LOCAL MIXING ON ABELIAN COVERS OF HYPERBOLIC SURFACES WITH CUSPS

WENYU PAN

ABSTRACT. We prove the local mixing theorem for geodesic flows on abelian covers of finite volume hyperbolic surfaces with cusps, which is a continuation of the work [17]. We also describe applications to counting problems and the prime geodesic theorem.

1. Introduction

1.1. Setting and main results. Let $M$ be a hyperbolic surface. We can present $M$ as the quotient $\Gamma \backslash \mathbb{H}^2$ of the hyperbolic 2-space for some torsion-free discrete subgroup $\Gamma$ in $\text{PSL}_2(\mathbb{R})$. The unit tangent bundle $T^1(M)$ is isomorphic to $\Gamma \backslash G$ and the geodesic flow on $T^1(M)$ corresponds to the right translation action of $a_t = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}$ on $\Gamma \backslash G$.

Set $m_{\text{Haar}}^\Gamma$ to be the $G$-invariant measure on $\Gamma \backslash G$, which equals the hyperbolic volume measure on $T^1(M)$ when identifying $\Gamma \backslash G$ with $T^1(M)$. When there is no ambiguity about the group $\Gamma$, we write $m_{\text{Haar}}$ for simplicity. For any $\psi_1, \psi_2 \in C_c(\Gamma \backslash G)$, consider the matrix coefficient

$$\langle a_t \cdot \psi_1, \psi_2 \rangle = \int_{\Gamma \backslash G} \psi_1(xa_t)\psi_2(x)m_{\text{Haar}}(x).$$

One of the central questions in homogeneous dynamics is what we can say about the asymptotic behavior of $\langle a_t \cdot \psi_1, \psi_2 \rangle$ as $t \to \infty$. When $m_{\text{Haar}}(\Gamma \backslash G) < \infty$, a classical result by Howe-Moore [11] states

$$\lim_{t \to \infty} \langle a_t \cdot \psi_1, \psi_2 \rangle = \frac{1}{m_{\text{Haar}}(\Gamma \backslash G)}m_{\text{Haar}}(\psi_1)m_{\text{Haar}}(\psi_2).$$

So the geodesic flow on $T^1(\Gamma \backslash \mathbb{H}^2)$ satisfies the strong mixing property. When $m_{\text{Haar}}(\Gamma \backslash G) = \infty$, Howe-Moore’s result states

$$\lim_{t \to \infty} \langle a_t \cdot \psi_1, \psi_2 \rangle = 0.$$

In view of these, the question of understanding $\langle a_t \cdot \psi_1, \psi_2 \rangle$ in infinite volume setting can be formulated more precisely as whether there exists a renormalization function $\alpha : \mathbb{R} \to \mathbb{R}_{>0}$ such that for any $\psi_1, \psi_2$, we have

$$\alpha(t)\langle a_t \cdot \psi_1, \psi_2 \rangle \to \text{non-trivial as } t \to \infty.$$

If such renormalization function exists, the $a_t$-action on $\Gamma \backslash G$ (or the geodesic flow on $T^1(\Gamma \backslash \mathbb{H}^2)$) is said to satisfy the local mixing property, a property introduced in [17] to substitute the strong mixing property in infinite volume setting.
In this paper, we investigate the matrix coefficients/ local mixing property of abelian covers of finite volume hyperbolic surfaces with cusps. It is a follow-up work of [17] and the extensive paper by Ledrappier and Sarig [12] on this kind of surfaces has laid a solid foundation for us. Throughout the paper, set $\Gamma_0$ to be a torsion-free non-uniform lattice in $G$ and

$$\Gamma \triangleleft \Gamma_0$$

to be a normal subgroup with $\mathbb{Z}^d$-quotient. Then $M = \Gamma \backslash \mathbb{H}^2$ is a regular cover of $M_0 = \Gamma_0 \backslash \mathbb{H}^2$ whose group of deck transformations is isomorphic to $\mathbb{Z}^d$. Our main result is as follows:

**Theorem 1.1.** For any $\psi_1, \psi_2 \in C_c(\Gamma \backslash G)$,

$$\lim_{t \to +\infty} t^{p+h/2} \int_{\Gamma \backslash G} \psi_1(xa_t)\psi_2(x)dm^{\text{Haar}}(x) = cm^{\text{Haar}}(\psi_1)m^{\text{Haar}}(\psi_2),$$

where $p, h \in \mathbb{N}$ and $c > 0$ are some constants given in (1.5) and (1.7) and $p + h = d$.

**Example 1.2.** Let $\Gamma(2) = \{(a b \ c d) \in \text{PSL}_2(\mathbb{Z}) : (a b \ c d) \equiv (1 0 \ 0 1) (\text{mod } 2)\}$, which is a principle congruent subgroup of the modular group PSL$_2(\mathbb{Z})$. It is a free group generated by $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$. The quotient $\Gamma(2) \backslash \mathbb{H}^2$ is the thrice punctured sphere equipped with the hyperbolic metric. Let $\Gamma \backslash \mathbb{H}^2$ be the homology cover of $\Gamma(2) \backslash \mathbb{H}^2$, that is to say, the group of deck transformations is isomorphic to $H_1(\Gamma(2) \backslash \mathbb{H}^2; \mathbb{Z})$. As $H_1(\Gamma(2) \backslash \mathbb{H}^2; \mathbb{Z}) \cong \mathbb{Z}^2$, $\Gamma \backslash \mathbb{H}^2$ is a $\mathbb{Z}^2$-cover of $\Gamma(2) \backslash \mathbb{H}^2$. Applying Theorem 1.1 to $\Gamma \backslash G$, we have $p = 2, h = 0$ and

$$\lim_{t \to \infty} t^{2} \int_{\Gamma \backslash G} \psi_1(xa_t)\psi_2(x)dm^{\text{Haar}}(x) = cm^{\text{Haar}}(\psi_1)m^{\text{Haar}}(\psi_2)$$

for any $\psi_1, \psi_2 \in C_c(\Gamma \backslash G)$.

In finite volume case, strong mixing property of geodesic flow and its effective refinement have been successfully applied to address some counting and equidistribution problems [13, 8, 7]. In abelian cover case, we obtain similar applications using local mixing property.

1.2. **Orbit counting.**

**Theorem 1.3.** For any $x, y \in \mathbb{H}^2$, we have

$$\#\{\gamma \in \Gamma : d(x, \gamma y) < T\} \sim c\frac{e^{T}}{T^{p+h/2}}$$

With Theorem 1.1 available, this theorem can be shown using the argument in [13] Theorem 7.16 (see also [15] Section 7 and [16] Section 6]) by setting the group $H$ in their statement to be $\text{SO}_2(\mathbb{R})/\{\pm I\}$. When $\Gamma_0$ is a uniform lattice, an analogous result is due to Pollicott-Sharp [19] (see also [17]).

\footnote{We write $f(T) \sim g(T)$ if $\lim_{T \to \infty} f(T)/g(T) = 1.$}
1.3. Prime geodesic theorem. For $T > 0$, let $\mathcal{P}_T$ be the collection of all primitive closed geodesics in $T^1(M)$ of length at most $T$. Note that $\# \mathcal{P}_T = \infty$ if $d \geq 1$.

**Theorem 1.4.** Let $\Omega \subset T^1(M)$ be a compact subset with Haar-negligible boundary. Then as $T \to \infty$,

$$\# \{ C \in \mathcal{P}_T : C \cap \Omega \neq \emptyset \} \sim \frac{ce^T}{T^{p+h/2+1}} m_{\text{Haar}}(\Omega).$$

Given Theorem [14, Theorem 5.1] (see also [17, Theorem 7.9]) applies in the same way. Another formulation of the prime geodesic theorem would be studying the distribution of closed geodesics in $T^1(M)$ satisfying some homological constraints. An explicit main term in our setting was obtained by Epstein [6]. By applying Theorem 1.4 to a proper covering space $T^1(M)$ of $T^1(M_0)$, we recover Epstein’s result.

1.4. Constants in the main theorem. Let $\mathcal{H} \subset H^1(M_0; \mathbb{R})$ be the linear subspace of cohomology classes which vanish on projections of cycles in $H_1(M_0; \mathbb{R})$ to $H_1(M; \mathbb{R})$. Since $M$ is a $\mathbb{Z}^d$-cover of $M_0$, the dimension of $\mathcal{H}$ is $d$.

We describe a basis for $\mathcal{H}$. Given a loop $c$ in $M_0$, let $\tilde{c}$ be a lift of $c$ in $M$ and $g_c \in \Gamma_0$ the isometry mapping the beginning of $\tilde{c}$ to its endpoint. This defines a map

$$\text{Frob} : \{\text{loops in } M_0\} \to \Gamma \backslash \Gamma_0, \ c \mapsto \Gamma g_c.$$  

Observe that $\text{Frob}(\cdot)$ depends only on the homotopy classes. So $\text{Frob}(\cdot)$ is a homomorphism from $\pi_1(M_0)$ to $\Gamma \backslash \Gamma_0$. Since $\Gamma \backslash \Gamma_0$ is abelian, $\text{Frob}(\cdot)$ can be regarded as a homomorphism from $H_1(M_0; \mathbb{Z})$ to $\Gamma \backslash \Gamma_0$. Using the isomorphism $\Gamma \backslash \Gamma_0 \cong \mathbb{Z}^d$, the maps

$$[c] \in H_1(M_0; \mathbb{Z}) \mapsto \langle \text{Frob}([c]), e_i \rangle \ (\{e_i\} \text{ is the standard basis in } \mathbb{R}^d)$$

yield $d$ linearly independent cohomology classes in $\mathcal{H}$, which form a basis of $\mathcal{H}$.

We represent the elements in $\mathcal{H}$ by real harmonic 1-forms with at most simple poles at the cusps (this is possible, see [10], for example). The residue of a 1-form at a cusp is the integral of that form on a loop which is homotopic to the cusp.

Decompose $\mathcal{H}$ into a direct sum

$$\mathcal{H} = \mathcal{H}_p \oplus \mathcal{H}_h,$$

where

$$\mathcal{H}_h := \{ w \in \mathcal{H} : \text{the residues of } w \text{ at the cusps are all zero} \}$$

$$\mathcal{H}_p := \mathcal{H} / \mathcal{H}_h.$$

The constants $p$ and $h$ in Theorem 1.1 are given by

$$p = \dim \mathcal{H}_p \text{ and } h = \dim \mathcal{H}_h.$$  

Set $m_{\text{Haar}}^0$ to be the measure on $M_0$ induced by $m_{\text{Haar}}^0$ and

$$m_0 = m_{\text{Haar}}^0(M_0).$$

Assume $M_0$ is a genus $g$ surface with $t+1$ cusps. For $w \in \mathcal{H}$, denote by $\lambda_1(w), \ldots, \lambda_{t+1}(w)$ the residues of $w$ at $t+1$ cusps of $M_0$. Denote by $\|\cdot\|$ the norm in the cotangent
bundle. Equip $\mathcal{H}_p, \mathcal{H}_h$ with the following norms

$$
\|w\|_p := \frac{1}{m_0} \sum_{j=1}^{t+1} |\lambda_j(w)| \quad (w \in \mathcal{H}_p),
$$

$$
\|w\|_h := \left( \frac{1}{m_0} \int_{M_0} \|w\|^2 dm_{\text{Haar}}^{\Gamma_0} \right)^{1/2} \quad (w \in \mathcal{H}_h).
$$

We identify $\mathcal{H}$ with $\mathbb{R}^d$ using the basis for $\mathcal{H}$ described above. Let $\mathbb{E}_p, \mathbb{E}_h$ be the linear subspaces in $\mathbb{R}^d$ corresponding to $\mathcal{H}_p$ and $\mathcal{H}_h$ respectively under the this identification. The norm on $\mathcal{H}_p$ (resp. $\mathcal{H}_h$) induces a norm on $\mathbb{E}_p$ (resp. $\mathbb{E}_h$), which we still denote by $\|\cdot\|_p$ (resp. $\|\cdot\|_h$). The constant $c$ in Theorem 1.1 is given by

$$
(1.7) \quad c = \frac{1}{(2\pi)^d \cdot m_0} \int_{\mathbb{E}_p} e^{-\|x\|_p} dx \int_{\mathbb{E}_h} e^{-\|y\|_h^2} dy.
$$

When $M$ is the homology cover of $M_0$, we have $p = t$, $h = d - p = 2g$ and $c$ can be explicitly expressed in terms of $g$ and $p$ (see [6] for an explicit formula).

1.5. **Heuristic proof for the renormalization function in Theorem [1.1]**. We consider the special case: let $\Gamma$ be the normal subgroup of $\Gamma(2)$ such that $\Gamma$ is the kernel of the homomorphism $\varphi : \Gamma(2) \to \mathbb{Z}$ mapping $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ to 1 and $\begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}$ to 0. The fundamental domain of $\Gamma$ in $\mathbb{H}^2$ is the following shaded region (Figure 1). Let $A$ and $B$ be two flow boxes in $T^1(\Gamma(2) \backslash \mathbb{H}^2)$ (Figure 1). Assume that the dimensions of $A$ along the stable direction, unstable direction and flow direction are $s_A, u_A$ and $\delta_A$ respectively. Assume the dimensions of $B$ are $s_B, u_B$ and $\delta_B$. Suppose $\delta_B \ll \delta_A$.

![Figure 1. Fundamental domains of $\Gamma$ and $\Gamma(2)$](image)

Observe that $A \cap B_{at}$ consists of finitely many thin cubes (see Figure 2). The dimensions of a thin cube are roughly $e^{-t} s_B, u_A$ and $\delta_B$. Hence we have

$$
\mathcal{M}_{\Gamma(2)}^{\text{Haar}}(A \cap B_{at}) = \int \chi_A(xa_t) \chi_B(x) dm_{\Gamma(2)}^{\text{Haar}}(x) \approx e^{-t} s_B u_A \delta_B \#(A \cap B_{at}).
$$

As the geodesic flow on $T^1(\Gamma(2) \backslash \mathbb{H}^2)$ is strong mixing, as $t \to \infty$, the above yields

$$
\mathcal{M}_{\Gamma(2)}^{\text{Haar}}(A) \mathcal{M}_{\Gamma(2)}^{\text{Haar}}(B) / \mathcal{M}_{\Gamma(2)}^{\text{Haar}}(\Gamma(2) \backslash G) \approx e^{-t} s_B u_A \delta_B \#(A \cap B_{at}).
$$
Let $\tilde{A}$ and $\tilde{B}$ be one of the preimages of $A$ and $B$ respectively in $T^1(\Gamma \backslash \mathbb{H}^2)$. Similarly, we have

$$m^\text{Haar}_\Gamma(\tilde{A} \cap \tilde{B}) = \int \chi_{\tilde{A}}(x)a_t)\chi_{\tilde{B}}(x)dm^\text{Haar}_\Gamma(x) \approx e^{-ts_B}u_A\delta_B\#(\tilde{A} \cap \tilde{B}).$$

We compare $\#(A \cap B_{a_t})$ with $\#(\tilde{A} \cap \tilde{B}_{a_t})$. The key is to see where $\tilde{B}_{a_t}$ lies relative to $\tilde{B}$. Note that $\Gamma \backslash \mathbb{H}^2$ is obtained from $\Gamma(2) \backslash \mathbb{H}^2$ by cutting a geodesic winding around a cusp of $\Gamma(2) \backslash \mathbb{H}^2$. The position of $\tilde{B}_{a_t}$ is related to the study of winding process of a closed 1-form supported on a neighborhood of a cusp of $\Gamma(2) \backslash \mathbb{H}^2$ [10].

Loosely speaking, the result states that “most” of $\tilde{B}_{a_t}$ lies within distance $ct$ from $\tilde{B}$ for some constant $c > 0$ [10]. So we expect $\#(A \cap B_{a_t})/\#(\tilde{A} \cap \tilde{B}_{a_t}) \approx ct$. Based on these, a candidate for the renormalization function for the matrix coefficients of $\Gamma \backslash G$ is (a multiple of) $t$.

1.6. About the proof of Theorem 1.1. Theorem 1.1 is proved following the strategy developed in [17], which is inspired by the work of Guivarc’h and Hardy [9]. We model the geodesic flow on $T^1(M)$ by the suspension flow on the suspension space

$$\Sigma^{f,\tau} := \Sigma \times \mathbb{Z}^d \times \mathbb{R}/\sim,$$

where $(\Sigma, \sigma)$ is a topologically mixing two-sided Markov shift of countable states and the equivalence relation is defined via the left shift map $\sigma$, the first return time map $\tau : \Sigma \to \mathbb{R}$ and the $\mathbb{Z}^d$-coordinate map $f : \Sigma \to \mathbb{Z}^d$ (Lemma 2.2). Using this model, the matrix coefficients of the geodesic flow on $T^1(M)$ can be understood through studying the matrix coefficients of the suspension flow on $\Sigma^{f,\tau}$. Applying the unfolding technique and the Fourier transformation to the matrix coefficients of $\Sigma^{f,\tau}$, it turns out that the key is to estimate the asymptotic behavior of a symbolic sum

$$\sum_{n=0}^{\infty} \sum_{\sigma^n y = x} e^{-r_n(y)(\Phi \cdot \psi)(y)}\delta(x)(f_n(y))u(r_n(y) - t)$$

as $t \to \infty$ (see Section 3.1 for meaning of the notations). The estimate is obtained using two-parameter twisted transfer operator. This is the most computation-heavy part in the whole argument and it is essentially done in the work of Ledrappier and Sarig [12]. It is worth to point out that the presence of cusps causes the loss of the regularity of the transfer operator so the computation is much more involved compared to [17].
1.7. On the generalization of Theorem 1.1. Let $\mathbb{H}^n$ be the hyperbolic $n$-space with $n \geq 1$. Let $\Gamma_0 < \text{Isom}_+ (\mathbb{H}^n)$ be a Schottky group and $\Gamma$ be a normal subgroup of $\Gamma_0$ satisfying $\Gamma \backslash \Gamma_0 \cong \mathbb{Z}^d$. We expect that the geodesic flow on $T^1 (\Gamma \backslash \mathbb{H}^n)$ can be shown to satisfy the local mixing property using the argument of Theorem 1.1 and the knowledge of the corresponding twisted transfer operator obtained by Babillot and Peigné [5].

1.8. Acknowledgement. I would like to thank my advisor Hee Oh, who was a source of inspiration and encouragement throughout.

2. Preparations: coding and transfer operator

2.1. Symbolic model for the geodesic flow. Let $\Omega_0 \subset T^1 (M_0)$ be the collection of unit tangent vectors which do not escape to infinity under the geodesic flow. Classically, the geodesic flow $G^t : \Omega_0 \to \Omega_0$ can be coded using the geodesic cutting sequences (see Series’s chapter in [4]). But this model is difficult to use in our setting. One of its drawbacks is that the roof function, which geometrically represents the lengths of the geodesic segments lying in the fundamental domain of $\Gamma_0 (= \pi_1 (M_0))$ in $\mathbb{H}^2$, is not Hölder or even bounded. To resolve the problem, Ledrappier and Sarig developed a modified coding [12]. Using a lifting argument, we obtain a coding for the geodesic flow on $T^1 (M)$.

To describe the coding, we first recall some basic notions regarding countable Markov shifts. Suppose $S$ is a countable set of states and $A = (t_{ab})_{S \times S}$ is a matrix of zeros and ones without rows which are made solely of zeros.

**Definition 2.1** (TMS). The two-sided topological Markov shift (TMS) with the set of states $S$ and transition matrix $A$ is the set

$$
\Sigma := \{ (\ldots, x_{-1}, x_0, x_1, \ldots) \in S^\mathbb{Z} : \forall i \in \mathbb{Z}, t_{x_i, x_{i+1}} = 1 \}
$$

equipped with the topology generated by the two sided cylinders

$$[a_m, \ldots, a_n] := \{ x \in \Sigma : (x_m, \ldots, x_n) = (a_m, \ldots, a_n) \} \quad (m, n \in \mathbb{Z} \text{ and } m < n)$$

and the action of the left shift $\sigma : \Sigma \to \Sigma, (\sigma x)_i = x_{i+1}$.

We say $(\Sigma, \sigma)$ has the Big Images and Preimages (BIP) property if there is a finite set of states $b_1, \ldots, b_N$ s.t. $\forall a \in S, \exists 1 \leq i, j \leq N$ s.t. $t_{b_i a t_{ab_j}} = 1$.

Given a TMS $(\Sigma, \sigma)$, a function $g : \Sigma \to \mathbb{R}$ is called locally Hölder continuous if $\exists C > 0, 0 < \theta < 1$ s.t.

$$x^n_i = g^n_i \quad (n \geq 0) \implies |g(x) - g(y)| < C \theta^n.$$
Lemma 2.2. \[\text{[12] Lemma 2.2}\] There exist a topologically mixing TMS $(\Sigma, \sigma)$ of countable states, a positive continuous function $\tau: \Sigma \to \mathbb{R}$ and a continuous function $f: \Sigma \to \mathbb{Z}^d$ s.t. $f(x) = f(x_0)$ with the following properties:

1. $(\Sigma, \sigma)$ has the BIP property.
2. The quotient space $\Sigma^{f, \tau} := \Sigma \times \mathbb{Z}^d \times \mathbb{R}/(x, \xi, t) \sim (\sigma x, \xi + f(x), t - \tau(x))$ is homeomorphic to $\Omega$. Denote the homeomorphism by $\pi$.
3. The suspension flow over $\Sigma^{f, \tau}$ is conjugate to the geodesic flow over $\Omega$.
4. There exist a locally Hölder continuous function $r: \Sigma \to \mathbb{R}$ which depends only on nonnegative coordinates and a uniformly continuous function $h: \Sigma \to \mathbb{R}$ such that $\tau = r - (h \circ \sigma)$.
5. There exist $C > 0$ and $K > 0$ s.t. $r + r \circ \sigma + \cdots + r \circ \sigma^{n-1} \geq C$ for all $n \geq 1$.
6. If we set $Q_{\xi_0, t_0}(x, \xi, t) = (x, \xi + \xi_0, t + t_0)$, then $\pi \circ Q_{\xi_0, t_0} = (g^{t_0} \circ D_{\xi_0}) \circ \pi$ for all $(\xi_0, t_0) \in \mathbb{Z}^d \times \mathbb{R}$.

2.2. The transfer operator and the Haar measure. In this subsection, we characterize the Haar measure $m_{\text{Haar}}$ on $T^1(M)$ in the symbolic model $\Sigma^{f, \tau}$.

Let $(\Sigma^+, \sigma)$ be the one-sided version of the Markov shift $(\Sigma, \sigma)$ given in Lemma 2.2. Here we use $\sigma$ to denote the left-shift map on $\Sigma^+$ by abusing notation. As the function $r$ on $\Sigma$ depends only on nonnegative coordinates, we regard it as a function on $\Sigma^+$. For a map $g$ on $\Sigma^+$ and $n \geq 1$, we write

$$g_n(x) = g(x) + g(\sigma(x)) + \cdots + g(\sigma^{n-1}(x)).$$

The Gurevich topological pressure of $-r$ is given by

$$P_{\text{top}}(-r) := \lim_{n \to \infty} \frac{1}{n} \sum \sigma^n z \in \mathbb{Z}^d \quad e^{-r_n(z)} \mathbb{1}_{\sigma^n z}(z),$$

for some state $a_0 \in \mathbb{Z}$.

The limit exists and is independent of the choice of $a_0 \in \mathbb{Z}$ [22 Proposition 3.2].

Denote by $C_B(\Sigma^+)$ the set of bounded continuous functions. The Ruelle transfer operator $L: C_B(\Sigma^+) \to C_B(\Sigma^+)$ is defined by

$$(LF)(x) = \sum \sigma(y) = x e^{-r(y)} F(y) \quad \text{for any } F \in C_B(\Sigma^+).$$

A prior, $LF$ may not lie in $C_B(\Sigma^+)$. In our current setting, Ledrappier and Sarig considered the symbolic local strong stable manifolds, which are small pieces of strong stable manifolds. They showed that the lengths of such manifolds can be described in terms of a positive continuous function $\psi$ on $\Sigma^+; \text{ moreover, } \psi$ satisfies

$$(2.3) \quad L\psi = \psi$$

\[\text{[12] Equation (3)}\]. Note that there exists $M > 0$ such that for any $a_0 \in \mathbb{Z}, x \in [a_0]$ and $n \in \mathbb{N}$, we have $\sum \sigma^n z \in \mathbb{Z}^d \quad e^{-r_n(z)} \mathbb{1}_{[a_0]}(z) = M^{-1} L^n(1_{[a_0]})(x)$. This relation together with (2.3) yields $P_{\text{top}}(-r) < \infty$. It follows that $\psi$ is locally Hölder continuous and bounded away from zero and infinity [21 Corollary 2]. From this, we deduce that the operator $L: C_B(\Sigma^+) \to C_B(\Sigma^+)$ is well-defined.

Lemma 2.4. \[\text{[12] Lemma 3.1}\] There exists a unique finite measure $\rho$ on $\Sigma^+$ such that $\rho(LF) = \rho(F)$ for all non-negative $F \in C_B(\Sigma^+)$, and such that $d\nu := \rho dF$ is a shift invariant probability measure on $\Sigma^+$. Let $\nu$ be the natural extension to $\Sigma$. Then the Haar measure $m_{\text{Haar}}$ equals $m_0 \int_{\tau^+} d\nu(x) \mathbb{1}_{\{(x, \xi, t) \in \Sigma^+ : 0 \leq t < \tau(x)\}} \circ \pi^{-1}$, where $m_0$ is the constant given in (1.6).
The existence of $\rho$ follows from the general theory of BIP shifts [21]. Using the argument in [2] Proposition 6, it can be shown that $d\nu dt|_{((x,t):0<t<\tau(x))}$ gives a measure on $\Gamma_0/G$ which is invariant under the horocycle flow. It follows that this measure is proportional to the Haar measure on $\Gamma_0 \backslash G$ [5].

3. Local mixing and matrix coefficients for local functions

3.1. Asymptotic analysis of symbolic sum. For any $x, y \in \Sigma^+$, let $t(x, y) := \min\{n \geq 0 : x_n \neq y_n\}$. Recall that $r : \Sigma^+ \to \mathbb{R}$ is a locally Hölder continuous function, i.e., there are constants $C > 1$ and $0 < \theta < 1$ such that

$$|r(x) - r(y)| \leq C r^{\theta}(x, y)$$

for any $x, y$ with $x_0 = y_0$.

Define for $\Phi : \Sigma^+ \to \mathbb{R}$, $\text{Lip}(\Phi) := \sup \{\|\Phi(x) - \Phi(y)\| : x_0 = y_0\}$. Set

$$\mathcal{L}(\Sigma^+) := \{\Phi : \Sigma^+ \to \mathbb{R} : \|\Phi\| := \text{Lip}(\Phi) + \|\Phi\|_{\infty} < \infty\}.$$

This is in fact a Banach space.

For any $t > 1$, $\Phi \in \mathcal{L}(\Sigma^+)$ and $u \in C_c(\mathbb{R})$, consider the symbolic sum (which defines a function on $\Sigma^+ \times \mathbb{Z}^d$):

$$Q_t(\Phi \otimes u)(x, \xi) := \sum_{n=0}^{\infty} \sum_{\sigma^n y = x} e^{-rt_n(y)}(\Phi \cdot \psi)(y)\delta_{f_n(y)}(r_n(y) - t).$$

**Theorem 3.1.** [12] Lemma 5.1 For any $(x, \xi) \in \Sigma^+ \times \mathbb{Z}^d$, we have

$$\lim_{t \to \infty} t^{p + h/2} Q_t(\Phi \otimes u)(x, \xi) = \frac{\mathbf{c}m_0}{\tau d} \psi(x)\varphi(\Phi) \int u(t) dt,$$

where $\mathbf{c}$ and $m_0$ are the constants given in (1.7) and (1.6) and the convergence is uniform on compact sets in $\Sigma^+ \times \mathbb{Z}^d$.

Following Lalley’s idea, Theorem 3.1 is proved using two-parameter twisted transfer operator. It is shown in [12] for special functions and the way they formulated the result is different than ours. To state their result, recall the decomposition $\mathbb{R}^d = \mathbb{E}_p \oplus \mathbb{E}_h$. Define functions $F_p$ on $\mathbb{E}_p$ and $F_h$ on $\mathbb{E}_h$ by

$$F_p(\xi_p) = \frac{1}{(2\pi)^p} \int_{\mathbb{E}_p} e^{i\langle \xi, \xi \rangle} d\xi_p,$$

$$F_h(\xi_h) = \frac{1}{(2\pi)^h} \int_{\mathbb{E}_h} e^{i\langle \xi, \xi \rangle} d\xi_h.$$

Set $F(\xi_{2p} + \xi_{2h}) = F_p(\xi_{2p})F_h(\xi_{2h})$. For every $\epsilon > 0$, there are positive, bounded, smooth and Lipschitz functions $F_{\epsilon}^\pm$ with polynomial decay at infinity such that for all $\xi = \xi_{2p} + \xi_{2h} \in \mathbb{R}^d$ and $e^{-\epsilon} < t_1, t_2 < e^\epsilon$,

$$F_{\epsilon}^-(\xi_{2p} + \xi_{2h}) \leq F(t_1 \xi_{2p} + t_2 \xi_{2h}) \leq F_{\epsilon}^+(\xi_{2p} + \xi_{2h})$$

and such that $F_{\epsilon}^+/F_{\epsilon}^- \to 1$ as uniformly on compact sets as $\epsilon \to 0^+$, $\epsilon \mapsto F_{\epsilon}^+$ is decreasing and $\epsilon \mapsto F_{\epsilon}^-$ is increasing. Let $1_{[\varepsilon]}$ be the characteristic function of a cylinder set $[\xi_0]$ in $\Sigma^+$ and $1_{[-a/2, a/2]}$ be the characteristic function of an interval $[-a/2, a/2]$ in $\mathbb{R}$. It is shown in [12] Lemma 5.1 that for any $\epsilon > 0$, there exists
Let \( t_0 > 0 \) such that for any \( t > t_0, x \in \Sigma^+ \) and \( \xi \in \mathbb{Z}^d \), if \( \xi = \xi_p + \xi_n \) where \( \xi_p \in \mathbb{E}_p \) and \( \xi_n \in \mathbb{E}_h \), then

\[
\begin{align*}
t_{p+h/2}Q_t(1_{[1]} \otimes 1_{[-a/2,a/2]})(x,\xi) & \leq e^t \left[ F_+^p \left( \frac{\xi_p}{t} \right) \right] + \epsilon \frac{\nu[0,a]}{\int \tau d\nu} \psi(x); \\
t_{p+h/2}Q_t(1_{[1]} \otimes 1_{[-a/2,a/2]})(x,\xi) & \leq e^{-\epsilon} \left[ F_-^p \left( \frac{\xi_p}{t} \right) \right] + \epsilon \frac{\nu[0,a]}{\int \tau d\nu} \psi(x).
\end{align*}
\]

We deduce from the properties of \( F^p_+ \) that

\[
\lim_{t \to \infty} t^{p+h/2}Q_t(1_{[1]} \otimes 1_{[-a/2,a/2]})(x,\xi) = \frac{c \mu_0}{\int \tau d\nu} \psi(x) \nu[y_0].
\]

Theorem 3.1 follows from a verbatim repetition of the proof of [12, Lemma 5.1].

### 3.2. Correlation functions for \((\Gamma \setminus G, a, \mu^{\text{Haar}})\)

We write

\[
\begin{align*}
\tilde{X} & := \Sigma \times \mathbb{Z}^d \times \mathbb{R}, \\
\tilde{X}^+ & := \Sigma^+ \times \mathbb{Z}^d \times \mathbb{R}.
\end{align*}
\]

Consider the product measure on \( \tilde{X} \):

\[
d\tilde{M} := \mu_0 \int \tau d\nu (dvd\xi ds).
\]

The Haar measure \( \mu^{\text{Haar}} \) on \( \Gamma \setminus G \) corresponds to the measure on \( \Sigma^{\Gamma \setminus \tau} \) induced by \( M \). For \( \Psi_1, \Psi_2 \in C_c(\tilde{X}^+) \), define

\[
I_t(\Psi_1, \Psi_2) := \sum_{n=0}^{\infty} \int_{\tilde{X}} \Psi_1 \circ \tilde{\zeta}_n(x,\xi,s+t) \cdot \Psi_2(x,\xi,s) d\tilde{M}(x,\xi,s)
\]

where \( \tilde{\zeta}_n(x,\xi,s) := (\sigma^n x, \xi + f_n(x), s - \tau_n(x)) \). It follows from the property of \( r \) (Lemma 2.2 (5)) that, for any \( x, \xi, s \in \text{supp} \, \Psi_2 \),

\[
\Psi_1 \circ \tilde{\zeta}_n(x,\xi,s+t) = 0
\]

for \( n \) large enough. So \( I_t(\Psi_1, \Psi_2) \) is in fact a finite sum.

**Definition 3.2.** Let \( \mathcal{F}_0 \) be the family of functions on \( \tilde{X}^+ \) of the form

\[
\Psi(x,\xi,s) = \Phi(x)\delta_{\xi_0}(\xi)u(s)
\]

where \( \Phi \in \mathcal{L}(\Sigma^+), u \in C_c(\mathbb{R}) \) and \( \xi_0 \in \mathbb{Z}^d \). Denote by \( \mathcal{F} \) the space of functions which are finite linear combinations of functions from \( \mathcal{F}_0 \).

**Lemma 3.3.** Let \( \Psi_2(x,\xi,s) = \Phi(x)\delta_{\xi_0}(\xi)u(s) \in \mathcal{F}_0 \). Then for any \( \Psi_1 \in C_c(\tilde{X}^+) \), we have

\[
I_t(\Psi_1, \Psi_2) = \frac{\mu_0}{\int \tau d\nu} \int_{\tilde{X}^+} \Psi_1(x,\xi,s) \cdot Q_t(\Phi \otimes u)(x,\xi,\xi_0) d\rho(x)d\xi ds.
\]

**Proof.** Since \( d\rho \) is an eigenmeasure of \( L \) with eigenvalue 1, for any \( F, G \in \mathcal{L}(\Sigma^+) \), we have the following equality

\[
\int_{\Sigma^+} F \circ \sigma \cdot G d\rho = \int_{\Sigma^+} L(F \circ \sigma \cdot G)d\rho = \int_{\Sigma^+} F \cdot (LG)d\rho.
\]
Using this and the Fubini theorem, we obtain

\[ I_t(\Psi_1, \Psi_2) \]

\[ = \int_{\mathcal{X}^+} \Psi_1(x, \xi, s) \sum_{n=0}^{\infty} \sum_{s' \in \mathcal{F}} e^{-r_n(y)} (\Phi \cdot \psi)(y) \delta_{\xi_0}(\xi - f_n(y)) u(s - t + r_n(y)) d\rho(x) d\xi ds \]

\[ = \frac{m_0}{\tau d\nu} \int_{\mathcal{X}^+} \Psi_1(x, \xi, s) \mathcal{Q}_{t-s}(\Phi \otimes u)(x, \xi - \xi_0) d\rho(x) d\xi ds. \]

Proof of Theorem 1.1. Without loss of generality, we may suppose that \( \psi_1 \) and \( \psi_2 \) are defined on the suspension space \( \Sigma^{f, \tau} \), continuous and compactly supported. For each \( i = 1, 2 \), let \( \Psi_i \in C_c(\mathcal{X}) \) be the lift of \( \psi_i \) to \( \mathcal{X} \) satisfying

\[ \Psi_i[(x, \xi, s)] = \sum_{n \in \mathbb{Z}} \Psi_i \circ \zeta^n(x, \xi, s), \]

with \( \zeta(x, \xi, s) := (\sigma x, \xi + f(x), s - \tau(x)) \).

Using the unfolding,

\[ \int_{\mathcal{X}} \Psi_1(xa) \psi_2(x) d\text{Haar}(x) \]

\[ = \sum_{n \in \mathbb{Z}} \int_{\mathcal{X}} \Psi_1 \circ \zeta^n(x, \xi, s + t) \cdot \Psi_2(x, \xi, s) d\tilde{\mathbb{M}} \]

\[ = \sum_{n=0}^{\infty} \int_{\mathcal{X}} \Psi_1 \circ \zeta^n(x, \xi, s + t) \cdot \Psi_2(x, \xi, s) d\tilde{\mathbb{M}} \]

\[ + \sum_{n=1}^{\infty} \int_{\mathcal{X}} \Psi_1 \circ \zeta^{-n}(x, \xi, s + t) \cdot \Psi_2(x, \xi, s) d\tilde{\mathbb{M}}. \]

For the second term in the above equation, note that for \( (x, \xi, s) \in \text{supp}(\tilde{\Psi}_2) \),

\[ \Psi_1 \circ \zeta^{-n}(x, \xi, s + t) = \Psi_1(\sigma^{-n} x, \xi - f_n(x), s + t + \tau_n(x)), \]

which is 0 when \( t \) is large enough, as \( \tau_n(x) > 0 \). Define

\[ \Psi_i(x, \xi, s) := \tilde{\Psi}_i(x, \xi, s - h(x)) \text{ for } i = 1, 2. \]

As \( \tilde{\Psi}_i \) is compactly supported, the new function \( \tilde{\Psi}_i \) is also a continuous function with compact support. Moreover, we have \( \tilde{\mathbb{M}}(\Psi_1) = \tilde{\mathbb{M}}(\Psi_1) \). The first term of the last equation above equals \( I_t(\Psi_1, \Psi_2) \). Theorem 1.1 follows if this following holds

\[ \lim_{t \to \infty} t^{p+h/2} I_t(\Psi_1, \Psi_2) = \epsilon \tilde{\mathbb{M}}(\tilde{\Psi}_1) \tilde{\mathbb{M}}(\tilde{\Psi}_2). \]

The proof of (3.4) is divided into the following three cases.

Case 1. Assume \( \Psi_1 \in C_c(\mathcal{X}^+) \) and \( \Psi_2 \in F \). It suffices to consider the case where

\[ \Psi_2(x, \xi, s) = \Phi(x) \delta_{\xi_0}(\xi) u(s) \in \mathcal{F}_0. \]

By Lemma 3.3

\[ I_t(\Psi_1, \Psi_2) = \frac{m_0}{\tau d\nu} \int_{\mathcal{X}^+} \Psi_1(x, \xi, s) \cdot \mathcal{Q}_{t-s}(\Phi \otimes u)(x, \xi - \xi_0) d\rho(x) d\xi ds. \]
Applying Theorem 3.1 to the integrand, we obtain
\[
\lim_{t \to \infty} t^{p+h/2} \Psi_1(x, \xi, s) : Q_{t-s}(\Phi \otimes u)(x, \xi - \xi_0) = \frac{c}{\int \tau d\nu} \hat{M}(\Psi_2) \Psi_1(x, \xi, s) \psi(x).
\]
As the above convergence is uniform on compact sets, we have
\[
\lim_{t \to \infty} t^{p+h/2} I_1(\Psi_1, \Psi_2) = c \hat{M}(\Psi_1) \hat{M}(\Psi_2).
\]

**Case 2.** Assume $\Psi_1, \Psi_2 \in C_c(\hat{X}^+)$.
Approximate $\Psi_2$ by functions in $\mathcal{F}$: it follows from the Stone-Weierstrass theorem that for any $\epsilon > 0$, there exist $F_2 \in \mathcal{F}$ such that for any $(x, \xi, s) \in \hat{X}^+$
\[
|\Psi_2(x, \xi, s) - F_2(x, \xi, s)| < \epsilon.
\]
We can find $w_2 \in \mathcal{F}$ satisfying
- $|\Psi_2(x, \xi, s) - F_2(x, \xi, s)| < \epsilon w_2(x, \xi, s)$ for any $(x, \xi, s) \in \hat{X}^+$;
- $\hat{M}(w_2)$ is bounded by the size of the $\mathbb{Z}^d \times \mathbb{R}$-support of $\Psi_2$.
As (3.4) holds for the pairs $(\Psi_1, F_2), (\Psi_1, w_2)$ and $\epsilon$ is arbitrary, it also holds for the pair $(\Psi_1, \Psi_2)$.

**Case 3.** Let $\Psi_1, \Psi_2 \in C_c(\hat{X})$.
For any $\epsilon > 0$, there exist $k \in \mathbb{N}$ and $F_i \in C_c(\hat{X}^+)$ such that
\[
|\tilde{\Psi}_i \circ \check{\zeta}^k(x, \xi, t) - F_i(x, \xi + f_k(x), t - r_k(x))| < \epsilon.
\]
We can find $w_i \in \mathcal{F}$ satisfying
- $|\Psi_i \circ \check{\zeta}^k(x, \xi, t) - F_i(x, \xi + f_k(x), t - r_k(x))| < \epsilon \cdot w_i(x, \xi + f_k(x), s - r_k(x));$
- $\hat{M}(w_i)$ are bounded by the size of $\mathbb{Z}^d \times \mathbb{R}$-support of $\Psi_i$.
Since $\hat{M}$ is invariant under the action of $\check{\zeta}$, we can replace $\Psi_i$ by $\Psi_i \circ \check{\zeta}^k$ in (3.4).
For (3.4) holds for the pairs $(F_i, F_j), (F_i, w_j), (w_i, w_j)$ with $i, j \in \{1, 2\}$, it is valid for the pair $(\Psi_1 \circ \check{\zeta}^k, \Psi_2 \circ \check{\zeta}^k)$.

\[\square\]

References

[1] M. Babillot and F. Ledrappier. *Lalley’s theorem on periodic orbits of hyperbolic flows*. Ergodic Theory Dynam. Systems 18 (1998), no. 1, 17-39.

[2] M. Babillot and F. Ledrappier. *Geodesic paths and horocycle flow on abelian covers*. Proceedings of the International Colloquium on Lie Groups and Ergodic Theory, Mumbai 1996, Narosa Pub. House, New Delhi, (1998), 1-32.

[3] M. Babillot and M. Peigné. *Homologie des geodesiques fermes sur des varits hyperboliques avec bouts cuspdaux*. (French) Ann. Sci. École Norm. Sup. (4) 33 (2000), no. 1, 81-120.

[4] T. Bedford, M. Keane and C. Series (editors). *Ergodic theory, symbolic dynamics, and hyperbolic spaces*, Papers from the Workshop on Hyperbolic Geometry and Ergodic Theory held in Trieste, April 17-28, 1989, Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1991.

[5] S. G. Dani and J. Smillie. *Uniform distribution of horocycle orbits for Fuchsian groups*. Duke Math. J., 51 (1984), 185-194.

[6] C. Epstein. *Asymptotics for closed geodesics in a homology class, the finite volume case*. Duke Math. J. 55 (1987): 717-757.

[7] W. Duke, Z. Rudnick and P. Sarnak. *Density of integer points on affine homogeneous varieties*. Duke Math. J. 71.1 (1993): 143-179.

[8] A. Eskin and C. McMullen. *Mixing, counting, and equidistribution in Lie groups*. Duke Math. J. 71.1 (1993): 181-209.
[9] Y. Guivarc’h and J. Hardy. Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d’Anosov. Ann. I.H.P. 1 (1988): 73-98.
[10] Y. Guivarc’h and Y. Le Jan. Asymptotic winding of the geodesic flow on modular surfaces and continued fractions. Ann. Sci. Éc. Norm. Sup., 26 (1993), 23-50; Errata: Ann. Sci. Éc. Norm. Sup. (4), 29 (1996), 811-814.
[11] R. Howe and C. Moore. Asymptotic properties of unitary representations. Journal of Functional Analysis, Vol 32 (1), 72-96, 1979.
[12] F. Ledrappier and O. Sarig. Fluctuations of ergodic sums for horocycle flows on $\mathbb{Z}^d$-covers of finite volume surfaces. Discrete Contin. Dyn. Syst. 22 (2008), no. 1-2, 247-325.
[13] G. Margulis. On some aspects of the theory of Anosov systems. Springer Berlin Heidelberg, 2004. 1-71.
[14] G. Margulis, A. Mohammadi and H. Oh. Closed geodesics and holonomies for Kleinian manifolds. Geom. Funct. Anal. 24 (2014), no. 5, 1628-1636.
[15] A. Mohammadi and H. Oh. Matrix coefficients, counting and primes for orbits of geometrically finite groups. Journal of EMS 17 (2015), no. 4, 837-897.
[16] H. Oh. Harmonic analysis, ergodic theory and counting for thin groups. Thin groups and superstrong approximation, 179-210, Math. Sci. Res. Inst. Publ., 61, Cambridge Univ. Press, Cambridge, 2014.
[17] H. Oh and W. Pan. Local mixing and invariant measures for horospherical subgroups on abelian covers. Preprint, arXiv: 1701.02772v3. To appear in IMRN.
[18] H. Oh and N. Shah. Equidistribution and counting for orbits of geometrically finite hyperbolic groups. J. Amer. Math. Soc. 26 (2013), no. 2, 511-562.
[19] M. Pollicott and R. Sharp. Orbital counting for some discrete groups acting on simply connected manifolds with negative curvature. Inventiones 117, 275-301 (1994).
[20] T. Roblin. Ergodicité et équidistribution en courbure négative. Mém. Soc. Math. Fr. (N.S.), Vol 95 (2003).
[21] O. Sarig. Existence of Gibbs measures for countable Markov shifts. Proc. Amer. Math. Soc. 131 (2003), no. 6, 1751-1758.
[22] O. Sarig. Lecture notes on thermodynamic formalism for topological Markov shifts.
[23] P. Tukia. On discrete groups of the unit disk and their isomorphisms. Ann. Acad. Sci. Fenn. Ser., AI, 504 (1972).

MATHEMATICS DEPARTMENT, YALE UNIVERSITY, NEW HAVEN, CT 06511
E-mail address: wenyu.pan@yale.edu