COLLOQUIUM MATHEMATICUM

VOL. 149 2017 NO. 1

EFFECTIVE EQUIDISTRIBUTION OF PERIODIC ORBITS FOR SUBSHIFTS OF FINITE TYPE

BY

SHIRALI KADYROV (Astana)

Abstract. We study equidistribution of certain subsets of periodic orbits for subshifts of finite type. Our results solely rely on the growth of these subsets. As a consequence, effective equidistribution results are obtained for both hyperbolic diffeomorphisms and expanding maps on compact manifolds.

1. Introduction. For a given $s \in \mathbb{N}$ and an $s \times s$ transition matrix A with entries zero or one we let (Σ_A^+, σ) denote the one-sided subshift of finite type where Σ_A^+ is the symbolic space given by

$$\Sigma_A^+ = \Big\{x = (x_n)_{n\geq 0} \in \prod_{n=0}^{\infty}\{1,\dots,s\} : A(x_n, x_{n+1}) = 1,\, \forall n \in \mathbb{N}\Big\},$$

and $\sigma : \Sigma_A^+ \to \Sigma_A^+$ is the shift map by $(\sigma(x))_n = x_{n+1}$. For a given $\theta \in (0,1)$ define a metric d_θ on Σ_A^+ by $d_\theta(x,y) = \theta^{t(x,y)}$ where $t(x,y) = \max\{n \geq 0 : x_i = y_i,\, 0 \leq i < n\}$. We similarly define the two-sided subshift of finite type (Σ_A, σ) where

$$\Sigma_A = \Big\{x = (x_n)_{n\geq 0} \in \prod_{n=-\infty}^{\infty}\{1,\dots,s\} : A(x_n, x_{n+1}) = 1, \forall n \in \mathbb{Z}\Big\}$$

and the metric is given by $d_\theta(x,y) = \theta^{t(x,y)}$ where $t(x,y) = \max\{n \geq 0 : x_i = y_i, |i| < n\}$. Also, for any continuous function g on Σ_A^+ we let

$$|g|_\theta = \sup_{n\geq 0}\Big\{\frac{|g(x) - g(y)|}{\theta^n} : x_i = y_i,\, 0 \leq i \leq n\Big\}.$$

We similarly define $|\cdot|_\theta$ on Σ_A with $0 \leq i \leq n$ replaced by $|i| \leq n$. In particular, $|g|_\theta < \infty$ implies that g is a Lipschitz function with the least Lipschitz constant $|g|_\theta$. Consider a norm $\|\cdot\|_\theta = |\cdot|_\theta + |\cdot|_\infty$ where $|\cdot|_\infty$ is the

2010 *Mathematics Subject Classification*: Primary 37A35, 37D20; Secondary 28D20, 37D05.

Key words and phrases: closed orbits, maximal entropy, symbolic dynamics, equidistribution.

Received 1 June 2015; revised 25 August 2016.

Published online 21 April 2017.

DOI: 10.4064/cm6653-9-2016

supremum norm and let $\mathcal{F}_\theta^+$ denote the space of all continuous functions f on Σ_A^+ with $\|f\|_\theta < \infty$. Analogously we define $\mathcal{F}_\theta$ on Σ_A. For both (Σ_A^+, σ) and (Σ_A, σ) we let $h(\sigma)$ denote the topological entropy and m the measure of maximal entropy so that $h(\sigma) = h_m(\sigma)$. Let $\xi = \{C_1, \dots, C_s\}$ denote the generating partition of Σ_A^+ (or Σ_A), where $C_i = \{x : x_0 = i\}$. To simplify the notation we let $\xi_\ell^n := \bigvee_{i=\ell}^n \sigma^{-i}\xi$. In [K] we studied the effective uniqueness of m, the measure of maximal entropy. By effective uniqueness we mean a statement that specifies how close a given measure is to the measure of maximal entropy if its metric entropy is close to the maximal entropy. See results from [Po, R] similar to [K].

In this paper we obtain the following improvement of [K, Theorem 1.1], which will lead to effective equidistribution statements. The matrix A is said to be *irreducible* if for each pair (i,j) there exists $n \geq 1$ such that $A^n(i,j) > 0$. We say that A is *aperiodic* if $A(i,i) = 1$ for all $i = 1, \dots, s$.

THEOREM 1.1. *Assume that A is irreducible and aperiodic. Then there exists a constant $c > 0$ such that for any $N \in \mathbb{N} \cup \{\infty\}$, any σ-invariant probability measure μ on Σ_A^+ and any Lipschitz function f we have*

$$(1.1) \qquad \left| \int f \, d\mu - \int f \, dm \right| \leq c \|f\|_\theta \left(\theta^{N/2} + 2\sqrt{2} \left(h(\sigma) - \frac{1}{N} H_\mu(\xi_0^{N-1}) \right)^{1/2} \right),$$

where m is the measure of maximal entropy on Σ_A^+. Moreover, the same result holds for the two-sided subshift (Σ_A, σ) with the exponent $N/2$ of θ replaced by $N/4$.

This generalizes and improves [Po, Theorems 4.1.2 and 4.1.3].

We now want to discuss how Theorem 1.1 can be applied to show the effective equidistribution of periodic orbits. In other words, we want to obtain a rate of convergence of the distribution of periodic orbits. For any $n \in \mathbb{N}$ we let Fix_n denote the set of periodic points of period n,

$$\mathrm{Fix}_n = \{x \in \Sigma_A^+ : \sigma^n(x) = x\}.$$

By abuse of notation we let Fix_n also denote the periodic orbits of Σ_A of order n. For any nonempty finite set I in Σ_A^+ (or Σ_A) we let μ_I denote the uniform probability measure supported on I,

$$\mu_I = \frac{1}{|I|} \sum_{x \in I} \delta_x.$$

Clearly each nonempty element of ξ_0^{n-1} contains exactly one element from Fix_n. More precisely, for any $x \in \mathrm{Fix}_n$ we have

$$x \in P(x,n) := C_{x_0} \cap \sigma^{-1} C_{x_1} \cap \dots \cap \sigma^{-(n-1)} C_{x_{n-1}} \in \xi_0^{n-1}.$$

Thus, for any $P \in \xi_0^{n-1}$ we have $\mu_I(P) = 1/|I|$ if $P = P(x,n)$ for some

$x \in I$, and $\mu_I(P) = 0$ otherwise. Thus,

$$H_{\mu_I}(\xi_0^{n-1}) = \log|I|.$$

Consequently, applying Theorem 1.1 we see that for any Lipschitz f,

$$\left|\int f\,d\mu - \int f\,dm\right| \le c\|f\|_\theta\left(\theta^{n/4} + 2\sqrt{2}\left(h_m(\sigma) - \frac{1}{n}\log|I|\right)^{1/2}\right),$$

which proves the following.

THEOREM 1.2. *Fix $n \in \mathbb{N}$ and let I be a nonempty invariant subset of* Fix_n. *Then there exists $c > 0$ such that for any Lipschitz function f we have*

$$\left|\int f\,dm - \int f\,d\mu_I\right| \le c\|f\|_\theta\left(\theta^{n/4} + 2\sqrt{2}\sqrt{h(\sigma) - \frac{1}{n}\log|I|}\right).$$

As an immediate consequence we get

THEOREM 1.3. *If $\{I_n\}$ is a sequence of invariant sets with $I_n \subset \mathrm{Fix}_n$ and $\varphi(n) := h(\sigma) - \frac{1}{n}\log|I_n| \to 0$ as $n \to \infty$ then there exists a constant $c > 0$ such that for any $n \in \mathbb{N}$ and any Lipschitz function f we have*

$$\left|\int f\,dm - \int f\,d\mu_{I_n}\right| \le c\|f\|_\theta\big(\theta^{n/4} + 2\sqrt{2}\sqrt{\varphi(n)}\big).$$

A similar result was studied in a different context in [AE]. We note that our methods are completely different from those of [AE]. It is well known (see e.g. [PY, Sublemma 4.10.1]) that $|\mathrm{Fix}_n| \sim e^{h(\sigma)n}$, i.e., $\lim_{n\to\infty}|\mathrm{Fix}_n|/e^{h(\sigma)n} = 1$. In fact, it is easy to see that $|\mathrm{Fix}_n| = \mathrm{tr}(A^n) = \lambda_1^n + \cdots + \lambda_s^n$, where λ_i's are the eigenvalues of A with $\lambda_1 > |\lambda_i|$ for all $i \ne 1$. Thus, there exists $\delta > 0$ such that $|\mathrm{Fix}_n| = \lambda_1^n(1 + O(e^{-\delta n}))$, and since $h(\sigma) = \log\lambda_1$ we deduce that

$$h(\sigma) - \frac{1}{n}\log|\mathrm{Fix}_n| = O(e^{-\delta n}).$$

Hence, as a particular case of Theorem 1.3 we obtain the effective equidistribution of periodic orbits:

THEOREM 1.4. *There exist constants $c, \delta > 0$ such that for any Lipschitz function f and any $n \in \mathbb{N}$ we have*

$$\left|\int f\,dm - \int f\,d\mu_{\mathrm{Fix}_n}\right| \le c\|f\|_\theta\, e^{-\delta n}.$$

It is well known that repellers and Axiom A diffeomorphisms admit Markov partitions. We refer to [B1, B2, S1, S2, K] for more details. Consequently, we can realize a repeller (J, T) and an Axiom A diffeomorphism (Ω, T) as a factor of a subshift of finite type. By abuse of notation let Fix_n denote the set of closed orbits x with $T^n x = x$. Using this link together with Theorem 1.1 we can obtain the effective equidistribution statements for closed orbits.

Theorem 1.5. *Let (J,T) be a mixing repeller or $(\Omega(T),T)$ be a mixing Axiom A diffeomorphism. Let $\{I_n\}$ be a sequence of invariant sets with $I_n \subset \mathrm{Fix}_n$ and $\varphi(n) := h(T) - \frac{1}{n}\log|I_n| \to 0$ as $n\to\infty$ where $h(T)$ is the topological entropy. Then for any Lipschitz function f there exist constants $C(f)>0$ and $\theta\in(0,1)$ such that*

$$\left|\int f\,dm - \int f\,d\mu_{I_n}\right| \le C(f)\big(\theta^n + 2\sqrt{2}\sqrt{\varphi(n)}\big)$$

for any $n\in\mathbb{N}$. Moreover, there exists a constant $\delta>0$ such that for any Lipschitz function f we have

$$\left|\int f\,dm - \int f\,d\mu_{\mathrm{Fix}_n}\right| \le C(f)e^{-\delta n}.$$

As stated in [K, Lemma 4.1], any invariant measure on J or $\Omega(T)$ can be lifted to an invariant measure on a subshift of finite type. This is sufficient to derive the above theorem from Theorem 1.1. We skip the details of the proof and refer to [K] for more information about using the standard arguments. We note that defining suitable norms on J or Ω one can make the dependence of $C(f)$ on f precise. Equidistribution of closed orbits of expanding maps and of hyperbolic diffeomorphisms was obtained by M. Misiurewicz [M] and by R. Bowen [B3, B1], respectively. Our results in Theorem 1.5 generalize and improve these results with exponential error terms.

We can also consider a subset Fix'_n of Fix_n consisting of primitive periodic orbits, that is, orbits with the least period n. When $T:\mathbb{T}^d\to\mathbb{T}^d$ is a linear hyperbolic automorphism of the d-torus, the equidistribution of Fix'_n was obtained in [DI]. From [DI, Proposition 2.3] it follows that $\mathrm{Fix}'_n \sim e^{h(T)n}/n$, which implies that $h(T)-(\log|\mathrm{Fix}'_n|)/n = O((\log n)/n)$. Thus, Theorem 1.5 in this special case gives the equidistribution of Fix'_n with the error term $O(\sqrt{(\log n)/n})$, improving [DI, Proposition 2.4].

2. Proof of Theorem 1.1. We first note that the deduction of the second half of Theorem 1.1 from (1.1) is standard as we sketch it now. It follows from [B2, PP] that if $f\in\mathcal{F}_\theta$ in Σ_A then it is cohomologous to $f'\in\mathcal{F}_{\theta^{1/2}}$ with $f'(x)=f'(y)$ for all $x,y\in\Sigma_A$ satisfying $x_i=y_i$ for all $i\ge 0$. More precisely, there exist $f',u\in\mathcal{F}_{\theta^{1/2}}$ and a constant $C>0$ independent of f such that $\|f'\|_{\theta/2}\le C\|f\|_\theta$ and $f+u\circ\sigma-u=f'$. Then f' can be considered as a function in $\mathcal{F}^+_{\theta/2}$. Also, any invariant measure on Σ_A can be considered as an invariant measure on Σ^+_A. Thus, using $\int f d\mu - \int f\,dm = \int f' d\mu - \int f'\,dm$ and (1.1) we obtain the second half of the theorem. Consequently, it suffices to prove Theorem 1.1 for the one-sided subshift.

We recall that $\xi_0^{N-1} = \bigvee_{i=0}^{N-1}\sigma^{-i}\xi$ where N is a natural number or $N=\infty$. We first state the properties of the measure m of maximal entropy, known as Parry measure [Pa]. This helps us to study the information function $I_m(\xi|\xi_1^\infty)$ and see that $\int I_m(\xi|\xi_1^\infty)d\mu = h_m(\sigma)$ for any invariant measure μ. We then use the Pinsker inequality to relate the difference $I_m(\xi|\xi_1^\infty) - I_\mu(\xi|\xi_1^N)$ to the difference of entropies of partitions. What remains to do is to relate the difference of information functions to $\int f\,dm - \int f\,d\mu$, and this is done by constructing the sequence of functions $f_n = \mathcal{L}^n f$ using the transfer operator $\mathcal{L}$ for subshifts of finite type.

Let A be an $s \times s$ irreducible and aperiodic transition matrix and $\lambda > 0$ its largest eigenvalue. It follows from Perron–Frobenius theory [W, §0.9] that there are strictly positive left and right eigenvectors $(u_0, u_1, \ldots, u_{s-1})$ and $(v_0, v_1, \ldots, v_{s-1})$ respectively with $\sum_{i=0}^{s-1} u_i v_i = 1$. We set $p_i = u_i v_i$ and $p_{ij} = a_{ij} v_j / \lambda v_i$. Then the Markov measure m given by the probability vector $\mathbf{p} = (p_0, p_1, \ldots, p_{s-1})$ and the stochastic matrix (p_{ij}) is the unique measure of maximal entropy [W, Theorem 8.10]. It is easy to see that for any admissible $(i_0, i_1, \ldots, i_k)$, the $(k+1)$-cylinder set $C(i_0, i_1, \ldots, i_k) := \{x \in \Sigma_A^+ : x_0 = i_0, \ldots, x_k = i_k\} = \bigcap_{n=0}^{k} \sigma^{-n} C_{i_n}$ satisfies

$$m(C(i_\ell, \ldots, i_{\ell+k})) = \frac{u_{i_\ell} v_{i_{\ell+k}}}{\lambda^k}. \tag{2.1}$$

For any partition ζ of Σ_A^+, let $[x]_\zeta := \bigcap_{x \in B \in \zeta} B$ denote the atom of ζ containing x, and let m_x^ζ denote the conditional measure with respect to ζ supported on $[x]_\zeta$. For more information on conditional measures we refer to [EW, §5]. It follows from (2.1) that

$$m_x^{\xi_1^\infty}([x]_{\xi_0^\infty}) = \lim_{N\to\infty} \frac{m([x]_{\bigvee_{i=0}^{N-1}\sigma^{-i}\xi})}{m([x]_{\bigvee_{i=1}^{N-1}\sigma^{-i}\xi})} = \frac{u_{x_0}}{\lambda u_{x_1}}. \tag{2.2}$$

Thus, $m_x^{\xi_1^\infty}([x]_{\xi_0^\infty})$ is defined everywhere and for any $x \in \Sigma_A^+$ the information function I_m satisfies

$$I_m(\xi|\xi_1^\infty)(x) = -\log m_x^{\xi_1^\infty}([x]_\xi) = \log\lambda + g(\sigma x) - g(x),$$

where $g(y) = \log u_{y_0}$. So, we immediately get

LEMMA 2.1. *For any σ-invariant probability measure μ on Σ_A^+, we have*

$$\int I_m(\xi|\xi_1^\infty)\,d\mu = h_m(\sigma) = \log\lambda.$$

We now state the Pinsker inequality. Consider the n-dimensional simplex Δ_n of probability vectors $q = (q_1, \ldots, q_n)$. For a given $p \in \Delta_n$ with strictly positive entries we define the function

$$\phi_p : \Delta_n \to \mathbb{R} \quad \text{by} \quad \phi_p(q) = -\sum_{i=1}^{n} q_i \log\frac{p_i}{q_i},$$

with the convention $0 \log \frac{p_i}{0} = 0$. Fix the norm $\|q\| = \sum_i |q_i|$ on $\mathbb{R}^n$. We have [CT, Lemma 12.6.1]

Lemma 2.2 (Pinsker inequality). *ϕ_p is nonnegative and has a unique zero at p. Moreover, for any $q \in \Delta_n$ we have*

$$\|q-p\| \leq \sqrt{2\phi_p(q)}.$$

Let $p(x), q(x) \in \Delta_s$ be given by $p_i = p_i(x) = m_x^{\xi_1^\infty}(C_i)$ and $q_i = q_i(x) = \mu_x^{\xi_1^N}(C_i)$. Then

$$\int \big(I_m(\xi|\xi_1^\infty)(y) - I_\mu(\xi|\xi_1^N)(y)\big)\, d\mu_x^{\xi_1^N}(y) \\ = -\sum_{i=1}^{s} \mu_x^{\xi_1^N}(C_i) \log \frac{m_x^{\xi_1^\infty}(C_i)}{\mu_x^{\xi_1^N}(C_i)} = \phi_{p(x)}(q(x)).$$

It is easy to see that $p_i = m_x^{\xi_1^\infty}(C_i) = 0$ for some i if and only if $C_i \cap [x]_{\xi_1^\infty} = \emptyset$ if and only if $C_i \cap [x]_{\xi_1^N} = \emptyset$ for any $N \in \mathbb{N}$. So, we must have $\mu_x^{\xi_1^N}(C_i) = 0$, in which case we simply drop the ith term in the definition of ϕ_p. Now, applying Lemma 2.1 together with the fact $\int\int I_\mu(\xi|\xi_1^N)\, d\mu_x^{\xi_1^N}\, d\mu(x) = H_\mu(\xi|\xi_1^N)$ we obtain

Lemma 2.3. *For any invariant probability measure μ on Σ_A^+, we have*

$$\int \phi_{p(x)}(q(x))\, d\mu(x) = h_m(\sigma) - H_\mu(\xi|\xi_1^N).$$

Now we are in a position to introduce the sequence $(f_n)_{n\geq 0}$ of functions using the transfer operator. Let $\mathcal{L} : L^1(\Sigma_A^+, \xi_0^\infty, m) \to L^1(\Sigma_A^+, \xi_0^\infty, m)$ denote the transfer operator given by

$$\mathcal{L}f = \frac{dm_f \circ \sigma^{-1}}{dm} \quad \text{where} \quad dm_f = f dm.$$

The following is classical (see e.g. [B2, Lemma 1.10] and [PP, Theorem 2.2]).

Lemma 2.4. *There exist constants $C > 0$ and $\rho \in (0,1)$ such that for any Lipschitz function g on Σ_A^+ with $\int g\, dm = 0$ we have*

$$\|\mathcal{L}^n g\|_\theta \leq C\rho^n \|g\|_\theta \quad \textit{for any } n \geq 0.$$

We have

Lemma 2.5. *For any $f \in \mathcal{F}_A^+$, any probability invariant measure μ on Σ_A^+, and any $n, N \in \mathbb{N}$ we have*

$$\left| \int f_{n+1}\, d\mu - \int f_n\, d\mu \right| \leq \|f_n\|_\theta \big(\theta^{N+1} + \sqrt{2}(h_m(\sigma) - H_\mu(\xi|\xi_1^N))^{1/2}\big),$$

where $f_n := \mathcal{L}^n f = \mathcal{L} f_{n-1}$.

Proof. It is easy to see that $(\mathcal{L}f)\circ\sigma = E_m(f|\xi_1^\infty)$. Hence, using σ-invariance of μ we have

$$\int f_{n+1}\,d\mu - \int f_n\,d\mu = \int E_m(f_n|\xi_1^\infty)\,d\mu - \int E_\mu(f_n|\xi_1^N)\,d\mu. \tag{2.3}$$

Clearly $C_i \cap [x]_{\xi_1^\infty}$ is $\{y^{(i)} = ix_1x_2\cdots\}$ or empty. In any case we have

$$E_m(f_n|\xi_1^\infty)(x) = \int f_n(y)\,dm_x^{\xi_1^\infty}(y) = \sum_{i\in\Lambda} f_n(y^{(i)})m_x^{\xi_1^\infty}(C_i).$$

Also, for any $y \in C_i \cap [x]_{\xi_1^N}$ we have $d(y, y_i) \le \theta^{N+1}$. Thus, for μ-a.e. x,

$$\begin{aligned}\Big|E_\mu(f_n|\xi_1^N)(x) - \sum_{i\in\Lambda} f_n(y^{(i)})\mu_x^{\xi_1^N}(C_i)\Big|\\ = \Big|\sum_{i\in\Lambda}\int_{C_i}(f_n(y) - f_n(y^{(i)})\,d\mu_x^{\xi_1^N}(y)\Big| \le \theta^{N+1}|f_n|_\theta.\end{aligned}$$

Consequently,

$$\begin{aligned}&|E_m(f_n|\xi_1^\infty)(x) - E_\mu(f_n|\xi_1^N)(x)|\\ &\quad\le \theta^{N+1}|f_n|_\theta + \sum_{i\in\Lambda}|f_n(y^{(i)})|\,|m_x^{\xi_1^\infty}(C_i) - \mu_x^{\xi_1^N}(C_i)|\\ &\quad\le \theta^{N+1}|f_n|_\theta + |f_n|_\infty\|p(x) - q(x)\|.\end{aligned}$$

Using Lemma 2.2 and the Cauchy–Schwarz inequality we deduce

$$\begin{aligned}\Big|\int f_{n+1}\,d\mu - \int f_n\,d\mu\Big| &\le \int(\theta^{N+1}|f_n|_\theta + |f_n|_\infty\|p(x)-q(x)\|)\,d\mu\\ &\le \theta^{N+1}|f_n|_\theta + |f_n|_\infty\int\sqrt{2\phi_{p(x)}(q(x))}\,d\mu(x)\\ &\le \theta^{N+1}|f_n|_\theta + |f_n|_\infty\sqrt{2\int\phi_{p(x)}(q(x))\,d\mu(x)}\\ &= \theta^{N+1}|f_n|_\theta + \sqrt{2}\,|f_n|_\infty\sqrt{h_m(\sigma) - H_\mu(\xi|\xi_1^N)}. \quad\blacksquare\end{aligned}$$

We need one more lemma before we prove Theorem 1.1.

LEMMA 2.6. *Let $(a_n)_{n\ge0}$ be a decreasing sequence of nonnegative integers and set $A_n = (a_0 + a_1 + \cdots + a_{n-1})/n$. Then, for any $n\in\mathbb{N}$ and $h \ge a_0$,*

$$2(h - A_n) \ge h - a_{\lfloor n/2\rfloor}.$$

Proof. It suffices to prove the lemma for $h = a_0$, and in this case the conclusion follows from the inequality

$$A_n \le \frac{1}{n}\bigg(\Big\lfloor\frac{n}{2}\Big\rfloor a_0 + \Big(n - \Big\lfloor\frac{n}{2}\Big\rfloor\Big)a_{\lfloor n/2\rfloor}\bigg) \le \frac{1}{2}(a_0 + a_{\lfloor n/2\rfloor}). \quad\blacksquare$$

Proof of Theorem 1.1. It suffices to prove the conclusion for Lipschitz functions f with $\int f\,dm = 0$. As before we set $f_n = \mathcal{L}^n f$ for $n \geq 0$. From Lemma 2.4 we see that $\int f_n\,d\mu$ converges to $0 = \int f\,dm$, which gives

$$\left|\int f\,d\mu - \int f\,dm\right| = \lim_{n\to\infty}\left|\int f\,d\mu - \int f_n\,d\mu\right| \leq \sum_{n=0}^{\infty}\left|\int f_{n+1}\,d\mu - \int f_n\,d\mu\right|.$$

Now, the estimate from Lemma 2.5 together with Lemma 2.4 yields

$$\begin{aligned}\left|\int f\,d\mu - \int f\,dm\right| &\leq \sum_{n=0}^{\infty}\|f_n\|_\theta\big(\theta^{N+1} + \sqrt{2}(h_m(\sigma) - H_\mu(\xi|\xi_1^N))^{1/2}\big)\\ &\leq \frac{C}{1-\rho}\|f\|_\theta\big(\theta^{N+1} + \sqrt{2}(h_m(\sigma) - H_\mu(\xi|\xi_1^N))^{1/2}\big).\end{aligned}$$

Now, we wish to replace $H_\mu(\xi|\xi_1^N)$ by $\frac{1}{N}H_\mu(\xi_0^{N-1})$. We know that

$$\frac{1}{N}H_\mu(\xi_0^{N-1}) = \frac{1}{N}H_\mu(\xi_0^{N-1}) = \frac{1}{N}\sum_{n=0}^{N-1}H_\mu(\xi|\xi_1^n). \tag{2.4}$$

It follows from Lemmas 2.2 and 2.3 that $H_\mu(\xi|\xi_1^n) \leq h_m(\sigma)$ for any $n \in \mathbb{N}$, and in particular $\frac{1}{N}H_\mu(\xi_0^{N-1}) \leq h_m(\sigma)$. Thus, applying Lemma 2.6 for $h = h_m(\sigma)$ and $a_n = H_\mu(\xi|\xi_1^n)$ we get

$$2\left(h_m(\sigma) - \frac{1}{N}H_\mu(\xi_0^{N-1})\right) \geq h_m(\sigma) - H_\mu\Big(\xi|\bigvee_{i=1}^{\lfloor N/2\rfloor}\sigma^{-i}\xi\Big).$$

Hence, for any $N \in \mathbb{N}$,

$$\left|\int f\,d\mu - \int f\,dm\right| \leq \frac{C}{1-\rho}\|f\|_\theta\left(\theta^{N/2} + 2\sqrt{2}\left(h_m(\sigma) - \frac{1}{N}H_\mu(\xi_0^{N-1})\right)^{1/2}\right),$$

which finishes the proof. ∎

REFERENCES

[AE] M. Aka and M. Einsiedler, *Duke's theorem for subcollections*, Ergodic Theory Dynam. Systems 36 (2016), 335–342.

[B1] R. Bowen, *Markov partitions and Axiom A diffeomorphisms*, Amer. J. Math. 92 (1970), 725–747.

[B2] R. Bowen, *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, Lecture Notes in Math. 470, Springer, New York, 1975.

[B3] R. Bowen, *Periodic points and measures for Axiom A diffeomorphisms*, Trans. Amer. Math. Soc. 154 (1971), 377–397.

[CT] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.

[DI] M. Degli Eposti and S. Isola, *Distribution of closed orbits for linear automorphisms of tori*, Nonlinearity 8 (1995), 827–842.

[EW] M. Einsiedler and T. Ward, *Ergodic Theory with a View towards Number Theory*, Grad. Texts in Math. 259, Springer London, London, 2011.

[K] S. Kadyrov, *Effective uniqueness of Parry measure and exceptional sets in ergodic theory*, Monatsh. Math. 178 (2015), 237–249.

[M] M. Misiurewicz, *On expanding maps of compact manifolds and local homeomorphisms of a circle*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. 18 (1970), 725–732.

[Pa] W. Parry, *Intrinsic Markov chains*, Trans. Amer. Math. Soc. 112 (1964), 55–66.

[PP] W. Parry and M. Pollicott, *Zeta functions and the periodic orbit structure of hyperbolic dynamics*, Astérisque 187-188 (1990), 1–268.

[PY] M. Pollicott and M. Yuri, *Dynamical Systems and Ergodic Theory*, Cambridge Univ. Press, 1998.

[Po] F. Polo, *Equidistribution in Chaotic Dynamical Systems*, thesis, Ohio State Univ., https://etd.ohiolink.edu (2011).

[R] R. Rühr, *Effectivity of uniqueness of the maximal entropy measure on p-adic homogeneous spaces*, Ergodic Theory Dynam. Systems 36 (2016), 1972–1988.

[S1] Ya. G. Sinai, *Markov partitions and C-diffeomorphisms*, Funct. Anal. Appl. 2 (1968), 61–82.

[S2] Ya. G. Sinai, *Construction of Markov partitions*, Funct. Anal. Appl. 2 (1968), 245–253.

[W] P. Walters, *An Introduction to Ergodic Theory*, Grad. Texts in Math. 79, Springer, New York, 1982.

Shirali Kadyrov
Department of Mathematics
Nazarbayev University
Astana, Kazakhstan
E-mail: shirali.kadyrov@nu.edu.kz