estimate says

$$C_1 d^\beta(x, y) \leq \Phi(x, y) \leq C_2 d^\beta(x, y) \tag{48}$$

for any $x, y \in B(o, r)$, where $\beta := 2(\nu + 1) > 1$,

$$C_1 := \left(\frac{4}{\pi^2(1 - \tilde{h}^2)}\right)^{\nu+1} \quad \text{and} \quad C_2 := \left(\frac{1}{2(h^2 - \tilde{h}^2)}\right)^{\nu+1}.$$

5.1. Variance Inequality.

PROPOSITION 49 (p-variance inequality). *Suppose* (Y, d) *and* $\mu \in \mathcal{P}(Y)$ *are as in Assumption* 47. *Let* $b^p(\mu) \in B(o, r)$ *be the* p*-barycenter of* μ *for* $p \geq 2$. *Then*

$$F_\mu^p(y) - F_\mu^p(b^p(\mu)) \geq c \cdot d^{\max\{p, \alpha\}}(y, b^p(\mu))$$

holds for any $y \in B(o, r)$, *where* $c > 0$ *is a constant depending only on* κ, r *and* p *and* $\alpha > 2$ *is from Proposition* 38.

For the proof, we need

LEMMA 50 (cf. Ohta–Palfia [OP]). *For any* $\kappa > 0$, $r < R_\kappa/2$ *and* $p > 1$, *there exists a constant* $K_p \leq 0$ *with*

$$d^p(x, \gamma_{yz}(t)) \leq (1-t)d^p(x, y) + td^p(x, z) - \frac{K_p}{2} t(1-t)d^2(y, z)$$

for any $x, y, z \in B(o, r)$ *with* $o \in Y$ *and* $t \in [0, 1]$.

PROOF. It suffices to prove this when (Y, d) is isometric to $(\mathbf{S}^2, d_{\mathbf{S}^2})$. The proposition follows from the C^2 property of $d_{\mathbf{S}^2}^p(x, \cdot)$ on $B(x, \pi) \subset \mathbf{S}^2$ if $p \geq 2$ and from Proposition 9 and the C^2 property of $d_{\mathbf{S}^2}^p(x, \cdot)$ on $B(x, \pi) \setminus \{x\} \subset \mathbf{S}^2$ if $p < 2$. □

PROOF OF PROPOSITION 49. We fix $p \geq 2$ and put $z := b^p(\mu)$. We choose small $\varepsilon > 0$ with

$$k_p(1-\varepsilon) + K_p(R_\kappa)^{2-p}\varepsilon \geq k_p/2, \tag{51}$$

where $k_p > 0$ and $K_p \leq 0$ are the constants from Proposition 9 and Lemma 50 respectively.

Since

$$a^{p/2} - b^{p/2} \geq \frac{p}{2} b^{(p/2)-1}(a-b) \quad \text{for any } a \geq b \geq 0,$$

Inequality (40) yields

$$\begin{aligned}\int_Y d^p(y, \cdot)\, d\mu - \int_Y d^p(z, \cdot)\, d\mu &\geq \frac{p}{2} E_\mu^{p-2}(z) \cdot cd^\alpha(y, z) \\ &\geq \frac{p}{2(R_\kappa)^2} \int_Y d^p(z, \cdot)\, d\mu \cdot cd^\alpha(y, z).\end{aligned}$$

If $\int_Y d^p(z,\cdot)\,d\mu \geq \varepsilon^{p+1}$, we derive the desired inequality from this one. Otherwise, Chebyshev's inequality yields $\mu(B(z,\varepsilon)) > 1-\varepsilon$. Then

$$\int_Y d^p(y,\cdot)\,d\mu > (1-\varepsilon)\varepsilon^p > \int_Y d^p(z,\cdot)\,d\mu + (1-2\varepsilon)\left(\frac{\varepsilon}{R_\kappa}d(y,z)\right)^p$$

holds for any $y \in B(o,r)\backslash B(z,2\varepsilon)$. The combination of Proposition 9, Lemma 50 and Inequality (51) yields

$$\begin{aligned}\int_Y & d^p(x,\gamma_{yz}(t))\,d\mu(x) \\ &< (1-t)\int_Y d^p(x,y)\,d\mu(x) + t\int_Y d^p(x,z)\,d\mu(x) - \frac{k_p}{4}t(1-t)d^p(y,z)\end{aligned}$$

for any $y \in B(o,r)\cap B(z,2\varepsilon)$. We then divide this inequality by $1-t$ and let $t\to 1$ to obtain

$$\int_Y d^p(z,\cdot)\,d\mu \leq \int_Y d^p(y,\cdot)\,d\mu - \frac{k_p}{4}d^p(y,z).$$

Now the proof is complete. □

Remark 52. In the situation of Proposition 49, Hölder's inequality yields

$$\frac{\int_Y d^{p-2}(z,\cdot)\,d\mu}{(\int_Y d^p(z,\cdot)\,d\mu)^{(p-2)/p}} \geq \frac{1}{(R_\kappa)^2}\left(\int_Y d^p(z,\cdot)\,d\mu\right)^{2/p} \geq \frac{1}{(R_\kappa)^2}\int_Y d^2(z,\cdot)\,d\mu$$

and hence Inequality (40) yields a useful inequality

$$\begin{aligned}&\left(\int_Y d^p(y,\cdot)\,d\mu\right)^{2/p} - \left(\int_Y d^p(z,\cdot)\,d\mu\right)^{2/p} \\ &\qquad \geq \frac{c}{(R_\kappa)^2}\int_Y d^2(z,\cdot)\,d\mu\cdot d^\alpha(y,z)\end{aligned} \tag{53}$$

for any $y\in B(o,r)$, where $c>0$ and $\alpha>2$ are the constants in Proposition 38 and hence independent of p.

5.2. Continuity of p-Barycenter. We here investigate the behaviour of p-barycenter when the probability measure and p vary.

For probability measures $\mu,\nu \in \mathscr{P}_p(X)$ on a metric space (X,d),

$$W_p(\mu,\nu) := \inf_\pi\left(\int_{X\times X} d^p(x,y)\,d\pi(x,y)\right)^{1/p}$$

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

It is known that $W_p(\mu_n, \mu) \to 0$ as $n \to \infty$ if and only if $(\mu_n)_{n \in \mathbf{N}}$ converges weakly to μ and $F^p_{\mu_n}(x) \to F^p_\mu(x)$ as $n \to \infty$ for any $x \in X$ on a complete separable metric space (X, d). In general we still have

$$\int_X d^p(x, y)\, d\mu(y) \leq (1 + \varepsilon) \int_X d^p(x, z)\, d\nu(z) + C_\varepsilon \int_{X \times X} d^p(y, z)\, d\pi(y, z)$$

for any $\varepsilon > 0$ with some $C_\varepsilon < \infty$, $x \in X$ and any coupling $\pi \in \mathscr{P}(X \times X)$ of μ and $\nu \in \mathscr{P}_p(X)$. This implies that $F^p_{\mu_n}(x) \to F^p_\mu(x)$ for all $x \in X$ and $p \geq 1$ if $W_p(\mu_n, \mu) \to 0$ as $n \to \infty$, cf. Villani [Vi, Theorem 6.9].

THEOREM 54. *Let (Y, d) and $\mu \in \mathscr{P}(Y)$ be as in Assumption 47. Suppose sequences $(\mu_n)_{n \in \mathbf{N}} \subset \mathscr{P}(Y)$ and $(p_n)_{n \in \mathbf{N}} \subset [1, \infty)$ of probability measures concentrated on $B(o, r)$ and of real numbers satisfy $W_1(\mu_n, \mu) \to 0$ and $p_n \to p$ as $n \to \infty$ for some $p \in [1, \infty)$. Then any sequence $(z_n)_{n \in \mathbf{N}}$ of p_n-barycenter of μ_n has a subsequence which converges to a p-barycenter of μ. In particular, if in addition μ admits a unique p-barycenter $b^p(\mu) \in Y$, the original sequence $(z_n)_{n \in \mathbf{N}}$ converges to $b^p(\mu)$.*

PROOF. Our proof is similar to that of Theorem 25. We set $F_n := F^{p_n}_{\mu_n}$.

CLAIM 55. *If $F_n(z_n) \to 0$ as $n \to \infty$, then μ is a Dirac measure centered at a point $z \in B(o, r)$ and $(z_n)_{n \in \mathbf{N}}$ converges to $z = b^p(\mu)$.*

PROOF. The triangle inequality yields

$$\begin{aligned} d(z_m, z_n) &\leq \int_{Y \times Y} [d(x, y) + d(z_m, x) + d(z_n, y)]\, d\pi(x, y) \\ &= \int_{Y \times Y} d(\cdot, \cdot)\, d\pi + \int_Y d(z_m, \cdot)\, d\mu_m + \int_Y d(z_n, \cdot)\, d\mu_n \end{aligned}$$

for any coupling $\pi \in \mathscr{P}(Y \times Y)$ of μ_m and μ_n. Since Hölder's inequality yields

$$\left(\int_Y d(z_n, \cdot)\, d\mu_n \right)^{p_n} \leq p_n F_n(z_n) \to 0 \quad \text{as } n \to \infty,$$

$(z_n)_{n\in\mathbf{N}}$ is a Cauchy sequence and hence converges to a point $z \in Y$. It follows that $\mu = \delta_z$ and hence $b^p(\mu) = z$. This confirms the claim. □

Claim 55 allows us to assume $\liminf_{n\to\infty} F_n(z_n) > 0$. We set $p_n^i := p_n + (1/i)$ and $F_n^i := F_{\mu_n}^{p_n^i}$ for $i, n \in \mathbf{N}$. Then Hölder's inequality yields

$$F_n^i(z_n) - \inf_Y F_n^i < \frac{1}{p_n^i}\left[(R_\kappa)^{1/i} - \left(\int_Y d^{p_n}(z_n,\cdot)\, d\mu_n\right)^{1/ip_n}\right]\int_Y d^{p_n}(z_n,\cdot)\, d\mu_n \le D_i$$

for some $D_i < \infty$ with $D_i \to 0$ as $i \to \infty$.

We fix $\varepsilon_i > 0$ with $\varepsilon_i \to 0$ and $D_i/\varepsilon_i \to 0$ as $i \to \infty$. By appealing to Lemma 32, we find $z_n^i \in B(o,r)$ with $d(z_n^i, z_n) \le D_i/\varepsilon_i$ and

$$F_n^i(y) \ge F_n^i(z_n^i) - \varepsilon_i d(y, z_n^i)$$

for any $y \in Y$ and $i, n \in \mathbf{N}$.

Lemma 26 states that for any $i \in \mathbf{N}$ there exist an infinite subset $\mathcal{N}_i \subset \mathbf{N}$ with $\mathcal{N}_{i+1} \subset \mathcal{N}_i \backslash \{\min \mathcal{N}_i\}$ and $z_\infty^i \in Y$ with

$$\liminf_{\mathcal{N}_i \ni n\to\infty} \Phi(z_n^i, y) - \Phi(z_n^i, z_\infty^i) \ge 0$$

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

We fix small $\varepsilon > 0$ and $\delta > 0$. For any $x, y \in B(o,r)$, the convexity of Φ and Fact 6 yield

$$\begin{aligned}\Phi(x,y) - \Phi(z_n^i, y_\varepsilon) &\ge D\Phi[\log_{(z_n^i, y_\varepsilon)}(x,y)]\\ &= D\Phi(\cdot, y_\varepsilon)[\log_{z_n^i} x] + D\Phi(z_n^i, \cdot)[\log_{y_\varepsilon} y],\end{aligned}$$

where $y_\varepsilon \in [y, z_\infty^i]$ is the point with $d(y_\varepsilon, z_\infty^i) = \varepsilon d(y, z_\infty^i)$. We also reuse the symbol $d_\delta(\cdot,\cdot)$ used in our proof of Theorem 25 above.

In what follows, $C < \infty$ is a constant depending on κ, r and p similar to the one in our proof of Theorem 25. For example we have

$$\int_{B(z_n^i,\delta)} D\Phi(\cdot, y_\varepsilon)[\log_{z_n^i} x] d^{p_n^i - 2}(z_n^i, x)\, d\mu_n(x) \le C\mu_n(B(z_n^i, \delta))\delta^{p_n^i - 1}.$$

We put $M_n^i := \max\{\delta^{p_n^i-2}, (R_\kappa)^{p_n^i-2}\} < \infty$ and fix couplings $\pi_n \in \mathscr{P}(Y\times Y)$ of μ_n and μ with $\int_{Y\times Y} \Phi(\cdot,\cdot)\, d\pi_n \to 0$ as $n \to \infty$.

Then we have

$$\int_{Y\times Y} D\Phi(\cdot\,, y_\varepsilon)[\log_{z_n^i} x] d_\delta^{p_n^i-2}(z_n^i, x)\, d\pi_n(x,y) + C(M_n^i\varepsilon + \delta^{p_n^i-1})$$
$$\geq \int_Y D\Phi(\cdot\,, z_\infty^i)[\log_{z_n^i} x] d^{p_n^i-2}(z_n^i, x)\, d\mu_n(x)$$
$$\geq -\int_Y \langle \nabla_{z_n^i}^- \Phi(\cdot\,, z_\infty^i), \log_{z_n^i} x\rangle d^{p_n^i-2}(z_n^i, x)\, d\mu_n(x)$$
$$\geq -\varepsilon_i |\nabla_{z_n^i}^- \Phi(\cdot\,, z_\infty^i)|$$
$$\geq -C\varepsilon_i$$

and

$$\int_{Y\times Y} D\Phi(z_n^i, \cdot)[\log_{y_\varepsilon} y] d_\delta^{p_n^i-2}(z_n^i, x)\, d\pi_n(x,y)$$
$$\geq \frac{1-\varepsilon}{\varepsilon}\int_{Y\times Y} (\Phi(z_n^i, y_\varepsilon) - \Phi(z_n^i, z_\infty^i)) d(z_\infty^i, y) d_\delta^{p_n^i-2}(z_n^i, x)\, d\pi_n(x,y)$$
$$\geq \frac{1-\varepsilon}{\varepsilon} M_n^i \int_Y \min\{\Phi(z_n^i, y_\varepsilon) - \Phi(z_n^i, z_\infty^i), 0\} d(z_\infty^i, y)\, d\mu(y),$$

with which the dominated convergence theorem yields

$$\liminf_{\mathcal{N}_i \ni n\to\infty} \int_{Y\times Y} D\Phi(z_n^i, \cdot)[\log_{y_\varepsilon} y] d_\delta^{p_n^i-2}(z_n^i, x)\, d\pi_n(x,y) \geq 0.$$

As $\varepsilon > 0$ is taken arbitrarily, we obtain

$$\limsup_{\mathcal{N}_i \ni n\to\infty} \Phi(z_n^i, z_\infty^i) \int_Y d_\delta^{p_n^i-2}(z_n^i, x)\, d\mu_n(x) \leq C\varepsilon_i + C\delta^{p-1+(1/i)}.$$

Then, since $\delta > 0$ is taken arbitrarily and

$$\int_Y d_\delta^{p_n^i-2}(z_n^i, \cdot)\, d\mu_n \geq \frac{1}{(R_\kappa)^{2-(1/i)}}\left(\int_Y d^{p_n}(z_n^i, \cdot)\, d\mu_n - \mu(B(z_n^i, \delta))\delta^{p_n}\right),$$

we have $\limsup_{\mathcal{N}_i \ni n\to\infty} \Phi(z_n^i, z_\infty^i) \to 0$ as $i \to \infty$.

Since

$$d(z_m, z_n) \leq d(z_m, z_m^i) + d(z_m^i, z_\infty^i) + d(z_n, z_n^i) + d(z_n^i, z_\infty^i)$$

for any $m, n \in \mathcal{N}_i$ and $i \in \mathbf{N}$, we conclude that $(z_{\min \mathcal{N}_i})_{i\in\mathbf{N}}$ is a Cauchy sequence and hence the limit $z_\infty := \lim_{i\to\infty} z_{\min \mathcal{N}_i}$ exists. It follows that z_∞ is a p-barycenter of μ. Now the proof is complete. □

Proposition 56 (cf. Al-Salman–Hajja [AH]). *If* (Y, d) *and* $\mu \in \mathscr{P}(Y)$ *are as in Assumption* 47, *then* $d(b^p(\mu), b^\infty(\mu)) \to 0$ *as* $p \to \infty$.

Proof. We may assume that μ is not a Dirac measure. Lemma 29 states that any sequence $(z_n)_{n\in\mathbf{N}}$ of p_n-barycenter $z_n := b^{p_n}(\mu) \in B(o, r)$ of μ with $p_n \to \infty$ as $n \to \infty$ has a subsequence, still denoted $(z_n)_{n\in\mathbf{N}}$, which converges weakly to a point $z_\infty \in \bar{B}(o, r)$. We put

$$\|f(\cdot)\|_p := \left(\int_Y |f(\cdot)|^p \, d\mu\right)^{1/p} \quad \text{and} \quad \|f(\cdot)\|_\infty := \operatorname*{ess\,sup}_Y |f(\cdot)|$$

for a function $f : Y \to \mathbf{R}$ and $d_-(\cdot, \cdot) := \min\{d(\cdot, \cdot), R_\kappa/2\}$.

The combination of Hölder's inequality, Fatou's lemma and Fact 31 yields

$$\begin{aligned}
\liminf_{n\to\infty} \|d(z_n, \cdot)\|_{p_n} &\geq \liminf_{n\to\infty} \|d(z_n, \cdot)\|_p \\
&\geq \|\liminf_{n\to\infty} \; d(z_n, \cdot)\|_p \\
&\geq \|\liminf_{n\to\infty} \; d_-(z_n, \cdot)\|_p \geq \|d_-(z_\infty, \cdot)\|_p
\end{aligned}$$

for any $p \in (1, \infty)$. Since $\|d_-(z_\infty, \cdot)\|_p \to \|d_-(z_\infty, \cdot)\|_\infty$ as $p \to \infty$, we have

$$\liminf_{n\to\infty} \|d(z_n, \cdot)\|_{p_n} \geq \|d_-(z_\infty, \cdot)\|_\infty \geq \|d(b^\infty(\mu), \cdot)\|_\infty.$$

On the other hand, Inequality (53) states

$$\|d(b^\infty(\mu), \cdot)\|_{p_n}^2 - \|d(z_n, \cdot)\|_{p_n}^2 \geq c(\mu) d^\alpha(b^\infty(\mu), z_n),$$

where $c(\mu) > 0$ and $\alpha > 2$ are constants independent of n.

We conclude $z_n \to b^\infty(\mu)$ as $n \to \infty$ and hence $b^p(\mu) \to b^\infty(\mu)$ as $p \to \infty$. Now the proof is complete. □

5.3. Convex Hull Property of p-Barycenter. It is known that the barycenter of a probability measure $\mu \in \mathscr{P}_1(Y)$ on a complete CAT(0)-space Y lies in the closed convex hull of a subset on which μ is concentrated, e.g. Sturm [St, Proposition 6.1]. This was also proved in [Yo] for barycenter of probability measures on CAT(κ)-spaces as in Theorem B. We prove that this is the case for p-barycenter on CAT(κ)-spaces.

Theorem 57. *Let* (Y, d) *be a complete CAT*(κ)*-space with* $\kappa > 0$ *and* $p \geq 1$. *Suppose* $\mu \in \mathscr{P}(Y)$ *is concentrated on a subset* $S \subset Y$ *with* $C := \overline{\mathrm{conv}}(S) \subset B(o, r)$

for some $o \in Y$ *and* $r < R_\kappa/2$. *Then*

$$F_\mu^p(x) > \inf_{x \in C} F_\mu^p(x)$$

holds for any $x \in Y \backslash C$. *In particular, any* p*-barycenter of* μ *lies in* C.

We first prove a weaker inequality. For possible future application, we state and prove it in general form.

Proposition 58. *Suppose* (Y, d), $\mu \in \mathscr{P}(Y)$ *and* $C \subset B(o, r)$ *are as in Theorem* 57. *Let* $U : [0, \infty) \to [0, \infty)$ *be a nondecreasing continuous function. Then*

$$\int_Y U(d(x, \cdot))\, d\mu \geq \inf_{x \in C} \int_Y U(d(x, \cdot))\, d\mu$$

holds for any $x \in Y$.

Proof. We set $F(x) := \int_Y U(d(x, \cdot))\, d\mu$ for $x \in Y$.

Claim 59 (cf. Claim 33). $F(x) \geq \inf_{B(o,r)} F$ *for any* $x \in Y$.

Proof. If $x \in Y \backslash B(o, 2r)$, we have $F(x) \geq U(r) \geq F(o)$.

If $x \in B(o, 2r) \backslash \bar{B}(o, r)$, we choose $x' \in [o, x]$ with $d(x, x') = 2(d(x, o) - r)$. Then we have $d(x', y) < d(x, y)$ for any $y \in B(o, r)$ and thus $F(x) \geq F(x') \geq \inf_{B(o,r)} F$, cf. [Af, Yo]. This verifies the claim. □

We fix small $\delta > 0$ and define a sequence $(C_\delta^n)_{n=0}^\infty$ of closed R_κ-convex subsets of Y as follows:

$$C_\delta^0 := C \quad \text{and} \quad C_\delta^{n+1} := \left\{ x \in \bar{B}(o, r) : \inf_{y \in C_\delta^n} \Phi(x, y) \leq \delta \right\}$$

for $n \geq 0$.

We fix $x \in B(o, r) \backslash C$. Then there exists a minimum number $N \in \mathbf{N} \cup \{0\}$ for which $x \in C_\delta^N$. Since

$$\bar{B}\left(C_\delta^n, \left(\frac{\delta}{C_1} \right)^{1/\beta} \right) \subset C_\delta^{n+1} \subset \bar{B}\left(C_\delta^n, \left(\frac{\delta}{C_2} \right)^{1/\beta} \right),$$

we have $N \leq (C_2/\delta)^{1/\beta} d(x, C) < \infty$, where C_1 and C_2 are the constants in Inequality (48). We then define a sequence $(x_\delta^n)_{n=0}^N$ of points as follows:

$$x_\delta^N := x \quad \text{and} \quad x_\delta^n := \pi_{C_\delta^n}(x_\delta^{n+1}) \in C_\delta^n$$

for $n = 0, \dots, N-1$. We have

$$\sum_{n=1}^{N} d(x_\delta^{n-1}, x_\delta^n) \le N\left(\frac{\delta}{C_1}\right)^{1/\beta} \le \left(\frac{C_2}{C_1}\right)^{1/\beta} d(x, C) =: D < \infty.$$

Since $\tilde{\angle}_\kappa(x_\delta^{n-1}; x_\delta^n, y) \ge \pi/2$ and

$$d(x_\delta^{n-1}, y) + d(x_\delta^n, y) + d(x_\delta^{n-1}, x_\delta^n) < 4r < 2R_\kappa$$

for any $y \in C$ we have

$$\begin{aligned} d(x_\delta^{n-1}, y) &< d(x_\delta^n, y) && \text{if } d(x_\delta^n, y) < R_\kappa/2; \\ d(x_\delta^{n-1}, y) &\le d(x_\delta^n, y) + \varepsilon d(x_\delta^{n-1}, x_\delta^n) && \text{if } d(x_\delta^n, y) \ge R_\kappa/2, \end{aligned}$$

where $\varepsilon = \varepsilon(\delta; r) > 0$ is a constant with $\varepsilon \to 0$ as $\delta \to 0$, and hence

$$\begin{aligned} d(x_\delta^0, y) &< d(x, y) && \text{if } d(x, y) < R_\kappa/2; \\ d(x_\delta^0, y) &\le d(x, y) + D\varepsilon && \text{if } d(x, y) \ge R_\kappa/2. \end{aligned}$$

Now the dominated convergence theorem yields

$$\inf_C F \le \limsup_{\delta \to 0} F(x_\delta^0) \le \lim_{\varepsilon \to 0} \int_Y U(d(x, \cdot) + D\varepsilon)\, d\mu = F(x).$$

Combined with Claim 59, this finishes the proof. □

Proof of Theorem 57. We set $F := F_\mu^p$ and assume that there is a point $x_0 \in Y \backslash C$ with $F(x_0) = \inf_Y F$. By Claims 33 and 34, we know $x_0 \in B(o, r) \backslash C$. We repeat the argument in our proof of Proposition 58 with $U(s) := (1/p)s^p$ to obtain a sequence $(x_n)_{n \in \mathbf{N}}$ of points $x_n := x_{1/n}^0 \in C$ for which

$$\limsup_{n \to \infty} d(x_n, y) \le d(x_0, y) \quad \text{for any } y \in C.$$

Theorem 25 states that a subsequence of $(x_n)_{n \in \mathbf{N}}$ converges to a point $x_\infty \in C$ where $F(x_\infty) = F(x_0) = \inf_Y F$ and

$$d(x_\infty, y) = d(x_0, y) \quad \text{for } \mu\text{-a.e. } y \in Y.$$

We use the convexity of Φ in Theorem A and Fact 6 to derive for any $y \in C$

$$\begin{aligned} \Phi(y, y) - \Phi(x_0, x_\infty) &\ge D\Phi[\log_{(x_0, x_\infty)}(y, y)] \\ &= D\Phi(\cdot, x_\infty)[\log_{x_0} y] + D\Phi(x_0, \cdot)[\log_{x_\infty} y] \\ &\ge -\langle \nabla_{x_0}^- \Phi(\cdot, x_\infty), \log_{x_0} y \rangle - \langle \nabla_{x_\infty}^- \Phi(x_0, \cdot), \log_{x_\infty} y \rangle. \end{aligned}$$

We integrate this inequality with the measure $d^{p-2}(x_0,\cdot)\mu$ to obtain

$$\begin{aligned}
&-\Phi(x_0,x_\infty)\int_Y d^{p-2}(x_0,y)\,d\mu(y)\\
&\quad\geq -\int_Y \langle\nabla^-_{x_0}\Phi(\cdot,x_\infty),\log_{x_0} y\rangle d^{p-2}(x_0,y)\,d\mu(y)\\
&\qquad-\int_Y \langle\nabla^-_{x_\infty}\Phi(x_0,\cdot),\log_{x_\infty} y\rangle d^{p-2}(x_0,y)\,d\mu(y)\\
&\quad= DF[\nabla^-_{x_0}\Phi(\cdot,x_\infty)]+DF[\nabla^-_{x_\infty}\Phi(x_0,\cdot)]\geq 0.
\end{aligned}$$

Since

$$\int_Y d^{p-2}(x_0,\cdot)\,d\mu\geq\min\{d^{p-2}(x_0,C),(R_\kappa)^{p-2}\}>0,$$

we conclude $x_0=x_\infty\in C$. This completes the proof. □

REMARK 60. In [Ku2], a minimizer of the restriction of the function $x\mapsto\int_X d^p(\cdot,x)-d^p(\cdot,x_0)\,d\mu$, with $x_0\in(X,d)$ being fixed, on the closed convex hull of the support of $\mu\in\mathscr{P}_{p-1}(X)$ is called a *pure p-barycenter* of μ. The *support* of a measure μ 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 μ is concentrated. Theorem 57 states that p-barycenter and pure p-barycenter coincide for $\mu\in\mathscr{P}(Y)$ as in the theorem on a complete separable CAT(κ)-space (Y,d) with $\kappa>0$.

5.4. Jensen's Inequality. Jensen's inequality is also one of the properties that we expect to hold for barycenter, cf. Kuwae [Ku, Ku2]. The following is a direct consequence of Proposition 39 and Jensen's inequality proved for barycenter in [Yo, Proposition 10 and Theorem 25]. Due to the subtlety of Jensen's inequality for p-barycenter, also pointed out by Kell [Kel2], this is the best that we can prove now.

PROPOSITION 61 (Jensen's inequality). *Let (Y,d) be a complete CAT(κ)-space with $\kappa>0$, $\mu\in\mathscr{P}(Y)$, $p\geq 2$ and $\varphi:Y\to\mathbf{R}\cup\{\infty\}$ be a lower-semicontinuous convex function. Suppose either μ is concentrated on a ball of radius $<R_\kappa/2$ in Y*

and hence it admits a unique p-barycenter $b^p(\mu) \in Y$ or φ is locally Lipschitz at a p-barycenter $b^p(\mu)$ of μ and μ is concentrated on $B(b^p(\mu), R_\kappa)$. Then

$$\varphi(b^p(\mu)) \le \int_Y \varphi \, d\tilde{\mu}.$$

Here, $\tilde{\mu} \in \mathscr{P}(Y)$ is the probability measure defined in Proposition 39.

6. Banach–Saks Property of CAT(κ)-Spaces

In this section, we establish analogues of the Banach–Saks–Kakutani type result formulated with p-barycenter on CAT(κ)-spaces. They generalize the theorems of Jost [Jo, Theorem 2.2] and the author [Yo, Theorem C].

Kakutani [Ka] proved the *Banach–Saks property* of uniformly convex Banach spaces: any bounded sequence $(x_n)_{n\in\mathbf{N}}$ of points of an uniformly convex Banach space B has a subsequence, still denoted $(x_n)_{n\in\mathbf{N}}$, for which the sequence $(m_n)_{n\in\mathbf{N}}$ of the arithmetic means $m_n := (1/n)\sum_{i=1}^n x_i \in B$ converges to a point of B. The following theorems formulate this property with p-barycenter on CAT(κ)-spaces.

Theorem C. *Let (Y, d) be a complete CAT(κ)-space with $\kappa \in \mathbf{R}$ and $(x_n)_{n\in\mathbf{N}}$ be a sequence of points in $B(o, r)$ with $o \in Y$ and $r < R_\kappa/2$. Then it has a subsequence, still denoted $(x_n)_{n\in\mathbf{N}}$, for which any sequence $(m_n^p)_{n\in\mathbf{N}}$ of p-barycenter of finitely and uniformly supported probability measures $(1/n)\sum_{i=1}^n \delta_{x_i} \in \mathscr{P}(Y)$ converges to a point $x_\infty \in Y$ for all $p \in [2, \infty)$.*

Theorem D. *There exists $h_0 \in (1/4, 1/2)$ which satisfies the following: Let (Y, d) be a complete CAT(κ)-space with $\kappa \in \mathbf{R}$ and $(x_n)_{n\in\mathbf{N}}$ be a sequence of points in $B(o, r)$ with $o \in Y$ and $r < h_0 R_\kappa$. Then it has a subsequence, still denoted $(x_n)_{n\in\mathbf{N}}$, for which any sequence $(m_n^p)_{n\in\mathbf{N}}$ of p-barycenter of finitely and uniformly supported probability measures $(1/n)\sum_{i=1}^n \delta_{x_i} \in \mathscr{P}(Y)$ converges to a point $x_\infty \in Y$ for all $p \in [1, \infty)$.*

In particular, Theorem D holds for any bounded sequence in complete CAT(0)-spaces. It might be interesting if Theorems C and D could be generalized as a theorem. Namely it is not clear now whether we can take $h_0 = 1/2$ in Theorem D. Our proof of Theorems C and D uses only a few properties of CAT(κ)-spaces and it also works for more general convex spaces, cf. Kell [Kel].

Now we begin our proof of Theorems C and D. They share several initial steps in the proof.

PROOF OF THEOREMS C AND D. We may assume that $\kappa > 0$ because the proof of the theorems for nonpositive $\kappa \leq 0$ is reduced to that for positive $\kappa > 0$.

Lemma 29 states that $(x_n)_{n\in\mathbf{N}}$ has a subsequence, still denoted $(x_n)_{n\in\mathbf{N}}$, which converges weakly to a point $x_\infty \in \bar{B}(o,r)$. By Fact 31, we may further assume that the limit $\rho := \lim_{n\to\infty} d(x_n, x_\infty) \leq r$ exists and

$$\lim_{n\to\infty} \inf_{m\geq n} d(x_m, [x_n, x_\infty]) = \rho. \tag{62}$$

We put

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

for a finite subset $I \subset \mathbf{N}$ of cardinality $\#I < \infty$. We notice that $2\Lambda^p(I\cup J) \geq \Lambda^p(I) + \Lambda^p(J)$ for any $I, J \subset \mathbf{N}$ with $\#I = \#J$ and $I \cap J = \emptyset$.

The following observation is the key.

CLAIM 63. *For each $k, N \in \mathbf{N}$, we put $I_k^N := \{(k-1)2^N + 1, \ldots, k2^N\} \subset \mathbf{N}$. If $(x_n)_{n\in\mathbf{N}}$ satisfies*

$$\sup\left\{\liminf_{k\to\infty} \Lambda^q(I_k^N) : N \in \mathbf{N}\right\} = \rho^q \tag{64}$$

for some $q \geq 1$ and p-barycenter m_n^p satisfies $m_n^p \in B(x_\infty, \underline{r})$ with $r + \underline{r} < R_\kappa/2$ for all $n \in \mathbf{N}$ if $p \in [q, 2)$, then the sequence $(m_n^p)_{n\in\mathbf{N}}$ converges to x_∞ for all $p \in [q, \infty)$.

PROOF. Hölder's inequality yields

$$\rho \geq \liminf_{k\to\infty} (\Lambda^p(I_k^N))^{1/p} \geq \liminf_{k\to\infty} (\Lambda^q(I_k^N))^{1/q}$$

for any $p > q$ and $N \in \mathbf{N}$. This means that Equation (64) for some $q \geq 1$ implies the same equation for all $p > q$.

We fix $p \in [q, \infty)$. By assumption, there exists $N \in \mathbf{N}$ for any $\varepsilon > 0$ with

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

and hence we have

$$\rho^p = \lim_{n\to\infty}\left[\frac{1}{n}\sum_{i=1}^n d^p(x_i, x_\infty)\right] = \lim_{n\to\infty}\left[\frac{1}{n}\sum_{i=1}^n d^p(x_i, m_n^p)\right]. \tag{65}$$

If $p \geq 2$, Proposition 49 states

$$\frac{1}{n}\sum_{i=1}^{n}(d^p(x_i, x_\infty) - d^p(x_i, m_n^p)) \geq c \cdot d^{\max\{p,\alpha\}}(m_n^p, x_\infty)$$

for $n \in \mathbf{N}$. If $1 < p < 2$, Corollary 16 gives a smilar variance inequality on $B(x_\infty, \underline{r})$. We then infer that $d(m_n^p, x_\infty) \to 0$ as $n \to \infty$ if $p > 1$.

We now consider the case $p = 1$ and suppose $\limsup_{n\to\infty} d(m_n^1, x_\infty) > 0$. For $i \leq n$, we define $\varepsilon_i^n \geq 0$ by

$$d(m_n^1, x_\infty) = |d(x_i, m_n^1) - d(x_i, x_\infty)| + \varepsilon_i^n \mathscr{M}_1(d(x_i, m_n^1), d(x_i, x_\infty)),$$

where $\mathscr{M}_1(\cdot, \cdot)$ is defined in Definition 18. With $w_n := m(m_n^1, x_\infty)$, it implies

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n} d(x_i, w_n) &\leq \frac{1}{n}\sum_{i=1}^{n}(1 - \rho(\varepsilon_i^n))\mathscr{M}_1(d(x_i, m_n^1), d(x_i, x_\infty)) \\ &\leq \frac{1}{2n}\sum_{i=1}^{n}(d(x_i, m_n^1) + d(x_i, x_\infty)),\end{aligned}$$

where $\rho(\cdot) = \rho_1(\cdot) > 0$ is the constant in Proposition 21 with $\rho(0) := 0$. Hence Equation (65) with $p = 1$ gives $\#I(\varepsilon; n)/n \to 1$ as $n \to \infty$ as well as $\liminf_{n\to\infty} \#I_\pm(\varepsilon; n)/n < 1$ for any $\varepsilon > 0$, where

$$\begin{aligned}I(\varepsilon; n) &:= \{i \in \{1, \ldots, n\} : \varepsilon_i^n \leq \varepsilon\}; \\ I_\pm(\varepsilon; n) &:= \{i \in I(\varepsilon; n) : \pm(d(x_i, m_n^1) - d(x_i, x_\infty)) \geq 0\}.\end{aligned}$$

We choose an infinite subset $\mathscr{N} \subset \mathbf{N}$ with

$$\rho' := \lim_{\mathscr{N} \ni n\to\infty} d(m_n^1, x_\infty) = \limsup_{n\to\infty} d(m_n^1, x_\infty) > 0$$

and $i(n), j(n) \in I_-(\varepsilon(n); n)$ for some $\varepsilon(n) > 0$ with $i(n) < j(n)$, $i(n) \to \infty$ and $\varepsilon(n) \to 0$ as $n \to \infty$. We pick $x'_{i(n)} \in [x_{i(n)}, x_\infty]$ for $n \in \mathscr{N}$ with $\lim_{\mathscr{N} \ni n\to\infty} d(x'_{i(n)}, x_\infty) = \rho'$. Then we have $\lim_{\mathscr{N} \ni n\to\infty} d(m_n^1, x'_{i(n)}) = 0$,

$$\begin{aligned}\lim_{\mathscr{N} \ni n\to\infty} d(x_{i(n)}, m_n^1) &= \lim_{\mathscr{N} \ni n\to\infty} d(x_{i(n)}, x_\infty) - d(m_n^1, x_\infty) \\ &= \lim_{\mathscr{N} \ni n\to\infty} d(x_{j(n)}, m_n^1) \\ &= \lim_{\mathscr{N} \ni n\to\infty} d(x_{j(n)}, x_\infty) - d(m_n^1, x_\infty) = \rho - \rho'\end{aligned}$$

and

$$\rho > \rho - \rho' = \lim_{\mathcal{N} \ni n \to \infty} d(x_{j(n)}, x'_{i(n)}) \geq \limsup_{\mathcal{N} \ni n \to \infty} d(x_{j(n)}, [x_{i(n)}, x_\infty]).$$

This contradicts Equation (62). The claim is confirmed. □

PROOF OF THEOREM C. To prove Theorem C, we find a subsequence $(x_n)_{n \in \mathbf{N}}$ with

$$\inf_{k \in \mathbf{N}} \Lambda^2(I_k^N) \nearrow \rho^2 \quad \text{as } N \nearrow \infty.$$

This was done in the proof of [Yo, Theorem C] by using Fact 31 and Proposition 38. Then Theorem C follows from Claim 63. □

PROOF OF THEOREM D. There exist $h_0 \in (1/4, 1/2)$ and $\theta_0 > 0$ with

$$\tilde{\angle}_1(x; y, z) \leq \pi/2 - \theta_0$$

for any $x, y, z \in (\mathbf{S}^2, d_{\mathbf{S}^2})$ with $d_{\mathbf{S}^2}(x, z) \in [((1/2) - h_0)\pi, h_0\pi]$, $d_{\mathbf{S}^2}(y, z) \leq h_0\pi$ and $d_{\mathbf{S}^2}(x, y) \geq \pi/8$.

We put $\underline{r} := r$ if $r < R_\kappa/4$ and $\underline{r} := ((1/2) - h_0)R_\kappa$ if $R_\kappa/4 \leq r < h_0 R_\kappa$. Then $r + \underline{r} < R_\kappa/2$. We notice

$$d(x, x_n) \leq d(x, x_\infty) + d(x_n, x_\infty) \leq r + \underline{r}$$

for any $x \in \bar{B}(x_\infty, \underline{r})$ and Fact 31 implies that we may assume that the set $\{B(x_n, \rho/2)\}_{n \in \mathbf{N}}$ of balls is mutually disjoint.

For any probability measure $\nu \in \mathscr{P}(Y)$ which is finitely and uniformly supported on $\{x_n : n \in \mathbf{N}\} \subset B(o, r)$, if $\#(\mathrm{supp}[\nu]) \in \mathbf{N}$ is large enough, we have

$$\begin{aligned} DF_\nu^p[\uparrow_x^{x_\infty}] &= -\int_Y \cos \angle_x(y, x_\infty) d^{p-1}(x, y)\, d\nu(y) \\ &\leq -\int_Y \cos \tilde{\angle}_\kappa(x; y, x_\infty) d^{p-1}(x, y)\, d\nu(y) < 0 \end{aligned}$$

for any $x \in \bar{B}(x_\infty, r) \backslash B(x_\infty, \underline{r})$ and hence $b^p(\nu) \in B(x_\infty, \underline{r})$. Then Corollary 16 states that the p-variance inequality holds for such $\nu \in \mathscr{P}(Y)$ on $B(x_\infty, \underline{r})$ and $p \in (1, 2]$.

To prove Theorem D, we find a subsequence $(x_n)_{n \in \mathbf{N}}$ for which

$$\Lambda^{q_i}(I_k^N) > \rho_i^{q_i} \quad \text{for any } k \in \mathbf{N} \text{ and } N > N_i.$$

holds for any $i \in \mathbf{N}$ with some $q_i \searrow 1$, $\rho_i \nearrow \rho$ and $N_i \nearrow \infty$ as $i \nearrow \infty$. This is done in a way similar to the proof of [Yo, Theorem C] by using Fact 31 and Corollary 16. Then Theorem D follows from Claim 63. $\square$

Now the proof of Theorems C and D is complete. $\square$

We conclude this paper with several remarks.

Remark 66. It is not known now whether the condition $p \geq 2$ is optimal for the uniqueness of the p-barycenter in Theorem B, cf. Example 24.

Buss–Fillmore [BF] proved that any finitely supported probability measure $\mu \in \mathscr{P}(\mathbf{S}^n)$ which is concentrated on $\bar{B}(o, \pi/2)$ but not on the boundary $\partial \bar{B}(o, \pi/2)$ for some $o \in \mathbf{S}^n$ admits a unique barycenter. The author does not know whether this can be generalized to p-barycenter of probability measures on general CAT(1)-spaces.

Ohta–Pálfia [OP] recently studied gradient flow on CAT(1)-spaces. It would be interesting to establish convergence of gradient flow or some algorithm to a p-barycenter, cf. Afsari–Tron–Vidal [ATV].

Appendix A. Proof of Proposition 9 for $p > 2$

In this appendix, we prove the following proposition, which might be of independent interest. Proposition 9 for $p > 2$ follows from a similar argument. Recall the definition of p-uniformly convex spaces in Definition 11.

Proposition 67. *Any p-uniformly convex space (X, d) for some $p \geq 2$ is a q-uniformly convex space for all $q > p$.*

Proof. We fix $x \in X$, a geodesic $\gamma : [0,1] \to X$, $t \in [0,1]$ and $q > p$ then put $y := \gamma(0)$, $z := \gamma(1)$ and $w := \gamma(t)$. We start our proof with the following observation.

Claim 68. *If $d(x, w) \geq \varepsilon d(y, z)$ for some $\varepsilon \geq 0$, we have*

$$d^q(x, w) \leq (1-t)d^q(x, y) + td^q(x, z) - \frac{q}{p}\varepsilon^{q-p}c_p \cdot t(1-t)d^q(y, z).$$

In particular, the function $d^q(x, \cdot)$ is convex on X for any $x \in X$.

PROOF. To see this, we let $J(s) := s^{q/p}$ be the increasing convex function on $[0, \infty)$. We have

$$J(d^p(x,y)) - J(d^p(x,w)) \geq J'(d^p(x,w))(d^p(x,y) - d^p(x,w));$$
$$J(d^p(x,z)) - J(d^p(x,w)) \geq J'(d^p(x,w))(d^p(x,z) - d^p(x,w))$$

and hence

$$\begin{aligned}(1-t)d^q(x,y) &+ td^q(x,z) - d^q(x,w)\\ &\geq J'(d^p(x,w))[(1-t)d^p(x,y) + td^p(x,z) - d^p(x,w)]\\ &\geq \frac{q}{p}\varepsilon^{q-p}c_p \cdot t(1-t)d^q(y,z).\end{aligned}$$

This verifies the claim. □

We put $c_q := (q/15^q p)c_p > 0$. Now we suppose $d(x,w) < (1/5)d(y,z)$. We may also assume $t \in [1/2, 1)$ and put $y' := \gamma(t/3)$ and $y'' := \gamma(2t/3)$.

Since $d(x,y') \geq d(w,y') - d(x,w) \geq (1/5)d(y,y'')$, Claim 68 implies

$$\begin{aligned}A &:= \frac{d^q(x,y) - d^q(x,y')}{t} - \frac{d^q(x,y') - d^q(x,y'')}{t}\\ &\geq \frac{q}{5^{q-p}p}\frac{c_p}{2t}d^q(y,y'')\\ &\geq c_q d^q(y,z)\end{aligned}$$

as well as

$$B := \frac{d^q(x,y') - d^q(x,y'')}{t} - \frac{d^q(x,y'') - d^q(x,w)}{t} \geq 0;$$
$$C := \frac{d^q(x,y'') - d^q(x,w)}{t/3} - \frac{d^q(x,w) - d^q(x,z)}{1-t} \geq 0.$$

Now we gather

$$\begin{aligned}\frac{d^q(x,y) - d^q(x,w)}{t} - \frac{d^q(x,w) - d^q(x,z)}{1-t} &= A + 2B + C\\ &\geq c_q d^q(y,z),\end{aligned}$$

which is equivalent to the desired inequality. This completes the proof. □

Proposition 67 implies that CAT(0)-spaces are p-uniformly convex spaces for all $p \geq 2$. In literature, e.g. Naor–Silberman [NS], Kuwae [Ku2, Ku3], this fact is stated as a consequence of an isometric embedding of the Euclidean plane $\mathbf{R}^2$ into L^p-space.

Acknowledgements

The author is grateful to Professor Kazuhiro Kuwae for introducing Kendall's convex function to him. He also thanks the anonymous referee for perusing the manuscript very carefully. This work is partly supported by JSPS KAKENHI (No. 26800035).

References

[Af] B. Afsari, Riemannian L^p center of mass: existence, uniqueness, and convexity. Proc. Amer. Math. Soc. **139** (2011), no. 2, 655–673.

[ATV] B. Afsari; R. Tron; R. Vidal, On the convergence of gradient descent for finding the Riemannian center of mass. SIAM J. Control Optim. **51** (2013), no. 3, 2230–2260.

[AH] A. Al-Salman; M. Hajja, Towards a well-defined median. J. Math. Inequal. **1** (2007), no. 1, 23–30.

[BL] A. Balser; A. Lytchak, Centers of convex subsets of buildings. Ann. Global Anal. Geom. **28** (2005), no. 2, 201–209.

[BBI] D. Burago; Y. Burago; S. Ivanov, A course in metric geometry. Graduate Studies in Mathematics, 33. American Mathematical Society, Providence, RI, 2001.

[BF] S. R. Buss and J. P. Fillmore, Spherical averages and application to spherical splines and interpolation, ACM Transactions on Graphics **20** (2001), no. 2, 95–126.

[Ek] I. Ekeland, Nonconvex minimization problems. Bull. Amer. Math. Soc. **1** (1979), 443–474.

[EF] R. Espínola; A. Fernández-León, CAT(k)-spaces, weak convergence and fixed points. J. Math. Anal. Appl. **353** (2009), no. 1, 410–427.

[Fo] T. Foertsch, Ball versus distance convexity of metric spaces. Beiträge Algebra Geom. **45** (2004), no. 2, 481–500.

[Jo] J. Jost, Equilibrium maps between metric spaces. Calc. Var. Partial Differential Equations **2** (1994), no. 2, 173–204.

[Jo2] J. Jost, Generalized harmonic maps between metric spaces. Geometric analysis and the calculus of variations, 143–174, Int. Press, Cambridge, MA, 1996.

[Ka] S. Kakutani, Weak convergence in uniformly convex spaces. Tôhoku Math. J. **45** (1938), 188–193.

[Kel] M. Kell, Uniformly convex metric spaces. Anal. Geom. Metr. Spaces **2** (2014), 359–380.

[Kel2] M. Kell, Symmetric orthogonality and contractive projections in metric spaces. Preprint, arXiv:1604.01993.

[Ke] W. Kendall, Probability, convexity, and harmonic maps with small image. I. Uniqueness and fine existence. Proc. London Math. Soc. (3) **61** (1990), no. 2, 371–406.

[Ke2] W. Kendall, Convexity and the hemisphere. J. London Math. Soc. (2) **43** (1991), no. 3, 567–576.

[Ku] K. Kuwae, Jensen's inequality over CAT(κ)-space with small diameter. Potential theory and stochastics in Albac, 173–182, Theta Ser. Adv. Math., 11, Theta, Bucharest, 2009.

[Ku2] K. Kuwae, Jensen's inequality on convex spaces. Calc. Var. Partial Differential Equations **49** (2014), no. 3–4, 1359–1378.

[Ku3] K. Kuwae, Resolvent flows for convex functionals and p-harmonic maps. Anal. Geom. Metr. Spaces **3** (2015), 46–72.
[Ly] A. Lytchak, Open map theorem for metric spaces. Algebra i Analiz **17** (2005), no. 3, 139–159.
[NS] A. Naor; L. Silberman, Poincaré inequalities, embeddings, and wild groups. Compos. Math. **147** (2011), no. 5, 1546–1572.
[Oh] S.-i. Ohta, Convexities of metric spaces. Geom. Dedicata **125** (2007), 225–250.
[Oh2] S.-i. Ohta, Barycenters in Alexandrov spaces of curvature bounded below. Adv. Geom. **12** (2012), 571–587.
[OP] S.-i. Ohta; M. Pálfia, Gradient flows and a Trotter–Kato formula of semi-convex functions on CAT(1)-spaces. To appear in Amer. J. Math.
[Sa] T. Sakai, Riemannian geometry. Translations of Mathematical Monographs, 149. American Mathematical Society, Providence, RI, 1996.
[St] 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.
[Vi] C. Villani, Optimal transport. Old and new. Grundlehren der Mathematischen Wissenschaften, 338. Springer-Verlag, Berlin, 2009.
[Ya] L. Yang, Riemannian median and its estimation. LMS J. Comput. Math. **13** (2010), 461–479.
[Yo] T. Yokota, Convex functions and barycenter on CAT(1)-spaces of small radii. J. Math. Soc. Japan **68** (2016), no. 3, 1297–1323.

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