J. Korean Math. Soc. **59** (2022), No. 3, pp. 549–570
https://doi.org/10.4134/JKMS.j210239
pISSN: 0304-9914 / eISSN: 2234-3008

BIRKHOFF'S ERGODIC THEOREMS IN TERMS OF WEIGHTED INDUCTIVE MEANS

Byoung Jin Choi and Jae Hun Kim

Abstract. In this paper, we study the Birkhoff's ergodic theorem on geodesic metric spaces, especially on Hadamard spaces, using the notion of weighted inductive means. Also, we study a deterministic weighted sequence for the weighted Birkhoff's ergodic theorem in Hadamard spaces.

1. Introduction

Let $(X, \mathcal{F}, \mu)$ be a measure space and $f : X \to \mathbb{R}$ be a measurable function. Discrete-time evolution of X can be considered as an operator T on X so that if x is the starting point in X, then $T(x) \in X$ is the point after one time step.

To measure some physical situation, we usually repeat same measurements in (discrete) time and consider their time average of f for T given by

$$\frac{1}{n}\sum_{i=0}^{n} f(T^i(x)),$$

where $x \in X$ is a starting point and $f : X \to \mathbb{R}$ is a measurable function. Actually, the measurements of f at discrete-time units are given by

$$f(x), f(T(x)), \ldots, f(T^n(x)) \quad \text{for } n \in \mathbb{N}.$$

On the other hand, we can define *the space average of f* as follows:

$$\int_X f(x) d\mu.$$

The well-known Birkhoff's ergodic theorem in probability spaces says that the sequence of the time average of f for T converges to the space average of f if T is an ergodic measure preserving map. More precisely, we have:

Received April 8, 2021; Revised September 22, 2021; Accepted February 10, 2022.
2020 *Mathematics Subject Classification.* Primary 37A30, 28D05.
Key words and phrases. Birkhoff's ergodic theorem, weighted inductive means, barycenter, Hadamard space.
B. J. Choi was supported by the 2022 scientific promotion program funded by Jeju National University.

Theorem 1.1 (Birkhoff's ergodic theorem)**.** *Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $T : \Omega \to \Omega$ be a ergodic measure preserving map. Then for any $f \in L^1(\Omega, \mathcal{F}, P)$ we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} f(T^{i-1}(\omega)) = \int_\Omega f(\omega)dP(\omega)$$

holds for P-a.e. $\omega \in \Omega$.

In [6], the authors studied the ergodic theorem for a weighted average. The authors concentrated the almost everywhere (a.e.) convergence of a weighted arithmetic mean and studied on the conditions of weighted sequences so that the weighted average converges a.e.. Especially the authors considered the case that a weighted sequence is an arithmetical function in the number theory. On the other hand, the Birkhoff's ergodic theorem is related to the strong law of large numbers in probability theory, i.e., the strong law of large numbers is the special case of the Birkhoff's ergodic theorem. In [5], the authors proved that the law of large numbers in terms of the weighted inductive mean in Hadamard spaces, so their results have influenced us.

The Birkhoff's ergodic theorem can be extended to general metric spaces, for example, the Halpern's iteration method can be considered as the Birkhoff's ergodic theorem in geodesic metric spaces, e.g., [4, 15, 16], etc. In particular, Saejung [16] proved that a strong convergence of the Halpern's iteration method in a complete CAT(0)-space. A similar result in CAT(κ)-space, $\kappa > 0$ has been established by Piątek [15]. Also several authors extended the Birkhoff's ergodic theorem to geodesic metric space using the notion of inductive means and barycenter maps, e.g., [1, 3, 9, 13], etc.

The main purpose of this paper is to study weighted Birkhoff's ergodic theorems on geodesic metric spaces, specially on Hadamard spaces using the notion of weighted inductive means. More precisely, let $\mathbf{x} = \{x_i\}$ be a sequence in a Hadamard space N and $\{a_{ni}\}$ be a positive weighted sequence with $\sum_{i=1}^n a_{ni} = 1$ for any $n \in \mathbb{N}$. The new weighted sequence $\{y_{\ell i}\}_{i=1}^{\ell}$ for $\ell = 1, 2, \ldots, n-1$ is defined as follows:

$$y_{\ell i} = \frac{a_{ni}}{\sum_{k=1}^{\ell} a_{nk}} \quad \text{for all} \quad \ell = 1, \ldots, n-1.$$

Then by definition, we have

$$y_{\ell\ell} = \frac{a_{n\ell}}{\sum_{k=1}^{\ell} a_{nk}} \quad \text{for} \quad \ell = 1, \ldots, n-1.$$

The weighted inductive mean, denoted by $\{S_n(\mathbf{x})\}$, is defined as follows ([5,11]):

$$S_1(\mathbf{x}) = x_1,$$
$$S_n(\mathbf{x}) = z_{n-1}(\mathbf{x})\#_{a_{nn}} x_n \quad (n \geq 2),$$

where $z_1 = x_1$, $z_\ell = z_{\ell-1}\#_{y_{\ell\ell}} x_\ell$ for $\ell = 2, \ldots, n-1$. If we take $\{a_{ni} = \frac{1}{n}\}$, $i = 1, 2, \ldots, n$, then we have the inductive mean. In Hilbert spaces, the weighted inductive mean is exactly equal to the weighted sums as following:

$$S_n = \sum_{i=1}^{n} a_{ni} x_i.$$

So our results are the weighted version of Birkhoff's ergodic theorems in geodesic metric spaces. Also we give some deterministic weighted sequence for the weighted Birkhoff's ergodic theorem in Hadamard spaces (Theorems 3.19 and 3.24). Actually, a deterministic convergence result concerning with inductive means in geodesic metric spaces is firstly studies by Holbrook in [10], and then in [11], Lim and Pálfia studied a deterministic convergence result for weighted inductive means, so their results also have influenced us.

This paper is organized as follows. In Section 2, we firstly recall the notions of Hadamard spaces, and several properties for our study. Secondly, we recall the notion of weighted inductive means. In Section 3, we prove the several Birkhoff's ergodic theorems concerning with the weighted inductive means in Hadamard spaces. Also, we study a deterministic weighted sequence for the weighted Birkhoff's ergodic theorem in Hadamard spaces.

2. Preliminaries

2.1. Hadamard spaces

Let (N, d) be a geodesic metric space. (N, d) is called a Hadamard space if (N, d) is a complete metric space satisfying the semiparallelogram law, i.e., for any $x, y \in N$ there exists $m \in N$ such that

$$d(m, z)^2 \leq \frac{1}{2} d(x, z)^2 + \frac{1}{2} d(y, z)^2 - \frac{1}{4} d(x, y)^2 \tag{2.1}$$

for all $z \in N$. If we take $z = x$ and $z = y$ in the inequality (2.1), then it holds that $d(m, x) = d(m, y) = \frac{1}{2} d(x, y)$. The point m is called the midpoint between x and y. Also, by the inequality (2.1), the midpoint is unique. Actually, given two points $a, b \in N$ using the existence and uniqueness of midpoints, we can construct the unique geodesic $\gamma_{a,b} : [0, 1] \to N$. In this paper, we use the notation $a \#_t b$ for the geodesic joining a and b. The following properties of $a \#_t b$ are useful for our study.

Theorem 2.1 ([17]). *Let (N, d) be a Hadamard space. Then for any $x, y, z \in N$ and $t \in [0, 1]$ we have*

$$d(x \#_t y, z)^2 \leq (1 - t) d(x, z)^2 + t d(y, z)^2 - t(1 - t) d(x, y)^2.$$

Corollary 2.2 ([17]). *Let (N, d) be a Hadamard space and let $x, y, x', y' \in N$. Define a function $f : [0, 1] \to \mathbb{R}$ by*

$$f(t) = d(x \#_t x', y \#_t y').$$

Then the following holds:

$$f(t) \leq (1 - t) d(x, y) + t d(x', y').$$

That is, f is convex on $[0,1]$.

2.2. Probability measures and barycenters on Hadamard spaces

Let (N,d) be a Hadamard space. Let $\mathcal{B}(N)$ be the Borel σ-algebra on N and $\mathcal{P}(N)$ be the set of all probability measures on $\mathcal{B}(N)$ with separable support. For $1 \le q < \infty$, let $\mathcal{P}^q(N)$ be the set of all $\mu \in \mathcal{P}(N)$ satisfying

$$\int_N d(x,y)^q d\mu(x) < \infty$$

for some (and hence all) $y \in N$. Denote by $\mathcal{P}^\infty(N)$ the set of all $\mu \in \mathcal{P}(N)$ such that support of μ is bounded.

For $\mu \in \mathcal{P}(N)$ the value $\mathrm{var}(\mu) := \inf_{z\in N} \int_N d(z,x)^2 d\mu(x)$ is called the variance of μ. By definition, $\mu \in \mathcal{P}^2(N)$ if and only if $\mathrm{var}(\mu) < \infty$.

Proposition 2.3 ([17])**.** *Let (N,d) be a Hadamard space and fix $y \in N$. For $\mu \in \mathcal{P}(N)$ the function $z \mapsto \int_N [d(z,x)^2 - d(y,x)^2] d\mu(x)$ is uniformly convex and continuous. Also there exists a unique point $z \in N$ which minimizes the function $z \mapsto \int_N [d(z,x)^2 - d(y,x)^2] d\mu(x)$.*

The unique point $z \in N$ in Proposition 2.3 is called the *barycenter of* μ and is denoted by $b(\mu)$, that is,

$$b(\mu) := \operatorname*{arg\,min}_{z\in N} \int_N [d(z,x)^2 - d(y,x)^2] d\mu(x).$$

If $\mu \in \mathcal{P}^2(N)$, then

$$b(\mu) = \operatorname*{arg\,min}_{z\in N} \int_N d(z,x)^2 d\mu(x) \quad \text{and} \quad \mathrm{var}(\mu) = \int_N d(b(\mu),x)^2 d\mu(x).$$

The following proposition is very useful for our study.

Proposition 2.4 (Variance inequality ([17]))**.** *Let (N,d) be a Hadamard space. Then for any $\mu \in \mathcal{P}^1(N)$ and any $z \in N$ we have*

$$\int_N [d(z,x)^2 - d(b(\mu),x)^2] d\mu(x) \ge d(z,b(\mu))^2.$$

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Following [17], for a strong measurable function $A : \Omega \to N$, P_A is called a push-forward measure of P under A and is defined by

$$P_A(B) := P(A^{-1}(B)) \quad (B \in \mathcal{B}(N)).$$

Obviously, $P_A \in \mathcal{P}(N)$. Also $P_A \in \mathcal{P}^q(N)$ if and only if $A \in L^q(\Omega, N)$ for $1 \le q < \infty$ and $A \in L^\infty(\Omega, N)$ if and only if the function $x \mapsto d(A(x), y)$ is essentially bounded for some (and hence all) $y \in N$. The barycenter of A is defined by

$$b(A) := \operatorname*{arg\,min}_{z\in N} \int_\Omega [d(z,A(\omega))^2 - d(y,A(\omega))^2] dP(\omega)$$

$$= \underset{z\in N}{\arg\min} \int_N [d(z,x)^2 - d(y,x)^2] dP_A(x). \tag{2.2}$$

Proposition 2.5 (Variance inequality ([17])). *For any $A \in L^1(\Omega, N)$ we have*

$$\int_\Omega [d(z, A(\omega))^2 - d(b(A), A(\omega))^2] dP(\omega) \geq d(z, b(A))^2.$$

If K is a closed and convex subset of N, then there exists a unique map $\pi_K : N \to K$ such that $d(\pi_K(a), a) = \inf_{x\in K} d(x, a)$. The map π_K is called a (metric) projection onto K. Note that π_K is orthogonal, i.e., for any $z \in N$ and $w \in K$

$$d(z,w)^2 \geq d(z, \pi_K(z))^2 + d(\pi_K(z), w)^2, \tag{2.3}$$

(see [17]). An element in N can be identified as a constant map in $L^2(\Omega, N)$, then we can write $N \subset L^2(\Omega, N)$. So, N can be considered as a closed and convex subset of $L^2(\Omega, N)$ and the map $\pi_N : L^2(\Omega, N) \to N, \pi_N(A) = b(A)$ is a projection onto N (see [17]). Let $d_{L^2} : L^2(\Omega, N) \times L^2(\Omega, N) \to [0, \infty)$ be the metric on $L^2(\Omega, N)$ defined by

$$d_{L^2}(A, B) = \left(\int_\Omega d(A(\omega), B(\omega))^2 dP(\omega) \right)^{\frac{1}{2}}.$$

Since π_N is orthogonal, by (2.3) it holds that

$$d_{L^2}(A, B)^2 \geq d_{L^2}(A, b(A))^2 + d_{L^2}(b(A), B)^2,$$

that is,

$$\begin{aligned} &\int_\Omega d(A(\omega), B(\omega))^2 dP(\omega) \\ &\geq \int_\Omega d(A(\omega), b(A))^2 dP(\omega) + \int_\Omega d(b(A), B(\omega))^2 dP(\omega). \end{aligned} \tag{2.4}$$

2.3. Weighted inductive mean

Following [5] and [11], we introduce the notation of two types of weighted inductive mean. But depending on the case, we use a different notation for the convenience.

Let $\mathbf{x} = \{x_i\}$ be a sequence in N and $\{a_{ni}\}$ be a positive weighted sequence with $\sum_{i=1}^n a_{ni} = 1$ for any $n \in \mathbb{N}$. The new weighted sequence $\{y_{\ell i}\}_{i=1}^{\ell}$ for $\ell = 1, 2, \ldots, n-1$ is defined as follows:

$$y_{\ell i} = \frac{a_{ni}}{\sum_{k=1}^{\ell} a_{nk}} \quad \text{for all} \quad \ell = 1, \ldots, n-1. \tag{2.5}$$

Then by definition,

$$y_{\ell\ell} = \frac{a_{n\ell}}{\sum_{k=1}^{\ell} a_{nk}} \quad \text{for} \quad \ell = 1, \ldots, n-1.$$

The weighted inductive mean, denoted by $\{S_n(\mathbf{x})\}$, is defined as follows [5]:

$$S_1(\mathbf{x}) = x_1,$$
$$S_n(\mathbf{x}) = z_{n-1}(\mathbf{x})\#_{a_{nn}} x_n \quad (n \geq 2), \tag{2.6}$$

where $z_1 = x_1$, $z_\ell = z_{\ell-1}\#_{y_{\ell\ell}} x_\ell$ for $\ell = 2, \ldots, n-1$.

In [11], the deterministic weighted inductive mean, denoted by $\{D_n(\mathbf{x})\}$, is defined as follows: for a probability vector $(w_1, \ldots, w_n)$,

$$D_1(\mathbf{x}) = x_1,$$
$$D_{k+1}(\mathbf{x}) = D_k(\mathbf{x})\#_{d_{k+1}} x_{\overline{k+1}},$$

where $d_k = w_{\overline{k}}/l(k)$ with $l(k) = \sum_{i=1}^k w_{\overline{i}}$. The notation $\overline{k}$ means the residue of k modulo n and we set $\overline{n} = n$ for the convenience.

3. Ergodic theorem for weighted inductive means in Hadamard spaces

Let $(G, +)$ be a compact, abelian and metrizable topological group and let m be a Haar measure on G. For the ergodic theorem, we consider an ergodic automorphism $\tau(h) = h + g$ for some $g \in G$. We denote d_G by a shift invariant metric on G. Also, we assume that m is normalized, i.e., (G, m) is a probability space.

Remark 3.1. Following [18], we say that a topological dynamical system $(X, \mathcal{F}, T)$ is isometric if there exists a metric d on X such that the shift maps $T^n : X \to X$ are all isometries, i.e., $d(T^n x, T^n y) = d(x, y)$ for all $n \in \mathbb{N}$ and all $x, y \in X$. Also $(X, \mathcal{F}, T)$ is called a Kronecker system if it is isomorphic to a system of the form $(K, \mathcal{K}, S)$, where $(K, +)$ is a compact abelian metrizable topological group and $S : K \to K$, $S(x) = x + \alpha$ is a group rotation for some $\alpha \in K$. Note that every topological Kronecker system can be converted into a compact measure preserving system (see the paragraph under Corollary 2.11.9 in [18]).

By Propositions 2.6.7 and 2.6.9 in [18], every minimal isometric system is isomorphic to a Kronecker system and so any isometric system can be represented as the union of disjoint Kronecker systems.

In this section we always assume that (G, τ, m) is a Kronecker system.

3.1. L^∞ case

Following the notation in [5], we prove the ergodic theorem of the weighted inductive mean for the L^∞ case.

Lemma 3.2 ([5]). *Let $\{y_{\ell i}\}_{i=1}^{\ell}$ be the weighted sequence for $\ell = 1, \ldots, n-1$ in* (2.5). *Then we have*

$$\sum_{i=1}^n a_{ni}^2 = (1 - a_{nn})^2 \sum_{k=1}^{n-2} \left[\left(\prod_{i=1}^k (1 - y_{(n-i)(n-i)})^2 \right) y_{(n-(k+1))(n-(k+1))}{}^2 \right]$$

$$+ (1 - a_{nn})^2 {y_{(n-1)(n-1)}}^2 + a_{nn}^2.$$

Using the inequality (2.4), we have the following lemma.

Lemma 3.3. *Let $A, B \in L^2(G, N)$ and let $t \in [0, 1]$. Then the following inequality holds.*

$$\int_G d(B(g)\#_t A(g), b(A))^2 dm(g) \leq (1-t)^2 \int_G d(B(g), b(A))^2 dm(g) + t^2 \int_G d(A(g), b(A))^2 dm(g),$$

where $b(A)$ is the barycenter of A defined as (2.2).

Proof. For the proof, see Lemma 3.2 in [5]. □

Theorem 3.4. *Let $A \in L^2(G, N)$ and $g \in G$. Put $\mathbf{A}^{\tau,g} = \{A(\tau^{i-1}(g))\}_{i=1}^n$. Let $\{a_{ni}\}$ be a positive weighted sequence satisfying $\sum_{i=1}^n a_{ni} = 1$. Suppose $\{a_{ni}\}$ satisfies the following condition*:

$$\lim_{n\to\infty} \sum_{i=1}^n a_{ni}^2 = 0.$$

Then the weighted inductive mean $S_n(\mathbf{A}^{\tau,g})$ converges to $b(A)$ in measure.

Proof. The proof is a simple modification of the proof of Theorem 3.3 in [5]. Let $\{y_{\ell i}\}_{i=1}^{\ell}$ for $\ell = 1, 2, \ldots, n-1$ be the weighted sequence defined in (2.5). By Lemma 3.3, we have

$$\begin{aligned}
&\int_G d(S_n(\mathbf{A}^{\tau,g}), b(A))^2 dm(g) \\
&\leq (1 - a_{nn})^2 \prod_{i=1}^{n-2} (1 - y_{(n-i)(n-i)})^2 \int_G d(b(A), A(g))^2 dm(g) \\
&\quad + (1 - a_{nn})^2 \prod_{i=1}^{n-3} (1 - y_{(n-i)(n-i)})^2 {y_{22}}^2 \int_G d(b(A), A(\tau(g)))^2 dm(g) \\
&\quad \vdots \\
&\quad + (1 - a_{nn})^2 (1 - y_{(n-1)(n-1)})^2 {y_{(n-2)(n-2)}}^2 \int_G d(b(A), A(\tau^{n-3}(g)))^2 dm(g) \\
&\quad + (1 - a_{nn})^2 {y_{(n-1)(n-1)}}^2 \int_G d(b(A), A(\tau^{n-2}(g)))^2 dm(g) \\
&\quad + a_{nn}^2 \int_G d(b(A), A(\tau^{n-1}(g)))^2 dm(g).
\end{aligned}$$

Since τ is measure preserving, for all $i \in \mathbb{N}$,

$$\int_G d(b(A), A(\tau^{i-1}(g)))^2 dm(g) = \int_G d(b(A), A(g))^2 dm(g).$$

Therefore the above inequality is equivalent to

$$\begin{aligned}
&\int_G d(S_n(\mathbf{A}^{\tau,g}), b(A))^2 dm(g) \\
\leq\ & \left((1-a_{nn})^2 \sum_{k=1}^{n-2} \left[\left(\prod_{i=1}^{k} (1-y_{(n-i)(n-i)})^2 \right) y_{(n-(k+1))(n-(k+1))}{}^2 \right] \right. \\
&\qquad \left. + (1-a_{nn})^2 y_{(n-1)(n-1)}{}^2 + a_{nn}^2 \right) \int_G d(b(A), A(g))^2 dm(g) \\
=\ & \left(\sum_{i=1}^{n} a_{ni}^2 \right) \int_G d(b(A), A(g))^2 dm(g),
\end{aligned}$$

where the last equality holds by Lemma 3.2. From the condition

$$\lim_{n\to\infty} \sum_{i=1}^{n} a_{ni}^2 = 0,$$

we see that $S_n(\mathbf{A}^{\tau,g})$ converges to $b(A)$ in L^2. Since L^2 convergence implies the convergence in measure, the proof is complete. □

Remark 3.5. In fact, Theorem 3.4 holds when (G, m, τ) is a measure preserving dynamical system. Indeed, in the proof of Theorem 3.4, we only use the fact that τ is measure preserving for m. In this case, under the same conditions in Theorem 3.4 we have a following corollary.

Corollary 3.6 ([5, Theorem 3.3])**.** *Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\mathbf{X} = \{X_i\}$ be a sequence of independent identically distributed random variables in $L^2(\Omega, N)$. Let $\{a_{ni}\}$ be a positive weighted sequence such that $\sum_{i=1}^{n} a_{ni} = 1$ and*

$$\lim_{n\to\infty} \sum_{i=1}^{n} a_{ni}^2 = 0.$$

Then $S_n(\mathbf{X})$ converges to EX_1 in probability.

Proof. The proof is similar to the proof of Theorem 11.16 in [7]. Define $\phi : \Omega \to N^{\mathbb{N}}$ by $\phi(\omega) = (X_1(\omega), X_2(\omega), \ldots)$ and $m = P \circ \phi^{-1}$. Then m is a probability distribution on $N^{\mathbb{N}}$. Since $\{X_i\}$ are independent, m is of the form $m = \nu_1 \times \nu_2 \times \cdots$. Also since $\{X_i\}$ are identically distributed, if we let $\nu = P \circ X_1^{-1}$, then $\nu = \nu_i$ for all $i \in \mathbb{N}$ and $m = \Pi_{i\in\mathbb{N}}\nu$.

Let $G = N^{\mathbb{N}}$. Then (G, m) is a probability space. Define two functions $A : G \to N$ and $\tau : G \to G$ by

$$A((x_1, x_2, \ldots)) = x_1, \quad \tau((x_1, x_2, \ldots)) = (x_2, x_3, \ldots).$$

Then $A \in L^2(G, N)$ (since $\{X_i\} \subseteq L^2(\Omega, N)$), $\{X_i(\omega)\} = \{A(\tau^{i-1}(\phi(\omega)))\}$, and (G, m, τ) is a ergodic measure preserving dynamical system. Hence by Theorem 3.4, we have

$$S_n(\mathbf{X}) = S_n(\mathbf{A}^{\tau,\phi(\omega)}) \to b(A) = EX_1$$

in probability. □

So, the weighted version of weak law of large numbers can be regarded as the special case of some ergodic theorem (Theorem 3.4). It can be considered that the independence condition can be replaced by a measure preserving ergodic condition.

From Remark 3.6 and Lemma 3.8 in [5], we have the following two lemmas.

Lemma 3.7. *Let $A \in L^2(G,N)$ and let $\{a_{ni}\}$ be a positive weighted sequence satisfying $\sum_{i=1}^n a_{ni} = 1$. Then*

$$\int_G d(z_{n-i}(\mathbf{A}^{\tau,g}), b(A))^2 dm(g) \le \frac{\sum_{k=1}^{n-i} {a_{nk}}^2 \int_G d(A(\tau^{k-1}(g)), b(A))^2 dm(g)}{\left(\sum_{k=1}^{n-i} a_{nk}\right)^2},$$

where $\{z_{n-i}(\mathbf{A}^{\tau,g})\}_{i=1}^{n-1}$ is the sequence defined in (2.6).

Lemma 3.8. *Let $g \in G$ and $\mathbf{X} = \{X_i\}$ be a sequence in $L^\infty(G,N)$. Let $\{a_{ni}\}$ be a positive weighted sequence satisfying $\sum_{i=1}^n a_{ni} = 1$. Put $\mathbf{X}^g = \{X_i(g)\}$. Then there exists $R > 0$ such that for all $n \in \mathbb{N}$, $i = 1, 2, \dots, n-1$, and all $z \in N$,*

$$d(S_n(\mathbf{X}^g), z) \le \left(\sum_{k=1}^n a_{nk} d(X_k(g), z)\right) < R,$$

$$d(z_{n-i}(\mathbf{X}^g), z) \le \left(\frac{\sum_{k=1}^{n-i} a_{nk} d(X_k(g), z)}{\sum_{k=1}^{n-i} a_{nk}}\right) < R.$$

Note that for $A \in L^\infty(G,N)$ and all $i \in \mathbb{N}$, $A \circ \tau^{i-1}$ is in $L^\infty(G,N)$, where the operation $\circ$ is the composition of operators.

Theorem 3.9. *Let $A \in L^\infty(G,N)$ and let $\{a_{ni}\}$ be a positive weighted sequence satisfying $\sum_{i=1}^n a_{ni} = 1$. Suppose that there exists a constant $C \ge 1$ such that*

$$\max_{1 \le i \le n} a_{ni} \le C \min_{1 \le i \le n} a_{ni}.$$

Then

$$\lim_{n \to \infty} S_n(\mathbf{A}^{\tau,g}) = b(A) \quad a.e.$$

Proof. The proof is a simple modification of the proof of Theorem 3.7 in [5]. By triangle inequality, it holds that for all $k, n \in \mathbb{N}$,

$$d(S_n(\mathbf{A}^{\tau,g}), b(A)) \le d(S_n(\mathbf{A}^{\tau,g}), z_{k^2}(\mathbf{A}^{\tau,g})) + d(z_{k^2}(\mathbf{A}^{\tau,g}), b(A)).$$

Suppose that $k, n \in \mathbb{N}$ such that $k^2 < n \le (k+1)^2$. First, we will show that

$$d(S_n(\mathbf{A}^{\tau,g}), z_{k^2}(\mathbf{A}^{\tau,g})) \to 0 \quad \text{a.e.} \quad \text{as} \quad k \to \infty.$$

Since $A \in L^\infty(G,N)$, $A \circ \tau^{i-1} \in L^\infty(G,N)$ for all $i \in \mathbb{N}$ and so there exist $z \in N$ and $R > 0$ such that $d(A(\tau^{i-1}(g)), z) < R$ a.e.. By Lemma 3.8, for all $n \in \mathbb{N}$ and $i = 1, \dots, n-1$

$$d(S_n(\mathbf{A}^{\tau,g}), z) < R \quad \text{a.e.},$$

$$d(z_{n-i}(\mathbf{A}^{\tau,g}), z) < R \quad \text{a.e..}$$

Because the function $x \mapsto d(x,z)$ is convex, we have

$$d(S_n(\mathbf{A}^{\tau,g}), z_{n-1}(\mathbf{A}^{\tau,g})) \leq 2a_{nn}R \quad \text{a.e.} \tag{3.1}$$

and

$$d(z_{n-i}(\mathbf{A}^{\tau,g}), z_{n-(i+1)}(\mathbf{A}^{\tau,g})) \leq 2y_{(n-i)(n-i)}R \quad \text{a.e.} \tag{3.2}$$

for $i = 1, \ldots, n-2$. By the inequalities (3.1) and (3.2), we obtain that

$$\begin{aligned} d(S_n(\mathbf{A}^{\tau,g}), z_{k^2}(\mathbf{A}^{\tau,g})) &\leq 2(y_{(k^2+1)(k^2+1)} + y_{(k^2+2)(k^2+2)} + \cdots + a_{nn})R \\ &\leq \frac{\sum_{j=k^2+1}^{n} a_{nj}}{\sum_{j=1}^{k^2+1} a_{nj}} 2R \\ &\leq \frac{\sum_{j=k^2+1}^{n} \max_{1\leq i\leq n} a_{ni}}{\sum_{j=1}^{k^2+1} \min_{1\leq i\leq n} a_{ni}} 2R. \end{aligned}$$

Since $\max_{1\leq i\leq n} a_{ni} \leq C \min_{1\leq i\leq n} a_{ni}$ and $k^2 < n \leq (k+1)^2$, it holds that

$$\begin{aligned} d(S_n(\mathbf{A}^{\tau,g}), z_{k^2}(\mathbf{A}^{\tau,g})) &\leq \frac{(n-k^2)C \min_{1\leq i\leq n} a_{ni}}{(k^2+1)\min_{1\leq i\leq n} a_{ni}} 2R \\ &\leq \frac{2k+1}{k^2+1} 2RC \\ &\to 0 \quad \text{as} \quad k \to \infty. \end{aligned}$$

To end the proof, we show that $d(z_{k^2}(\mathbf{A}^{\tau,g}), b(A)) \to 0$ a.e. as $k \to \infty$. Let $\epsilon > 0$ be given. Note that since τ is measure-preserving, we have

$$\int_G d(A(\tau^{i-1}(g)), b(A))^2 dm(g) = \int_G d(A(g), b(A))^2 dm(g)$$

for all $i \in \mathbb{N}$. Since $\max_{1\leq i\leq n} a_{ni} \leq C \min_{1\leq i\leq n} a_{ni}$, by Chebyshev's inequality and Lemma 3.7

$$\begin{aligned} &\sum_{k=1}^{\infty} m(d(z_{k^2}(\mathbf{A}^{\tau,g}), b(A)) > \epsilon) \\ &\leq \frac{1}{\epsilon^2} \sum_{k=1}^{\infty} \int_G d(z_{k^2}(\mathbf{A}^{\tau,g}), b(A))^2 dm(g) \\ &\leq \frac{1}{\epsilon^2} \sum_{k=1}^{\infty} \frac{\sum_{i=1}^{k^2} {a_{ni}}^2}{\left(\sum_{i=1}^{k^2} a_{ni}\right)^2} \left(\int_G d(A(g), b(A))^2 dm(g)\right) \\ &\leq \frac{1}{\epsilon^2} \sum_{k=1}^{\infty} \frac{\sum_{i=1}^{k^2} \max_{1\leq j\leq n} {a_{ni}}^2}{\left(\sum_{i=1}^{k^2} \min_{1\leq j\leq n} a_{nj}\right)^2} \left(\int_G d(A(g), b(A))^2 dm(g)\right) \\ &\leq \frac{1}{\epsilon^2} \sum_{k=1}^{\infty} \frac{k^2C^2(\min_{1\leq j\leq n} a_{nj})^2}{k^4(\min_{1\leq j\leq n} a_{nj})^2} \left(\int_G d(A(g), b(A))^2 dm(g)\right) \end{aligned}$$

$$= \frac{C^2}{\epsilon^2} \sum_{k=1}^{\infty} \frac{1}{k^2} \left(\int_G d(A(g), b(A))^2 dm(g) \right) < \infty.$$

Hence $d(z_{k^2}(\mathbf{A}^{\tau,g}), b(A)) \to 0$ a.e. as $k \to \infty$ by Borel-Cantelli lemma. □

3.2. Continuous case: deterministic weighted inductive means

We follow the notation for the deterministic weighted inductive mean in Section 2. From the proof of Theorem 3.4 in [11], we have following two lemmas.

Lemma 3.10. *For a probability vector* $(a_{n1}, a_{n2}, \ldots, a_{nn})$ *and* $\mathbf{x} = \{x_i\} \subseteq N$, *we obtain*

$$d(D_{k+n}(\mathbf{x}), z)^2 \le \frac{l(k)}{l(k+n)} d(D_k(\mathbf{x}), z)^2 + \frac{1}{l(k+n)} \sum_{j=0}^{n-1} a_{n\overline{k+j+1}} d(x_{k+j+1}, z)^2 - \frac{l(k)}{l(k+n)^2} \sum_{j=0}^{n-1} a_{n\overline{k+j+1}} d(D_{k+j}(\mathbf{x}), x_{k+j+1})^2.$$

Lemma 3.11. *For a probability vector* $(a_{n1}, a_{n2}, \ldots, a_{nn})$ *and* $\mathbf{x} = \{x_i\} \subseteq N$ *with* $\triangle(\mathbf{x}) < \infty$, *we obtain*

$$\sum_{j=0}^{n-1} a_{n\overline{k+j+1}} d(D_k(\mathbf{x}), x_{k+j+1})^2 \le \frac{1+2l(k+1)}{l(k+1)^2} \triangle(\mathbf{x})^2 + \sum_{j=0}^{n-1} a_{n\overline{k+j+1}} d(D_{k+j}(\mathbf{x}), x_{k+j+1})^2,$$

where $\triangle(\mathbf{x}) = \sup_{n,m \in \mathbb{N}} d(x_n, x_m)$.

Let (X, ν) be a σ-finite measure space. A linear operator $T : L^1(\nu) \to L^1(\nu)$ is called a Dunford-Schwartz contraction if T is an L^1-contraction and it satisfies

$$\|Tf\|_\infty \le \|f\|_\infty \quad \text{for} \quad f \in L^1(\nu) \cap L^\infty(\nu).$$

For a Dunford-Schwartz contraction T, the limit $\mathcal{A}(T)f = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} T^k f$ exists a.e. for $f \in L^p(\nu)$, $1 \le p < \infty$ and in L^p norm for $p > 1$. Note that if (X, ν) is a probability space, then $\mathcal{A}(T)f$ exists in L^1 norm also.

The following result is a generalization of Theorem 3.1 in [12].

Theorem 3.12. *Let* $\{a_{ni}\}$ *be a sequence of positive real numbers and be triangular array with* $\sum_{k=1}^{n} a_{ni} = 1$. *Suppose that* $\lim_{n\to\infty} a_{ni} = 0$ *for any* $i \in \mathbb{N}$ *and*

$$\limsup_{n\to\infty} \left(\sum_{k=1}^{n-1} k|a_{nk} - a_{n(k+1)}| + n a_{nn} \right) < \infty.$$

Let (X, μ) *be a probability space and* $\{f_k\}$ *be a sequence in* $L^p(\mu)$. *Let* $\{\gamma_k\}$ *be a non-null sequence of nonnegative numbers satisfying* $\sup_{n\in\mathbb{N}} \frac{1}{n} \sum_{k=1}^{n} \gamma_k < \infty$.

Assume that $\frac{\sum_{k=1}^n f_k}{\sum_{k=1}^n \gamma_k} \to f$ *almost everywhere* (*in norm*) *as* $n \to \infty$. *Then it holds that*

$$\frac{\sum_{k=1}^n a_{nk} f_k}{\sum_{k=1}^n a_{nk}\gamma_k} \to f \tag{3.3}$$

almost everywhere (*in norm*) *as* $n \to \infty$.

Proof. The proof is similar to the proof in [19]. Let $F_n = \sum_{k=1}^n f_k$ and $G_n = \sum_{k=1}^n \gamma_k$. For a function $g \in L^p(\mu)$, denote $|g|$ by $|g(x)|$ given point x or $\|g\|_p$. By Abel's summation formula (see e.g, [2, 14]), we have

$$F_n^* := \sum_{k=1}^n a_{nk} f_k = \sum_{k=1}^{n-1} (a_{nk} - a_{n(k+1)}) F_k + a_{nn} F_n,$$
$$G_n^* := \sum_{k=1}^n a_{nk} \gamma_k = \sum_{k=1}^{n-1} (a_{nk} - a_{n(k+1)}) G_k + a_{nn} G_n.$$

Using these formulas, we obtain that

$$\left|\frac{F_n^*}{G_n^*} - f\right| \le \sum_{k=1}^{n-1} |a_{nk} - a_{n(k+1)}| \left|\frac{F_k}{G_k} - f\right| \frac{G_k}{G_n^*} + a_{nn}\left|\frac{F_n}{G_n} - f\right| \frac{G_n}{G_n^*}. \tag{3.4}$$

First, we assume that $\gamma_k \to \infty$. Then $G_n^* \to \infty$ by Theorems 3.2.2 and 3.6.10 in [8]. Let $\epsilon > 0$ be given. By assumption, there exists $N \in \mathbb{N}$ such that $\left|\frac{F_k}{G_k} - f\right| < \epsilon$ and $\frac{1}{G_n^*} < \epsilon$ for all $k \ge N$. Then the right hand side of (3.4) is equal to

$$\begin{aligned} &\sum_{k=1}^{N-1} |a_{nk} - a_{n(k+1)}| \left|\frac{F_k}{G_k} - f\right| \frac{G_k}{G_n^*} \\ &+ \sum_{k=N}^{n-1} k|a_{nk} - a_{n(k+1)}| \left|\frac{F_k}{G_k} - f\right| \frac{G_k}{kG_n^*} + na_{nn}\left|\frac{F_n}{G_n} - f\right| \frac{G_n}{nG_n^*}. \end{aligned} \tag{3.5}$$

The first summation term of (3.5) tends to 0 as $n \to \infty$. For the remainder terms of (3.5), since $\gamma := \sup_{n\in\mathbb{N}} \frac{1}{n} G_n < \infty$ by assumption, we have

$$\begin{aligned} &\sum_{k=N}^{n-1} k|a_{nk} - a_{n(k+1)}| \left|\frac{F_k}{G_k} - f\right| \frac{G_k}{kG_n^*} + na_{nn}\left|\frac{F_n}{G_n} - f\right| \frac{G_n}{nG_n^*} \\ &< \gamma\epsilon^2 \left(\sum_{k=N}^{n-1} k|a_{nk} - a_{n(k+1)}| + na_{nn}\right). \end{aligned}$$

Hence by taking the limit superior on both sides of (3.4), we get

$$\limsup_{n\to\infty} \left|\frac{F_n^*}{G_n^*} - f\right| \le \gamma\epsilon^2 \limsup_{n\to\infty} \left(\sum_{k=N}^{n-1} k|a_{nk} - a_{n(k+1)}| + na_{nn}\right) = C\gamma\epsilon^2,$$

where

$$C = \limsup_{n\to\infty} \left(\sum_{k=N}^{n-1} k|a_{nk} - a_{n(k+1)}| + na_{nn} \right) < \infty.$$

Suppose that the limit of γ_k is finite. Then the limit of G_n^* is also finite by Theorems 3.2.2 and 3.6.10 in [8], say $\lim_{n\to\infty} G_n^* = G$. Similarly to above computations, we have

$$\limsup_{n\to\infty} \left| \frac{F_n^*}{G_n^*} - f \right| \le \frac{\gamma\epsilon}{G} \limsup_{n\to\infty} \left(\sum_{k=N}^{n-1} k|a_{nk} - a_{n(k+1)}| + na_{nn} \right) = \frac{C\gamma}{G}\epsilon.$$

Finally, suppose that the limit of γ_n diverges. Since since $a_{nk}\gamma_k, k = 1, 2, \ldots, n$ is positive for all $n \in \mathbb{N}$, the limit of G_n^* is finite or ∞, so by previous computations, both cases are satisfied (3.3). The proof is completed. ☐

If we consider a sequence $\gamma_k \equiv 1$ in Theorem 3.12, we have the following corollary.

Corollary 3.13. *Let $\{a_{ni}\}$ be a sequence of positive real numbers and be triangular array with $\sum_{k=1}^n a_{ni} = 1$. Suppose that $\lim_{n\to\infty} a_{ni} = 0$ for any $i \in \mathbb{N}$ and*

$$\limsup_{n\to\infty} \left(\sum_{k=1}^{n-1} k|a_{nk} - a_{n(k+1)}| + na_{nn} \right) < \infty.$$

Let (X, μ) be a probability space and $\{f_k\}$ be a sequence in $L^p(\mu)$. Assume that $\frac{1}{n}\sum_{k=1}^n f_k \to f$ almost everywhere (in norm) as $n \to \infty$. Then it holds that

$$\sum_{k=1}^{n} a_{nk} f_k \to f$$

almost everywhere (in norm) as $n \to \infty$.

If ζ is a measure preserving map on Ω, then an operator T defined by $Tf := f \circ \zeta$ is a contraction on $L^\infty(\mu)$ and an isometry on $L^p(\mu)$, $1 \le p < \infty$. That is, the operator T is a Dunford-Schwartz contraction.

In our setting, let K be a compact subset of (N, d) and let $A : G \to N$ be continuous. If we define $F_n : G \times K \to \mathbb{R}$ by

$$F_n(g, x) = \frac{1}{n} \sum_{j=1}^{n} d(A(\tau^{j-1}(g)), x)^2$$

for all $n \in \mathbb{N}$, then $\{F_n\}_{n\in\mathbb{N}}$ is equicontinuous, so by Arzelà-Ascoli and Birkhoff ergodic theorem F_n converges to $\int_G d(A(\gamma), x)^2 dm(\gamma)$. This convergence is almost uniformly in $g \in G$ (see e.g., Lemma 3.2 and Proposition 3.3 in [1]). Especially, we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} d(A(\tau^{j-1}(g)), b(A))^2 = \int_G d(A(\gamma), b(A))^2 dm(\gamma).$$

Since (G, m) is a probability space and τ is measure preserving, an operator T defined by $Tf = f \circ \tau$ for $f \in L^1(G)$ is a Dunford-Schwartz contraction. For any $x \in K$, if we denote $f : G \to \mathbb{R}$ by $f(g) = d(A(g), x)^2$, then $f \in L^1(G)$ and by Corollary 3.13 we have the following proposition:

Proposition 3.14. *Let K be a compact subset of (N, d), $A : G \to N$ be a continuous function and $\{a_{ni}\}$ be a positive weighted sequence satisfying $\sum_{i=1}^n a_{ni} = 1$ and $\lim_{n\to\infty} a_{ni} = 0$ for any $i \in \mathbb{N}$. Suppose that*

$$\limsup_{n\to\infty} \left(na_{nn} + \sum_{k=1}^{n-1} k|a_{nk} - a_{n(k+1)}| \right) < \infty.$$

Then it holds that

$$\lim_{n\to\infty} \sum_{i=1}^{n} a_{ni} d(A(\tau^{i-1}(g)), x)^2 = \int_G d(A(\gamma), x)^2 dm(\gamma),$$

and the convergence is almost uniformly in $g \in G$.

Suppose that $A : G \to N$ is continuous. If we set

$$\alpha = \min_{x\in N} \int_G d(A(g), x)^2 dm(g),$$

then this value can be attained by $b(A)$, i.e.,

$$\alpha = \int_G d(A(g), b(A))^2 dm(g).$$

From now on, unless otherwise noted, we always assume that for any $n \in \mathbb{N}$, let $\{a_{ni}\}$ be a positive weighted sequence satisfying $\sum_{i=1}^n a_{ni} = 1$ and $\lim_{n\to\infty} a_{ni} = 0$ for any $i \in \mathbb{N}$, and

$$\limsup_{n\to\infty} \left(na_{nn} + \sum_{k=1}^{n-1} k|a_{nk} - a_{n(k+1)}| \right) < \infty.$$

Lemma 3.15. *Let $\{a_{ni}\}$ be a positive weighted sequence satisfying $\sum_{i=1}^n a_{ni} = 1$ and $\lim_{n\to\infty} a_{ni} = 0$ for any $i \in \mathbb{N}$. Let $\{b_m\}$ be a bounded sequence. Fix $k \in \mathbb{N}$. If $\lim_{n\to\infty} \sum_{i=1}^n a_{ni} b_i = \alpha$, then*

$$\lim_{n\to\infty} \sum_{i=1}^{n} a_{n\overline{k+i}} b_{k+i} = \alpha.$$

Proof. Since $\{b_m\}$ is bounded, there exists a constant $M > 0$ such that $|b_m| \le M$ for all $m \in \mathbb{N}$. Also since $\lim_{n\to\infty} a_{ni} = 0$ for any $i \in \mathbb{N}$, we have

$$\begin{aligned} \left| \sum_{i=1}^{n} a_{n\overline{k+i}} b_{k+i} - \alpha \right| &= \left| \sum_{i=1}^{n} a_{n\overline{k+i}} b_{k+i} + \sum_{i=1}^{k} a_{ni} b_i - \sum_{i=1}^{k} a_{ni} b_i - \alpha \right| \\ &= \left| \sum_{i=1}^{n} a_{ni} b_i - \alpha + \sum_{i=1}^{k} a_{ni} b_{n+i} - \sum_{i=1}^{k} a_{ni} b_i \right| \end{aligned}$$

$$
\begin{aligned}
&\leq \left|\sum_{i=1}^{n} a_{ni} b_i - \alpha\right| + \sum_{i=1}^{k} |a_{ni}||b_{n+i} - b_i| \\
&\leq \left|\sum_{i=1}^{n} a_{ni} b_i - \alpha\right| + 2M \sum_{i=1}^{k} |a_{ni}| \\
&\to 0 \quad \text{as} \quad n \to \infty. \qquad \square
\end{aligned}
$$

Lemma 3.16. *For any $\epsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ and all $k \in \mathbb{N}$, we have*

$$
\begin{aligned}
d(D_{k+n}(\mathbf{A}^{\tau,g}), b(A))^2 \leq{}& \frac{l(k)}{l(k+n)} d(D_k(\mathbf{A}^{\tau,g}), b(A))^2 + \frac{1}{l(k+n)}(\alpha + \epsilon) \\
&- \frac{l(k)}{l(k+n)^2} \sum_{j=0}^{n-1} a_{n\overline{k+j+1}} d(D_{k+j}(\mathbf{A}^{\tau,g}), \mathbf{A}^{\tau,g}_{k+j+1})^2,
\end{aligned}
$$

where $\mathbf{A}^{\tau,g}_{k+j+1}$ is the $(k+j+1)$-th term of the sequence $\mathbf{A}^{\tau,g}$, i.e., $\mathbf{A}^{\tau,g}_{k+j+1} = A(\tau^{(k+j+1)-1}(g))$.

Proof. Let $\epsilon > 0$ be given. By Lemma 3.15 and Proposition 3.14 there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$,

$$
\left|\sum_{j=0}^{n-1} a_{n\overline{k+j+1}} d(\mathbf{A}^{\tau,g}_{k+j+1}, b(A))^2 - \alpha\right| < \epsilon, \text{ almost everywhere } g \in G. \tag{3.6}
$$

Also by Lemma 3.10,

$$
\begin{aligned}
&d(D_{k+n}(\mathbf{A}^{\tau,g}), b(A))^2 \\
&\leq \frac{l(k)}{l(k+n)} d(D_k(\mathbf{A}^{\tau,g}), b(A))^2 + \frac{1}{l(k+n)} \sum_{j=0}^{n-1} a_{n\overline{k+j+1}} d(\mathbf{A}^{\tau,g}_{k+j+1}, b(A))^2 \\
&\quad - \frac{l(k)}{l(k+n)^2} \sum_{j=0}^{n-1} a_{n\overline{k+j+1}} d(D_{k+j}(\mathbf{A}^{\tau,g}), \mathbf{A}^{\tau,g}_{k+j+1})^2.
\end{aligned}
$$

Combining this inequality with (3.6), we obtain that for all $n \geq n_0$,

$$
\begin{aligned}
d(D_{k+n}(\mathbf{A}^{\tau,g}), b(A))^2 \leq{}& \frac{l(k)}{l(k+n)} d(D_k(\mathbf{A}^{\tau,g}), b(A))^2 + \frac{1}{l(k+n)}(\alpha + \epsilon) \\
&- \frac{l(k)}{l(k+n)^2} \sum_{j=0}^{n-1} a_{n\overline{k+j+1}} d(D_{k+j}(\mathbf{A}^{\tau,g}), \mathbf{A}^{\tau,g}_{k+j+1})^2. \qquad \square
\end{aligned}
$$

Note that we denote a diameter of a sequence $\mathbf{x} = \{x_i\}$ by

$$
\triangle(\mathbf{x}) = \sup_{n,m \in \mathbb{N}} d(x_n, x_m).
$$

If $A : G \to N$ is continuous, then by the compactness of G we have

$$\beta = \sup_{g\in G} \triangle(\mathbf{A}^{\tau,g}) < \infty.$$

Lemma 3.17. *For any $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ and all $k \in \mathbb{N}$,*

$$d(D_k(\mathbf{A}^{\tau,g}), b(A))^2 - \epsilon + \alpha - Q_{n,k} \leq \sum_{j=0}^{n-1} a_{n\overline{k+j+1}} d(D_{k+j}(\mathbf{A}^{\tau,g}), \mathbf{A}^{\tau,g}_{k+j+1})^2,$$

where $Q_{n,k} = \frac{1+2l(k+1)}{l(k+1)^2}\beta^2$.

Proof. Let K be the closure of the set of all convex combinations of $\{D_k(\mathbf{A}^{\tau,g}) : k \in \mathbb{N}\}$, where the convex hull is in the geodesic sense. Let $\epsilon > 0$ be given. By the variance inequality (Proposition 2.5) and Proposition 3.14, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$,

$$\begin{aligned} d(D_k(\mathbf{A}^{\tau,g}), b(A))^2 &\leq \int_G d(D_k(\mathbf{A}^{\tau,g}), A(\gamma))^2 dm(\gamma) - \alpha \\ &\leq \epsilon + \sum_{j=0}^{n-1} a_{n\overline{k+j+1}} d(D_k(\mathbf{A}^{\tau,g}), \mathbf{A}^{\tau,g}_{k+j+1})^2 - \alpha. \end{aligned}$$

By Lemma 3.11, we have

$$\begin{aligned} d(D_k(\mathbf{A}^{\tau,g}), b(A))^2 \leq& \frac{1+2l(k+1)}{l(k+1)^2} \triangle(\mathbf{A}^{\tau,g})^2 \\ &+ \sum_{j=0}^{n-1} a_{n\overline{k+j+1}} d(D_{k+j}(\mathbf{A}^{\tau,g}), \mathbf{A}^{\tau,g}_{k+j+1})^2 + \epsilon - \alpha \\ \leq& \frac{1+2l(k+1)}{l(k+1)^2}\beta^2 + \sum_{j=0}^{n-1} a_{n\overline{k+j+1}} d(D_{k+j}(\mathbf{A}^{\tau,g}), \mathbf{A}^{\tau,g}_{k+j+1})^2 \\ &+ \epsilon - \alpha. \end{aligned}$$

□

Lemma 3.18. *For any $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $t \in \mathbb{N}$,*

$$d(D_{tn_0}(\mathbf{A}^{\tau,g}), b(A))^2 \leq \frac{L}{t} + \epsilon \tag{3.7}$$

uniformly in $g \in G$, where $L = \alpha + 3\beta^2$.

Proof. Let $\epsilon > 0$ be given. By Lemma 3.16 and Lemma 3.17, there exists $n_0 \in \mathbb{N}$ such that for all $k \in \mathbb{N}$

$$\begin{aligned} d(D_{k+n_0}(\mathbf{A}^{\tau,g}), b(A))^2 \leq& \frac{l(k)}{l(k+n_0)} d(D_k(\mathbf{A}^{\tau,g}), b(A))^2 + \frac{1}{l(k+n_0)}(\alpha+\epsilon) \\ &- \frac{l(k)}{l(k+n_0)^2}(d(D_k(\mathbf{A}^{\tau,g}), b(A))^2 - \epsilon + \alpha - Q_{n_0,k}). \end{aligned}$$

Since $l(k+n_0) = l(k)+1$, the above inequality is equivalent to

$$d(D_{k+n_0}(\mathbf{A}^{\tau,g}), b(A))^2 \leq \frac{l(k)}{l(k+n_0)^2} d(D_k(\mathbf{A}^{\tau,g}), b(A))^2 + \frac{2l(k)+1}{l(k+n_0)^2}\epsilon + \frac{l(k)}{l(k+n_0)^2} Q_{n_0,k} + \frac{1}{l(k+n_0)^2}\alpha,$$

where $Q_{n_0,k} = \frac{1+2l(k+1)}{l(k+1)^2}\beta^2$. Put $k = tn_0$. Then

$$\begin{aligned} Q_{n_0,tn_0} &= \frac{1+2l(tn_0+1)}{l(tn_0+1)^2}\beta^2 \\ &= \left(\frac{1}{l(tn_0+1)^2} + \frac{2}{l(tn_0+1)}\right)\beta^2 \\ &\leq \left(\frac{1}{t^2} + \frac{2}{t}\right)\beta^2 \leq \left(\frac{1}{t} + \frac{2}{t}\right)\beta^2 = \frac{3}{t}\beta^2. \end{aligned}$$

Consequently,

$$d(D_{(t+1)n_0}(\mathbf{A}^{\tau,g}), b(A))^2 \leq \frac{t^2}{(t+1)^2} d(D_{tn_0}(\mathbf{A}^{\tau,g}), b(A))^2 + \frac{2t+1}{(t+1)^2}\epsilon + \frac{3}{(t+1)^2}\beta^2 + \frac{1}{(t+1)^2}\alpha.$$

Using this inequality, we prove this lemma using induction on (3.7). For $t = 1$, we have

$$d(D_{n_0}(\mathbf{A}^{\tau,g}), b(A))^2 \leq \beta^2 \leq L.$$

Suppose that the inequality (3.7) holds for $t = m$. If $t = m+1$, then we have

$$\begin{aligned} d(D_{(m+1)n_0}(\mathbf{A}^{\tau,g}), b(A))^2 &\leq \frac{1}{(m+1)^2}\left[m^2\left(\frac{L}{m}+\epsilon\right) + \alpha + (2m+1)\epsilon + 3\beta^2\right] \\ &= \frac{1}{(m+1)^2}(mL + (m+1)^2\epsilon + L) = \frac{L}{m+1} + \epsilon. \end{aligned}$$ □

Theorem 3.19. *Let $A : G \to N$ be a continuous operator. Then*

$$\lim_{n\to\infty} D_n(\mathbf{A}^{\tau,g}) = b(A)$$

almost uniformly in $g \in G$.

Proof. Let $\epsilon > 0$ be given. By Lemma 3.18, there exists $n_0 \in \mathbb{N}$ such that for all $t \in \mathbb{N}$,

$$d(D_{tn_0}(\mathbf{A}^{\tau,g}), b(A))^2 \leq \frac{L}{t} + \frac{\epsilon^2}{8}.$$

Since $\frac{1}{t} \to 0$ as $t \to \infty$, we can take $t_0 \in \mathbb{N}$ such that for all $t \geq t_0$,

$$d(D_{tn_0}(\mathbf{A}^{\tau,g}), b(A))^2 \leq \frac{\epsilon^2}{4}.$$

Since

$$d(D_n(\mathbf{A}^{\tau,g}), b(A)) \leq d(D_n(\mathbf{A}^{\tau,g}), D_{tn_0}(\mathbf{A}^{\tau,g})) + d(D_{tn_0}(\mathbf{A}^{\tau,g}), b(A)),$$

if we prove that $d(D_n(\mathbf{A}^{\tau,g}), D_{tn_0}(\mathbf{A}^{\tau,g})) \to 0$ for sufficiently large n, then the proof is complete. Let $n = tn_0 + k$ for $t \geq t_0$, $k \in \{1, \ldots, n_0 - 1\}$. Then

$$\begin{aligned} d(D_{tn_0+k}(\mathbf{A}^{\tau,g}), D_{tn_0}(\mathbf{A}^{\tau,g})) &\leq \sum_{i=0}^{k-1} d(D_{tn_0+i}(\mathbf{A}^{\tau,g}), D_{tn_0+i+1}(\mathbf{A}^{\tau,g})) \\ &\leq \sum_{i=0}^{k-1} \frac{a_{n_0\overline{tn_0+i+1}}}{l(tn_0+i+1)} d(D_{tn_0+i}(\mathbf{A}^{\tau,g}), \mathbf{A}^{\tau,g}_{tn_0+i+1}) \\ &< \sum_{i=0}^{k-1} \frac{a_{n_0\overline{tn_0+i+1}}}{l(tn_0+i+1)} \beta \leq \sum_{i=0}^{k-1} \frac{1}{l(tn_0)} a_{n_0\overline{tn_0+i+1}} \beta \\ &= \frac{1}{t} \sum_{i=0}^{k-1} a_{n_0\overline{tn_0+i+1}} \beta \leq \frac{1}{t}\beta \to 0 \end{aligned}$$

as $t \to \infty$. □

Considering the notation of weighted inductive mean in [5] again, by the compactness of G, a continuous function from G into N is in $L^\infty(G, N)$. Hence we get the following theorem.

Theorem 3.20. *For any $n \in \mathbb{N}$, let $\{a_{ni}\}$ be a positive weighted sequence satisfying $\sum_{i=1}^n a_{ni} = 1$. Let $A : G \to N$ be continuous. Suppose that there exists a constant $C \geq 1$ such that*

$$\max_{1 \leq i \leq n} a_{ni} \leq C \min_{1 \leq i \leq n} a_{ni}.$$

Then

$$\lim_{n\to\infty} S_n(\mathbf{A}^{\tau,g}) = b(A) \quad a.e..$$

3.3. L^1-case

In this section, we prove the ergodic theorem for $A \in L^1(G, N)$, which is one of main results in this paper. To prove the result, we use L^1-approximation by continuous functions introduced in [1].

For $1 \leq p < \infty$, let $A \in L^p(G, N)$ and define $\phi : G \to [0, \infty)$ by

$$\phi(h) = \int_G d(A(g), A(g+h))^p dm(g).$$

Then ϕ is continuous on G.

Let $\epsilon > 0$ be given and let O_ϵ be a neighborhood of the identity of G such that $m(O_\epsilon) < \epsilon$ and diam $O_\epsilon = \sup_{x,y\in O_\epsilon} d(x, y) < \epsilon$. For fixed $y \in N$, let

$$A_\epsilon(g_0) = \arg\min_{z\in N} \int_{O_\epsilon} [d(z, A(g+g_0))^2 - d(y, A(g+g_0))^2] dm(g).$$

That is, $A_\epsilon(g_0)$ is the barycenter of the pushforward by A of the Haar measure m restricted to $g_0 + O_\epsilon$. Then A_ϵ is continuous by the continuity of ϕ and it holds that for $A \in L^1(G, N)$

$$\lim_{\epsilon \to 0+} \int_G d(A(g), A_\epsilon(g))dm(g) = 0.$$

It means that A_ϵ converges to A in L^1. For more details, see [1].

Lemma 3.21 ([1, Lemma 3.8]). *Given $A, B \in L^1(G, N)$, we have*

$$d(b(A), b(B)) \leq \int_G d(A(g), B(g))dm(g).$$

By simple computation, we obtain that the following lemma:

Lemma 3.22. *Let $g \in G$ and $\mathbf{X}^g = \{X_i(g)\}, \mathbf{Z}^g = \{Z_i(g)\} \subset L^1(G, N)$. For a positive weighted sequence $\{a_{ni}\}$ with $\sum_{i=1}^n a_{ni} = 1$ we have*

$$d(T_n(\mathbf{X}^g), T_n(\mathbf{Z}^g)) \leq \sum_{i=1}^n a_{ni} d(X_i(g), Z_i(g)),$$

where T_n is a weighted inductive mean, i.e., $T_n = S_n$ or $T_n = D_n$.

Lemma 3.23. *Let $A, B \in L^1(G, N)$. Given $\epsilon > 0$, for almost every $g \in G$ there exists n_0 such that for all $n \geq n_0$*

$$d(T_n(\mathbf{A}^{\tau,g}), T_n(\mathbf{B}^{\tau,g})) \leq \epsilon + \int_G d(A(g), B(g))dm(g),$$

where T_n is a weighted inductive mean, i.e., $T_n = S_n$ or $T_n = D_n$. Note that the number n_0 depends on g.

Proof. Since $A, B \in L^1(G, N)$, by the Birkhoff's ergodic theorem

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^n d(A(\tau^{i-1}(g)), B(\tau^{i-1}(g))) = \int_G d(A(\gamma), B(\gamma))dm(\gamma) \quad \text{a.e.}.$$

Also by Corollary 3.13

$$\lim_{n \to \infty} \sum_{i=1}^n a_{ni} d(A(\tau^{i-1}(g)), B(\tau^{i-1}(g))) = \int_G d(A(\gamma), B(\gamma))dm(\gamma) \quad \text{a.e.}.$$

Let $\epsilon > 0$ be given. Then there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ and for almost all $g \in G$, by Lemma 3.22, we have

$$\begin{aligned} d(T_n(\mathbf{A}^{\tau,g}), T_n(\mathbf{B}^{\tau,g})) &\leq \sum_{i=1}^n a_{ni} d(A(\tau^{i-1}(g)), B(\tau^{i-1}(g))) \\ &\leq \epsilon + \int_G d(A(\gamma), B(\gamma))dm(\gamma). \end{aligned}$$

□

Now, we prove the ergodic theorem for $A \in L^1(G,N)$. Firstly, we use the notation for the weighted inductive mean in [11]. This notation means that a weighted sequence is deterministic.

Theorem 3.24. *Let $A \in L^1(G,N)$. Then for almost all $g \in G$*

$$\lim_{n\to\infty} D_n(\mathbf{A}^{\tau,g}) = b(A).$$

Proof. The proof is a simple modification of Theorem 3.7 in [1]. Let $\epsilon > 0$ be given. Firstly, for each $k \in \mathbb{N}$ we can find a continuous function $A_{\frac{1}{k}}$ from G to N such that

$$\int_G d(A(g), A_{\frac{1}{k}}(g))dm(g) \le \frac{1}{k}.$$

By Lemma 3.23, we take a measure zero set $G_0 \subset G$ and take $g \in G \backslash G_0$ and it holds that there exist a constant $k > 0$ and n_0 such that for all $n \ge n_0$

$$d(D_n(\mathbf{A}^{\tau,g}), D_n(\mathbf{A}^{\tau,g,k})) \le \frac{\epsilon}{6} + \int_G d(A(g), A_{\frac{1}{k}}(g))dm(g), \tag{3.8}$$

where $\mathbf{A}^{\tau,g,k} = \{A_{\frac{1}{k}}(\tau^{i-1}(g))\}_{i=1}^{\infty}$. Since $A_{\frac{1}{k}}$ is continuous, $D_n(\mathbf{A}^{\tau,g,k}) \to b(A_{\frac{1}{k}})$ as $n \to \infty$ by Theorem 3.19. So there exists $n_1 \in \mathbb{N}$ such that for all $n \ge n_1$,

$$d(D_n(\mathbf{A}^{\tau,g,k}), b(A_{\frac{1}{k}})) < \frac{\epsilon}{3}. \tag{3.9}$$

By Lemma 3.21, we have

$$d(b(A), b(A_{\frac{1}{k}})) \le \int_G d(A(g), A_{\frac{1}{k}}(g))dm(g) \le \frac{1}{k}. \tag{3.10}$$

Take $k \in \mathbb{N}$ such that $\frac{1}{k} \le \frac{\epsilon}{6}$. Then by (3.8) and (3.10), we have

$$d(D_n(\mathbf{A}^{\tau,g}), D_n(\mathbf{A}^{\tau,g,k})) \le \frac{\epsilon}{6} + \frac{\epsilon}{6} = \frac{\epsilon}{3}, \tag{3.11}$$

$$d(b(A), b(A_{\frac{1}{k}})) \le \int_G d(A(g), A_{\frac{1}{k}}(g))dm(g) \le \frac{1}{k} \le \frac{\epsilon}{6} < \frac{\epsilon}{3}. \tag{3.12}$$

Consequently, combining (3.9), (3.11), and (3.12), we have

$$\begin{aligned} d(D_n(\mathbf{A}^{\tau,g}), b(A)) &\le d(D_n(\mathbf{A}^{\tau,g}), D_n(\mathbf{A}^{\tau,g,k})) \\ &\quad + d(D_n(\mathbf{A}^{\tau,g,k}), b(A_{\frac{1}{k}})) + d(b(A_{\frac{1}{k}}), b(A)) \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon \end{aligned}$$

for sufficiently large n. □

The notation for the weighted inductive mean in [5] means that the weighted sequence is random. The ergodic theorem for this case can be proved as Theorem 3.24 similarly.

Theorem 3.25. *Let $\{a_{ni}\}$ be a positive weighted sequence satisfying $\sum_{i=1}^{n} a_{ni} = 1$ and $\lim_{n\to\infty} a_{ni} = 0$ for any $i \in \mathbb{N}$. Let $A \in L^1(G, N)$. Suppose that there exists a constant $C \geq 1$ such that*

$$\max_{1\leq i\leq n} a_{ni} \leq C \min_{1\leq i\leq n} a_{ni}.$$

Then we have

$$\lim_{n\to\infty} S_n(\mathbf{A}^{\tau,g}) = b(A)$$

for almost all $g \in G$.

References

[1] J. Antezana, E. Ghiglioni, and D. Stojanoff, *Ergodic theorem in Hadamard spaces in terms of inductive means*, preprint, 2018.

[2] T. M. Apostol, *Mathematical Analysis*, second edition, Addison-Wesley Publishing Co., Reading, MA, 1974.

[3] T. Austin, *A* CAT(0)*-valued pointwise ergodic theorem*, J. Topol. Anal. **3** (2011), no. 2, 145–152. https://doi.org/10.1142/S1793525311000544

[4] Y. J. Cho, L. Ćirić, and S. Wang, *Convergence theorems for nonexpansive semigroups in* CAT(0) *spaces*, Nonlinear Anal. **74** (2011), no. 17, 6050–6059. https://doi.org/10.1016/j.na.2011.05.082

[5] B. J. Choi, J. Heo, and U. C. Ji, *Law of large numbers for weighted inductive means in a Hadamard space*, Math. Nachr. **290** (2017), no. 17-18, 2775–2787. https://doi.org/10.1002/mana.201600504

[6] C. Cuny and M. Weber, *Ergodic theorems with arithmetical weights*, Israel J. Math. **217** (2017), no. 1, 139–180. https://doi.org/10.1007/s11856-017-1441-y

[7] T. Eisner, B. Farkas, M. Haase, and R. Nagel, *Operator Theoretic Aspects of Ergodic Theory*, Graduate Texts in Mathematics, 272, Springer, Cham, 2015. https://doi.org/10.1007/978-3-319-16898-2

[8] G. H. Hardy, *Divergent Series*, Oxford, at the Clarendon Press, 1949.

[9] F. Hiai and Y. Lim, *Convergence theorems for barycentric maps*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **22** (2019), no. 3, 1950016, 40 pp. https://doi.org/10.1142/S0219025719500164

[10] J. Holbrook, *No dice*: *a deterministic approach to the Cartan centroid*, J. Ramanujan Math. Soc. **27** (2012), no. 4, 509–521.

[11] Y. Lim and M. Pálfia, *Weighted deterministic walks for the squares mean on Hadamard spaces*, Bull. Lond. Math. Soc. **46** (2014), no. 3, 561–570. https://doi.org/10.1112/blms/bdu008

[12] M. Lin and M. Weber, *Weighted ergodic theorems and strong laws of large numbers*, Ergodic Theory Dynam. Systems **27** (2007), no. 2, 511–543. https://doi.org/10.1017/S0143385706000769

[13] 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), no. 2, 609–623. https://doi.org/10.1017/S0143385711001015

[14] C. P. Niculescu and M. M. Stănescu, *A note on Abel's partial summation formula*, Aequationes Math. **91** (2017), no. 6, 1009–1024. https://doi.org/10.1007/s00010-017-0504-9

[15] B. Piątek, *Halpern iteration in* CAT(κ) *spaces*, Acta Math. Sin. (Engl. Ser.) **27** (2011), no. 4, 635–646. https://doi.org/10.1007/s10114-011-9312-7

[16] S. Saejung, *Halpern's iteration in* CAT(0) *spaces*, Fixed Point Theory Appl. **2010** (2010), Art. ID 471781, 13 pp. https://doi.org/10.1155/2010/471781

[17] K.-T. Sturm, *Probability measures on metric spaces of nonpositive curvature*, in Heat kernels and analysis on manifolds, graphs, and metric spaces (Paris, 2002), 357–390, Contemp. Math., 338, Amer. Math. Soc., Providence, RI, 2003. https://doi.org/10.1090/conm/338/06080

[18] T. Tao, *Poincaré's legacies, pages from year two of a mathematical blog. Part I*, American Mathematical Society, Providence, RI, 2009. https://doi.org/10.1090/mbk/066

[19] M. Weber, *Dynamical systems and processes*, IRMA Lectures in Mathematics and Theoretical Physics, 14, European Mathematical Society (EMS), Zürich, 2009. https://doi.org/10.4171/046

Byoung Jin Choi
Department of Mathematics Education
Jeju National University
Jeju 63243, Korea
Email address: choibj@jejunu.ac.kr

Jae Hun Kim
Department of Mathematics
Chungbuk National University
Cheongju 28644, Korea
Email address: jaehunkim@chungbuk.ac.kr