

RATES OF MIXING FOR POTENTIALS OF SUMMABLE VARIATION

MARK POLLICOTT

ABSTRACT. It is well known that for subshifts of finite type and equilibrium measures associated to Hölder potentials we have exponential decay of correlations. In this article we derive explicit rates of mixing for equilibrium states associated to more general potentials.

0. INTRODUCTION

In this paper we shall consider the rate of mixing of subshifts of finite type with respect to equilibrium states for potentials of summable variation.

Let $\sigma : X_A \rightarrow X_A$ denote a transitive two-sided subshift of finite type and let $g : X_A \rightarrow \mathbb{R}$ be a function of summable variation. Thus there exists a unique equilibrium state μ for g [7]. Assume that $f : X_A \rightarrow \mathbb{R}$ is Hölder, then we will study the behaviour as $N \rightarrow +\infty$ of the *correlation function*

$$\rho(N) := \int f \circ \sigma^N \cdot f d\mu - \left(\int f d\mu \right)^2.$$

It is a well known result that if the n th variation $\text{var}_n(g)$ tends to zero exponentially fast, then $\rho(N)$ tends to zero exponentially fast [1] (i.e., if there exists $0 < \beta < 1$ and $C > 0$ such that $\text{var}_n(g) \leq C\beta^n$, $n \geq 0$, then there exists $0 < \theta < 1$ and $D > 0$ such that $|\rho(N)| \leq D\theta^N$, $N \geq 0$).

In this paper we shall consider the rate at which $\rho(N)$ tends to zero when $\text{var}_n(g)$ tends to zero at a sub-exponential rate. To help our exposition we shall concentrate on a number of particular cases.

Our first main result is the following.

Theorem 1. *Let $\sigma : X_A \rightarrow X_A$ denote a transitive two-sided subshift of finite type.*

- (1) *Polynomially decay: If $\exists r > 2$, $\exists C > 0$ such that $\text{var}_n(g) \leq C \left(\frac{1}{n^r}\right)$, then $|\rho(N)| = O\left(\frac{1}{N^{r-2-\epsilon}}\right)$ for any $\epsilon > 0$.*
- (2) *Intermediate decay: If $\exists 0 < \beta < 1$, $\exists C > 0$, $\exists \gamma > 1$ such that $\text{var}_n(g) \leq C(\beta^{(\log n)^\gamma})$, then $\exists D > 0$, $\exists 0 < \theta < 1$ such that $|\rho(N)| = D\left(\theta^{(\log N)^{\gamma-\epsilon}}\right)$ for any $\epsilon > 0$.*
- (3) *Stretched exponential decay: If $\exists 0 < \theta < 1$, $\exists C > 0$ such that $\text{var}_n(g) \leq C(\theta^{n^{1/2}})$, then $\exists D > 0$, $\exists 0 < \beta < 1$ such that $|\rho(N)| = D\left(\beta^{N^{1/3}}\right)$ for any $\epsilon > 0$.*

Received by the editors September 22, 1997.

1991 *Mathematics Subject Classification.* Primary 58Fxx.

In the case that $\text{var}_n(g)$ tends to zero exponentially fast it is possible to show that an associated Ruelle-Perron-Frobenius transfer operator has a spectral gap when acting on the Banach space of Hölder functions [1]. This is the usual proof of exponential mixing in this context. In this more general setting we shall use a different approach motivated by a method of Chernov [2] and Liverani [6], for a somewhat different problem. Related results have been obtained independently by Kondah et al. [4] and Fisher and Lopes [9].

Let us consider two simple examples where such conditions arise.

Example 1. Let $X = \prod_{-\infty}^{\infty} \{0, 1\}$ denote the shift space associated to a full shift on two symbols. We can denote, for each $k \geq 1$, the subset

$$[0^k 1] := \{x \in X : x_i = 0, 0 \leq i \leq k - 1, \text{ and } x_k = 1\}$$

(which is both an open and a closed set). Consider a monotone decreasing sequence of real numbers $\alpha_n \rightarrow 0$ and define a continuous map $g : X \rightarrow \mathbb{R}$ by

$$g(x) = \begin{cases} \alpha_k & \text{if } x \in [0^k 1], \\ 0 & \text{if } x = (0, 0, 0, \dots), \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to see that $\text{var}_n(g) = \alpha_n$. Thus, by varying the choice of sequence $\{\alpha_n\}$ we can arrange the different hypotheses on $\text{var}_n(g)$ in the theorem. For example, if we take $\alpha_n = \frac{1}{n^r}$, then the hypothesis of part (1) of Theorem 1 is satisfied.

Example 2. Consider a C^∞ transitive Anosov diffeomorphism $T : M \rightarrow M$ on a compact manifold M . One can associate to this diffeomorphism a subshift of finite type $\sigma : X_A \rightarrow X_A$ and a continuous map $\pi : X_A \rightarrow M$ such that $\pi\sigma = T\pi$. Moreover, $\pi : X_A \rightarrow M$ is Hölder continuous with respect to the metric $d(x, y) = \frac{1}{2^{N(x,y)}}$, where $N(x, y) = \sup\{N \geq 0 : x_i = y_i, i = 0, \dots, N\}$.

If $G : M \rightarrow \mathbb{R}$ is a continuous function such that

$$|G(x) - G(y)| \leq \frac{C}{|\log d_M(x, y)|^\beta},$$

for some $C, \beta > 0$, where d_M is a Riemannian metric on M , then it is easy to see that $\text{var}(G \circ \pi) \leq C_0 \left(\frac{1}{n^\beta}\right)$ and, in particular, there is a unique equilibrium state ν for G . Moreover, $\pi_*(\mu) = \nu$, where μ is an equilibrium state on X_A for $G \circ \pi$.

For $\beta > 2$ we can apply part 1 of the theorem to see that for any Hölder function $F : M \rightarrow \mathbb{R}$ we have

$$\int F \circ T^N F d\nu - \left(\int F d\nu\right)^2 = O\left(\frac{1}{N^{\beta-2-\epsilon}}\right).$$

By way of demonstrating the usefulness of estimates on the rate of mixing we consider the following two results which are consequences of Theorem 1.

Theorem 2 (Central limit theorem). *Assume that $\text{var}_n(g) = O\left(\frac{1}{n^\beta}\right)$ and $f : X_A \rightarrow \mathbb{R}$ is a Hölder continuous function. If μ is the equilibrium state for g , then the limit*

$$\rho^2 = \lim_{n \rightarrow +\infty} \frac{1}{n} \left(\sum_{l=0}^{n-1} f \circ \sigma^l - n \int f d\mu \right)^2$$

exists and

$$\lim_{n \rightarrow +\infty} \mu \left\{ x : \sum_{l=0}^{n-1} f \circ \sigma^l - n \int f d\mu < tn^{1/2} \right\} = \frac{1}{\rho(2\pi)^{1/2}} \int_{-\infty}^t \exp\left(-\frac{x^2}{2\rho^2}\right) dx.$$

The derivation of the central limit from the estimates in Theorem 1 is described in Section 5. (An alternative proof appears in [4, §4]. For g a Hölder continuous potential this result can be found in [1].)

Theorem 3. Assume that $\exists r > 2, \exists C > 0$ such that $\text{var}_n(g) \leq C \left(\frac{1}{n^r}\right)$. For $\epsilon > 0$ there exists $\delta > 0$ such that almost all $x \in X_A^+$ (with respect to (μ))

$$\frac{1}{N} \sum_{i=0}^{N-1} f(T^i x) - \int f d\mu = \begin{cases} O\left(\frac{(\log N)^{1/2}(\log \log N)^{1/2+\delta}}{N^{(r-2-\epsilon)/2}}\right) & \text{for } 2 < r < 3, \\ O\left(\frac{(\log N)^{3/2}(\log \log N)^{1/2+\delta}}{N}\right) & \text{for } r > 3. \end{cases}$$

Theorem 3 follows by combining part 1 of Theorem 1 and Theorem 16 from [4].

We now give a brief summary of the organization of this paper. In Section 1 we recall the basic results on equilibrium states and subshifts of finite type. In Section 2 we recall the basic properties of the transfer operator and its relation to $\rho(N)$.

In Section 3, we prove the corresponding decay results for $\rho(N)$ in the context of one-sided subshifts of finite type. In Section 4 we translate these into the statements in Theorem 1.

In Section 5 we prove the central limit theorem, making essential use of part 1 of Theorem 1.

1. DEFINITIONS

Let A be a $k \times k$ irreducible matrix with entries either 0 or 1. We define a (two-sided) shift space of finite type by

$$X_A = \left\{ x = (x_n)_{n=-\infty}^{+\infty} \in \prod_{n=-\infty}^{\infty} \{1, \dots, k\} : A(x_n, x_{n+1}) = 1 \text{ for } n \in \mathbb{Z} \right\}.$$

This has the Tychonov product topology with sub-basis given by cylinders (of length $n - m$) of the form

$$[z_m, \dots, z_{n-1}] = \{x \in X_A : x_i = z_i \text{ for } i = m, \dots, n - 1\}$$

for $n > m$. We can define a metric by

$$d(x, y) = \sum_{n=-\infty}^{\infty} \frac{(1 - \delta(x_n, y_n))}{2^{|n|}} \text{ where } \delta(i, j) = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j, \end{cases}$$

and $x = (x_n)_{n=-\infty}^{\infty}, y = (y_n)_{n=-\infty}^{\infty}$.

The subshift of finite type $\sigma : X_A \rightarrow X_A$ is the homeomorphism defined by $(\sigma x)_n = x_{n+1}$.

Let $g : X_A \rightarrow \mathbb{R}$ be a continuous function. For each $k \geq 1$ we define

$$\text{var}_k(g) = \sup \{|g(x) - g(y)| : x_i = y_i \text{ for } |i| \leq k\}.$$

It is easy to see from the uniform continuity of g that $\text{var}_k(g)$ is a monotone decreasing sequence. Moreover, if g is a Hölder continuous function satisfying $|g(x) - g(y)| \leq C(d(x, y))^\alpha$, for some $C > 0, 0 < \alpha < 1$, then we have that $\text{var}_k(g) \leq C \left(\frac{1}{2^{\alpha(k-1)}}\right)$.

More generally, we say that a function $g : X_A \rightarrow \mathbb{R}$ has *summable variation* if $\sum_{k=0}^{\infty} \text{var}_k(g) < +\infty$.

Definition. We define the *pressure* $P(g)$ of g by

$$P(g) = \sup \left\{ h(\nu) + \int g d\nu : \nu \text{ is a } \sigma\text{-invariant probability} \right\}$$

where $h(\nu)$ is the entropy of ν .

When g has summable variation, there is a unique measure μ , called the *equilibrium state* of g such that $P(g) = h(\mu) + \int g d\mu$. It is easy to see that if $g = g' + u \circ \sigma - u + C$, for some $u \in C^0(X_A)$ and $C > 0$, then g and g' have the same equilibrium state.

2. THE RUELLE-PERRON-FROBENIUS OPERATOR

The key ingredient in the analysis of the correlation function $\rho(n)$ is the Ruelle-Perron-Frobenius operator. In this section we recall this approach. We define a *one-sided shift space* by

$$X_A^+ = \left\{ x = (x_n)_{n=0}^{+\infty} \in \prod_{n=0}^{\infty} \{1, \dots, k\} : A(x_n, x_{n+1}) = 1 \text{ for } n \geq 0 \right\}.$$

This has the Tychonov product topology with sub-basis given by *cylinders* (of length m) of the form

$$[z_0, \dots, z_{m-1}] = \{x \in X_A^+ : x_i = z_i \text{ for } i = 0, \dots, m-1\}.$$

The *(one-sided) subshift of finite type* $\sigma : X_A \rightarrow X_A$ is the local homeomorphism defined by $(\sigma x)_n = x_{n+1}$. We use the same notation as for a two-sided subshift of finite type, the context allowing the reader to distinguish.

Let $g : X_A \rightarrow \mathbb{R}$ be a continuous function. For each $k \geq 1$ we define

$$\text{var}_k(g) = \sup\{|g(x) - g(y)| : x_i = y_i \text{ for } i = 0, \dots, k-1\}.$$

As in the case of the two-sided subshift, the uniform continuity of g ensures that $\text{var}_k(g)$ is a monotone decreasing sequence. A function $g : X_A^+ \rightarrow \mathbb{R}$ has *summable variation* if $\sum_{k=0}^{\infty} \text{var}_k(g) < +\infty$.

Definition. We define the *pressure* of g by

$$P(g) = \sup \left\{ h(\nu) + \int g d\nu : \mu \text{ is a } \sigma\text{-invariant probability} \right\}$$

where $h(\nu)$ is the entropy of ν .

The definition of $P(g)$ for both two-sided and one-sided shifts are consistent under the natural inclusion $C(X_A^+) \subset C(X_A)$.

When g has summable variation there is unique measure μ , called the *equilibrium state* of g , such that $P(g) = h(\mu) + \int g d\mu$. It is easy to see that if $g = g' + u \circ \sigma - u + C$, for some $u \in C^0(X_A^+)$ and $C > 0$, then g and g' have the same equilibrium state.

We now come to a central definition.

Definition. We define a Ruelle operator $L_g : C(X_A^+) \rightarrow C(X_A^+)$ by

$$L_g h(x) = \sum_{\sigma y = x} e^{g(y)} h(y).$$

A standard approach when g is Hölder is to use spectral theory for L_g to understand the convergence of $\rho(N)$. However, we shall consider a broader class of functions and use an alternative method.

- Proposition 1.** (1) *If g has summable variation, then there exists $h \in C^0(X_A^+)$ with $h > 0$ such that $L_g h(x) = e^{P(g)} h(x)$.*
 (2) *If we let $g_0(x) = g(x) + \log h(x) - \log h(\sigma x)$, then $0 < g_0 < 1$ satisfies $L_{g_0} 1 = 1$ and $L_{g_0}^* \mu = \mu$.*
 (3) *If $C_n(g_0) = \sum_{k=n}^\infty \text{var}_k(g_0)$, then $\text{var}_n(\log h) \leq C_n(g_0)$, for $n \geq 0$.*

This result can be found in [7, Theorem 3.3]. Although the estimate in (3) on $\text{var}_n(\log h)$ doesn't appear explicitly in the statement of the theorem there, it is shown in the proof that:

- (i) there exists a probability measure ν on X_A^+ such that $L_{g_0}^* \nu = e^{P(g)} \nu$ (i.e., $\int L_{g_0} k d\nu = \int k d\nu$ for all $k \in C^0(X_A)$),
- (ii) h lies in the space

$$\Lambda = \{f \in C(X_A^+) : \nu(f) = 1, f > 0 \text{ and } f(x) \leq f(y)e^{C_n(g)}, \\ x_i = y_i \text{ for } i = 1, \dots, n\}.$$

Thus the result follows.

Lemma 1. *If μ is an equilibrium state for g_0 , then there exists $0 < \gamma < 1$ such that*

$$\gamma \leq \frac{\mu[x_0, \dots, x_{n-1}]}{e^{g_0^n(x)}} \leq \frac{1}{\gamma}$$

where $g_0^n(x) = g_0(x) + g_0(\sigma x) + \dots + g_0(\sigma^{n-1} x)$ [1].

We let $\|f\|_1 = \int |f| d\mu$ denote the L^1 -norm on integrable functions on X_A^+ , with respect to the measure μ . The next lemma plays an important role in that it relates $\rho(N)$ to $L_{g_0}^N$.

- Lemma 2.** (1) $\|L_{g_0} k\|_1 \leq \|k\|_1$ for $k \in C^0(X_A^+)$;
 (2) $\rho(N) = \int f(L_{g_0}^N f) d\mu - (\int f d\mu)^2$.

Proof. For the first part, observe that $\|L_{g_0} k\|_1 = \int |L_{g_0} k| d\mu \leq \int L_{g_0} |k| d\mu = \int |k| d\mu$ for any $k \in C^0(X_A)$. For the second part, we have that $\int f \circ \sigma^N \cdot f d\mu = \int L_{g_0}^N (f \circ \sigma^N \cdot f) d\mu = \int f L_{g_0}^N f d\mu$. □

3. CONDITIONAL EXPECTATION

Let $\alpha = \{[i] : i = 1, \dots, k\}$ be the partition by 1-cylinders and denote $\alpha^{(n)} = \bigvee_{i=0}^{n-1} \sigma^{-i} \alpha$. The finite sigma algebra $\alpha^{(n)}$ consists of n -cylinders. We can write the conditional expectation of $f \in L^1(X, \mu)$ corresponding to $\alpha^{(n)}$ by

$$(3.1) \quad E(f|\alpha^{(n)}) = \sum_{C \in \alpha^{(n)}} \left(\frac{\int_C f d\mu}{\mu(C)} \right) \chi_C(x)$$

where $\chi_C(x)$ is the characteristic function for the cylinder C (i.e., $E(f|\alpha^{(n)})$ is the unique function constant on n -cylinders which integrates to the same value as f over sets in $\alpha^{(n)}$).

For $n > 0$ and $m > 0$ let $N = nm$ and we shall consider $\|L_{g_0}^N f\|_1$ using the inequality

$$(3.2) \quad \|L_{g_0}^N f\|_1 \leq \left\| \left(L_{g_0}^N - \left(L_{g_0}^n \circ E(\cdot|\alpha^{(n)}) \right)^m \right) f \right\|_1 + \left\| \left(L_{g_0}^n \circ E(\cdot|\alpha^{(n)}) \right)^m f \right\|_1.$$

The following lemma gives important properties of $E(f|\alpha^{(n)})$. Henceforth, we shall assume for simplicity that $\int f d\mu$.

Lemma 3. (i) We can write $L_{g_0}^n E(f|\alpha^{(n)})(x) = \int K(x, z) f(z) d\mu(z)$ with

$$K(x, z) = \sum_{C \in \alpha^{(n)}} \left(\frac{e^{g_0^n(x_C)}}{\mu(C)} \right) \chi_C(z)$$

where $C = [i_0, \dots, i_{n-1}]$ and $x_C = (i_0, \dots, i_{n-1}, x_0, x_1, \dots)$ (and the summand is zero if $A(i_{n-1}, x_0) = 1$).

(ii) There exists $0 < \gamma < 1$ (independent of n) such that $K(x, z) \geq \gamma$.

(iii) $\|L_{g_0}^n \circ E(\cdot|\alpha^{(n)})\|_1 \leq (1 - \gamma)$ for all $n \geq 1$.

Proof. Part (i) follows from the definitions. For part (ii) we observe that since μ is an equilibrium state for g_0 , we have by Lemma 1 that there exists $0 < \gamma < 1$ such that $\mu[x_0, \dots, x_{n-1}] \leq \frac{e^{g_0^n(x)}}{\gamma}$. Finally, for part (iii) we begin by defining $\Omega_0 := \{x : f(x) \geq 0\}$ and $\Omega_1 := \{x : L_{g_0}^n \circ E(f|\alpha^{(n)})(x) \geq 0\}$ and observing that $\int |L_{g_0}^n \circ E(f|\alpha^{(n)})(x)| d\mu(x) = 2 \int_{\Omega_1} L_{g_0}^n \circ E(f|\alpha^{(n)})(x) d\mu(x)$. We note, using Lemma 3 (i), that for $f \in L^1(X_A^+, \mu)$.

$$\begin{aligned} \|L_{g_0}^n \circ E(f|\alpha^{(n)})\|_1 &= \int |L_{g_0}^n \circ E(f|\alpha^{(n)})(x)| d\mu(x) \\ &= 2 \int_{\Omega_1} L_{g_0}^n \circ E(f|\alpha^{(n)})(x) d\mu(x) \\ &= 2 \int_{\Omega_1} \left(\int f(z) K(x, z) d\mu(z) \right) d\mu(x) \\ &= 2 \int_{\Omega_1} \left(\int f(z) K(x, z) d\mu(z) \right) d\mu(x) - 2\mu(\Omega_1)\gamma \int f(x) d\mu(x) \end{aligned}$$

since $\int f(x) d\mu(x) = 0$. Therefore,

$$\begin{aligned} \|L_{g_0}^n \circ E(f|\alpha^{(n)})\|_1 &= 2 \int \left(\int_{\Omega_1} f(z) (K(x, z) - \gamma) d\mu(x) \right) d\mu(z) \\ &= 2 \int f(z) \left(\int_{\Omega_1} (K(x, z) - \gamma) d\mu(x) \right) d\mu(z) \\ (3.3) \quad &\leq 2 \int f(z) \left(\int (K(x, z) - \gamma) d\mu(x) \right) d\mu(z) \\ &\leq 2 \int_{\Omega_0} f(z) \left(\int (K(x, z) - \gamma) d\mu(x) \right) d\mu(z) \\ &= 2 \int_{\Omega_0} f(z) (1 - \gamma) d\mu(z), \end{aligned}$$

where we have used that

$$\begin{aligned} \int K(x, z)d\mu(x) &= \sum_C \left(\int L_{g_0}^n \chi_C(x)d\mu(x) \right) \frac{\chi_C(z)}{\mu(C)} \\ &= \sum_C \left(\int \chi_C(x)d\mu(x) \right) \frac{\chi_C(z)}{\mu(C)} = 1. \end{aligned}$$

Finally, we can bound

$$\begin{aligned} \|L_{g_0}^n \circ E(f|\alpha^{(n)})\|_1 &\leq 2(1 - \gamma) \int_{\Omega_0} f(z)d\mu(z) \\ &= (1 - \gamma) \int |f(z)|d\mu(z) \\ &= (1 - \gamma)\|f\|_1. \end{aligned}$$

Corollary 3.1. $\| (L_{g_0}^n \circ E(\cdot|\alpha^{(n)}))^m \|_1 \leq (1 - \gamma)^m$.

We continue with the estimates needed with the following lemma.

Lemma 4. (i) For all $n \geq 0$ we have that $\|f - E(f|\alpha^{(n)})\|_1 \leq \text{var}_n(f)$.

(ii) If $\text{var}_n(f) < +\infty, \forall n \geq 0$, then

$$\text{var}_n(L_{g_0}^l f) \leq C_n(g)\|f\|_\infty + \text{var}_{n+l}(f), \text{ for all } l \geq 0,$$

$$\text{where } C_n(g) = \sum_{k=n}^\infty \text{var}_k(g).$$

Proof. Part (i) is the easy consequence of the identity (3.1) for $E(f|\alpha^{(n)})$. In particular, $\|f - E(f|\alpha^{(n)})\|_1 \leq |E(f|\alpha^{(n)}) - f|_\infty \leq \text{var}_n(f)$.

Part (ii) is shown in [7, pp. 379-380].

Using Corollary 3.1 we can estimate the second term in (3.2) by

$$(3.4) \quad \| (L_{g_0}^n \circ E(\cdot|\alpha^{(n)}))^m f \|_1 \leq (1 - \gamma)^m \|f\|_1$$

We can now estimate the first term on the right-hand side of (3.2) by

$$\begin{aligned} (3.5) \quad &\|L_{g_0}^N f - (L_{g_0}^n \circ E(f|\alpha^{(n)}))^m\|_1 = \|(L_{g_0}^n)^m f - (L_{g_0}^n \circ E(f|\alpha^{(n)}))^m\|_1 \\ &\leq \sum_{k=0}^{m-1} \| (L_{g_0}^n \circ E(\cdot|\alpha^{(n)}))^k (L_{g_0}^n \circ E(\cdot|\alpha^{(n)}) - L_{g_0}^n) (L_{g_0}^n)^{m-k-1} f \|_1 \\ &\leq \sum_{k=0}^{m-1} \| (L_{g_0}^n \circ E(\cdot|\alpha^{(n)}))^k \|_1 \|L_{g_0}^n\|_1 \| (1 - E(\cdot|\alpha^{(n)})) (L_{g_0}^{m-k-1} f) \|_1 \\ &\leq \sum_{k=0}^{m-1} (1 - \gamma)^k \text{var}_n(L_{g_0}^{m-k-1} f) \\ &\leq \sum_{k=0}^{m-1} (1 - \gamma)^k (C_n(g)\|f\|_\infty + \text{var}_{n(m-k)}(f)) \end{aligned}$$

using Lemma 3 (ii) and Lemma 4, parts (i) and (ii). We can now take the trivial bound

$$(3.6) \quad \sum_{k=0}^{m-1} (1 - \gamma)^k (C_n(g)\|f\|_\infty + \text{var}_{n(m-k)}(f)) = O(C_n(g))$$

since the Hölder continuity of f ensures $\text{var}_n(f) = O(\theta^n)$. (We use the standard notation $g(n) = O(f(n))$ if $\overline{\lim}_{n \rightarrow +\infty} g(n)/f(n) < +\infty$.)

Comparing (3.2) with (3.4) and (3.5) we have that

$$(3.7) \quad \|L_{g_0}^N f\|_1 = O(\max\{C_n(g), (1 - \gamma)^m\}).$$

4. DECAY OF CORRELATION FOR ONE-SIDED SHIFTS

In this section we prove the analogue of Theorem 1 for one-sided subshifts. In Section 5 we shall relate this to two-sided subshifts.

(1) Polynomial decay of variances. If $\text{var}_n(g) = O(\frac{1}{n^r})$, then

$$C_n(g) = \sum_{l=n}^{\infty} \text{var}_l(g) = O\left(\int_{(n-1)}^{\infty} \frac{1}{x^r} dx\right) = O\left(\frac{1}{n^{r-1}}\right).$$

Given $N > 0$, we choose $n = [N^\beta]$ and $m = [N^{1-\beta}]$. Clearly, $n \cdot m \leq N$, but then using (3.7) we can write

$$\begin{aligned} \rho(N) &\leq \|L_{g_0}^N f\|_1 \leq \|L_{g_0}^{nm} f\|_1 \\ &= O\left(\max\left\{\frac{1}{N^{\beta(r-1)}}, (1 - \gamma)^{N^{1-\beta}}\right\}\right) \\ &= O\left(\frac{1}{N^{\beta(r-1)}}\right). \end{aligned}$$

By taking β sufficiently close to 1 we see that $\rho(N) = O(\frac{1}{N^{r-1-\epsilon}})$ for any $\epsilon > 0$.

(2) Intermediate decay of variances. If $\text{var}_n(g) = O(\theta^{(\log n)^\gamma})$, then we can estimate

$$\begin{aligned} C_n(g) &= \sum_{k=n}^{\infty} \text{var}_k(g) = O\left(\int_{(n-1)}^{\infty} \theta^{(\log x)^\gamma} dx\right) \\ &= O\left(\int_{(n-1)}^{\infty} \left(\frac{x}{(\log x)^{\gamma-1}}\right) \theta^{(\log x)^{\gamma-\epsilon}} dx\right) \\ &= O\left(\theta^{(\log n)^{\gamma-\epsilon}}\right). \end{aligned}$$

Given $N > 0$, we can choose $n = [N^{1/2}]$ and $m = [N^{1/2}]$. Thus

$$\begin{aligned} \rho(N) &= \|L_{g_0}^N f\|_1 \leq \|L_{g_0}^{nm} f\|_1 \\ &= O\left(\max\left\{\theta^{(\log N)^{\gamma-\epsilon}}, (1 - \gamma)^{N^{1/2}}\right\}\right) \\ &= \theta^{(\log N)^{\gamma-\epsilon}}, \end{aligned}$$

for each $\epsilon > 0$.

(3) Stretched exponential decay of variances. If $\text{var}_n(g) = O(\theta^{n^{1/2}})$, then we can estimate

$$C_n(g) = \sum_{k=n}^{\infty} \text{var}_k(g) = O\left(\int_{(n-1)}^{\infty} \theta^{x^{1/2}} dx\right) = O\left(n^{1/2} \theta^{n^{1/2}}\right).$$

We can choose $n = \lfloor N^{2/3} \rfloor$ and $m = \lfloor N^{1/3} \rfloor$. Thus

$$\rho(N) = 0 \left(\theta^{N^{1/3}} N^{1/3}, (1 - \gamma)^{N^{1/3}} \right) = O(\beta^{N^{1/3}})$$

for any $\beta > \max\{\theta, (1 - \gamma)\}$.

Proposition 2. For any $\epsilon > 0$:

- (1) If $\text{var}_n(g) = O\left(\frac{1}{n^r}\right)$ ($r \geq 2$), then $\rho(N) = O\left(\frac{1}{N^{r-1-\epsilon}}\right)$.
- (2) If $\text{var}_n(g) = O\left(\theta^{(\log n)^\gamma}\right)$, then $\rho(N) = O\left(\rho^{(\log N)^{\gamma-\epsilon}}\right)$.
- (3) If $\text{var}_n(g) = O(\theta^{n^{1/2}})$, then $\rho(N) = O\left(\beta^{N^{1/3-\epsilon}}\right)$.

Remark. We have assumed for convenience that $f : X_A^+ \rightarrow \mathbb{R}$ is Hölder continuous. If we relax this condition so that $\text{var}_n(f) \leq C_n(g)$ remains valid, then we see that (3.6) remains valid, and so does Proposition 2.

5. DECAY OF CORRELATIONS FOR TWO-SIDED SUBSHIFTS

We want to convert the estimate on one-sided subshifts of finite type to two-sided subshifts of finite type.

Given a function $g : X_A \rightarrow \mathbb{R}$ of bounded variation, we can write $g = \sum_{l=0}^\infty g_l$ where

- (a) $g_l(x)$ depends only on the co-ordinates x_{-l}, \dots, x_l ;
- (b) $\|g_l\|_\infty \leq \text{var}_l(g)$ for $l \geq 0$.

We let $g' = \sum_{l=0}^\infty g_l \circ \sigma^{-l}$. In particular, g' is a function depending only on co-ordinates $x_n, n \geq 0$, and thus interpreted as a function on X_A^+ . Moreover, since we can write

$$(5.1) \quad g - g' = \left(\sum_{l=0}^\infty \sum_{r=0}^{l-1} g_l \circ \sigma^{r-l} \right) - \left(\sum_{l=0}^\infty \sum_{r=0}^{l-1} g_l \circ \sigma^{r-l} \right) \circ \sigma,$$

we see that g' and g differ by a coboundary, and thus share the same equilibrium state (cf. Section 2). Finally, $\text{var}_l(g') \leq \text{var}_{\lfloor l/2 \rfloor}(g)$.

If $\text{var}_n(f) = O(\theta^n)$ for some $0 < \theta < 1$, then we can choose f_l depending only on the co-ordinates x_{-l}, \dots, x_l such that $\|f - f_l\|_\infty = O(\theta^l)$.

We can consider $f_{\lfloor n/3 \rfloor} \circ \sigma^{-\lfloor n/3 \rfloor}$ as an element of $C(X_A^+)$. Moreover, this function is Hölder continuous and applying the results from Section 4 we can write

$$\begin{aligned} \int f_{\lfloor n/3 \rfloor} \circ \sigma^n \cdot f_{\lfloor n/3 \rfloor} d\mu &= \int \left(f_{\lfloor n/3 \rfloor} \circ \sigma^{-\lfloor n/3 \rfloor} \right) \circ \sigma^n \cdot \left(f_{\lfloor n/3 \rfloor} \circ \sigma^{-\lfloor n/3 \rfloor} \right) d\mu \\ &= \int L_{g_0}^n \left(f_{\lfloor n/3 \rfloor} \circ \sigma^{-\lfloor n/3 \rfloor} \right) \cdot \left(f_{\lfloor n/3 \rfloor} \circ \sigma^{-\lfloor n/3 \rfloor} \right) d\mu \\ &= \left(\int f_{\lfloor n/3 \rfloor} d\mu \right)^2 + O\left(\max\{(1 - \gamma)^n C_n(g), (1 - \gamma)^m\}\right). \end{aligned}$$

We can estimate,

$$(5.2) \quad \left| \left(\int f_{\lfloor n/3 \rfloor} d\mu \right)^2 - \left(\int f d\mu \right)^2 \right| \leq 2\|f\|_\infty \|f - f_{\lfloor n/3 \rfloor}\|_\infty = O(\theta^{\lfloor n/3 \rfloor})$$

and

$$(5.3) \quad \left| \int f_{[n/3]} \circ \sigma^n \cdot f_{[n/3]} d\mu - \int f \circ \sigma^n \cdot f d\mu \right| \leq O(\theta^n).$$

The following relates the properties of g and g' .

- Lemma 5.** (i) If $\text{var}_n(g) = O\left(\frac{1}{n^r}\right)$, $r \geq 2$, then $\text{var}_n(g') = O\left(\frac{1}{n^{r-1}}\right)$.
(ii) If $\text{var}_n(g) = O\left(\beta^{(\log n)^\gamma}\right)$, then $\text{var}_n(g') = O\left(\beta^{(\log n)^\gamma}\right)$.
(iii) If $\text{var}_n(g) = O\left(\theta^{n^{1/2}}\right)$, then $\text{var}_n(g') = O\left(\theta^{n^{1/2}}\right)$.

Proof. We can estimate $\text{var}_n(g_1) \leq \sum_{l=2n}^{+\infty} \text{var}_l(g)$. In case (i) we see that

$$\text{var}_n(g_1) = O\left(\sum_{l=2n}^{+\infty} 1/l^r\right) = O\left(\frac{1}{n^{r-1}}\right).$$

The estimates in cases (ii) and (iii) follow similarly.

From equations (5.1), (5.2) and (5.3) we see that Theorem 1 follows from Proposition 2.

6. CENTRAL LIMIT THEOREMS

In this section we present the proof of Theorem 2. Consider the equilibrium state μ associated with the potential g with summable variation such that $\text{var}_n(g) = O\left(\frac{1}{n^3}\right)$. Let $f : X_A \rightarrow \mathbb{R}$ be a Hölder continuous function, then we can consider for each $t > 0$ the values

$$S_n(t) := \mu \left\{ x \in X_A : \frac{\sum_{i=0}^n f(T^i x)}{n} - \int f d\mu < t \cdot n^{1/2} \right\}.$$

In [3] Gordin gives sufficient conditions for

$$\lim_{n \rightarrow +\infty} S_n(t) = \frac{1}{\rho(2\pi)^{1/2}} \int_{-\infty}^t e^{-x^2/(2\rho^2)} dx,$$

which we shall describe below.

Let \mathcal{B}_0 be the sigma algebra of sets in X_A which correspond to the Borel sigma algebra for X_A^+ . We denote $\mathcal{B}_n = \sigma^{-n}\mathcal{B}_0$ for $n \in \mathbb{Z}$.

If we write $H_n = L^2(X_A, \mathcal{B}_n, \mu)$, then for $n \geq 0$ we have that the sigma-algebras are nested $\mathcal{B}_n \subset \mathcal{B}_{n+1}$ and so $H_n \subset H_{n+1}$. For $n \geq 0$ the orthogonal projection from H_0 to H_n takes the form $E(\cdot|\mathcal{B}_n) : H_0 \rightarrow H_n$.

Gordin's condition involves showing that the following series converge:

$$(6.1) \quad \sum_{n=0}^{\infty} \|E(f|\mathcal{B}_n)\|_2 + \sum_{n=0}^{\infty} \|f - E(f|\mathcal{B}_{-n})\|_2.$$

We first observe that

$$(6.2) \quad \begin{aligned} \sum_{n=0}^{\infty} \|E(f|\mathcal{B}_n)\|_2 &= \sum_{n=0}^{\infty} \|(L_{g_0}^n f) \circ \sigma^n\|_2 = \sum_{n=0}^{\infty} \|(L_{g_0}^n f)\|_2 \\ &= \|f\|_{\infty} \sum_{n=0}^{\infty} \|(L_{g_0}^n f)\|_1 = O\left(\sum_{n=1}^{\infty} \frac{1}{n^{r-2-\epsilon}}\right) < +\infty. \end{aligned}$$

We can also estimate

$$(6.3) \quad \sum_{n=0}^{\infty} \|f - E(f|\mathcal{B}_{-n})\|_2 \leq \sum_{n=0}^{\infty} \text{var}_n(f) = O\left(\sum_{n=0}^{\infty} \frac{1}{n^r}\right) < +\infty.$$

Comparing (6.2) and (6.3) shows that the sum in (6.1) is finite. Applying [3, Theorem 2] (with $\delta = 0$) the result follows.

REFERENCES

1. R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Lecture Notes in Mathematics 470, Springer, Berlin, 1975. MR **56**:1364
2. N. Chernov, *Markov approximations and decay of correlations for Anosov flows*, Ann. of Math. (2) **147** (1998), 269–324. MR **99d**:58101
3. M. Gordin, *The central limit theorem for stationary processes*, Soviet Math. Dokl. **10** (1969), 1174–1176. MR **40**:5012
4. A. Kachurovskii, *Rate of convergence in ergodic theorems*, Russian Math. Surveys **51**:4 (1996), 73–124. CMP 97:05
5. A. Kondah, V. Maume and B. Schmitt, *Vitesse de convergence vers l'état d'équilibre pour des dynamiques markoviennes non hölderiennes*, Ann. Inst. H. Poincaré Probab. Statist. **33** (1997), 675–695. MR **99e**:28032
6. C. Liverani, *Flows, random perturbations and the rate of mixing*, Ergod. Th. and Dynam. Sys., to appear.
7. P. Walters, *Ruelle's operator theorem and g-measures*, Trans. Amer. Math. Soc. **214** (1975), 375–387. MR **54**:515
8. P. Walters, *An Introduction to Ergodic Theory*, Springer, Berlin, 1982. MR **84e**:28017
9. A. Fisher and A. Lopes, *Polynomial decay of correlation and the central limit theorem for the equilibrium state of a non-Hölder potential*, Preprint.

DEPARTMENT OF MATHEMATICS, MANCHESTER UNIVERSITY, OXFORD ROAD, MANCHESTER M13 9PL, ENGLAND

E-mail address: mp@ma.man.ac.uk