ISSN 1055-1344, Siberian Advances in Mathematics, 2018, Vol. 28, No. 3, pp. 187–206. © Allerton Press, Inc., 2018.
Original Russian Text © I.V. Podvigin, 2017, published in Matematicheskie Trudy, 2017, Vol. 20, No. 2, pp. 90–119.

Estimates for Correlation in Dynamical Systems: From Hölder Continuous Functions to General Observables

I. V. Podvigin[1,2*]

[1] *Sobolev Institute of Mathematics, Novosibirsk, 630090 Russia*

[2] *Novosibirsk State University, Novosibirsk, 630090 Russia*

Received January 30, 2017

Abstract—For many dynamical systems that are popular in applications, estimates are known for the decay of correlation in the case of Hölder continuous functions. In the present article, we suggest an approach that allows us to obtain estimates for correlation in dynamical systems in the case of arbitrary functions. This approach is based on approximation and estimates are obtained with the use of known estimates for Hölder continuous functions. We apply our approach to transitive Anosov diffeomorphisms and derive the central limit theorem for the characteristic functions of certain sets with boundary of zero measure.

DOI: 10.3103/S1055134418030045

Keywords: *correlation, the best approximation, approximation spaces, Anosov diffeomorphisms, central limit theorem.*

1. INTRODUCTION

Let (M, d) be a metric space, let μ be a Borel measure on M, and let $T : M \to M$ be a transform that preserves the measure μ, i.e., we have $\mu(A) = \mu(T^{-1}A)$ for every Borel set A. Let $1 \le p, q \le \infty$ and let $1/p + 1/q = 1$. Consider functions $f \in L_p(M)$ and $g \in L_q(M)$. The *paired correlation coefficients* (or simply the *correlation function*) is defined by the rule

$$c_n(f, g) = \int_M f(x)g(T^n x)\, d\mu(x) - \int_M f(x)\, d\mu(x) \int_M g(x)\, d\mu(x), \quad n \in \mathbb{N}.$$

A dynamical system (T, μ) is said to be *mixing* if we have

$$c_n(f, g) \to 0 \ \text{ as } \ n \to \infty \ \text{ for all } \ f, g \in L_2(M).$$

An important statistical property of a dynamical system (T, μ) that characterizes its chaoticity is quite rapid decay of the correlation function in the case of regular functions (we call them *observables*). In applications, it is usual to regard H'older continuous functions (or functions from an extension of this class) as these regular observables. As is well known, in the case of irregular observables, the correlation function may decay as slowly as desired. For example, as is shown in [23], for every measure preserving mixing mapping, every numerical sequence $\{a_n\}$ that may decrease to zero as slowly as desired, and every nonzero function $g \in L_2^0(M)$, there exists a function $f \in L_2(M)$ such that $c_n(f, g) \ne \mathcal{O}(a_n)$ as $n \to \infty$. Moreover, such a behavior, i.e., as slow as desired decay of the correlation function, is typical for functions in $L_2(M)$ (see, for example, [9]).

Nevertheless, for each particular pair of observables, it is interesting to know the exact rate (even if it is very slow) of decay of their correlation function. To the best of author's knowledge, there are no examples of systems with estimates for the correlation function for all observables among publications on mixing

*E-mail: ipodvigin@math.nsc.ru

dynamical systems.[1] Examples are known of estimates for the correlation function for observables that are not H'older continuous (see, for example, [22, 28] and the references in these articles) such that the dependence of the rate of decay of the correlation function on regularity of observables is reflected in coefficients that are similar to the modulus of continuity. The aim of the present article is to use known estimates for regular (for example, Hölder continuous) functions and a suitable approximation by such functions and to obtain estimates for *all* other observables. It turns out that the rate of decay of the correlation function and regularity of observables are connected by the best approximation. We mention a feature of our approach. Namely, dynamical properties of systems that influence on the decay of the correlation function are already taken into account in estimates for regular (Hölder continuous) observables, while the rates of decay of additional terms that arise in approximation depend on the rate of approximation only.

1.1. *Initial estimates for correlation.* Let

$$\mathfrak{F}_p \subseteq L_p(M) \text{ and } \mathfrak{G}_q \subseteq L_q(M)$$

be normed spaces of complex-valued functions that are defined on M. We assume that, for all $f \in \mathfrak{F}_p$, $g \in \mathfrak{G}_q$, and $n \in \mathbb{N}$, the following estimate for the correlation function holds:

$$\big|c_n(f,g)\big| \le C(f,g)\Phi(n), \tag{1}$$

where $C(f,g)$ is a nonnegative constant and $\Phi(n) \to 0$ as $n \to \infty$. We assume that Φ depends on the whole spaces $\mathfrak{F}_p$ and $\mathfrak{G}_q$ instead of particular functions $f \in \mathfrak{F}_p$ and $g \in \mathfrak{G}_q$.

We consider the least possible constant C such that estimate (1) is valid. Then C possesses additional properties. We fix a function $g \in L_q(M)$. Notice that, for all $a \in \mathbb{C}$, $f_1, f_2 \in \mathfrak{F}_p$, and $n \in \mathbb{N}$, we have

$$c_n(af_1,g) = ac_n(f_1,g), \quad c_n(f_1+f_2,g) = c_n(f_1,g) + c_n(f_2,g).$$

Since the constant is minimal, we conclude that

$$C(f_1+f_2,g) \le C(f_1,g) + C(f_2,g),$$
$$C(af,g) \le |a|C(f,g) = |a|C\left(\frac{1}{a}af,g\right) \le C(af,g).$$

Moreover, if $f \equiv 0$ then $C(f,g) = 0$ in the obvious way. We conclude that $C = C(f,g)$ is a seminorm with respect to f. Similar arguments show that $C = C(f,g)$ is a seminorm with respect to g. In the sequel, we assume that these seminorms are proper norms of the corresponding spaces and C admits a representation of the form

$$C(f,g) = A\|f\|_{\mathfrak{F}_p}\|g\|_{\mathfrak{G}_q},$$

where A is a positive constant that is independent of f and g, $\|\cdot\|_{\mathfrak{F}_p}$ is a norm in $\mathfrak{F}_p$, and $\|\cdot\|_{\mathfrak{G}_q}$ is a norm in $\mathfrak{G}_q$. In [8, Theorem B.1], sufficient conditions can be found guaranteeing that the constant $C = C(f,g)$ can be represented as the product of norms (up to a constant factor) in the corresponding spaces.

Thus, estimate (1) is uniform on balls of the spaces $\mathfrak{F}_p$ and $\mathfrak{G}_q$ and has the form

$$\big|c_n(f,g)\big| \le A\|f\|_{\mathfrak{F}_p}\|g\|_{\mathfrak{G}_q}\Phi(n), \ \ n \in \mathbb{N}. \tag{2}$$

If inequality (2) holds then we say that *the correlation function for $\mathfrak{F}_p$-observables with respect to $\mathfrak{G}_q$-observables decays at the rate Φ.*

Numerous dynamical systems are known that admit estimates for the correlation function of this type (see, for example, the monograph [2]). Among them, we mention classical transitive Anosov diffeomorphisms [6] and a large class of systems admitting the Gibbs–Markov–Young structure [26, 27] that includes several popular billiards [10].

[1]During the preparation of the final version, the author found article [12], where estimates are obtained for the correlation function in the case of all L_2-observables for expanding endomorphisms of a torus.

1.2. *Approximation by the spaces* $\mathfrak{F}_p$ *and* $\mathfrak{G}_q$. In the sequel, we assume that the space $\mathfrak{F}_p$ is everywhere dense in $L_p(M)$ and the space $\mathfrak{G}_q$ is everywhere dense in $L_q(M)$. This assumption leads to the following natural questions. Is it possible to use approximation and obtain meaningful estimates for the correlation function for a class of functions that is wider than $\mathfrak{F}_p$ and $\mathfrak{G}_q$? How the rate of their decay depends on the rate of approximation? Is it possible to obtain estimates of the same type as in (2), i.e., with a constant that is representable as the product of norms in the corresponding spaces? By a meaningful estimate we mean that it allows us to derive statistical laws (for example, the central limit theorem) following [25] or to obtain estimates for the rate of convergence in the von Neumann ergodic theorem, see [14, 17].

The idea to use approximation in estimates for the correlation function is not new. It was successfully employed, for example, in the remarkable article [9], where approximation in $L_2(M)$ was considered for the case in which $\mathfrak{F}_2$ is the set of functions measurable with respect to a Markov partition of the phase space M and $\mathfrak{G}_2 = \mathfrak{F}_2$. Smooth approximation was used in [15] in the case of a periodic Lorentz gas and estimates for the correlation function were constructed for the characteristic functions of sets with rectifiable boundaries. The present article is dedicated to development of the approach from [15] in a more general situation. One more new feature of our estimates for the correlation function is as follows. We introduce normed approximation spaces such that the character of estimates for approximation (and, as we will see, for the correlation function) is the same for observables in these spaces.

The best approximation serves as a parameter of approximation. For $f \in L_p(M)$ and $t \geq 0$, the *best approximation* of order t of the function f by the class $\mathfrak{F}_p$ is defined as follows:

$$\tau_f(t) = essinf \left\{ \|f - h\|_p : \ h \in \mathfrak{F}_p, \ \|h\|_{\mathfrak{F}_p} \leq t \right\}. \tag{3}$$

The definition of the best $\mathfrak{G}_q$-approximation of order t of a function $g \in L_q$ is similar. We denote it by $\tau_g(t)$.

The article is organized as follows.

In Section 2, we formulate and prove the main result, i.e., Theorem 2.1 and its corollaries.

In Section 3, we show that the sets of functions associated with various estimates for the best approximation form normed spaces and establish certain their properties.

In Section 4, we consider transitive Anosov diffeomorphisms as an example of application of the main result. We show that the central limit theorem is valid for certain new observables, see Theorem 4.2.

The appendix is dedicated to Hölder approximation of the characteristic functions of sets with boundary of zero measure. We use results from the appendix in derivation of the central limit theorem.

2. ESTIMATES FOR THE CORRELATION FUNCTION FOR GENERAL OBSERVABLES

2.1. *Formulation of the main result.* We introduce sets of integrable functions that correspond to various estimates for the best approximation. Let Ξ be the set of all functions $\Theta : \mathbb{R}_+ \to \mathbb{R}_+$ that decrease to zero, i.e.,

$$\Theta(t_1) \geq \Theta(t_2) \quad \text{for } \ 0 \leq t_1 \leq t_2,$$

$$\Theta(t) \to 0 \qquad \text{as } \ t \to +\infty,$$

and let $\Xi^0 \subset \Xi$ be the subset of functions that eventually vanish.

Definition 2.1. For $\Theta \in \Xi$, we denote by $\mathfrak{F}_p(\Theta)$ the set of all functions $f \in L_p(M)$ such that, for a suitable constant $c \geq 0$, the best $\mathfrak{F}_p$-approximation satisfies the inequality

$$\tau_f(ct) \leq c\Theta(t) \tag{4}$$

for every $t \geq 0$. The set of all such constants is denoted by $C(\Theta, f)$. We denote by $\|f\|_{\mathfrak{F}_p(\Theta)}$ the greatest lower bound of this set, i.e.,

$$\|f\|_{\mathfrak{F}_p(\Theta)} = \underset{c \in C(\Theta, f)}{essinf} \ c. \tag{5}$$

In Proposition 3.3 below, we show that, for every function $\Theta \in \Xi$, the set $\mathfrak{F}_p(\Theta)$ forms a normed space with respect to the norm from (5).

The following assertion is the main result of the present article. It shows that the initial estimates for decay of the correlation function from (2) for $\mathfrak{F}_p$-observables with respect to $\mathfrak{G}_q$-observables allow us to obtain estimates of the same type in the case of $\mathfrak{F}_p(\Theta_1)$-observables and $\mathfrak{G}_q(\Theta_2)$-observables for arbitrary $\Theta_1, \Theta_2 \in \Xi$. Moreover, the rate of decay in the case of new observables is explicitly represented in terms of the rate of approximation and the rate of decay of the correlation function for $\mathfrak{F}_p$-observables with respect to $\mathfrak{G}_q$-observables. We denote by $\Theta_1 \vee \Theta_2$ the supremum of functions Θ_1 and Θ_2.

Theorem 2.1. *Assume that the correlation function for $\mathfrak{F}_p$-observables with respect to $\mathfrak{G}_q$-observables decays at rate Φ. Let*

$$\Theta_1, \Theta_2 \in \Xi.$$

Then, for every pair of functions $f \in \mathfrak{F}_p(\Theta_1)$, $g \in \mathfrak{G}_q(\Theta_2)$, there exist a number $n_0 \in \mathbb{N}$, a constant $A' > 0$, and a function Φ' with $\Phi'(n) \searrow 0$ as $n \to \infty$ such that, for every $n \geq n_0$, we have

$$\big|c_n(f, g)\big| \leq A' \|f\|_{\mathfrak{F}_p(\Theta_1)} \|g\|_{\mathfrak{G}_q(\Theta_2)} \Phi'(n). \tag{6}$$

If $\Theta_1 \vee \Theta_2 \notin \Xi^0$ then $\Phi'(n) = \Phi(n) \upsilon\big(\Phi(n)\big)$, where $\upsilon : \mathbb{R}_+ \to \mathbb{R}_+$ is the inverse of

$$\frac{1}{t}(\Theta_1 \vee \Theta_2)\big(\sqrt{t}\big), \quad t > 0,$$

and $n_0 \in \mathbb{N}$ is the least natural number such that the inequality

$$\Phi(n_0) \upsilon\big(\Phi(n_0)\big) \leq 1$$

holds.

If $\Theta_1 \vee \Theta_2 \in \Xi^0$ then $\Phi'(n) = \Phi(n)$ and $n_0 = 1$.

In applications, if the rate Φ of decay of the initial correlation function is either exponential or power function then the following problem often arises: Find new observables such that the same (exponential or power function) rate of decay is preserved, i.e., $\Phi'(n) = \Phi(n)\upsilon\big(\Phi(n)\big)$. We consider such situations. For specific rates of decay (see, for example, [13]), this question should be considered separately. We leave it to the reader interested in such dynamical systems. We recall that, for nonnegative values $a = a(t)$ and $b = b(t)$, the relation $a(t) = \mathcal{O}\big(b(t)\big)$ as $t \to t_0$ means that there exists a constant $c > 0$ such that the inequality $a(t) \leq cb(t)$ holds in a neighborhood of t_0. We put

$$\Theta = \Theta_1 \vee \Theta_2.$$

Corollary 2.1. *Let the conditions of Theorem 2.1 hold with an exponential estimate for the initial correlation function, i.e., $\Phi(n) = \theta^n$ for a suitable $\theta \in (0, 1)$. Then, for all $f \in \mathfrak{F}_p(\Theta_1)$ and $g \in \mathfrak{G}_q(\Theta_2)$, the following assertions are valid.*

(i) *If $\Theta(t) = \mathcal{O}(t^{-\beta})$ as $t \to +\infty$ then*

$$\big|c_n(f, g)\big| = \mathcal{O}(\theta^{\gamma n}) \ \text{ as } \ n \to +\infty,$$

where $\beta > 0$, $\gamma \in (0, 1)$, and $\beta = \frac{2\gamma}{1-\gamma}$.

(ii) *If $\Theta(t) \log^{\delta}_{\theta^{-1}} \big(\frac{t^2}{\Theta(t)}\big) = \mathcal{O}(1)$ as $t \to +\infty$ for a suitable $\delta > 0$ then*

$$\big|c_n(f, g)\big| = \mathcal{O}(n^{-\delta}) \ \text{ as } \ n \to +\infty.$$

Corollary 2.2. *Let the conditions of Theorem 2.1 hold with a power function estimate for the initial correlation function, i.e., $\Phi(n) = n^{-\alpha}$ for a suitable $\alpha > 0$. Then, for all $f \in \mathfrak{F}_p(\Theta_1)$ and $g \in \mathfrak{G}_q(\Theta_2)$, the following assertion is valid.*

(iii) *If $\Theta(t) = \mathcal{O}(t^{-\beta})$ as $t \to +\infty$ then*

$$\big|c_n(f, g)\big| = \mathcal{O}(n^{-\gamma}) \ \text{ as } \ n \to +\infty,$$

where $\beta > 0$, $\gamma \in (0, \alpha)$, and $\beta = \frac{2\gamma}{\alpha-\gamma}$.

Corollaries 2.1 and 2.2 allow us to claim that the result of Theorem 2.1 is meaningful. Indeed, we widen the set of functions that admit exponential and power function estimates for the correlation function. Moreover, on the basis of the exponential and power function estimates for $c_n(f,f)$ with new observables

$$f \in \mathfrak{F}_2(\Theta_1) \cap \mathfrak{G}_2(\Theta_2),$$

we obtain, with the use of [17, Theorem 7], estimates for the rate of convergence in the von Neumann ergodic theorem for the same observables. We also obtain certain properties of the spectral measure of $f - \int_M f\,d\mu$; namely, its behavior at zero [17, Theorem 3 and Remark 4] and an estimate for the Hausdorff dimension of this measure [19, Proposition 3.11 and Remark 3.12]. Using estimates for the rate of convergence in the von Neumann theorem, we may obtain estimates for the rate of convergence in the Birkhoff theorem [17, 24].

Let $\mathfrak{G}_\infty = L_\infty(M)$ and $\mathfrak{F}_1 \subseteq L_\infty(M)$. We use exponential and power function estimates for $c_n(f,g)$, where $f \in \mathfrak{F}_1(\Theta_1)$ and $g \in \mathfrak{G}_\infty$ and take into account [1, Theorem D] (see also [21, Theorem 1.2]). We immediately obtain the corresponding estimates for large deviations of the ergodic averages for $f \in \mathfrak{F}_1(\Theta_1) \cap L_\infty(M)$. In turn, this allows us to obtain estimates for the rate of convergence in the Birkhoff theorem [17, Theorem 13].

We present one more application of the obtained estimates (derivation of the central limit theorem for certain new observables) at the end of the article in discussion of particular dynamical systems (transitive Anosov diffeomorphisms).

2.2. *Proofs of the main results.* The following lemma shows how to take into account the quantitative characteristics of approximation in estimates for the correlation function.

Lemma 2.1. *Assume that the correlation function for $\mathfrak{F}_p$-observables with respect to $\mathfrak{G}_q$-observables decays at rate Φ. Then, for all $f \in L_p(M)$, $g \in L_q(M)$, $n \in \mathbb{N}$, and $t, s \geq 0$, we have*

$$\big|c_n(f,g)\big| \leq Ats\Phi(n) + R(t,s), \tag{7}$$

where

$$R(t,s) = 2\big(\tau_f(t)\|g\|_q + \tau_g(s)\|f\|_p + \tau_f(t)\tau_g(s)\big) \tag{8}$$

and $A > 0$ is the same constant as in inequality (2).

Proof. Let $h_f \in \mathfrak{F}_p$ be an approximation of f in $L_p(M)$ and let $h_g \in \mathfrak{G}_q$ be an approximation of g in $L_q(M)$. Simple calculations and estimate (2) show that, for every $n \in \mathbb{N}$, we have

$$\begin{aligned}\big|c_n(f,g)\big| &\leq \big|c_n(h_f,h_g)\big| + \big(\|f\|_p + \|h_f\|_p\big)\|g-h_g\|_q + \big(\|g\|_q + \|h_g\|_q\big)\|f-h_f\|_p \\ &\leq A\|h_f\|_{\mathfrak{F}_p}\|h_g\|_{\mathfrak{G}_q}\Phi(n) \\ &\quad + 2\big(\|f\|_p\|g-h_g\|_q + \|g\|_q\|f-h_f\|_p + \|g-h_g\|_q\|f-h_f\|_p\big).\end{aligned}$$

We consider the infimum over all $h_f \in \mathfrak{F}_p$ and $h_g \in \mathfrak{G}_q$ such that

$$\|h_f\|_{\mathfrak{F}_p} \leq t, \;\; \|h_g\|_{\mathfrak{G}_q} \leq s \;\text{ for }\; t, s \geq 0.$$

We obtain the required estimate (7). □

Choosing suitable $t = t_n$ and $s = s_n$ and substituting them into (7), we may obtain satisfiable estimates for the correlation function. It is clear that the choice of such sequences (cf. the proof of Theorem 2.1 below) depends on the rate of approximation $R(t,s)$.

Proof of Theorem 2.1. First we consider the case in which

$$\Theta = \Theta_1 \vee \Theta_2 \notin \Xi^0,$$

i.e., at least one of the functions Θ_1, Θ_2 does not eventually vanish. By the conditions of the theorem (see Definition 2.1 and (14)), we have

$$\tau_f(at) \leq a\Theta_1(t), \;\; \tau_g(bt) \leq b\Theta_2(t)$$

for every t with $t \geq 0$, where $a = \|f\|_{\mathfrak{F}_p(\Theta_1)}$ and $b = \|g\|_{\mathfrak{G}_q(\Theta_2)}$. Taking into account these inequalities and (8), we obtain

$$\begin{aligned} R(at, bt) &= 2\big(\|g\|_q \tau_f(at) + \|f\|_p \tau_g(bt) + \tau_f(at)\tau_g(bt)\big) \\ &\leq 2\big(a\|g\|_q \Theta_1(t) + b\|f\|_p \Theta_2(t) + ab\Theta_1(t)\Theta_2(t)\big) \\ &\leq 2\big(a\|g\|_q + b\|f\|_p + ab\big)(\Theta_1 \vee \Theta_2)(t) \end{aligned}$$

for every t such that $t \geq 0$ and $(\Theta_1 \vee \Theta_2)(t) \leq 1$.

We denote by υ the inverse of the monotone decreasing function $\frac{1}{t}(\Theta_1 \vee \Theta_2)\big(\sqrt{t}\,\big)$., i.e., we put

$$(\Theta_1 \vee \Theta_2)\big(\sqrt{\upsilon(s)}\,\big) = \upsilon(s)s, \quad s > 0. \tag{9}$$

It is clear that $\upsilon(s) \to +\infty$ as $s \to +0$. We consider an increasing sequence $\{t_n\}$ of real numbers such that

$$t_n^2 = \upsilon\big(\Phi(n)\big)$$

for every n with $n \geq n_0$. The number $n_0 \in \mathbb{N}$ is determined by the condition

$$(\Theta_1 \vee \Theta_2)(t_{n_0}) = (\Theta_1 \vee \Theta_2)\Big(\sqrt{\upsilon\big(\Phi(n_0)\big)}\,\Big) = \Phi(n_0)\upsilon\big(\Phi(n_0)\big) \leq 1,$$

where the latter equality is a consequence of representation (9). We use estimate (7) with $t = at_n$ and $s = bt_n$ and equality (9). For $n \geq n_0$, we obtain

$$\begin{aligned} \big|c_n(f, g)\big| &\leq Aabt_n^2\Phi(n) + R(at_n, bt_n) \\ &\leq Aabt_n^2\Phi(n) + 2\big(a\|g\|_q + b\|f\|_p + ab\big)(\Theta_1 \vee \Theta_2)(t_n) \\ &\leq Aab\Phi(n)\upsilon\big(\Phi(n)\big) + 2\big(a\|g\|_q + b\|f\|_p + ab\big)(\Theta_1 \vee \Theta_2)\Big(\sqrt{\upsilon\big(\Phi(n)\big)}\,\Big) \\ &= \big((A+2)ab + 2a\|g\|_q + 2b\|f\|_p\big)\Phi(n)\upsilon\big(\Phi(n)\big). \end{aligned}$$

It remains to notice that $\|f\|_p \leq a\Theta_1(0)$ and $\|g\|_q \leq b\Theta_2(0)$, see (15). We find that

$$\begin{aligned} (A+2)ab + 2a\|g\|_q + 2b\|f\|_p &\leq (A+2)ab + 2ab\Theta_2(0) + 2ab\Theta_1(0) \\ &= \big(A + 2 + 2\Theta_1(0) + 2\Theta_2(0)\big)ab = A'ab. \end{aligned}$$

Second we consider the case in which $\Theta_1, \Theta_2 \in \Xi^0$. There exist $t_0, s_0 > 0$ such that

$$\begin{aligned} &\tau_f\big(\|f\|_{\mathfrak{F}_p(\Theta_1)}t\big) = 0, \ \ t \geq t_0, \\ &\tau_g\big(\|g\|_{\mathfrak{G}_q(\Theta_2)}s\big) = 0, \ \ s \geq s_0; \end{aligned}$$

hence, we have $R(t, s) = 0$ if $t \geq t_0\|f\|_{\mathfrak{F}_p(\Theta_1)}$ and $s \geq s_0\|f\|_{\mathfrak{G}_q(\Theta_2)}$. We substitute $t = t_0\|f\|_{\mathfrak{F}_p(\Theta_1)}$ and $s = s_0\|g\|_{\mathfrak{G}_q(\Theta_2)}$ into estimate (7). For every $n \geq 1$, we obtain

$$\big|c_n(f, g)\big| \leq At_0\|f\|_{\mathfrak{F}_p(\Theta_1)}s_0\|g\|_{\mathfrak{G}_q(\Theta_2)}\Phi(n) = A'\|f\|_{\mathfrak{F}_p(\Theta_1)}\|g\|_{\mathfrak{G}_q(\Theta_2)}\Phi(n),$$

which is estimate (6) with the constant $A' = At_0s_0$ and the function $\upsilon \equiv 1$. □

Proofs of Corollaries 2.1 *and* 2.2. We prove the following equivalences:

(i) we have $\Theta(t) = \mathcal{O}(t^{-\beta})$ as $t \to +\infty$ if and only if $\Phi'(n) = \mathcal{O}(\theta^{\gamma n})$ as $n \to +\infty$, where $\beta > 0$, $\gamma \in (0, 1)$, and $\beta = \frac{2\gamma}{1-\gamma}$;

(ii) we have $\Theta(t)\log^{\delta}_{\theta^{-1}}\big(\frac{t^2}{\Theta(t)}\big) = \mathcal{O}(1)$ as $t \to +\infty$ if and only if $\Phi'(n) = \mathcal{O}(n^{-\delta})$ as $n \to +\infty$;

(iii) we have $\Theta(t) = \mathcal{O}(t^{-\beta})$ as $t \to +\infty$ if and only if $\Phi'(n) = \mathcal{O}(n^{-\gamma})$ as $n \to +\infty$, where $\beta > 0$, $\gamma \in (0, \alpha)$, and $\beta = \frac{2\gamma}{\alpha-\gamma}$.

Assertions (i) and (ii) imply Corollary 2.1 and assertion (iii) implies Corollary 2.2.

We prove assertion (ii) only. The proofs of assertions (i) and (iii) are similar. Assume that, for some $\delta > 0$, we have

$$\Theta(t)\log^{\delta}_{\theta^{-1}}\left(\frac{t^2}{\Theta(t)}\right) = \mathcal{O}(1) \text{ as } t \to +\infty.$$

We consider the inverse of υ, i.e., the mapping defined by the rule $\upsilon^{-1}(t) = \Theta\big(\sqrt{t}\,\big)/t$. This equality is equivalent to the condition

$$t\upsilon^{-1}(t)\log^{\delta}_{\theta^{-1}}\left(\frac{1}{\upsilon^{-1}(t)}\right) = \mathcal{O}(1) \text{ as } t \to +\infty.$$

We substitute $t = \upsilon(\theta^n)$. Since the functions under consideration are monotone, we obtain the following equivalent condition:

$$\theta^n\upsilon(\theta^n)\log^{\delta}_{\theta^{-1}}\left(\frac{1}{\theta^n}\right) = \mathcal{O}(1) \text{ as } n \to +\infty.$$

Simple transformations lead to the following condition:

$$\Phi'(n)n^{\delta} = \mathcal{O}(1) \text{ as } n \to +\infty. \qquad \square$$

3. APPROXIMATION SPACES $\mathfrak{F}_p(\Theta)$

In the present section, we study certain properties of sets of the form $\mathfrak{F}_p(\Theta)$. In particular, we show that they form normed spaces that are everywhere dense in $L_p(M)$ and extend $\mathfrak{F}_p$. The reader interested in applications of Theorem 2.1 may skip it and pass directly to Sec. 4.

3.1. *Properties of the best approximation.* It is obvious that, for each $f \in L_p(M)$, the best $\mathfrak{F}_p$-approximation τ_f is a function that decreases to zero and we have

$$\tau_f(0) = \|f\|_p. \tag{10}$$

Moreover, the function τ_f is continuous. Indeed, it is immediate from equality (3) that, for every pair of numbers $t \geq 0$ and $\varepsilon > 0$, there exist $n_0 = n_0(t,\varepsilon) \in \mathbb{N}$ and a sequence $h_n \in \mathfrak{F}_p$ with $\|h_n\|_{\mathfrak{F}_p \leq t}$ such that, for every $n \geq n_0$, we have

$$\|f - h_n\|_p - \tau_f(t) < \varepsilon$$

and, consequently,

$$\|h_n\|_p \leq \|f\|_p + \tau_f(t) + \varepsilon \leq 2\|f\|_p + \varepsilon.$$

If $t > 0$ then, for an arbitrary $\delta \in (0,t)$, we have $\|h_n - h_n\delta/t\|_{\mathfrak{F}_p} \leq t - \delta$. Taking into account this fact and the estimates above, we obtain

$$\begin{aligned}\tau_f(t) \leq \tau_f(t-\delta) &\leq \|f - h_n - h_n\delta/t\|_p \leq \|f - h_n\|_p + \frac{\delta}{t}\|h_n\|_p \\ &\leq \tau_f(t) + \varepsilon + \frac{\delta}{t}\big(2\|f\|_p + \varepsilon\big)\end{aligned}$$

for every $n \geq n_0$. Since $\varepsilon > 0$ was arbitrary, we have

$$\tau_f(t) \leq \tau_f(t-\delta) \leq \tau_f(t) + \frac{2\|f\|_p\delta}{t} \tag{11}$$

for every $t > 0$. We substitute $t + \delta$ for t in (11). For every $t \geq 0$, we obtain

$$\tau_f(t) - \frac{2\|f\|_p\delta}{t+\delta} \leq \tau_f(t+\delta) \leq \tau_f(t). \tag{12}$$

It remains to pass to the limit as $\delta \to 0+$ in inequalities (11) and (12).

In the sequel, we will need the following property of the best approximation (see [3, Lemma 7.1.1.] for a similar property).

Proposition 3.1. *If* $t \geq 0$, $a, b \in \mathbb{C}$, *and*

$$c = |a| + |b| \neq 0$$

then we have

$$\tau_{af_1+bf_2}(ct) \leq |a|\tau_{f_1}(t) + |b|\tau_{f_2}(t). \tag{13}$$

Proof. Inequality (13) is immediate from the following chain of inequalities:

$$\begin{aligned}\tau_{af_1+bf_2}(ct) &= essinf\,\big\{\|af_1+bf_2-h\|_p : h \in \mathfrak{F}_p,\ \|h\|_{\mathfrak{F}_p} \leq ct\big\} \\ &\leq essinf\,\big\{\|af_1+bf_2-h\|_p : h = ah_1+bh_2,\ h_j \in \mathfrak{F}_p,\ \|h\|_{\mathfrak{F}_p} \leq ct\big\} \\ &\leq essinf\,\big\{\|af_1+bf_2-ah_1-bh_2\|_p : h_j \in \mathfrak{F}_p,\ \|h_j\|_{\mathfrak{F}_p} \leq t\big\} \\ &\leq essinf\,\big\{\|af_1-ah_1\|_p : h_1 \in \mathfrak{F}_p,\ \|h_1\|_{\mathfrak{F}_p} \leq t\big\} \\ &\quad + essinf\,\big\{\|bf_2-bh_2\|_p : h_2 \in \mathfrak{F}_p,\ \|h_2\|_{\mathfrak{F}_p} \leq t\big\} \\ &= |a|\tau_{f_1}(t) + |b|\tau_{f_2}(t).\end{aligned}$$

□

3.2. *Geometrical meaning of the functional* $\|\cdot\|_{\mathfrak{F}_p(\Theta)}$. In practice, it is difficult to calculate (and even estimate) the values of the best $\mathfrak{F}_p$-approximation τ_f of an arbitrary function $f \in L_p(M)$ (see, for example, the classical result on approximation by integer-valued functions [3, Theorem 7.2.4]). If an estimate is known for τ_f of the form $\tau_f(t) \leq \Theta(t)$ for every $t \geq 0$ then this estimate need not be optimal in the following sense: There may be a tangible difference between the graphs of the functions τ_f and Θ. In this case, we can make these graphs closer by expanding the graph of τ_f along the axis OX (i.e., we pass to $\tau_f(ct)$) and simultaneously contracting the graph of Θ along the axis OY (i.e., we pass to $c\Theta(t)$). We perform such expansion and contraction while relation (4) is preserved for all $t \geq 0$. The optimal parameter for simultaneous expansion and contraction is $\|f\|_{\mathfrak{F}_p(\Theta)}$. The following assertion shows that inequality (4) holds for the constant $\|f\|_{\mathfrak{F}_p(\Theta)}$.

Proposition 3.2. *Let* $f \in \mathfrak{F}_p(\Theta)$, *where* $\Theta : \mathbb{R}_+ \to \mathbb{R}_+$ *decreases to zero. Then we have*

$$C(f, \Theta) = \big[\|f\|_{\mathfrak{F}_p(\Theta)}, +\infty\big). \tag{14}$$

Proof. It is clear that, for the zero function, we have $C(0, \Theta) = \mathbb{R}_+$. Let f be a nonzero function. Since τ_f is monotone and continuous, we obtain

$$\Theta(t) \geq \sup_{c \in C(f,\Theta)} \frac{1}{c}\tau_f(ct) = \frac{1}{\|f\|_{\mathfrak{F}_p(\Theta)}}\tau_f\big(\|f\|_{\mathfrak{F}_p(\Theta)}t\big)$$

for every $t \geq 0$. It remains to notice that if $c > \|f\|_{\mathfrak{F}_p(\Theta)}$ and $t \geq 0$ then we have

$$\tau_f(ct) \leq \tau_f\big(\|f\|_{\mathfrak{F}_p(\Theta)}t\big) \leq \|f\|_{\mathfrak{F}_p(\Theta)}\Theta(t) \leq c\Theta(t).$$

□

By Proposition 3.2 and equality (10), we find that the norm $\|\cdot\|_p$ is subordinate to the norm $\|\cdot\|_{\mathfrak{F}_p(\Theta)}$, i.e., for every function $f \in \mathfrak{F}_p(\Theta)$, the following inequality holds:

$$\|f\|_p \leq \|f\|_{\mathfrak{F}_p(\Theta)}\Theta(0). \tag{15}$$

3.3. *Norm in* $\mathfrak{F}_p(\Theta)$. Since the best approximation is monotone, we take property (13) into account and conclude that $\mathfrak{F}_p(\Theta)$ is a linear space.

Proposition 3.3. *Let* $\Theta : \mathbb{R}_+ \to \mathbb{R}_+$ *be a function that decreases to zero. Then* $\mathfrak{F}_p(\Theta)$ *is a normed space with respect to the norm* $\|\cdot\|_{\mathfrak{F}_p(\Theta)}$.

Proof. We prove that $\mathfrak{F}_p(\Theta)$ is a linear space. It is clear that it contains the zero function because $\tau_0(t) = 0 \leq \Theta(t)$ for every $t \geq 0$.

Let $f_1, f_2 \in \mathfrak{F}_p(\Theta)$. Then we have

$$\tau_{f_1}(c_1 t) \leq c_1\Theta(t), \quad \tau_{f_2}(c_2 t) \leq c_2\Theta(t)$$

for every $t \geq 0$, where $c_1, c_2 \geq 0$ are constants. We consider $a, b \in \mathbb{C}$ with $c = |a| + |b| \neq 0$. Since the best approximation is monotone and possesses property (13), we obtain

$$\begin{aligned}\tau_{af_1+bf_2}\big(c(c_1 \vee c_2)t\big) &\leq |a|\tau_{f_1}\big((c_1 \vee c_2)t\big) + |b|\tau_{f_2}\big((c_1 \vee c_2)t\big) \\ &\leq |a|\tau_{f_1}(c_1 t) + |b|\tau_{f_2}(c_2 t) \leq \big(|a|c_1 + |b|c_2\big)\Theta(t) \\ &\leq c(c_1 \vee c_2)\Theta(t)\end{aligned}$$

for every $t \geq 0$. By definition, this means that the function $af_1 + bf_2$ belongs to the set $\mathfrak{F}_p(\Theta)$. Thus, $\mathfrak{F}_p(\Theta)$ is a linear space. We check the properties of the norm.

It follows from inequality (15) that

$$\|f\|_{\mathfrak{F}_p(\Theta)} \geq \frac{\|f\|_p}{\Theta(0)}.$$

Hence, if $\Theta(0) \neq 0$ then we have $\|f\|_{\mathfrak{F}_p(\Theta)} = 0$ if and only if $f = 0$. It is also clear that $\Theta(0) = 0$ implies $f = 0$. If $f \neq 0$ then it follows from property (13) with $b = 0$ that

$$\tau_{af}\big(|a|ct\big) \leq |a|\tau_f(ct) \leq |a|c\Theta(t)$$

for all $t \geq 0$ and $c \in C(f, \Theta)$. We conclude that

$$\|af\|_{\mathfrak{F}_p(\Theta)} \leq |a|\|f\|_{\mathfrak{F}_p(\Theta)}.$$

The reverse inequality is a consequence of the above inequality and can be obtained by a well-known trick:

$$|a|\|f\|_{\mathfrak{F}_p(\Theta)} = |a| \left\|\frac{1}{a}af\right\|_{\mathfrak{F}_p(\Theta)} \leq |a|\frac{1}{|a|}\|af\|_{\mathfrak{F}_p(\Theta)} = \|af\|_{\mathfrak{F}_p(\Theta)}.$$

Thus, we have proven that the norm is positive uniform. It remains to check the triangle inequality. We again use property (13) and representation

$$f_1 + f_2 = \|f_1\|_{\mathfrak{F}_p(\Theta)} \frac{f_1}{\|f_1\|_{\mathfrak{F}_p(\Theta)}} + \|f_2\|_{\mathfrak{F}_p(\Theta)} \frac{f_2}{\|f_2\|_{\mathfrak{F}_p(\Theta)}}.$$

For all $t \geq 0$ and

$$c_1 \in C\left(\frac{f_1}{\|f_1\|_{\mathfrak{F}_p(\Theta)}}, \Theta\right), \quad c_2 \in C\left(\frac{f_2}{\|f_2\|_{\mathfrak{F}_p(\Theta)}}, \Theta\right),$$

we obtain

$$\begin{aligned}&\tau_{f_1+f_2}\Big(\big(\|f_1\|_{\mathfrak{F}_p(\Theta)} + \|f_2\|_{\mathfrak{F}_p(\Theta)}\big)(c_1 \vee c_2)t\Big) \\ &\quad\leq \|f_1\|_{\mathfrak{F}_p(\Theta)}\tau_{\frac{f_1}{\|f_1\|_{\mathfrak{F}_p(\Theta)}}}(c_1 t) + \|f_2\|_{\mathfrak{F}_p(\Theta)}\tau_{\frac{f_2}{\|f_2\|_{\mathfrak{F}_p(\Theta)}}}(c_2 t) \\ &\quad\leq \|f_1\|_{\mathfrak{F}_p(\Theta)} c_1\Theta(t) + \|f_2\|_{\mathfrak{F}_p(\Theta)} c_2\Theta(t) \\ &\quad\leq \big(\|f_1\|_{\mathfrak{F}_p(\Theta)} + \|f_2\|_{\mathfrak{F}_p(\Theta)}\big)(c_1 \vee c_2)\Theta(t).\end{aligned}$$

We conclude that

$$\begin{aligned}\|f_1 + f_2\|_{\mathfrak{F}_p(\Theta)} &\leq \big(\|f_1\|_{\mathfrak{F}_p(\Theta)} + \|f_2\|_{\mathfrak{F}_p(\Theta)}\big)\, essinf \left\{c_1 \vee c_2 : c_1 \in C\left(\frac{f_1}{\|f_1\|_{\mathfrak{F}_p(\Theta)}}, \Theta\right),\right. \\ &\qquad\qquad \left. c_2 \in C\left(\frac{f_2}{\|f_2\|_{\mathfrak{F}_p(\Theta)}}, \Theta\right)\right\} \\ &= \|f_1\|_{\mathfrak{F}_p(\Theta)} + \|f_2\|_{\mathfrak{F}_p(\Theta)},\end{aligned}$$

which is the required conclusion. This completes the proof of Proposition 3.3. □

3.4. *Partial order on the family of spaces* $\mathfrak{F}_p(\Theta)$ *with* $\Theta \in \Xi$. From the proof of Proposition 3.3 it follows that

$$\mathfrak{F}_p(0) = \{0\}.$$

It is also clear that

$$L_p(M) = \bigcup_{\Theta \in \Xi} \mathfrak{F}_p(\Theta).$$

Indeed, for an arbitrary function $f \in L_p(M)$, we may put $\Theta(t) = \tau_f(t)$. Then $f \in \mathfrak{F}_p(\Theta)$ in the obvious way. Taking into account (15), we obtain $\|f\|_{\mathfrak{F}_p(\Theta)} = 1$. Notice that from inequality (15) it follows that the embedding $\mathfrak{F}_p(\Theta) \subset L_p(M)$ is continuous for every function $\Theta : \mathbb{R}_+ \to \mathbb{R}_+$ that decreases to zero.

For $f \in L_p(M)$, we denote by Ξ_f the set of all $\Theta \in \Xi$ with $f \in \mathfrak{F}_p(\Theta)$. The following assertion shows that it is possible to introduce a partial order on the family of the spaces $\mathfrak{F}_p(\Theta)$ with $\Theta \in \Xi_f$. This family lacks the greatest element. If $f \not\equiv 0$ then it also lacks the least element.

Proposition 3.4. *If* $\Theta, \Theta' \in \Xi$ *and* $\Theta' \leq \Theta$ *then* $\mathfrak{F}_p(\Theta') \subseteq \mathfrak{F}_p(\Theta)$ *and, for every* $f \in \mathfrak{F}_p(\Theta')$, *we have*

$$\|f\|_{\mathfrak{F}_p(\Theta)} \leq \|f\|_{\mathfrak{F}_p(\Theta')}. \tag{16}$$

Moreover, if $0 \not\equiv f \in L_p(M)$ *then*

$$\operatorname*{\mathit{essinf}}_{\Theta \in \Xi_f} \Theta = 0, \qquad \sup_{\Theta \in \Xi_f} \|f\|_{\mathfrak{F}_p(\Theta)} = +\infty.$$

Proof. Since $\Theta' \leq \Theta$, we have

$$\tau_f(ct) \leq c\,\Theta'(t) \leq c\,\Theta(t)$$

for all $f \in \mathfrak{F}_p(\Theta')$, $t \geq 0$, and $c \in C(f, \Theta')$. Hence, we have $f \in \mathfrak{F}_p(\Theta)$ and inequality (16) holds.

We consider an arbitrary $\Theta \in \Xi_f$. It is not difficult to verify that, for every $\varepsilon > 0$, we have $\Theta_\varepsilon \in \Xi_f$, where

$$\Theta_\varepsilon(t) = \varepsilon\,\Theta\left(\frac{t}{\varepsilon}\right).$$

We conclude that

$$\operatorname*{\mathit{essinf}}_{\Theta \in \Xi_f} \Theta(t) \leq \operatorname*{\mathit{essinf}}_{\varepsilon > 0} \Theta_\varepsilon(t) = \operatorname*{\mathit{essinf}}_{\varepsilon > 0} \varepsilon\,\Theta\left(\frac{t}{\varepsilon}\right) = 0.$$

Assume that

$$\|f\|' = \sup_{\Theta \in \Xi_f} \|f\|_{\mathfrak{F}_p(\Theta)} < +\infty$$

for some $f \in L_p(M)$. Then, for all $\Theta \in \Xi_f$ and $t \geq 0$, we have

$$\tau_f\big(\|f\|' t\big) \leq \tau_f\big(\|f\|_{\mathfrak{F}_p(\Theta)} t\big) \leq \|f\|_{\mathfrak{F}_p(\Theta)} \Theta(t) \leq \|f\|'_p \Theta(t).$$

We find that

$$\tau_f\big(\|f\|'_p t\big) \leq \|f\|'_p \operatorname*{\mathit{essinf}}_{\Theta \in \Xi_f} \Theta(t) = 0$$

for every $t \geq 0$. We obtain $f \equiv 0$, which is a contradiction. Therefore, we have $\|f\|' = +\infty$. □

3.5. *Everywhere density of* $\mathfrak{F}_p(\Theta)$ *in* $L_p(M)$. For $a, b > 0$, we put

$$\Theta_{a,b}(t) = b(1 - t/a)\chi_{[0,a)}(t).$$

Lemma 3.1. *If* $a, b > 0$ *then* $\mathfrak{F}_p \subset \mathfrak{F}_p(\Theta_{a,b})$.

Proof. We consider a function $f \in \mathfrak{F}_p$ with $f \not\equiv 0$. It is not difficult to verify that

$$\tau_f(t) \leq \|f\|_p\big(1 - t/\|f\|_{\mathfrak{F}_p}\big)\chi_{[0,\|f\|_{\mathfrak{F}_p})}$$

for every $t \geq 0$. Indeed, it suffices, in the definition of $\tau_f(t)$, to consider the function $h \in \mathfrak{F}_p$ that coincides with f for $t \geq \|f\|_{\mathfrak{F}_p}$ and with $\frac{tf}{\|f\|_{\mathfrak{F}_p}}$ for $t < \|f\|_{\mathfrak{F}_p}$. We put $d = d(f) = \frac{\|f\|_p}{\|f\|_{\mathfrak{F}_p}}$. We find that

$$\tau_f\left(\frac{\|f\|_{\mathfrak{F}_p}}{\varepsilon}t\right) \leq \frac{\|f\|_{\mathfrak{F}_p}}{\varepsilon}\left(\varepsilon d(1 - t/\varepsilon)\chi_{[0,\varepsilon)}\right)$$

for all $\varepsilon > 0$ and $t \geq 0$. We choose $\varepsilon = \min\{a, b/d\}$. For every $t \geq 0$, we have

$$\tau_f\left(\frac{\|f\|_{\mathfrak{F}_p}}{\varepsilon}t\right) \leq \frac{\|f\|_{\mathfrak{F}_p}}{\varepsilon}\left(b(1 - t/a)\chi_{[0,a)}\right),$$

which is the required conclusion. Moreover, the inequality

$$\|f\|_{\mathfrak{F}_p(\Theta_{a,b})} \leq \frac{\|f\|_{\mathfrak{F}_p}}{\varepsilon} \tag{17}$$

holds for every $f \in \mathfrak{F}_p$. This completes the proof of the lemma. □

Proposition 3.5. *If* $\Theta \in \Xi$ *and* $\Theta(0+) = \lim\limits_{t\to 0+}\Theta(t) > 0$ *then we have*

$$\mathfrak{F}_p \subseteq \mathfrak{F}_p(\Theta).$$

Proof. In view of Proposition 3.4 and Lemma 3.1, it suffices to prove that there exists a number $a = a(\Theta) > 0$ such that

$$\Theta_{a,a}(t) \leq \Theta(t)$$

for every $t \geq 0$. We assume the contrary, i.e., assume that, for every $a > 0$, there exists $t_0 = t_0(a) > 0$ such that

$$\Theta_{a,a}(t_0) > \Theta(t_0). \tag{18}$$

By the definition of $\Theta_{a,a}$, we have $t_0 \in [0, a)$. In (18), we pass to the limit as $a \to 0+$. We find that $t_0(a) \to 0+$ and

$$0 < \Theta(0+) \leq \lim_{a\to 0+}\Theta_{a,a}(t_0) = \lim_{a\to 0+} a - t_0(a) = 0,$$

which is a contradiction. □

The following assertion is immediate from relations (15)–(17).

Corollary 3.1. *Assume that the conditions of Proposition* 3.5 *hold. Then there exists a number* $a = a(\Theta) > 0$ *such that, for every* $f \in \mathfrak{F}_p$, *we have*

$$\frac{\|f\|_{\mathfrak{F}_p}}{\Theta(0)} \leq \|f\|_{\mathfrak{F}_p(\Theta)} \leq \frac{\|f\|_p}{a} \ \ \textit{if} \ \ \|f\|_p \geq \|f\|_{\mathfrak{F}_p},$$
$$\frac{\|f\|_p}{\Theta(0)} \leq \|f\|_{\mathfrak{F}_p(\Theta)} \leq \frac{\|f\|_{\mathfrak{F}_p}}{a} \ \ \textit{if} \ \ \|f\|_p \leq \|f\|_{\mathfrak{F}_p}.$$

Notice that $\Theta(0+) = 0$ implies $\mathfrak{F}_p(\Theta) = \{0\}$. Thus, the nonzero space $\mathfrak{F}_p(\Theta)$ is an extension of the space $\mathfrak{F}_p$; hence, it is everywhere dense in $L_p(M)$.

4. TRANSITIVE ANOSOV DIFFEOMORPHISMS

In the present section, we consider a well-studied dynamical system (transitive Anosov diffeomorphism) and demonstrate the applications of Theorem 2.1 for deriving the central limit theorem. In the proof, we use approximation of the characteristic functions by Hölder continuous functions (see the appendix). We also use the notation from the appendix. For example, we denote by $H_\alpha(M)$ the set of Hölder continuous functions of class α with $\alpha \in (0,1)$, by $\|f\|_\alpha$ the Hölder norm, and by $\mathrm{Höld}_\alpha(f)$ the Hölder constant.

We begin with known results on Anosov diffeomorphisms that will be needed in the sequel.

4.1. *Initial estimates for the correlation function.* Let M be a compact C^∞-smooth Riemannian manifold and let T be a $C^{1+\alpha}$-smooth transitive Anosov diffeomorphism, i.e., the differential DT is a H'older continuous function of class α with $\alpha \in (0,1)$.

For an invariant measure μ we take the SRB-measure. A well known result of [5] says that, for H'older continuous functions, the correlation function decays at the exponential rate with respect to this measure. This result was first obtained for mixing topological Markov chains and then transferred to transitive Anosov diffeomorphisms with the use of methods of symbolic dynamics. It can be also proven with the use the construction of a Young tower with an exponential tail [26]. Here we use estimates with more precise constants which were obtained in [6] with the use of coupling methods.

Let $d_s = d_s(x,y)$ be a metric on the stable manifold W^s and let $d_u = d_u(x,y)$ be a metric on the unstable manifold W^u induced by the Riemannian metric on M. There exists $\nu \in (0,1)$ such that

$$d_s(Tx,Ty) \le \nu d_s(x,y), \quad x \in W^s(y),$$

$$d_u(T^{-1}x,T^{-1}y) \le \nu d_u(x,y), \quad x \in W^u(y).$$

We fix $\delta > 0$ and $\beta \in (0,1)$. For a measurable function $f : M \to \mathbb{R}$, we consider the following seminorms:

$$\|f\|_s = \|f\|_\infty + |f|_s, \quad |f|_s = \sup_{d_s(x,y)\le\delta} \frac{\big|f(x)-f(y)\big|}{d_s^\beta(x,y)},$$

$$\|f\|_u = \|f\|_1 + |f|_u, \quad |f|_u = \sup_{d_u(x,y)\le\delta} \frac{\big|f(x)-f(y)\big|}{d_u^\alpha(x,y)},$$

where the norm $\|f\|_1$ is taken with respect to the Riemannian volume on $\mathcal{M}$. Let C_s denote the set of all measurable functions $f : M \to \mathbb{R}$ with $\|f\|_s < \infty$. It is clear that $f \in H_\beta(M)$ implies

$$\|f\|_s \le \|f\|_\beta < \infty.$$

Similar arguments for $f \in H_\alpha(M)$ show that

$$\|f\|_u \le \max\{\|1\|_1, 1\}\|f\|_\alpha < \infty.$$

As is proven in [6, Corollary 2.1], there exist constants $0 < \vartheta < 1$ and $C_T > 0$ such that, for all $f \in H_\alpha(M)$, $g \in C_s$, and $n \in \mathbb{N}$, we have

$$\big|c_n(f,g)\big| \le C_T \|f\|_u \|g\|_s \vartheta^n.$$

In the sequel, we consider bounded functions only. Hence, we may use a similar estimate; namely,

$$\big|c_n(f,g)\big| \le C_T \max\big\{\|1\|_1, 1\big\} \|f\|'_u \|g\|_s \vartheta^n, \;\; n \in \mathbb{N},$$

where

$$\|f\|'_u = \|f\|_\infty + |f|_u.$$

It is clear that this inequality remains valid for complex-valued functions $f \in H_\alpha(M) + iH_\alpha(M)$ and $g \in C_s + iC_s$. We introduce the same norms as for real-valued functions. Simple calculations allow us to obtain a similar estimate; namely, we have

$$\big|c_n(f,g)\big| \le 4C_T \max\big\{\|1\|_1, 1\big\} \|f\|'_u \|g\|_s \vartheta^n, \;\; n \in \mathbb{N}. \tag{19}$$

Since, for every $\gamma \in (0,1)$, the set of functions of the form $H_\gamma(M) + iH_\gamma(M)$ is everywhere dense in $L_p(M)$, $1 \le p < \infty$, we may apply Theorem 2.1 (and Corollary 2.1) to

$$\mathfrak{F}_p = H_\alpha(M) + iH_\alpha(M) \text{ with the norm } \|\cdot\|'_u$$

and

$$\mathfrak{G}_q = C_s + iC_s \text{ with the norm } \|\cdot\|_s$$

for each $1 < p < \infty$ with $1/p + 1/q = 1$. Moreover, we have $\Phi(n) = \vartheta^n$.

4.2. *Central limit theorem.* In [25, Theorem 14], the exponential decay of the correlation function for every (real-valued) function $f \in H_\alpha(M) \cap C_s$ was used in the proof of validity of the central limit theorem (CLT) for the same observables. We generalize this result to a wider class of observables; namely, to the class of the characteristic functions of Borel sets $A \subseteq M$ such that, in a neighborhood of the boundary ∂A, there is a power function singularity of order $l > 0$, i.e.,

$$\gamma_A(\delta) + \gamma^A(\delta) \le C_A \delta^l \tag{20}$$

for every $\delta > 0$ and a suitable constant $C_A > 0$. For the definition of functions γ_A and γ^A, see formula (26) in the appendix. Let Σ_l denote the family of sets with boundary of zero measure that satisfy (20). Notice that, in [20], the principle of large deviations was studied for the sequence of times of return to a set from Σ_l.

Since, for every Borel set $A \subseteq M$, we have

$$\gamma_A(\delta) = \gamma^{M\setminus A}(\delta),$$

we find that $A \in \Sigma_l$ if and only if $M \setminus A \in \Sigma_l$. If $M \subseteq \mathbb{R}^N$, $N \ge 1$, and $\mu = m^N$ is an N-dimensional Lebesgue measure then each bounded Borel set A with rectifiable piecewise-smooth boundary belongs to the class Σ_1. A similar assertion is valid for measures μ with bounded density that are absolutely continuous with respect to m^N.

We recall the notation and facts that will be needed in the proof of the CLT. For a real-valued function f, we denote

$$S_n = \sum_{j=0}^{n-1} f \circ T^j, \quad \mathbb{E} f = \int_M f(x)\, dx, \quad \mathbb{D}\, S_n = \mathbb{E}(S_n^2) - \mathbb{E}^2(S_n).$$

The assertion of the CLT is as follows: The distribution of the sequence $(S_n - n\mathbb{E}f)/\sqrt{\mathbb{D}\, S_n}$ converges to the standard normal distribution, i.e., for every $v \in \mathbb{R}$, we have

$$\lim_{n\to\infty} \mu\left(\frac{S_n - n\mathbb{E}f}{\sqrt{\mathbb{D}\, S_n}} \le v\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{v} e^{-s^2/2}\, ds.$$

The following theorem (see [11, Sec. 7.8; 25, Theorem 1 and Corollary 3]) allows us to derive the CLT from the condition of rapid decay of the correlation function. Let $0 < b < a < 1/2$ and let

$$p = [n^a],\ \ q = [n^b],\ \ k = \big[n/(p+q)\big] \sim n^{1-a} \tag{21}$$

for every $n \in \mathbb{N}$

Theorem 4.1. *Let*

$$g = e^{ivf/\sqrt{k\mathbb{D}\, S_p}}, \quad w_1 = g \cdot g \circ T \cdots g \circ T^{p-1}.$$

For every $2 \le r \le k$*, put*

$$w_r = w_1 \circ T^{(p+q)(r-1)}, \quad W_r = w_1 w_2 \cdots w_{r-1}.$$

If the conditions

(i) $\sum_{n=1}^{\infty} n\big|c_n(f,f)\big| < \infty$,

(ii) $\lim\limits_{n\to\infty} \sum_{r=2}^{k} \big|c_{p+q}(w_1, W_r)\big| = 0$

hold for every $v \in \mathbb{R}$ *then the* CLT *is valid.*

On the basis of this result, we prove the following assertion.

Theorem 4.2. *The* CLT *is valid for every transitive Anosov diffeomorphism* T *with an SRB-measure* μ *and the characteristic function* χ_A*, where* $A \in \Sigma_l$ *and* $l > 0$.

4.3. *Proof of Theorem* 4.2. By Theorem 4.1, it suffices to prove that conditions (i) and (ii) hold for $f = \chi_A$, where $A \in \Sigma_l$. We verify condition (i).

We put

$$\mathfrak{F}_2 = H_\alpha(M) + iH_\alpha(M) \ \text{ with the norm } \ \|\cdot\|'_u$$

and

$$\mathfrak{G}_2 = C_s + iC_s \ \text{ with the norm } \ \|\cdot\|_s.$$

We describe spaces $\mathfrak{F}_2(\Theta_1)$ and $\mathfrak{G}_2(\Theta_2)$ containing χ_A. For arbitrary $a, b, c > 0$, we put

$$\Theta_{a,b,c}(t) = \min\left\{1, \frac{1}{|t-1|^{c/ab}}\right\}, \quad t \geq 0.$$

Lemma 4.1. *Let $A \in \Sigma_l$ for some $l > 0$ and let*

$$\Theta_1(t) = \Theta_{\alpha,2,l}(t), \ \ \Theta_2(t) = \Theta_{\beta,2,l}(t).$$

Then we have

$$\chi_A \in \mathfrak{F}_2(\Theta_1) \cap \mathfrak{G}_2(\Theta_2),$$
$$\max\left\{\|\chi_A\|_{\mathfrak{F}_2(\Theta_1)}; \|\chi_A\|_{\mathfrak{G}_2(\Theta_2)}\right\} \leq \max\left\{1; \sqrt{C_A}\,\right\}.$$

Proof of Lemma 4.1. We consider the case of $\mathfrak{F}_2$-approximations only. The case of $\mathfrak{G}_2$-approximations is similar. We choose a function $\varphi_A^{\alpha,\delta}$ as in Lemma A.2 below, where δ satisfies the equality $1 + 1/\delta^\alpha = t$ for $t > 1$. Then we have

$$1 < \|\varphi_A^{\alpha,\delta}\|'_u \leq \|\varphi_A^{\alpha,\delta}\|_\alpha \leq 1 + 1/\delta^\alpha = t$$

and, consequently,

$$\tau_{\chi_A}(t) \leq \left\|\chi_A - \varphi_A^{\alpha,\delta}\right\|_2 \leq \gamma_A^{1/2}(\delta) = \gamma_A^{1/2}\left(\frac{1}{(t-1)^{1/\alpha}}\right)$$

for every $t > 1$. Notice that, for $t \in \big(1, 1 + 1/\delta_*^\alpha(A)\big)$, we have

$$\gamma_A\left(\frac{1}{(t-1)^{1/\alpha}}\right) = \mu(\operatorname{int} A \setminus \varnothing) = \mu(A),$$

where $\delta_*(A)$ is defined in the appendix. Moreover, we find that $\tau_{\chi_A}(0) = \mu^{1/2}(A)$. Since the best approximation is continuous, we combine the above facts and conclude that

$$\tau_{\chi_A}(t) \leq \min\left\{\mu^{1/2}(A); \gamma_A^{1/2}\left(\frac{1}{|t-1|^{1/\alpha}}\right)\right\} \tag{22}$$

for every $t \geq 0$. Taking into account (22) and (20), we obtain

$$\begin{aligned}
\tau_{\chi_A}(t) &\leq \min\left\{\mu^{1/2}(A); \gamma_A^{1/2}\left(\frac{1}{|t-1|^{1/\alpha}}\right)\right\} \\
&\leq \min\left\{\mu^{1/2}(A); \frac{C_A^{1/2}}{|t-1|^{l/2\alpha}}\right\} \\
&\leq \max\{\mu^{1/2}(A); C_A^{1/2}\} \min\left\{1, \frac{1}{|t-1|^{l/2\alpha}}\right\}
\end{aligned}$$

for every $t \geq 0$. If $C_A \leq 1$ then

$$\max\left\{\mu^{1/2}(A); C_A^{1/2}\right\} \leq 1.$$

We obtain

$$\tau_{\chi_A}(t) \leq \Theta_1(t) \ \text{ for every } \ t \geq 0.$$

Hence, we have

$$\chi_A \in \mathfrak{F}_2(\Theta_1), \quad \|\chi_A\|_{\mathfrak{F}_2(\Theta_1)} \leq 1.$$

If $C_A > 1$ then

$$\max\{\mu^{1/2}(A); C_A^{1/2}\} = C_A^{1/2}.$$

We obtain

$$\tau_{\chi_A}(C_A^{1/2}t) \le \tau_{\chi_A}(t) \le C_A^{1/2}\Theta_1(t) \text{ for every } t \ge 0.$$

Hence, we have

$$\chi_A \in \mathfrak{F}_2(\Theta_1) \text{ and } \|\chi_A\|_{\mathfrak{F}_2(\Theta_1)} \le C_A^{1/2}.$$

□

By Lemma 4.1, we find that

$$\chi_A \in \mathfrak{F}_2(\Theta_1) \cap \mathfrak{G}_2(\Theta_2)$$

and

$$(\Theta_1 \vee \Theta_2)(t) = \mathcal{O}\Big(t^{(-l)/(2(\alpha\vee\beta))}\Big) \text{ as } t \to +\infty.$$

We use this asymptotic relation, estimate (19), and assertion (i) of Corollary 2.1. We obtain

$$\big|c_n(\chi_A, \chi_A)\big| = \mathcal{O}(\vartheta_0^n) \text{ as } n \to \infty,$$

where

$$\vartheta_0 = \vartheta^{l/(4(\alpha\vee\beta)+l)}. \tag{23}$$

Thus, condition (i) of Theorem 4.1 holds, i.e., we have

$$\sum_{n=1}^{\infty} n\big|c_n(\chi_A, \chi_A)\big| < \infty.$$

We verify condition (ii). Notice that

$$c_{p+q}(w_1, W_r) = c_q(w_1 \circ T^{-p}, W_r).$$

As above, we describe a space $\mathfrak{F}_2(\Theta_1)$ containing $w_1 \circ T^{-p}$ and a space $\mathfrak{G}_2(\Theta_2)$ containing W_r (see Lemmas 4.2 and 4.3 respectively).

Lemma 4.2. *Let $A \in \Sigma_l$ for some $l > 0$ and let*

$$\Theta_1(t) = \Theta_{\alpha,2,l}\big((1-\nu^\alpha)t\big).$$

Then we have $w_1 \circ T^{-p} \in \mathfrak{F}_2(\Theta_1)$ and

$$\|w_1 \circ T^{-p}\|_{\mathfrak{F}_2(\Theta_1)} \le \max\{1, 2^{1+1/\alpha}C_A^{1/2}p\}.$$

Lemma 4.3. *Let $A \in \Sigma_l$ for some $l > 0$ and let*

$$\Theta_2(t) = \Theta_{\beta,2,l}\big((1-\nu^\beta)t\big).$$

Then $W_r \in \mathfrak{G}_2(\Theta_2)$ and

$$\|W_r\|_{\mathfrak{G}_2(\Theta_2)} \le \max\{1, 2^{1+1/\beta}C_A^{1/2}p(r-1)\}.$$

These lemmas, estimate (19), Theorem 2.1, Corollary 2.1, and relations (21) imply that

$$\begin{aligned}
&\sum_{r=2}^{k}\big|c_q(w_1 \circ T^{-p}, W_r)\big| \\
&\quad= \sum_{r=2}^{k}\mathcal{O}\big(\|w_1 \circ T^{-p}\|_{\mathfrak{F}_2(\Theta_1)}\|W_r\|_{\mathfrak{G}_2(\Theta_2)}\Phi'(q)\big) \\
&\quad= \mathcal{O}\left(p^2\Phi'(q)\sum_{r=2}^{k}(r-1)\right) = \mathcal{O}\big(p^2k^2\vartheta_0^q\big) \\
&\quad= \mathcal{O}\big(n^{2a}n^{2-2a}\vartheta_0^{n^b}\big) = \mathcal{O}\big(n^2\vartheta_0^{n^b}\big)
\end{aligned}$$

as $n \to \infty$, where ϑ_0 is defined by equality (23). This yields condition (ii). Thus, it remains to prove lemmas 4.2 and 4.3, which will complete the proof of Theorem 4.2.

We put

$$a(n,v) = e^{iv/\sqrt{k\mathbb{D}\,S_p}}, \quad v \in \mathbb{R}.$$

Then

$$g = e^{iv\chi_A/\sqrt{k\mathbb{D}\,S_p}} = a(n,v)\chi_A + \chi_{M\setminus A}.$$

Proof of Lemma 4.2. Since

$$w_1 \circ T^{-p} = \prod_{j=1}^{p} g \circ T^{-j},$$

we find the corresponding $\mathfrak{F}_2$-approximation of the form $\prod_{j=1}^{p} h \circ T^{-j}$. Let $\varphi_A^{\alpha,\delta}$ and $\varphi_{M\setminus A}^{\alpha,\delta}$ be as in Lemma A.2, where δ satisfies the equality $1 + 2/\delta^\alpha = t$ for $t > 1$. For $h = a(n,v)\varphi_A^{\alpha,\delta} + \varphi_{M\setminus A}^{\alpha,\delta}$, we have

$$\begin{aligned} 1 < \|h\|'_u &\le 1 + \left|\varphi_A^{\alpha,\delta}\right|_u + \left|\varphi_{M\setminus A}^{\alpha,\delta}\right|_u \\ &\le 1 + \mathrm{Höld}_\alpha\big(\varphi_A^{\alpha,\delta}\big) + \mathrm{Höld}_\alpha\big(\varphi_{M\setminus A}^{\alpha,\delta}\big) \\ &\le 1 + 2/\delta^\alpha = t. \end{aligned}$$

From the proof of [25, Corollary 13] it follows that

$$\prod_{j=1}^{p} h \circ T^{-j} \in \mathfrak{F}_2$$

and

$$\left\|\prod_{j=1}^{p} h \circ T^{-j}\right\|'_u \le \frac{1}{1-\nu^\alpha}\|h\|'_u \le \frac{t}{1-\nu^\alpha}.$$

For $t > 1$, we obtain

$$\begin{aligned} \tau_{w_1\circ T^{-p}}\left(\frac{t}{1-\nu^\alpha}\right) &\le \left\|\prod_{j=1}^{p} g \circ T^{-j} - \prod_{j=1}^{p} h \circ T^{-j}\right\|_2 \\ &\le \sum_{i=1}^{p} \left\|\prod_{j=1}^{i-1} g \circ T^{-j}\right\|_\infty \|g\circ T^{-i} - h \circ T^{-i}\|_2 \left\|\prod_{j=i+1}^{p} h \circ T^{-j}\right\|_\infty \\ &\le \sum_{i=1}^{p} \|g-h\|_2 \le p\left(\left\|\chi_A - \varphi_A^{\alpha,\delta}\right\|_2 + \left\|\chi_{M\setminus A} - \varphi_{M\setminus A}^{\alpha,\delta}\right\|_2\right) \\ &\le 2p\gamma_A^{1/2}(\delta) \le 2pC_A^{1/2}\delta^{l/2} = \frac{2^{1+1/\alpha}C_A^{1/2}p}{(t-1)^{l/2\alpha}}. \end{aligned}$$

Since $\tau_{w_1\circ T^{-p}}(0) \le 1$, for every $t \ge 0$, we obtain

$$\begin{aligned} \tau_{w_1\circ T^{-p}}(t) &\le \min\left\{1; \frac{2^{1+1/\alpha}C_A^{1/2}p}{\left|(1-\nu^\alpha)t - 1\right|^{l/2\alpha}}\right\} \\ &\le \max\left\{1; 2^{1+1/\alpha}C_A^{1/2}p\right\}\Theta_{\alpha,2,l}\big((1-\nu^\alpha)t\big). \end{aligned}$$

□

Proof of Lemma 4.3. Notice that

$$W_r = \prod_{j=0}^{r-2}\left(\prod_{i=0}^{p-1} g\circ T^i\right)\circ T^{j(p+q)} = \prod_{j=0}^{r-2}\prod_{i=0}^{p-1} g\circ T^{i+j(p+q)};$$

hence, we find $\mathfrak{G}_2$-approximation of the same form, i.e.,

$$\prod_{j=0}^{r-2}\prod_{i=0}^{p-1} h\circ T^{i+j(p+q)}.$$

As in the proof of Lemma 4.2, we put

$$h = a(n,v)\varphi_A^{\beta,\delta} + \varphi_{M\setminus A}^{\beta,\delta} \quad \text{with} \quad \delta > 0$$

satisfying the equality $1 + 2/\delta^\beta = t$ for $t > 1$. Then we have

$$\begin{aligned} 1 < \|h\|_s &\le 1 + \left|\varphi_A^{\beta,\delta}\right|_s + \left|\varphi_{M\setminus A}^{\beta,\delta}\right|_s \\ &\le 1 + \text{Höld}_\beta\big(\varphi_A^{\beta,\delta}\big) + \text{Höld}_\beta\big(\varphi_{M\setminus A}^{\beta,\delta}\big) \\ &\le 1 + 2/\delta^\beta = t. \end{aligned}$$

From the proof of [25, Corollary 13] it follows that

$$\prod_{j=0}^{r-2}\prod_{i=0}^{p-1} h\circ T^{i+j(p+q)} \in \mathfrak{G}_2$$

and

$$\left\|\prod_{j=1}^{r-1}\prod_{i=0}^{p-1} h\circ T^{i+j(p+q)}\right\|_s \le \frac{1}{1-\nu^\beta}\|h\|_s \le \frac{t}{1-\nu^\beta}.$$

For every $t > 1$, we repeat calculations that are similar to those in the proof of Lemma 4.2 and obtain

$$\tau_{W_r}\left(\frac{t}{1-\nu^\beta}\right) \le \frac{2^{1+1/\beta} C_A^{1/2} p(r-1)}{(t-1)^{l/2\beta}}.$$

The remaining part of the proof is the same as the corresponding part of the proof of Lemma 4.2. □

APPENDIX. HÖLDER APPROXIMATION OF THE CHARACTERISTIC FUNCTIONS OF CONTINUITY SETS

For many dynamical systems, estimates for the correlation function are obtained for classes of regular observables that contain the class of H'older continuous functions. Therefore, it is useful to distinguish a class of functions such that observables in this class admit good Hölder approximations. In the present article, the following class $\mathcal{X}_\mu$ serves as an example of such a class. We put

$$\mathcal{X}_\mu = \big\{\chi_A : \mu(\partial A) = 0\big\},$$

i.e., the characteristic function χ_A of a Borel set $A \subseteq M$ belongs to $\mathcal{X}_\mu$ if the measure of the boundary ∂A of A is equal to zero. Sometimes such sets are called continuity sets of the measure μ. Functions in $\mathcal{X}_\mu$ are μ-almost everywhere continuous because the measure of the set of discontinuity points of χ_A (i.e., the boundary ∂A of A) is equal to zero. For μ-almost everywhere continuous functions, two-sided integral Hölder approximation is known which can be used, for example, in estimates for large deviations of the ergodic averages (see [16, 18]). In the present article, we use a more convenient construction. For the characteristic functions in $\mathcal{X}_\mu$, we consider the construction of two-sided Hölder approximation that involves neither supremal nor infimal convolutions, cf. [18].

We present auxiliary facts and then turn to the main result of the section.

A.1. *Hölder continuous observables.* We recall that $f : M \to \mathbb{R}$ is a H'older continuous function if there exist constants $\alpha \in (0,1]$ and $h \geq 0$ such that, for all $x, y \in M$, we have

$$\big|f(x) - f(y)\big| \leq h d^{\alpha}(x,y).$$

We denote by $\mathrm{Höld}_{\alpha}(f)$ the least such a constant h for f. We denote by $H_{\alpha}(M)$ the set of bounded Hölder continuous functions of class α with $\alpha \in (0,1]$. The norm in $H_{\alpha}(M)$ is defined by the rule

$$\|f\|_{\alpha} = \|f\|_{\infty} + \mathrm{Höld}_{\alpha}(f).$$

We consider a general construction that allows us to construct Hölder continuous functions on M.

Lemma A.1. *Let $\sigma : \mathbb{R}^{+} \to \mathbb{R}$ be an arbitrary Hölder continuous function of class α with $\alpha \in (0,1]$ and let $\varnothing \neq A \subseteq M$. Then, for all $h > 0$ and $\beta \in (0,1]$, the function $\sigma_{h,A,\beta}$, where*

$$\sigma_{h,A,\beta}(x) = \sigma\big(h \cdot d^{\beta}(x,A)\big), \ \ x \in M,$$

is a Hölder continuous function of class $\alpha\beta$ with $\alpha\beta \in (0,1]$; moreover, we have

$$\mathrm{Höld}_{\alpha\beta}(\sigma_{h,A,\beta}) \leq h^{\alpha}\mathrm{Höld}_{\alpha}(\sigma). \tag{24}$$

Proof. For every $\beta \in (0,1]$, the function d^{β} is a metric on M too. Therefore, is suffices to prove the assertion for $\beta = 1$.

For all $x, y \in M$, we have

$$\begin{aligned}\big|\sigma_{h,A,1}(x) - \sigma_{h,A,1}(y)\big| &= \Big|\sigma\big(h \cdot d(x,A)\big) - \sigma\big(h \cdot d(y,A)\big)\Big| \\ &\leq \mathrm{Höld}_{\alpha}(\sigma)\big|h \cdot d(x,A) - h \cdot d(y,A)\big|^{\alpha} \\ &= \mathrm{Höld}_{\alpha}(\sigma)h^{\alpha}\big|d(x,A) - d(y,A)\big|^{\alpha} \\ &\leq \mathrm{Höld}_{\alpha}(\sigma)h^{\alpha}d^{\alpha}(x,y).\end{aligned}$$

□

A.2. *Extensions and restrictions of sets.* Let $A \subseteq M$ and let A be a nonempty set. We put

$$(A)^{\delta} = \big\{x \in M : d(x,A) < \delta\big\}, \ \ \delta > 0,$$

and call this set the open δ-extension of A. We put

$$(A)_{\delta} = \big\{x \in A : d(x,\partial A) \geq \delta)\big\}, \ \ \delta > 0,$$

and call this set the closed δ-restriction of A. For definiteness, we assume that $(\varnothing)_{\delta} = (\varnothing)^{\delta} = \varnothing$ for each $\delta > 0$. It is clear from the definition that

$$(A)_{\delta} = (\mathrm{int}A)_{\delta} \subseteq \mathrm{int}A \subseteq A \subseteq \mathrm{cl}A \subseteq (\mathrm{cl}A)^{\delta} = (A)^{\delta}.$$

Moreover, the following equalities are valid (see, for example, [7, Sec. 2]):

$$(A)^{\delta} = M \setminus (M \setminus A)_{\delta}, \quad (A)_{\delta} = M \setminus (M \setminus A)^{\delta}. \tag{25}$$

It is not difficult to see that, for a nonempty set $A \subseteq M$, we have $(A)_{\delta} = \varnothing$ for every $\delta > \delta_{*}(A)$, where

$$\delta_{*}(A) = \sup_{x \in A} d(x,\partial A).$$

It is also clear that $(A)^{\delta} = M$ for every $\delta > \delta^{*}(A)$, where

$$\delta^{*}(A) = \sup_{x \in M} d(x,A).$$

In connection with extensions and restrictions, we consider concentration of the measure near the boundary. For every Borel set $A \subseteq M$, we consider functions γ_A and γ^A from $\mathbb{R}^{+}$ to $[0,1]$. They are defined by the equalities

$$\gamma_A(\delta) = \mu(\mathrm{int}A \setminus (A)_{\delta}), \quad \gamma^A(\delta) = \mu((A)^{\delta} \setminus \mathrm{cl}A) \tag{26}$$

for all $\delta > 0$. It is clear that $\gamma_A(\delta), \gamma^A(\delta) \to 0+$ as $\delta \to 0+$.

A.3. *Two-sided Hölder approximation for functions in $\mathfrak{X}_\mu$.* We use δ-extensions and δ-restrictions of a set A and construct Hölder continuous functions that approximate the characteristic function χ_A. We control over the growth of their Hölder constants as δ converges to zero.

Lemma A.2. *Let (M,d) be a metric space and let $\varnothing \neq A \subset M$. For all $\delta > 0$ and $\alpha \in (0,1]$, there exist Hölder continuous functions*

$$\varphi_A^{\alpha,\delta}, \psi_A^{\alpha,\delta} \in H_\alpha(M)$$

such that

$$0 \le \varphi_A^{\alpha,\delta} \le \chi_A \le \psi_A^{\alpha,\delta} \le 1, \tag{27}$$

$$\{\psi_A^{\alpha,\delta} \neq 0\} = (A)^\delta, \quad \{\psi_A^{\alpha,\delta} = 1\} = \operatorname{cl} A, \tag{28}$$

$$\{\varphi_A^{\alpha,\delta} \neq 0\} = \operatorname{int} A, \quad \{\varphi_A^{\alpha,\delta} = 1\} = (A)_\delta, \tag{29}$$

$$\text{Höld}_\alpha(\varphi_A^{\alpha,\delta}) \le \delta^{-\alpha}, \quad \text{Höld}_\alpha(\psi_A^{\alpha,\delta}) \le \delta^{-\alpha}; \tag{30}$$

moreover, for every $p \in [1,+\infty)$, the following inequalities hold:

$$\big\|\chi_A - \varphi_A^{\alpha,\delta}\big\|_p^p \le \gamma_A(\delta) + \mu(\partial A), \quad \big\|\chi_A - \psi_A^{\alpha,\delta}\big\|_p^p \le \gamma^A(\delta) + \mu(\partial A). \tag{31}$$

Proof. We follow the well-known construction from [4, Theorem 1.2]. We consider the function $\sigma : \mathbb{R}^+ \to \mathbb{R}$, where

$$\sigma(t) = \begin{cases} 1 - t, & t \in [0,1], \\ 0, & t > 1. \end{cases}$$

It is easy to verify that this is a Hölder continuous function. Indeed, for all $\alpha \in (0,1]$ and $t, s \ge 0$, we have

$$\big|\sigma(t) - \sigma(s)\big| \le |t-s|^\alpha.$$

For $\alpha \in (0,1]$ and $\delta > 0$, we put

$$\psi_A^{\alpha,\delta} = \sigma_{1/\delta,A,1}, \tag{32}$$

$$\varphi_A^{\alpha,\delta} = 1 - \sigma_{1/\delta,M\setminus A,1}, \tag{33}$$

where $\sigma_{1/\delta,A,1}$ and $\sigma_{1/\delta,M\setminus A,1}$ are constructed from σ as in Lemma A.1. Since the functions are independent of α, we omit such subscripts and superscripts in the sequel. We prove equality (28). We use the following equivalences:

$$\psi_A^\delta(x) = 0 \;\Leftrightarrow\; d(x,A) \ge \delta \;\Leftrightarrow\; x \notin (A)^\delta,$$
$$\psi_A^\delta(x) = 1 \;\Leftrightarrow\; d(x,A) = 0 \;\Leftrightarrow\; x \in \operatorname{cl} A.$$

We use similar relations in the proof of (29):

$$\varphi_A^\delta(x) = 0 \;\Leftrightarrow\; d(x, M\setminus A) = 0 \;\Leftrightarrow\; x \in \operatorname{cl}(M \setminus A) \;\Leftrightarrow\; x \notin \operatorname{int} A,$$
$$\varphi_A^\delta(x) = 1 \;\Leftrightarrow\; d(x, M\setminus A) \ge \delta \;\Leftrightarrow\; x \notin (M\setminus A)^\delta \;\Leftrightarrow\; x \in (A)_\delta.$$

In the last relation, we use property (25). Since $0 \le \sigma \le 1$, it is obvious that (28) and (29) imply (27). Inequality (30) for Hölder constants is exactly inequality (24) for the functions $\sigma_{1/\delta,A,1}$ and $\sigma_{1/\delta,M\setminus A,1}$. Estimates (31) are immediate from (26)–(29). We present the corresponding calculations, for example, for $\varphi^\delta(x)$:

$$\begin{aligned}
\big\|\chi_A - \varphi_A^\delta\big\|_p^p &= \int_A \big(\chi_A - \varphi_A^\delta\big)^p \, d\mu \\
&= \int_{\partial A} \big(1 - \varphi_A^\delta\big)^p \, d\mu + \int_{\operatorname{int} A} \big(1 - \varphi_A^\delta\big)^p \, d\mu \\
&= \mu(\partial A) + \int_{\operatorname{int} A \setminus (A)_\delta} \big(1 - \varphi_A^\delta\big)^p \, d\mu \\
&\le \mu(\partial A) + \gamma_A(\delta).
\end{aligned}$$

□

REFERENCES

1. J. F. Alves, J. M. Freitas, S. Luzzatto, and S. Vaienti, "From rates of mixing to recurrence times via large deviations," Adv. Math. **228** , 1203 (2011).
2. V. Baladi, *Positive Transfer Operators and Decay of Correlations* (World Scientific, Singapore, 2000).
3. J. Bergh and J. Löfström, *Interpolation Spaces. An Introduction* (Springer-Verlag, Berlin–Heidelberg–New York, 1976).
4. P. Billingsley, *Convergence of Probability Measures* (Wiley, Chichester, 1999).
5. R. Bowen, *Methods of Symbolic Dynamics. Collected Papers* (Mir, Moscow, 1979) [in Russian].
6. X. Bressaud and C. Liverani, "Anosov diffeomorphisms and coupling," Ergodic Theory Dynam. Systems **22**, 129 (2002).
7. V. I. Burenkov, "On the density of infinitely differentiable functions in Sobolev spaces for an arbitrary open set," Trudy Mat. Inst. Steklov **131**, 39 (1974) [Proc. Steklov Inst. Math. **131**, 39 (1974)].
8. J.-R. Chazottes, P. Collet, and B. Schmitt, "Statistical consequences of the Devroye inequality for processes. Applications to a class of non-uniformly hyperbolic dynamical systems," Nonlinearity **18**, 2341 (2005).
9. N. I. Chernov, "Limit theorems and Markov approximations for chaotic dynamical systems," Probab. Theory Relat. Fields **101**, 321 (1995).
10. N. Chernov, "Advanced statistical properties of dispersing billiards," J. Statist. Phys. **122**, 1061 (2006).
11. N. Chernov and R. Markarian, *Chaotic Billiards* (Izhevsk Inst. Comp. Sci., Izhevsk, 2006). [*Chaotic Billiards* (Amer. Math. Soc., Providence, RI, 2006)].
12. A.-H. Fan, "Decay of correlation for expanding toral endomorphisms," in *Dynamical Systems*, 29 (World Scientific, Singapore, 1999).
13. M. Holland, "Slowly mixing systems and intermittency maps," Ergodic Theory Dynam. Systems 25, 133 (2005).
14. A. G. Kachurovskiĭ, "The rate of convergence in ergodic theorems," Uspekhi Mat. Nauk **51**, no. 4, 73 (1996) [Russian Math. Surveys **51**, 653 (1996)].
15. A. G. Kachurovskiĭ and I. V. Podvigin, "Correlations, large deviations, and rates of convergence in ergodic theorems for characteristic functions," Dokl. Akad. Nauk **461**, 509 (2015) [Dokl. Math. **91**, 204 (2015)].
16. A. G. Kachurovskiĭ and I. V. Podvigin, "Large deviations and rates of convergence in the Birkhoff ergodic theorem: from Hölder continuity to continuity," Dokl. Akad. Nauk **466**, 12 (2016) [Dokl. Math. **93**, 6 (2016)].
17. A. G. Kachurovskiĭ and I. V. Podvigin, "Estimates of the rate of convergence in the von Neumann and Birkhoff ergodic theorems," Trudy Mosk. Mat. Obshch. **77**, 1 (2016) [Trans. Moscow Math. Soc. 1, (2016)].
18. A. G. Kachurovskiĭ and I. V. Podvigin, "Large deviations of the ergodic averages: from Hölder continuity to continuity almost everywhere," Mat. Trudy **20**, 97 (2017) [Siberian Adv. Math. **28**, 23 (2018)].
19. O. Knill, "Singular continuous spectrum and quantitative rates of weak mixing," Discrete Contin. Dynam. Systems **4**, 33 (1998).
20. R. Leplaideur and B. Saussol, "Large deviations for return times in non-rectangle sets for Axiom A diffeomorphisms," Discrete Contin. Dynam. Systems **22**, 327 (2008).
21. I. Melbourne, "Large and moderate deviations for slowly mixing dynamical systems," Proc. Amer. Math. Soc. **137**, 1735 (2009).
22. M. Ruziboev, "Decay of correlations for invertible maps with non-Hölder observables," Dynam. Systems **30**, 341 (2015).
23. O. Sarig, "Decay of correlations," in *Handbook of Dynamical Systems*, Vol. 1 B, 244 (Elsevier, Amsterdam, 2006).
24. V. V. Sedalishchev, "Interrelation between the convergence rates in von Neumann's and Birkhoff's ergodic theorems," Siberian Math. J. **55**, 336 (2014) [Sibirsk. Mat. Zh. **55**, 412 (2014)].
25. M. Stenlund, "A strong pair correlation bound implies the CLT for Sinai billiards," J. Statist. Phys. **140**, 154 (2010).
26. L.-S. Young, "Statistical properties of dynamical systems with some hyperbolicity," Ann. Math. **147**, 585 (1998).
27. L.-S. Young, "Recurrence times and rates of mixing," Israel J. Math. **110** , 153 (1999).
28. H.-K. Zhang, "Decay of correlations on non-Hölder observables," Internat. J. Nonlinear Sci. **10**, 359 (2010).