ISRAEL JOURNAL OF MATHEMATICS **160** (2007), 281–315
DOI: 10.1007/s11856-007-0064-0

INVARIANT MEASURES FOR THE HOROCYCLE FLOW ON PERIODIC HYPERBOLIC SURFACES

BY

François Ledrappier*

Department of Mathematics, University of Notre-Dame
Notre-Dame, IN 46556-4618, USA
e-mail: ledrappier.1@nd.edu

AND

Omri Sarig**

Mathematics Department, Pennsylvania State University
University Park PA 16802, USA
e-mail: sarig@math.psu.edu

Pour Martine

ABSTRACT

We classify the ergodic invariant Radon measures for the horocycle flow on geometrically infinite regular covers of compact hyperbolic surfaces. The method is to establish a bijection between these measures and the positive minimal eigenfunctions of the Laplacian of the surface. Two consequences arise: if the group of deck transformations G is of polynomial growth, then these measures are classified by the homomorphisms from G_0 to $\mathbb{R}$ where $G_0 \leq G$ is a nilpotent subgroup of finite index; if the group is of exponential growth, then there may be more than one Radon measure which is invariant under the geodesic flow and the horocycle flow. We also treat regular covers of finite volume surfaces.

* The first author was supported by NSF grant 0500630.
** The second author was supported by NSF grant 0400687.
Received July 19, 2005

1. Introduction

Let M be an oriented hyperbolic surface and $T^1(M)$ its unit tangent bundle. The **geodesic flow** is the flow $g^s\colon T^1(M) \to T^1(M)$ which moves a line element at unit speed along the geodesic it determines. The (stable) **horocycle** at a line element ω is the geometric locus of all $\omega' \in T^1(M)$ for which $d(g^s\omega, g^s\omega') \xrightarrow[s\to\infty]{} 0$. This is a smooth curve, and the collection of these curves can be oriented in a continuous fashion. The (stable) **horocycle flow** $h^t\colon T^1(M) \to T^1(M)$ moves line elements along the stable horocycle they determine at unit speed, in the positive direction.

A famous theorem of Furstenberg [F] says that if M is compact, then h has, up to normalization, a unique invariant Radon measure. Variants of this phenomena have been established for more general geometrically finite hyperbolic surfaces by Dani [D], Dani and Smillie [DS] and Burger [Bu], for compact Riemannian surfaces of variable negative curvature by Marcus [Mrc], and for more general actions, by Ratner [Rat]. The geometrically infinite case is still almost completely open.

We restrict our attention to the simplest possible geometrically infinite surfaces, the **periodic surfaces**. These are the surfaces of the form

$$M = \Gamma \setminus \mathbb{D} \quad \text{where } \{id\} \neq \Gamma \lhd \Gamma_0,\ \Gamma_0 \text{ is a torsion free lattice in } \text{Möb}(\mathbb{D}).$$

Here and throughout, $\mathbb{D}$ is the unit disc and Möb($\mathbb{D}$) is the group of Möbius transformations which preserve $\mathbb{D}$. M is a regular cover of the finite volume surface $M_0 = \Gamma_0 \setminus \mathbb{D}$, and is therefore made of some finite volume piece, which is repeated periodically according to some symmetry group. The **period** of M is M_0, and the **symmetry group** of M (relative to M_0) is the group G of deck transformations (which is isomorphic to Γ_0/Γ). A periodic surface is called **cocompact** if M_0 is compact.

The symmetry group G of a periodic surface is always finitely generated, and any finitely generated group can be realized as the symmetry group of some cocompact periodic surface. A periodic surface is called **abelian**, **nilpotent** etc. if its symmetry group is abelian, nilpotent etc.

It follows from the work of Dani and Smillie [DS] (see also [Rat]) that the horocycle flow has no finite invariant measures on periodic surfaces of infinite volume, other than measures supported on closed horocycles. But it has nontrivial invariant Radon measures, e.g. the volume measure on the unit tangent bundle.

So far, the ergodic invariant Radon measures (**e.i.r.m.'s**) for the horocycle flow on a periodic surface have only been classified for free abelian cocompact surfaces. In this case every homomorphism $\varphi\colon G \to \mathbb{R}$ determines a unique (ray of) e.i.r.m. m such that $m \circ dD = e^{\varphi(D)}m$ $(D \in G)$ [BL], and every e.i.r.m. arises this way [Sg].

Our aim here is to describe the e.i.r.m.'s for general periodic surfaces. One of our original aims was to understand how general is the situation that all h-e.i.r.m.'s are quasi-invariant under all deck transformations. We show below that although general abelian and even nilpotent cocompact surfaces have this property, polycyclic cocompact surfaces of exponential growth do not.

Indeed, we construct below a periodic surface with a Radon measure which is ergodic and invariant under the horocycle flow (and the geodesic flow!), but is not quasi-invariant under some of the deck transformations (Section 2, Example 7). The existence of such an example shows that the there is no hope in extending the methods of [Sg] to the general non-abelian case. The point of this paper is that we have found a way of bypassing the problem which makes the argument of that paper break down.[1]

We begin with some remarks on general hyperbolic surfaces $M = \Gamma\backslash\mathbb{D}$. Every h-e.i.r.m. on $T^1(M)$ lifts to some Γ-invariant h-invariant Radon measure on $T^1(\mathbb{D})$. The unit tangent bundle $T^1(\mathbb{D})$ can be identified with $(\partial\mathbb{D} \times \mathbb{R}) \times \mathbb{R}$ as follows: Let $o \in \mathbb{D}$ denote the origin. For every $e^{i\theta} \in \partial\mathbb{D}$ and $z \in \mathbb{D}$ let $\omega_\theta(z)$ be the line element based in z whose geodesic ends at $e^{i\theta}$. The identification is

$$(e^{i\theta}, s, t) \longmapsto (h^t \circ g^s)(\omega_\theta(o)).$$

We call $(e^{i\theta}, s, t)$ the KAN**-coordinates** of ω (this is the Iwasawa decomposition).

It is well-known that $g^s \circ h^t = h^{te^{-s}} \circ g^s$. Therefore, in these coordinates

$$h^t(e^{i\theta_0}, s_0, t_0) = (e^{i\theta_0}, s_0, t_0 + t);$$
$$g^s(e^{i\theta_0}, s_0, t_0) = (e^{i\theta_0}, s_0 + s, t_0 e^{-s}).$$

It follows that any h-invariant measure m must be of the form $d\mu(e^{i\theta}, s)dt$.

If, in addition, m is quasi–invariant with respect to the geodesic flow, then $m \circ g^s = e^{(\alpha-1)s}m$ for some α and all s,[2] and we can decompose m further into

1 The problem is that the homomorphism $\alpha\colon \mathbb{H}_m \to \mathbb{R}$ considered in [Sg, p. 540] does not always extend to a homomorphism on the entire group $\mathbb{R}\times$Deck Transformations, and as a result the Livsic Rigidity Theorem is inapplicable. In retrospect and given the results of this paper, we know that such an extension always exists (trivially) in the nilpotent case, but not always in the polycyclic case.

2 If m is h-e.i.r.m., then so is $m \circ g^s$ because $g^s \circ h^t = h^{te^{-s}} \circ g^s$. Since m, $m \circ g^s$ are ergodic and equivalent, they must be proportional. The constant must of the form $e^{\beta s}$. Set $\beta = \alpha - 1$.

$dm = e^{\alpha s} d\nu(e^{i\theta}) dsdt$ for some finite measure ν on $\partial\mathbb{D}$. The measure ν is then determined by the requirements that m be Γ-invariant and h-ergodic. These requirements turn out to be equivalent to ergodicity and conformality w.r.t. the Γ-action on $\partial\mathbb{D}$ (see [Ba] and below).

This is the approach used by Martine Babillot in [Ba] to classify h-e.i.r.m. which are quasi–invariant w.r.t. the geodesic flow (for a different approach, see [ASS]).

In general, it is not true that any h-e.i.r.m. is g-quasi-invariant: Take a non-cocompact periodic surface M with period M_0. Since M_0 is a non-compact hyperbolic surface of finite volume, it has cusps. Every cusp is encircled by closed horocycles of finite length. These horocycles lift to h-orbits on $T^1(M)$. The lifts are not necessarily of finite length, but they are always locally finite: The Lebesgue measure on them is a Radon measure on $T^1(M)$. This measure is h-ergodic and invariant, but is not g-quasi-invariant. We call these measures **trivial** h-e.i.r.m.'s.

Our contribution is to show that the trivial measures are the only obstruction to g-quasi–invariance:

THEOREM 1: *Let M be a periodic surface with period M_0. Any non-trivial h-e.i.r.m. on $T^1(M)$ is quasi-invariant w.r.t. the geodesic flow.*

Let Γ be a Fuchsian group, and ν some measure on $\partial\mathbb{D}$. We say that ν is **Γ-ergodic**, if any Γ-invariant function is constant on a set of full measure. We say that ν is **Γ-conformal** (with parameter α) if ν is finite, and $\frac{d\nu\circ g}{d\nu} = |g'|^\alpha$ for all $g \in \Gamma$ (see [Su3]). Theorem 1 allows us to complete Babillot's programme and show

THEOREM 2: *Let $M = \Gamma\backslash\mathbb{D}$ be a periodic surface. If ν is non-atomic, Γ-ergodic, and conformal with parameter α, then $e^{\alpha s} d\nu(e^{i\theta}) dsdt$ is a Γ-invariant measure on $T^1(\mathbb{D})$, which projects to a non-trivial h-e.i.r.m. on $T^1(\Gamma\backslash\mathbb{D})$. Any non-trivial h-e.i.r.m. on $T^1(\Gamma \setminus \mathbb{D})$ is of this form.*

Recall that the **hyperbolic Laplacian** of $\mathbb{D}$ is a second order differential operator on $C^2(\mathbb{D})$ s.t. $\Delta_{\mathbb{D}}(f \circ \varphi) = (\Delta_{\mathbb{D}} f) \circ \varphi$ for all $\varphi \in$ Möb$(\mathbb{D})$. This determines $\Delta_{\mathbb{D}}$ up to a constant, and this constant can be chosen to make $\Delta_{\mathbb{H}} = y^2(\partial^2/\partial x^2 + \partial^2/\partial y^2)$ in the upper half plane model. The invariance property of $\Delta_{\mathbb{D}}$ means that it descends to an operator Δ_M on $M = \Gamma \setminus \mathbb{D}$, called the **hyperbolic Laplacian** of M.

The collection of positive λ-eigenfunctions of Δ_M forms a cone. The extremal rays of this cone are directions generated by the *minimal* positive

λ-eigenfunctions: the λ-eigenfunctions F for which $\Delta_M G = \lambda G$, $0 \leq G \leq F \Rightarrow \exists c$ s.t. $G = cF$.

If $P(e^{i\theta}, z) := (1 - |z|^2)/|e^{i\theta} - z|^2$ (the Poisson kernel), then $P(e^{i\theta}, z)^\alpha$ is an $\alpha(\alpha - 1)$-positive eigenfunction of $\Delta_{\mathbb{D}}$ (see §5.1). Consequently, any Γ-invariant function of the form $\sum c_k P(e^{i\theta_k}, z)^\alpha$, $c_k \geq 0$ defines a positive eigenfunction of Δ_M. We call these eigenfunctions **trivial eigenfunctions** (see §6.1 for the connection with the Eisenstein series). In fact, $e^{i\theta_k}$ must all be fixed points of parabolic elements of Γ (see e.g. the proof of lemma 2 and the discussion below).

Following Babillot [Ba], we consider the assignment

$$(*) \qquad m = e^{\alpha s} d\nu(e^{i\theta}) ds dt \longmapsto F_m(z) := \int_{\partial\mathbb{D}} P(e^{i\theta}, z)^\alpha d\nu(e^{i\theta}).$$

THEOREM 3: *Let M be a periodic surface. The mapping $(*)$ is a bijection between the non-trivial e.i.r.m.'s of h on $T^1(M)$ and the non-trivial minimal positive eigenfunctions of Δ_M. This bijection satisfies:*

(1) $m \circ g^s = e^{(\alpha-1)s} m \Leftrightarrow \Delta_M F_m = \alpha(\alpha - 1) F_m$;

(2) $m \circ dD = cm \Leftrightarrow F_m \circ D = cF_m$ *for all D in the symmetry group of M.*

Remark: Cocompact periodic surfaces have no trivial h-e.i.r.m.'s, because compact surfaces do not admit closed horocycles. They admit no trivial positive eigenfunctions for the Laplacian, because uniform lattices have no parabolic fixed points. Therefore, for cocompact periodic surfaces, all h-e.i.r.m.'s are g-quasi-invariant, all h-e.i.r.m.'s have the form described in Theorem 2, and $(*)$ is a bijection between the collection of all h-e.i.r.m.'s and the collection of all minimal positive eigenfunctions of the Laplacian.

ACKNOWLEDGEMENT: The authors wish to thank J.-P Conze, Y. Coudene, M. Forester, L. Flaminio, Y. Guivarc'h, V. Kaimanovich, A. Raugi and P. Sarnak for helpful discussions.

2. Examples

We illustrate these results by examples. We remind the reader that any finitely generated group is the symmetry group of some cocompact periodic surface. The classes of examples described below are therefore not empty.

Example 1 (Furstenberg's Theorem [F])*:* The horocycle flow of a compact hyperbolic surface is uniquely ergodic.

Proof: This is the case when the symmetry group is trivial. Any e.i.r.m. m corresponds to a function F such that $\Delta_M F = \alpha(\alpha - 1)F$ where α satisfies

$m \circ g^s = e^{(\alpha-1)s}m$. Since m is finite (a Radon measure on a compact space), α must be equal to one. Therefore F is harmonic, hence (by compactness and the maximum principle) constant. The representing measure of the constant function is proportional to Haar's measure $d\lambda$. It follows that m is proportional to $e^s d\lambda(e^{i\theta})dsdt$ =volume measure. ∎

Example 2 (Dani–Smillie Theorem [DS])*:* The ergodic invariant Radon measures for the horocycle flow on a hyperbolic surface of finite area are all finite, and consist of trivial measures and measures proportional to the volume measure.

Proof: Dani and Smillie proved this by showing that non-periodic horocycle orbits are equidistributed. We deduce it from Theorem 3, and the fact that the minimal positive eigenfunctions in this case are either trivial, or constant (see §6.1). ∎

Example 3 (Kaimanovich's Theorem [Kai1])*:* The volume measure on a periodic surface is h-ergodic iff all bounded harmonic functions on the surface are constant (the Liouville property).

Proof: Kaimanovich proved this for all hyperbolic surfaces [Kai1]. We explain how his result fits with ours in the periodic case. Let M be a hyperbolic periodic surface with symmetry group G and period M_0. The volume measure on $T^1(\mathbb{D})$ is of the form $dm = e^s d\lambda(e^{i\theta})dsdt$, where λ is Haar's measure on $\partial\mathbb{D}$. Haar's measure is Γ-conformal of parameter 1. By Theorem 2, it is ergodic iff m is an e.i.r.m., in which case (by Theorem 3) $F_m(z) = \int_{\partial\mathbb{D}} P(e^{i\theta}, z)d\lambda(e^{i\theta}) \equiv 1$ is minimal. This shows: the volume measure is ergodic iff 1 is a minimal harmonic function. But 1 is minimal exactly when all bounded harmonic functions are constant. ∎

Example 4 (The strong Liouville property)*:* The volume measure on a periodic surface is the unique g-invariant h-e.i.r.m. on M iff all positive harmonic functions on the surface are constant (the strong Liouville property).

Proof: g-invariant h-e.i.r.m.'s are necessarily non-trivial, and therefore correspond to minimal positive harmonic functions. The volume measure corresponds to the constant function. ∎

Example 5 (Nilpotent surfaces)*:* Let M be a cocompact periodic surface with nilpotent symmetry group G. Every homomorphism $\varphi\colon G \to \mathbb{R}$ determines an

h-e.i.r.m. measure m (unique up to a constant) such that $m \circ dD = e^{\varphi(D)}m$ for every $D \in G$, and every h-e.i.r.m. is of this form.

Proof: This is because the minimal positive eigenfunctions for a cocompact nilpotent surface form a family $\{tF_\varphi : t > 0, \varphi\colon G \to \mathbb{R}$ is a homomorphism$\}$, where $F_\varphi \circ D = e^{\varphi(D)} F_\varphi$ for all $D \in G$ (see §6.2). This example strengthens the main result of [Ba] by removing the g-quasi-invariance assumption. ∎

Example 6 (Polynomial growth)*:* Let M be a cocompact periodic surface of polynomial growth[3]. The symmetry group of M contains a finitely generated normal nilpotent subgroup N of finite index, and the rays of h-e.i.r.m.'s on $T^1(M)$ are in bijection with the homomorphisms from N to $\mathbb{R}$.

Proof: Let M be a periodic cocompact surface of polynomial growth with period M_0 and symmetry group G. Let $F_0 \subset M$ be one of the connected preimages of M_0 under the covering group which project to M_0 bijectively. The collection $\{D \in G : \overline{F}_0 \cap D(\overline{F}_0) \neq \varnothing\}$ is a finite set of generators for G. Let $|\cdot|$ be the word metric w.r.t. to this set of generators. Then

$$\sharp\{D \in G : |D| \leq n\} \times \mathrm{vol}(F_0) \leq \mathrm{vol}\{p \in M : d(p, F_0) \leq (n+1)\cdot \mathrm{diam}(M_0)\}.$$

Therefore, G has polynomial growth. By Gromov's theorem [Gr], G contains a nilpotent subgroup N_0 of finite index. The group $N := \bigcap_{g\in G} g^{-1}N_0 g$ is normal and nilpotent. By Poincaré's theorem ([Ro], Theorem 1.3.12) N_0 has finite index in G, because the intersection which defines it has a finite number of different terms. Since N has finite index in G and G is finitely generated, N is finitely generated (see e.g. [Ro], Theorem 6.1.8).

We claim that there is a compact hyperbolic surface M_1 such that M is a nilpotent surface with period M_1 and symmetry group N (we thank Y. Coudene for this observation). This finishes the proof, by reducing Example 5 to Example 4.

Write $M_0 = \Gamma_0 \setminus \mathbb{D}$, $M = \Gamma \setminus \mathbb{D}$ and $G = \Gamma_0/\Gamma$. Since $N \lhd G$, $N = \Gamma_1/\Gamma$ for some $\Gamma \lhd \Gamma_1 \lhd \Gamma_0$. It follows that M is regular cover of $M_1 := \Gamma_1 \setminus \mathbb{D}$ and the group of deck transformations of this cover is $\Gamma_1/\Gamma \equiv N$. To see that M_1 is compact, note that it is a finite cover of M_0, because $|\Gamma_0/\Gamma_1| = |(\Gamma_0/\Gamma)/(\Gamma_1/\Gamma)| = |G/N| < \infty$. ∎

3 A Riemannian surface is said to be of polynomial growth, if the volume of balls of radius R is $O(R^\delta)$ for some δ as $R \to \infty$.

Remark: This shows that the h-e.i.r.m.'s on a cocompact periodic surface of polynomial growth can be naturally parameterized as a d-parameter family with $d = \mathrm{rank}(N/[N,N])$ (where the rank of the finitely generated abelian group $A = N/[N,N]$ is the d in $A/\mathrm{Tor}(A) \simeq \mathbb{Z}^d$).

Example 7 (Polycyclic surfaces)*:* Cocompact polycyclic surfaces which are not virtually nilpotent are Liouville, but not strongly Liouville.[4] Therefore,

(i) The volume measure on $T^1(M)$ is a g-invariant h-e.i.r.m.;

(ii) There are other g-invariant h-e.i.r.m.'s, and these measures are not quasi-invariant w.r.t. all deck transformations.

Proof: A cocompact polycyclic surface has the Liouville property (Kaimanovich [Kai2]), and we saw that this implies (i). If M is not virtually nilpotent, then it is not of polynomial growth. Polycyclic groups are linear, therefore the work of Bougerol and Élie [BE] provides a non-constant positive harmonic function F on M.

Any positive harmonic function is the barycenter of minimal positive harmonic functions, so it is possible to find a non-constant minimal positive harmonic function F_0.

The measure m_0 which corresponds to F_0 is g-invariant, because F_0 has eigenvalue zero. We claim that it cannot be quasi-invariant w.r.t. all deck transformations. The horocycle flow commutes with all deck transformations. If m were quasi-invariant w.r.t. all deck transformations, then $m_0 \circ dD = e^{\varphi(D)} m_0$ with $\varphi\colon G \to \mathbb{R}$ a homomorphism (equivalent ergodic invariant measure are proportional). Any homomorphism into $\mathbb{R}$ must vanish on $[G,G]$. Going back to F_0 we see that $F_0 \circ D = F_0$ for all $D \in [G,G]$. It follows that F_0 descends to a positive harmonic function on $M/[G,G]$. But this cocompact surface is abelian (its symmetry group is $G/[G,G]$) and all positive harmonic functions on abelian surfaces are constant [LS], a contradiction. ∎

Example 8 (The Thrice Punctured Sphere)*:* Working in the upper half plane $\mathbb{H}$, define $\Gamma(2) := \{\varphi(z) = \frac{az+b}{cz+d} : \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL(2,\mathbb{Z}),\ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \bmod 2\}$.

(1) $M_0 = \Gamma(2)\backslash\mathbb{H}$ is a finite volume hyperbolic surface, and is homeomorphic to the sphere minus three points, which correspond to three cusps;

4 G is **polycyclic** if $\exists G_i \lhd G$ s.t. $\{1\} = G_0 \lhd \cdots \lhd G_n = G$ and G_i/G_{i-1} are cyclic. Polycyclic groups are characterized as the solvable groups all of whose subgroups are finitely generated. G is **virtually nilpotent** if $\exists N \lhd G$ nilpotent such that $|G/N| < \infty$. Finitely generated virtually nilpotent groups are characterized as the groups of polynomial growth: if Λ is a finite set of generators, then $|\Lambda^n| = O(n^\alpha)$ for some α.

(2) suppose G is a group generated by two elements and $G \not\cong F_2$. There exists a periodic surface M_G with period M_0 and symmetry group $\cong G$;

(3) the e.i.r.m.'s for $h\colon T^1(M_G) \to T^1(M_G)$ consist of trivial measures and of the measures given by Theorems 2 and 3.

Remark: A theorem of B. H. Neumann says that there are uncountably many non-isomorphic groups with two generators, see e.g. [Ro].

Proof: The topological description of M_0 can be found in [Kat], page 141. It is a classical fact due to Klein that $\Gamma(2)$ is a free group on two generators. If G is generated by two elements, then there is a surjective homomorphism $H\colon \Gamma(2) \to G$, and $\Gamma := \ker(H)$ is a normal subgroup of $\Gamma(2)$. If $G \not\cong F_2$, then H is not an isomorphism, so $\Gamma \neq \{id\}$. The surface $M := \Gamma\backslash\mathbb{H}$ is then a periodic surface with symmetry group $\Gamma(2)/\ker(H) \cong \mathrm{Im}(H) = G$. Parts (2) and (3) follow. ∎

3. Generalities on Möbius transformations, Fuchsian groups and orbit cocycles

3.1. The Bowen–Series map. Fix a Fuchsian group Γ_0 s.t. $\Gamma_0\backslash\mathbb{D}$ has finite volume. Let $\mathrm{Par}(\Gamma_0)$ denote the collection of all fixed points of parabolic $g \in \Gamma_0$. Bowen and Series constructed in [BS] a countable partition $\{I_a\}_{a\in S}$ of $\partial\mathbb{D}$ into arcs with disjoint interiors, a generating set $\{g_a\}_{a\in S} \subset \Gamma_0$ and $f_{\Gamma_0}\colon \partial\mathbb{D} \to \partial\mathbb{D}$ with the following properties:

(Orb) f_{Γ_0} is (almost) orbit equivalent to Γ_0: For all except finitely many $(\xi,\eta) \in (\partial\mathbb{D})^2$, $\exists m,n > 0$ s.t. $f^m_{\Gamma_0}(\xi) = f^n_{\Gamma_0}(\eta) \Leftrightarrow \exists g \in \Gamma_0$ s.t. $\xi = g(\eta)$.

(Res) $f_{\Gamma_0}|_{\mathrm{int}(I_a)} = g_a|_{\mathrm{int}(I_a)}$ $(a \in S)$.

(Mar) $\{I_a\}_{a\in S}$ is a Markov partition: $f_{\Gamma_0}(I_a) \cap I_b \neq \varnothing \Rightarrow f(I_a) \supseteq I_b$.

(Tr) f_{Γ_0} is topologically transitive. In particular, for every $a, b \in S$ there exists some n such that $f^n_{\Gamma_0}(I_a) \supseteq I_b$.

(Fin) If $\mathrm{Par}(\Gamma_0) = \varnothing$, then S is finite. Otherwise, $\exists S_0 \subset S$ finite s.t. the forward f_{Γ_0}-orbit of every $x \in \partial\mathbb{D}\backslash\mathrm{Par}(\Gamma_0)$ enters $\bigcup_{a\in S_0} I_a\backslash\mathrm{Par}(\Gamma_0) =: \Lambda$ infinitely many times.

(BD) For any finite set S_0 as in (Fin), let $f_{S_0}\colon \Lambda \to \Lambda$ be the first return map: $f_{S_0}(x) = f^{\varphi(x)}_{\Gamma_0}(x)$ where $\varphi(x) := \min\{n \geq 1 : f^n_{\Gamma_0}(x) \in \Lambda\}$. There exists N such that $\inf|(f^N_{S_0})'| > 1$ and $\sup|f''_{S_0}/f'^2_{S_0}| < \infty$ (Adler's condition).

Every word $\underline{a} = (a_0, \ldots, a_{n-1}) \in S^n$ determines a set

$$[\underline{a}] := \bigcap_{k=0}^{n-1} f^{-k}_{\Gamma_0}(I_{a_k}).$$

This set, called a **cylinder (of length n)**, is an arc. If it is nonempty, we say that $\underline{a}$ is **admissible**.

Condition (Res) shows that any admissible word $\underline{a} = (a_0, \ldots, a_{n-1})$ determines an element $g_{\underline{a}} \in \Gamma_0$ such that $g_{\underline{a}} = f_{\Gamma_0}^{n-1}|_{[\underline{a}]} = g_{a_{n-2}} \circ \cdots \circ g_{a_0}$ if $n \geq 2$, or $g_{\underline{a}} := id$, if $n = 1$. Condition (Mar) implies that $g_{\underline{a}}$ maps $[\underline{a}]$ onto $I_{a_{n-1}}$. The content of (BD) is that if $a_0, a_{n-1} \in S_0$, then this is done with uniformly bounded distortion: There exists a modulus of continuity $\omega(\delta) \xrightarrow[\delta \to 0]{} 0$ such that for any cylinder $[\underline{a}]$,

$$(1) \qquad |\log g'_{\underline{a}}(\xi_1) - \log g'_{\underline{a}}(\xi_2)| \leq \omega(|g_{\underline{a}}(\xi_1) - g_{\underline{a}}(\xi_2)|) \quad \text{whenever } \xi_1, \xi_2 \in [\underline{a}].$$

(see [Ad]). In particular, there exists a constant B_0 (independent of $\underline{a}$ or n) s.t.

$$\frac{1}{B_0} \leq \left|\frac{g'_{\underline{a}}(x)}{g'_{\underline{a}}(y)}\right| = \left|\frac{(f_{\Gamma_0}^{n-1})'(x)}{(f_{\Gamma_0}^{n-1})'(y)}\right| \leq B_0 \quad \text{for all } x, y \in [\underline{a}].$$

Properties (Orb), (Res), (Mar) and (Tr) are proved in [BS]. Property (BD) is also proved in [BS], although it is stated there in a slightly weaker form. The proof of (Fin) is sketched in the appendix.

3.2. The Busemann function and the Poisson kernel. Define two functions $a_\theta(z_1, z_2), b_\theta(z_1, z_2)$ $(0 \leq \theta < 2\pi,\ z_1, z_2 \in \mathbb{D})$ such that

$$\omega_\theta(z_2) = h^{a_\theta(z_1, z_2)} \circ g^{b_\theta(z_1, z_2)}(\omega_\theta(z_1)).$$

The action of Möb($\mathbb{D}$) on $T^1(\mathbb{D})$ in the KAN-coordinates is then

$$(2) \quad g(e^{i\theta}, s, t) = (g(e^{i\theta}), s + b_\theta(g^{-1}o, o), t + e^{-s} a_\theta(g^{-1}o, o)) \quad (g \in \text{Möb}(\mathbb{D})).$$

The function $b_\theta(\cdot,\cdot)$ is called the **Busemann function** (some authors use this name for $-b_\theta(\cdot,\cdot)$). The function $a_\theta(\cdot,\cdot)$ is not important in our context.

The geometric meaning of the Busemann function is explained by the identity $|b_\theta(z_1, z_2)| = \lim_{s\to\infty} d(g^s\omega_\theta(z_1), g^s\omega_\theta(z_2))$. It satisfies $b_\theta(x, y) + b_\theta(y, z) = b_\theta(x, z)$ and $b_{g\cdot\theta}(g(z), g(w)) = b_\theta(z, w)$ for all $g \in$ Möb($\mathbb{D}$) (where $g \cdot \theta$ is an angle such that $g(e^{i\theta}) = e^{ig\cdot\theta}$).

We now explain the potential theoretic meaning of the Busemann function, following [Kai2] and [F]. The **harmonic measures** of $\mathbb{D}$ are

$$d\lambda_z(e^{i\theta}) = P(e^{i\theta}, z) d\lambda(e^{i\theta})$$

where λ is the normalized Haar measure of $\partial\mathbb{D}$ and $P(e^{i\theta}, z) = \frac{1-|z|^2}{|e^{i\theta}-z|^2}$ is the

Poisson kernel. The harmonic measures satisfy $\lambda_z \circ g^{-1} = \lambda_{g(z)}$ $(g \in \text{Möb}(\mathbb{D}))$.[5] The Busemann function satisfies $b_\theta(z_1, z_2) = -\log \frac{d\lambda_{z_1}}{d\lambda_{z_2}}(e^{i\theta})$.[6] In particular,

$$b_\theta(g^{-1}o, o) = -\log |g'(e^{i\theta})|. \tag{3}$$

3.3. The limit set. The **limit set** of a Fuchsian group Γ (acting on $\mathbb{D}$) is the set

$$\Lambda := \{z : z \text{ is an accumulation point of } \Gamma w \text{ for some } w \in \mathbb{D}\}.$$

The limit set is subset of $\partial\mathbb{D}$. It is characterized as the smallest closed non-empty Γ-invariant subset of $\partial\mathbb{D}$. In particular, Λ is equal to the set of accumulation points of any single Γ-orbit, and Γ acts minimally on its limit set (see [Be]).

A Fuchsian group is called **non-elementary** if its limit set contains more than two points. In this case it must be uncountable [Be]. A Fuchsian group is said to be of the **first kind** if $\Lambda = \partial\mathbb{D}$.

Any torsion free lattice in Möb($\mathbb{D}$) is of the first kind [Be]. Any non-trivial normal subgroup Γ of a group of the first kind Γ_0 is again of the first kind: The limit set of Γ is invariant under Γ_0, because $\Gamma \lhd \Gamma_0$. Since $\Gamma \neq \{id\}$, this set is non-empty, and therefore (being a closed Γ_0-invariant set), must contain the limit set of Γ_0. But this set is $\partial\mathbb{D}$, by assumption.

3.4. Translation lengths. Recall that $id \neq g \in \text{Möb}(\mathbb{D})$ is called **hyperbolic** if it has two fixed points in $\partial\mathbb{D}$. In this case one of these points is repelling, the other is attracting, and the geodesic which connects them — called the **axis** of g — is left invariant by g. A hyperbolic Möbius transformation moves the points on its axis a fixed (hyperbolic) distance, called the **translation length** of g; the translation length of g is related to the action of g on the boundary by the formula $\tau(g) := |\log |g'(p)||$, where p is one of the fixed points of g.[7]

---

5 Every $f \in C(\partial\mathbb{D})$ determines an harmonic function $F(z) = \int f d\lambda_z$ with boundary values f. Since $g \in \text{Möb}(\mathbb{D})$, $F \circ g$ is harmonic, with boundary values $f \circ g$. Thus, $\int f d\lambda_{g(z)} = F(g(z)) = \int f \circ g d\lambda_z = \int f d\lambda_z \circ g^{-1}$. Since f was arbitrary, the identity must hold.

6 The following argument is from [Kai2]: $h_\theta(z_1, z_2) := -\log \frac{d\lambda_{z_1}}{d\lambda_{z_2}}(e^{i\theta})$ satisfies $h_\theta(x, z) = h_\theta(x, y) + h_\theta(y, z)$ and $h_{g\theta}(gx, gy) = h_\theta(x, y)$ for all $g \in \text{Möb}(\mathbb{D})$. All such functions must be proportional to the Busemann function. Checking specific points we see that $h_\theta = b_\theta$.

7 One way to prove this is to note that any hyperbolic $g \in \text{Möb}(\mathbb{D})$ is conjugate to $z \mapsto kz$ $(k > 0)$ on $\{z : \text{Re}(z) > 0\}$. This isometry has translation length $|\log k|$.

Define for a torsion free Fuchsian group Γ,

$$\tau(\Gamma) := \{\tau(g) : g \in \Gamma \text{ is hyperbolic}\}.$$

This set is also called the **length spectrum** of $\Gamma \setminus \mathbb{D}$, because it is equal to the collection of lengths of closed geodesics on $\Gamma \setminus \mathbb{D}$. We need the following two properties of $\tau(\Gamma)$:

(FI) If $\Gamma_0 \backslash \mathbb{D}$ has finite volume, then $\tau(\Gamma_0)$ intersects any compact interval at most finitely many times (see §6.4). Clearly, every subgroup $\Gamma \leq \Gamma_0$ inherits this property.

(NA) If Γ is Fuchsian and non-elementary, then $\tau(\Gamma)$ generates a dense subgroup of $\mathbb{R}$ (Guivarc'h and Raugi [GR], Dal'bo [Da2]). This, in particular, is the case for non-trivial normal subgroups of lattices (which as mentioned above are groups of the first kind, whence non-elementary).

3.5. The orbit equivalence relation. Let X be a complete metric separable space, and suppose G is a countable discrete group acting continuously on X. The **orbit equivalence relation** of G is

$$\mathfrak{G} = \mathfrak{G}(G) := \{(x,y) \in X \times X : \exists g \in G \text{ s.t. } y = g(x)\}.$$

An **orbit cocycle** is a Borel function $\Phi\colon \mathfrak{G} \to \mathbb{R}$ with the cancellation property: $\Phi(x,y)+\Phi(y,z) = \Phi(x,z)$. Automatically, $\Phi(x,x) = 0$ and $\Phi(x,y) = -\Phi(y,x)$.

A **$\mathfrak{G}$-holonomy** is a bi-measurable bijection between Borel sets

$$\mathrm{dom}(\kappa), \mathrm{im}(\kappa) \subset X \text{ Borel} \quad \text{s.t. for all } x \in X,\ (x, \kappa(x)) \in \mathfrak{G}.$$

Such maps take the form $x \mapsto g_x(x)$ where $g_x \in G$ depends on x measurably. The following fact is standard: *If m is G-invariant, then $m \circ \kappa|_{\mathrm{dom}(\kappa)} = m|_{\mathrm{dom}(\kappa)}$ for all $\mathfrak{G}$-holonomies.*

More generally, let $(X, \mathcal{F})$ be a complete metric separable space with its Borel σ-algebra. A **countable Borel equivalence relation** is an equivalence relation $\mathfrak{G} \subset X \times X$ with countable equivalence classes which forms a Borel subset of $X \times X$. The $\mathfrak{G}$-holonomies are defined as before. A Borel measure on X is called **$\mathfrak{G}$-invariant** if it is invariant under all $\mathfrak{G}$-holonomies, and **$\mathfrak{G}$-ergodic** if every Borel function which is invariant under all holonomies is a.e. equal to a constant. In the case of the orbit equivalence relation of a countable group, these definitions coincide with the usual definition for ergodicity and invariance w.r.t. a group action. (In fact, any countable Borel equivalence relation is the orbit equivalence relation of some countable group of Borel automorphisms [FM].)

Suppose m is a Borel measure on $(X, \mathcal{F})$. Some care is needed in discussing 'almost everywhere' statements in $\mathfrak{G}$, because an equivalence relation usually has zero measure w.r.t. $m \times m$. A property $P(x,y)$ of pairs $(x,y) \in X \times X$ is called a **Borel property**, if $\{(x,y) \in \mathfrak{G} : P(x,y) \text{ holds}\}$ is a Borel subset of $X \times X$. A Borel property is said to **hold m-almost everywhere in $\mathfrak{G}$**, if the set

$$\{x \in X : P(x,y) \text{ holds for all } y \text{ s.t. } (x,y) \in \mathfrak{G}\}$$

has full measure. The Borel measurability of sets of this form is proved in [FM].

4. Proof of Theorem 1

4.1. Notation. Fix two Fuchsian groups Γ, Γ_0 such that $\{id\} \neq \Gamma \lhd \Gamma_0$ and Γ_0 is a lattice. Let m_0 be a Γ-invariant measure on $T^1(\mathbb{D})$ which descends to a non-trivial h-e.i.r.m. on $T^1(M)$ where $M = \Gamma \setminus \mathbb{D}$. We have already remarked that in the KAN-coordinates, any h-invariant measure is of the form $dm(e^{i\theta}, s)dt$. By (2), the measure m_0 is Γ-invariant iff the measure m is invariant under the following Γ action on $\partial\mathbb{D} \times \mathbb{R}$:

$$g\colon (e^{i\theta}, s) \mapsto (g(e^{i\theta}), s + b_\theta(g^{-1}o, o)) = (g(e^{i\theta}), s - \log|g'(e^{i\theta})|) \tag{4}$$

It is also easy to see that the condition that m_0 descends to an h-ergodic measure is equivalent to saying that m is ergodic with respect to the Γ action (4).

Abusing notation, we denote the action $g^s\colon (e^{i\theta_0}, s_0) \mapsto (e^{i\theta_0}, s_0 + s)$ by the symbol reserved for the geodesic flow, and set $H_m := \{s \in \mathbb{R} : m \circ g^s \sim m\}$. This is a closed subgroup of $\mathbb{R}$,[8] and our goal is to show that $H_m = \mathbb{R}$. This suffices, because $m \circ g^s \sim m$ iff $m_0 \circ g^s \sim m_0$.

4.2. Two Lemmas. Let $N_\varepsilon(\cdot)$ denote the ε-neighborhood of a set, and Γ, Γ_0, and m be as above. We assume throughout that m_0 projects to a non-trivial measure on $T^1(\Gamma \setminus \mathbb{D})$.

Lemma 1 (Holonomy Lemma): *Let $[\underline{a}] \subset \partial\mathbb{D}$ be a cylinder and I be a compact interval such that $m([\underline{a}] \times I) \neq 0$. For every $\tau_0 \in \tau(\Gamma)$ and $\varepsilon > 0$, there exists a $1-1$ measure preserving Borel $\overline{\kappa}$ such that $\overline{\kappa}([\underline{a}] \times I) \subset [\underline{a}] \times N_\varepsilon(I + \tau_0) \bmod m$.*

Remark: $\overline{\kappa}$ can be made a holonomy of the orbit relation of the action (4) on $\partial\mathbb{D} \times \mathbb{R}$.

---

8 $m \circ g^s \sim m \Leftrightarrow m \circ g^s \propto m$, so $H_m = \{s : \exists c \text{ s.t. } \forall f \in C_c(\partial\mathbb{D} \times \mathbb{R}), \int f \circ g^s = c \int f\}$. This can be easily checked to be a closed set.

Proof: The non-triviality of m_0 implies that $m(\mathrm{Par}(\Gamma_0)\times\mathbb{R})=0$; otherwise, by ergodicity, m is supported on a set of the form

$$\{(g(e^{i\theta_0}), s_0-\log|g'(e^{i\theta_0})|): g\in\Gamma\}$$

for some parabolic fixed point $e^{i\theta_0}$ and some $s_0\in\mathbb{R}$. This means that m_0 is carried by the Γ-images of a single horocycle whose line elements determine geodesics which terminate at $e^{i\theta_0}$. Such horocycles project to one closed horocycle on $\Gamma_0\backslash\mathbb{D}$, in contradiction to the non-triviality assumption.

Let $S_0\subset S$ be the finite set given by (Fin), and assume w.l.o.g. that S_0 contains the first symbol a_0 of $[\underline{a}]$ (otherwise add this symbol to S_0). We claim that $\exists b\in S_0$ such that the f_{Γ_0}-orbit of a.e. $\xi\in\partial\mathbb{D}$ enters I_b infinitely many times:[9] There exists certainly $b\in S_0$ such that this happens with positive measure, because, by (Fin) and the previous paragraph, a.e. orbit enters $\bigcup_{b'\in S_0} I_{b'}$ infinitely often, and this union is finite. Now, the event we are describing is f_{Γ_0}-invariant, therefore by (Orb) Γ_0-invariant, whence (since $\Gamma\subset\Gamma_0$) Γ-invariant. Since m is ergodic, this event must have full measure.

Now fix some $[\underline{a}], I, \tau_0, \varepsilon$ as in the statement. By the definition of $\tau(\Gamma)$, there is a hyperbolic element, $g\in\Gamma$ with attracting fixed point ξ^+ and repelling fixed point ξ^- such that $|g'(\xi^-)|=|g'(\xi^+)|^{-1}=e^{\tau_0}$. We may assume w.l.o.g. that $\xi^+\in\mathrm{int}(I_b)$. Otherwise, choose some $h\in\Gamma$ such that $h(\xi^+)\in\mathrm{int}(I_b)$ and work with $h\circ g\circ h^{-1}$ (such h exists because Γ is of the first kind, and such groups act minimally on $\partial\mathbb{D}$).

If the repelling fixed point of g also lies in $\mathrm{int}(I_b)$, divide I_b into two intervals I_b^+, I_b^- such that $\xi^\pm\in\mathrm{int}(I_b^\pm)$. Otherwise, set $I_b^+=I_b$, $I_b^-=\varnothing$. We can always make sure that the point p_b which separates I_b^+ from I_b^- satisfies $m(\{p_b\}\times\mathbb{R})=0$, because there are at most countably many p_b's for which this is false.

Observe that $g^{\pm1}(I_b^\pm)\subset I_b^\pm$ (any hyperbolic isometry contracts intervals which contain its attracting fixed point but not its repelling fixed point). Therefore, if

$$\gamma(\xi):=\begin{cases} g(\xi) & \xi\in I_b^+\\ g^{-1}(\xi) & \xi\in I_b^-,\end{cases}$$

then $\gamma(I_b)\subset I_b$ and $|\gamma'(\xi^\pm)|=e^{-\tau_0}$.

Fix ℓ (to be determined later) and set $[b_\pm^\ell]:=g^{\pm\ell}(I_b^\pm)$. We claim that almost every f_{Γ_0}-orbit enters $[b_+^\ell]\cup[b_-^\ell]=\gamma^\ell(I_b)$ infinitely many times.

9 More precisely: if $\Omega_b\subset\partial\mathbb{D}$ is the set of points with this property, then $m[(\Omega_b\times\mathbb{R})^c]=0$.

Assume by way of contradiction that this is not the case. In this case the function $N(\xi) := 1_{[\underline{a}]}(\xi) \max\{n : f_{\Gamma_0}^n(\xi) \in \gamma^\ell(I_b)\} \cup \{0\}$ is finite for m-a.e. (ξ, s).

By choice of b, the f_{Γ_0}-orbit of a.e. ξ enters I_b infinitely many times. Denote these times by $n_1(\xi) < n_2(\xi) < \cdots$. For every ξ, let $[\xi_0, \dots, \xi_{n_i(\xi)}]$ be the cylinder (c.f. §3.1) which contains ξ. The map $f_{\Gamma_0}^{n_i(\xi)}$ maps this cylinder onto $[\xi_{n_i(\xi)}] = I_b$, because of the Markov property of the Bowen–Series map. This allows us to define

$$\kappa_i(\xi) := (f_{\Gamma_0}^{n_i(\xi)}|_{[\xi_0,\dots,\xi_{n_i(\xi)}]})^{-1} \circ \gamma^\ell \circ f_{\Gamma_0}^{n_i(\xi)}|_{[\xi_0,\dots,\xi_{n_i(\xi)}]}(\xi) \quad (\xi \in [\xi_0, \dots, \xi_{n_i(\xi)}]).$$

For every ξ, $\kappa_i(\xi) = g_\xi(\xi)$ or $g_\xi^{-1}(\xi)$ for some $g_\xi \in \Gamma$ (which depends on i but is constant on $[\xi_0, \dots, \xi_{n_i(\xi)}]$), because of (Res) and the normality of Γ in Γ_0. Abusing notation, we define $\kappa_i'(\xi)$ to be $g_\xi'(\xi)$ or $(g_\xi^{-1})'(\xi)$ (depending on whether $\kappa_i(\xi) = g_\xi(\xi)$ or $g_\xi^{-1}(\xi)$), and define for i larger than the length of $[\underline{a}]$

$$\overline{\kappa}_i \colon (e^{i\theta}, s) \mapsto (\kappa_i(e^{i\theta}), s - \log|\kappa_i'(e^{i\theta})|).$$

(i) $\overline{\kappa}_i$ *is injective*, because κ_i is injective (it is piecewise injective and the images of the pieces are disjoint).

(ii) $\overline{\kappa}_i$ *is measure preserving*, because it is a holonomy of the orbit relation of the action of Γ on $\partial\mathbb{D} \times \mathbb{R}$.

(iii) $\exists M_0$ *such that* $\kappa_i([\underline{a}] \times I) \subset [\underline{a}] \times N_{M_0}(I)$, because the chain rule and (1) show that $|\kappa'(\xi)| = B_0^{\pm}|\gamma'(\eta)|$ for some $\eta \in I_b$, and this is uniformly bounded away from zero and infinity.

(iv) *For a.e.* $(\xi, s) \in [\underline{a}] \times I$, $\kappa_i(\xi, s) \in \{\eta \in \partial\mathbb{D} : N(\eta) \geq i\} \times N_{M_0}(I)$, because by construction, $N(\overline{\kappa}_i(\xi)) \geq n_i(\xi) \geq i$.

Now, $\{\xi \in \partial\mathbb{D} : N(\xi) \geq i\} \times N_{M_0}(I)$ is a decreasing sequence of sets whose intersection is negligible (because $N < \infty$ a.e.). These are subsets of the finite measure set $[\underline{a}] \times N_{M_0}(I)$, so their measure must tend to zero. By (iv), $(m \circ \overline{\kappa}_i)([\underline{a}] \times I) \xrightarrow[i\to\infty]{} 0$. But this contradicts (ii).

Therefore, for any ℓ, the orbit of a.e. $\xi \in [\underline{a}]$ enters $\gamma^\ell(I_b)$ infinitely often. It follows that $[\underline{a}]$ is (up to measure zero) of the form

$$[\underline{a}] = \biguplus_{i=1}^{\infty} [\underline{p}_i] \cap f_{\Gamma_0}^{-\ell_i}(\gamma^\ell I_b)$$

where $[\underline{p}_i]$ are cylinders of length $\ell_i + 1$ and $f_{\Gamma_0}^{\ell_i}[\underline{p}_i] = I_b$. Define a map κ on $[\underline{a}]$ by

$$\kappa|_{[\underline{p}_i] \cap f_{\Gamma_0}^{-\ell_i}(\gamma^\ell I_b)} = (f_{\Gamma_0}^{\ell_i}|_{[\underline{p}_i]})^{-1} \circ \gamma \circ f_{\Gamma_0}^{\ell_i}|_{[\underline{p}_i]}.$$

(i) *κ is injective and $\kappa[\underline{a}] \subset [\underline{a}]$*: Indeed, κ maps $[\underline{p}_i] \cap f_{\Gamma_0}^{-\ell_i}(\gamma^\ell I_b)$ bijectively onto $[\underline{p}_i] \cap f_{\Gamma_0}^{-\ell_i}(\gamma^{\ell+1} I_b) \subset [\underline{p}_i] \cap f_{\Gamma_0}^{-\ell_i}(\gamma^\ell I_b)$.

(ii) *κ is a holonomy of the Γ action on $\partial\mathbb{D}$*: This is because of (Res) and the normality of Γ in Γ_0.

(iii) $\sup|\log|\kappa'| + \tau_0| \xrightarrow[\ell\to\infty]{} 0$ on $[\underline{a}]$. See below.

Before checking (iii), we explain how it can be used to complete the construction. Fix, using (iii), ℓ large enough that $|\log|\kappa'| + \tau_0| < \varepsilon$. As before,

$$\overline{\kappa}\colon (e^{i\theta}, s) \mapsto (\kappa(e^{i\theta}), s - \log|\kappa'(e^{i\theta})|)$$

makes sense, is measure preserving, and maps $[\underline{a}] \times I$ into $[\underline{a}] \times N_\varepsilon(I + \tau_0)$.

We check (iii). Observe first that each $[\underline{p}_i]$ starts with a_0 and recall that $a_0 \in S_0$. By the chain rule and the fact that $|\gamma'(\xi^\pm)| = e^{-\tau_0}$, for every $\xi \in [\underline{p}_i] \cap f_{\Gamma_0}^{-\ell_i}(\gamma^\ell I_b^\pm)$ there are $\xi_1, \xi_2 \in [\underline{p}_i] \cap f^{-\ell_i}(\gamma^\ell I_b^\pm)$ (same sign for both) and $\xi_3 \in \gamma^\ell I_b$ such that

$$|\kappa'| = \frac{|(f^{\ell_i})'(\xi_1)|}{|(f^{\ell_i})'(\xi_2)|}\frac{|\gamma'(\xi_3)|}{|\gamma'(\xi^\pm)|}e^{-\tau_0}.$$

Now $|f_{\Gamma_0}^{\ell_i}(\xi_1) - f_{\Gamma_0}^{\ell_i}(\xi_2)| \le |\gamma^\ell I_b^\pm|$. Writing $\omega_\pm(\delta)$ for the moduli of continuity of $\log|\gamma'_\pm|$ and using (1) , we see that

$$|\log|\kappa'(\xi)| + \tau_0| \le \omega(|\gamma^\ell I_b^\pm|) + \omega_\pm(|\gamma^\ell I_b^\pm|)$$

(where the sign is decided according to the half of I_a which contains $f_{\Gamma_0}^{\ell_i}(\xi)$). Since $|\gamma^\ell(I_b^\pm)| \xrightarrow[\ell\to\infty]{} 0$, the result follows. ∎

Lemma 2: *For every $\xi \in \partial\mathbb{D}$, $m(\{\xi\} \times \mathbb{R}) = 0$.*

Proof: The non-triviality of m_0 implies that $m(\mathrm{Par}(\Gamma_0) \times \mathbb{R}) = 0$, because of the discussion at the beginning of the proof of Lemma 1. It is therefore enough to consider $\xi \in \partial\mathbb{D} \setminus \mathrm{Par}(\Gamma_0)$ and show $m(\{\xi\} \times \mathbb{R}) = 0$. Assume by way of contradiction that there is a $\xi \in \partial\mathbb{D} \setminus \mathrm{Par}(\Gamma_0)$ for which this is false.

Define

$$\tau(\xi) := \{\pm\tau(g) : g \in \Gamma,\ g(\xi) = \xi\} = \{\log|g'(\xi)| : g \in \Gamma,\ g(\xi) = \xi\}.$$

This is a proper definition because any $g \in \Gamma$ which fixes ξ is hyperbolic, otherwise $\xi \in \mathrm{Par}(\Gamma_0)$.

The set $\tau(\xi)$ forms a subgroup of $\mathbb{R}$. This subgroup is closed, because $\tau(\xi) \subseteq \tau(\Gamma)$, and $\tau(\Gamma)$ intersects any compact interval at most finitely many times (FI).

As mentioned in §3.4, $\tau(\Gamma)$ is not contained is a closed (proper) subgroup of $\mathbb{R}$. Therefore, $\tau(\xi) \subsetneq \tau(\Gamma)$.

Fix some $\tau_0 \in \tau(\Gamma)\setminus\tau(\xi)$, and let $\varepsilon := (1/4)d(\tau_0, \tau(\xi))$. Choose some compact interval I of length ε such that $m(\{\xi\} \times I) \neq 0$. Consider the sequence of cylinders $[\xi_0, \ldots, \xi_{n-1}]$ which contain ξ. By the holonomy lemma, there exist measure-preserving injections $\overline{\kappa}_n$ defined on $[\xi_0, \ldots, \xi_n]$ such that

$$\overline{\kappa}_n([\xi_0, \ldots, \xi_{n-1}] \times I) \subset [\xi_0, \ldots, \xi_{n-1}] \times N_\varepsilon(I + \tau_0).$$

The proof the holonomy lemma shows that we can choose $\overline{\kappa}_n$ to be of the form $(e^{i\theta}, s) \mapsto (\kappa_n(e^{i\theta}), s - \log|\kappa_n'(e^{i\theta})|)$ with κ_n piecewise hyperbolic Möbius transformation. As before, κ_n' can be defined unambiguously.

By construction, $\log|\kappa_n'(\xi)|$ is 2ε-close to $(-\tau_0)$, and therefore does not belong to $\tau(\xi)$. It follows that $\kappa_n(\xi) \neq \xi$. Since by construction $\kappa_n(\xi) \to \xi$, the set $\{\kappa_n(\xi)\}_{n\geq 1}$ is infinite.

Hence, there are infinitely many pairwise disjoint sets in the list $\{\overline{\kappa}_n(\{\xi\} \times I)\}_{n\geq 1}$. These sets have measure $m(\{\xi\} \times I) \neq 0$, because $\overline{\kappa}_n$ is measure preserving. But this is impossible, because they are all subsets of the set $\partial\mathbb{D} \times N_\varepsilon(I + \tau_0)$, and this set has finite measure because of the Radon property. ∎

4.3. Proof of Theorem 1.

Let m_0, m and H_m be as in §4.1.

Step 1: There exists a Borel measurable $u\colon \partial\mathbb{D} \to \mathbb{R}$ such that m is carried by the set $\{(e^{i\theta}, s) : s - u(e^{i\theta}) \in H_m\}$.

Proof: Let $\mathfrak{G}$ denote the orbit equivalence relation of the action of Γ on $\partial\mathbb{D}$:

$$\mathfrak{G} := \{(\xi, \eta) \in \partial\mathbb{D} \times \partial\mathbb{D} : \exists g \in \Gamma \text{ s.t. } \eta = g(\xi)\}.$$

Let $\Lambda_0 \subset \partial\mathbb{D}$ be the collection of all points which are fixed by some $id \neq g \in \Gamma$. Define $\Phi\colon \mathfrak{G} \to \mathbb{R}$ by

$$\Phi(e^{i\theta_1}, e^{i\theta_2}) := \begin{cases} b_{\theta_1}(g^{-1}o, o) & e^{i\theta_1} \notin \Lambda_0 \text{ and } e^{i\theta_2} = g(e^{i\theta_1}),\ g \in \Gamma \\ 0 & \text{otherwise.} \end{cases}$$

Using the various properties of the Busemann function, it is not difficult to see that this is a $\mathfrak{G}$-cocycle, i.e.

$$\Phi(x, y) + \Phi(y, z) = \Phi(x, z) \quad \text{for all } \mathfrak{G}\text{-equivalent } x, y, z \in \partial\mathbb{D}.$$

The set of fixed points $\Lambda_0 \times \mathbb{R}$ is Γ-invariant. It is clear that the orbit equivalence relation of Γ on $(\partial\mathbb{D} \times \mathbb{R}) \setminus (\Lambda_0 \times \mathbb{R})$ is the same as

$$\mathfrak{G}_\Phi := \{((x,s),(x',s')) : (x,x') \in \mathfrak{G} \text{ and } s'-s = \Phi(x,x')\}.$$

Since Λ_0 is countable, $m(\Lambda_0 \times \mathbb{R}) = 0$. Therefore, since m is Γ-invariant and ergodic, m is $\mathfrak{G}_\Phi$-invariant and ergodic.

The cocycle reduction theorem of [Sg] yields a Borel function u: $\partial\mathbb{D} \to \mathbb{R}$ s.t.

$$\Phi(e^{i\theta_1}, e^{i\theta_2}) + u(e^{i\theta_1}) - u(e^{i\theta_2}) \in H_m \ m\text{-a.e. in } \mathfrak{G}_\Phi.$$

This implies that F: $\partial\mathbb{D} \times \mathbb{R} \to \mathbb{R}/H_m$, $F(e^{i\theta}, s) := s - u(e^{i\theta}) + H_m$ is $\mathfrak{G}_\Phi$-invariant, and therefore (by ergodicity) constant almost everywhere. It now remains to modify u by a constant to ensure that $F = H_m$ almost everywhere.

STEP 2: The function $u(e^{i\theta})$ of the previous step can be made essentially bounded.

Proof: H_m is a closed subgroup of $\mathbb{R}$, so it is either equal to $\mathbb{R}$, $c\mathbb{Z}$ or $\{0\}$. In the first case there is nothing to prove. In the second case, one can choose u with range in $[0, c)$. It remains to treat the case $H_m = \{0\}$. In this case, m is supported on the graph of u, $\{(e^{i\theta}, u(e^{i\theta})) : 0 \le \theta < 2\pi\}$. We claim that u is then automatically essentially bounded.

Assume by way of contradiction that $\operatorname{ess\,sup}|u| = \infty$. In this case there are intervals I, J, $\tau_0 \in \tau(\Gamma)$, and $\varepsilon > 0$ s.t. $m(\partial\mathbb{D}\times I) \neq 0, m(\partial\mathbb{D}\times J) \neq 0$, $I \cap J = \varnothing$ and $N_\varepsilon(I + \tau_0) \subset J$.[10]

Define two measures on $\partial\mathbb{D}$ by $\mu_I(E) := m(E \times I)$, $\mu_J(E) = m(E \times J)$. These measures are mutually singular: Indeed, $s = u(e^{i\theta})$ m-a.e., so $u(e^{i\theta}) \in I$ μ_I-a.e. and $u(e^{i\theta}) \in J$ μ_J-a.e. Since $I \cap J = \varnothing$, $\mu_I \perp \mu_J$.

Since $\mu_I \perp \mu_J$, there exists some cylinder $[\underline{a}]$ such that $\mu_I[\underline{a}] > 2\mu_J[\underline{a}]$: Indeed, the collection of all Borel sets which satisfy the opposite inequality is a monotone class. If it contains all cylinders, then it must contain all Borel sets (because the cylinders generate the Borel sets). But this implies that $\mu_I \le 2\mu_J$ in contradiction to $\mu_I \perp \mu_J$.

10 To see this pick some $\tau \in \tau(\Gamma)$ and observe that the partition $\partial\mathbb{D} \times \mathbb{R} = \biguplus_{k\in\mathbb{Z}} \partial\mathbb{D} \times [k\tau, k\tau + \tau)$ contains at least two non-adjacent 'tiles' which carry some measure (otherwise m is supported inside a bounded set, which is impossible since m is carried by the graph an unbounded function). If these intervals are $[k_i\tau, k_i\tau + \tau)$ where $k_1 < k_2$, then take $\varepsilon = \tau/2$, $I = [k_1\tau, k_1\tau + \tau)$, $J = N_\varepsilon([k_2\tau, k_2\tau + \tau))$ and $\tau_0 := (k_2 - k_1)\tau$ (this is the translation length of the $(k_2 - k_1)$-power of the Γ-isometry with translation length τ).

By the definition of $[\underline{a}]$, μ_I and μ_J, $m([\underline{a}] \times I) > 2m([\underline{a}] \times J)$. We now obtain a contradiction: Let $\overline{\kappa}$ be a measure preserving injection

$$\overline{\kappa}\colon [\underline{a}] \times I \hookrightarrow [\underline{a}] \times N_\varepsilon(I + \tau_0) \subseteq [\underline{a}] \times J.$$

Then $2m([\underline{a}] \times J) < m([\underline{a}] \times I) = (m \circ \overline{\kappa})([\underline{a}] \times I) \leq m([\underline{a}] \times J)$, a contradiction.

Step 3: After the change of coordinates $\vartheta(e^{i\theta}, s) = (e^{i\theta}, s - u(e^{i\theta}))$, m takes the form $dm \circ \vartheta^{-1} = e^{\lambda s} d\nu(e^{i\theta}) dm_{H_m}(s)$ where $\lambda \in \mathbb{R}$, m_{H_m} is Haar's measure on H_m, and ν is a finite measure on $\partial\mathbb{D}$ which is equivalent to a Γ-ergodic Γ-conformal measure with parameter λ.

Proof: We have seen that m is supported on $\{(e^{i\theta}, s) : s - u(e^{i\theta}) \in H_m\}$ with $u(e^{i\theta})$ Borel. It follows that $m \circ \vartheta^{-1}$ is carried by $\partial\mathbb{D} \times H_m$. If we choose an essentially bounded version of u, then $m \circ \vartheta^{-1}$ is Radon.

Since m is ergodic and g^s commutes with the Γ-action, $m \circ g^s$ is also Γ-ergodic and invariant. It is therefore either proportional to m, or singular w.r.t. m. It follows from the definition of H_m that $\exists \lambda$ such that for all $s \in H_m$, $m \circ g^s = e^{\lambda s} m$. Since ϑ and g^s commute, we also have $m \circ \vartheta^{-1} \circ g^s = e^{\lambda s} m \circ \vartheta^{-1}$ for all $s \in H_m$. Consequently, $e^{-\lambda s} dm \circ \vartheta^{-1}$ is invariant w.r.t. translations in H_m. It is not difficult to deduce from this and the fact that $e^{-\lambda s} m \circ \vartheta^{-1}$ is supported in $\partial\mathbb{D} \times H_m$, that $e^{-\lambda s} dm \circ \vartheta^{-1} = \nu \times m_{H_m}$ with some measure ν on $\partial\mathbb{D}$.

This measure must be finite, because $m \circ \vartheta^{-1}$ is Radon. For every $g \in \Gamma$, m is g-invariant, and therefore $m \circ \vartheta^{-1}$ is $\vartheta \circ g \circ \vartheta^{-1}$-invariant. Comparing this with the formula, $m \circ \vartheta^{-1} = e^{\lambda s} \nu \times m_{H_m}$ we see that

$$\frac{d\nu \circ g}{d\nu}(e^{i\theta}) = |g'|^\lambda \frac{e^{-\lambda u}}{e^{-\lambda u \circ g}}.$$

Therefore $e^{\lambda u} \nu$ is Γ-conformal with parameter λ (this is a finite measure because $\operatorname{ess\,sup} |u| < \infty$ and $\nu(\partial\mathbb{D}) < \infty$.)

Step 4: $H_m = \mathbb{R}$, which proves the theorem.

Proof: Assume by way of contradiction that $H_m \neq \mathbb{R}$. Since this is a closed subgroup of $\mathbb{R}$, $H_m = c\mathbb{Z}$ for some c.

By the theorem of Guivarc'h, Raugi and Dal'bo, mentioned in §3.4, $\tau(\Gamma)$ generates a dense subgroup of $\mathbb{R}$, and therefore there must be some $\tau_0 \in \tau(\Gamma) \setminus c\mathbb{Z}$. Set $\varepsilon_0 := \frac{1}{2} d(\tau_0, c\mathbb{Z})$, and fix some u_0 such that $A := \{e^{i\theta} : |u(e^{i\theta}) - u_0| < \frac{\varepsilon_0}{6}\}$ has positive measure. We construct a Borel set A_0 and $\mathfrak{G}$-holonomy κ such that $A_0 \subseteq A$ and $\nu(E) \neq 0$, where

$$E := A_0 \cap \kappa^{-1} A_0 \cap [|\Phi(\xi, \kappa\xi) + u(\xi) - u(\kappa\xi) - \tau_0| < \varepsilon_0].$$

Sets of this form appear in the theory of essential values (see [Sch], [Kai1]).

Before constructing A_0, we show how to use its existence to derive the contradiction which proves the step. If $\overline{\kappa}(e^{i\theta}, s) = (\kappa(e^{i\theta}), s + \Phi(x, \kappa x))$, then

$$(\vartheta \circ \overline{\kappa} \circ \vartheta^{-1})(E \times \{0\}) \subseteq \partial\mathbb{D} \times N_{\varepsilon_0}(\tau_0) \subset (\partial\mathbb{D} \times H_m)^c.$$

Since $\vartheta \circ \overline{\kappa} \circ \vartheta^{-1}$ preserves the measure $m \circ \vartheta^{-1}$,

$$0 \neq \nu(E) = (m \circ \vartheta^{-1})(E \times \{0\}) \leq (m \circ \vartheta^{-1})[(\partial\mathbb{D} \times H_m)^c] = 0,$$

a contradiction.

THE CONSTRUCTION OF A_0: Fix $\varepsilon > 0$, to be determined later. There exists a cylinder of positive measure $[\underline{a}]$ such that $\nu(A \cap [\underline{a}]) \geq (1-\varepsilon)\nu[\underline{a}]$: Indeed, $\exists U \supseteq A$ open with $\nu(A) \geq (1-\varepsilon)\nu(U)$ (regularity of Borel measures). Now ν has no atoms (because $m(\{\xi\} \times \mathbb{R}) = 0$ for all $\xi \in \partial\mathbb{D}$). Therefore, every open set is a countable disjoint union of cylinders up to a set of measure zero. One of these sets must satisfy the desired inequality.

By the choice of u_0, $m([\underline{a}] \times N_{\varepsilon_0/6}(u_0)) \neq 0$. Construct a $\mathfrak{G}_\Phi$-holonomy $\overline{\kappa}$ s.t.

$$\overline{\kappa}([\underline{a}] \times N_{\varepsilon_0/6}(u_0)) \subseteq [\underline{a}] \times N_{\varepsilon_0/3}(u_0 + \tau_0).$$

Any $\mathfrak{G}_\Phi$-holonomy is of the form $(\xi, s) \mapsto (\kappa\xi, s + \Phi(\xi, \kappa\xi))$ where κ is a $\mathfrak{G}$-holonomy. We must have

$$\kappa[\underline{a}] \subseteq [\underline{a}] \quad \text{and} \quad |\Phi(x, \kappa x) - \tau_0| < \varepsilon_0/2.$$

Set $A_0 := [\underline{a}] \cap A$. We claim that if ε is small enough, then $\nu(A_0 \cap \kappa^{-1}A_0) \neq \varnothing$. We begin with an estimate of the Radon–Nikodym derivative of κ on A_0. The Γ-invariance of m is equivalent to its $\mathfrak{G}_\Phi$-invariance, and this translates to the $\mathfrak{G}_{\Phi_u}$-invariance of $m \circ \vartheta^{-1}$, where $\Phi_u(\xi_1, \xi_2) := \Phi(\xi_1, \xi_2) + u(\xi_1) - u(\xi_2)$. Since $m \circ \vartheta^{-1} = e^{\lambda s}\nu \times m_{H_m}$, this forces $((d\nu \circ \kappa)/d\nu)(\xi) = e^{-\lambda\Phi_u(\xi,\kappa\xi)} = e^{\pm 2\|\lambda u\|_\infty} e^{-\lambda\Phi(\xi,\kappa\xi)}$. Therefore, on $A_0 \subset [\underline{a}]$

$$(d\nu \circ \kappa)/d\nu \geq e^{-2\|\lambda u\|_\infty - \frac{|\lambda|\varepsilon_0}{2} - |\lambda|\tau_0} =: \delta_0.$$

It follows that $\nu(\kappa A_0) = \int_{A_0}((d\nu \circ \kappa)/d\nu)d\nu \geq \delta_0\nu(A_0) \geq \delta_0(1-\varepsilon)\nu[\underline{a}]$, since by construction $\nu(A_0) \geq (1-\varepsilon)\nu[\underline{a}]$. It follows that

$$\nu(A_0) + \nu(\kappa A_0) \geq (1-\varepsilon)(1+\delta_0)\nu[\underline{a}] \xrightarrow[\varepsilon \to 0]{} (1+\delta_0)\nu[\underline{a}],$$

so we can choose ε small enough such that the left hand side is strictly larger than $\nu[\underline{a}]$. But $A_0, \kappa(A_0) \subseteq [\underline{a}]$, so necessarily $\nu(A_0 \cap \kappa(A_0)) \neq 0$. Since κ is non-singular, $\nu(A_0 \cap \kappa^{-1}(A_0)) = \nu \circ \kappa^{-1}(A_0 \cap \kappa(A_0)) \neq 0$.

Finally, we observe that if $\xi \in A_0 \cap \kappa^{-1} A_0$, then $\xi, \kappa(\xi) \in A = [|u-u_0| < \varepsilon_0/6]$ and so $|\Phi(\xi,\kappa\xi) + u(\xi) - u(\kappa\xi) - \tau_0| \le |\Phi(\xi,\kappa\xi) - \tau_0| + |u-u_0| + |u_0 - u\circ\kappa| < \varepsilon_0$. It follows that $\nu(E) \ne 0$. ∎

5. Proof of Theorems 2 and 3

5.1. λ-Potential Theory. It is known that

$$\Delta_{\mathbb{D}} P(e^{i\theta}, z)^\alpha = \alpha(\alpha-1) P(e^{i\theta}, z)^\alpha \quad \text{for all } 0 \le \theta < 2\pi.^{11}$$

It turns out that if $\alpha \ge 1/2$, then this is a complete family of minimal eigenfunctions for the eigenvalue $\alpha(\alpha-1)$ ([Kar], [Su1]):

Theorem 4 (Karpelevich): *Any positive eigenfunction $F\colon \mathbb{D} \to \mathbb{R}$ of $\Delta_{\mathbb{D}}$ has eigenvalue $\lambda \ge -1/4$, and admits a unique representation of the form*

$$F(z) = \int_{\partial\mathbb{D}} P(e^{i\theta}, z)^\alpha d\nu(e^{i\theta})$$

where ν is a finite measure on $\partial\mathbb{D}$, $\alpha(\alpha-1) = \lambda$ and $\alpha \ge 1/2$. Any (positive) finite Borel measure on $\partial\mathbb{D}$ arises this way.

The following lemma is from [Su1] (see also [Ba]):

Lemma 3: *Let ν be a finite Borel measure on $\partial\mathbb{D}$, and set $dm = e^{\alpha s} d\nu(e^{i\theta}) ds dt$, $F(z) := \int_{\partial\mathbb{D}} P(e^{i\theta}, z)^\alpha d\nu(e^{i\theta})$. If $g \in$ Möb$(\mathbb{D})$ acts on $T^1(\mathbb{D})$ by (2) and on $\partial\mathbb{D}$ and $\mathbb{D}$ in the standard way, then*

(1) $\frac{d\nu\circ g}{d\nu} = |g'|^\alpha \iff m \circ g = m$;

(2) $\frac{d\nu\circ g}{d\nu} = |g'|^\alpha \Longrightarrow F \circ g = F$, *and if $\alpha \ge \frac{1}{2}$ then this is an $\Leftrightarrow$.*

Proof: By (3), $\frac{dm\circ g}{dm} = e^{\alpha b_\theta(g^{-1}o, o)} \frac{d\nu\circ g}{d\nu} = |g'(e^{i\theta})|^{-\alpha} \frac{d\nu\circ g}{d\nu}$. This proves part (1).

To prove part (2), we use the harmonic measures λ_z from §3.2. Writing for $g \in$ Möb$(\mathbb{D})$, $P(e^{i\theta}, gz) d\lambda \equiv \lambda_{gz} = \lambda_z \circ g^{-1} \equiv P(g^{-1}e^{i\theta}, z) d\lambda \circ g^{-1}$, we see that

$$(5) \qquad |(g^{-1})'(e^{i\theta})| = \frac{d\lambda \circ g^{-1}}{d\lambda}(e^{i\theta}) = \frac{P(e^{i\theta}, gz)}{P(g^{-1}e^{i\theta}, z)} \quad \text{for all } g \in \text{Möb}(\mathbb{D}).$$

---

11 $f(z) = \text{Im}(z)^\alpha$ is a $\alpha(\alpha-1)$-eigenfunction of $\Delta_{\mathbb{H}} = y^2(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2})$. Now $\varphi(z) = i\frac{1+z}{1-z}$ maps $\mathbb{D}$ isometrically onto $\mathbb{H}$, so $f \circ \varphi$ is a $\alpha(\alpha-1)$-eigenfunction of $\Delta_{\mathbb{D}}$. Calculating, we see that $f \circ \varphi = \text{Re}[\frac{1+z}{1-z}]^\alpha = P(1, z)^\alpha$, so $P(e^{i\theta}, \cdot)^\alpha$ is a $\alpha(\alpha-1)$-eigenfunction for $\theta = 0$. But for every θ $\exists \varphi_\theta \in$ Möb$(\mathbb{D})$ s.t. $P(e^{i\theta}, \cdot) = P(1, \cdot) \circ \varphi_\theta$ so $P(e^{i\theta}, \cdot)^\alpha$ is a $\alpha(\alpha-1)$-eigenfunction for all θ.

It follows that

$$\begin{aligned}F(gz) &= \int_{\partial\mathbb{D}} P(e^{i\theta}, gz)^\alpha d\nu(e^{i\theta}) = \int_{\partial\mathbb{D}} P(g^{-1}e^{i\theta}, z)^\alpha |(g^{-1})'|^\alpha d\nu(e^{i\theta}) \\ &= \int_{\partial\mathbb{D}} P(e^{i\theta}, z)^\alpha |(g^{-1})' \circ g|^\alpha d\nu \circ g(e^{i\theta}) \\ &= \int_{\partial\mathbb{D}} P(e^{i\theta}, z)^\alpha |g'|^{-\alpha} \frac{d\nu \circ g}{d\nu} d\nu(e^{i\theta})\end{aligned}$$

Comparing this with $F(z) = \int_{\partial\mathbb{D}} P(e^{i\theta}, z)^\alpha d\nu(e^{i\theta})$ we see (by the uniqueness part of Theorem 4) that when $\alpha \geq 1/2$, $F \circ g = F$ iff $\frac{d\nu\circ g}{d\nu} = |g'|^\alpha$. ∎

5.2. Proof of Theorem 2. We divide the proof into two parts:

Part 1: Any non-trivial h-e.i.r.m. lifts to a measure of the form $e^{\alpha s} d\nu ds dt$, where ν is non-atomic, Γ-ergodic and Γ-conformal with parameter α.

Proof: Let m_0 be a Γ-invariant measure on $T^1(\mathbb{D})$ which descends to a non-trivial h-e.i.r.m. on $T^1(\Gamma\backslash\mathbb{D})$. By Theorem 1, m_0 is quasi-invariant under the geodesic flow. As explained in the introduction, this forces m_0 to take the following form in the KAN-coordinates: $dm_0(e^{i\theta}, s, t) = e^{\alpha s} d\nu(e^{i\theta}) ds dt$.

Since m_0 is Radon, ν is finite. Lemma 2 shows that ν is non-atomic. Since m_0 is Γ-invariant, ν is Γ-conformal with parameter α (Lemma 3). Finally, ν is ergodic under the action of Γ on $\partial\mathbb{D}$: Any $F(e^{i\theta})$ is h-invariant on $T^1(\mathbb{D})$ (it is independent of t in the KAN-coordinates). If it is Γ-invariant, then it descends to an h-invariant function on $T^1(\Gamma\backslash\mathbb{D})$, and therefore must be constant (ergodicity on $T^1(\Gamma\backslash\mathbb{D})$).

Part 2: If ν is non-atomic, Γ-ergodic and Γ-conformal measure with parameter α on $\partial\mathbb{D}$, then $dm_0 := e^{\alpha s} d\nu ds dt$ descends to a non-trivial h-e.i.r.m. measure on $T^1(\Gamma\backslash\mathbb{D})$.

Proof: As before m_0 is h-invariant and Γ-invariant, and therefore descends to an h-invariant Radon measure on $T^1(\Gamma \setminus \mathbb{D})$. This measure is non-trivial, otherwise ν would have to be supported on $\mathrm{Par}(\Gamma)$ and would therefore have to be atomic. But h-ergodicity is not clear.

It is enough to show that $d\mu := e^{\alpha s} d\nu ds$ is ergodic w.r.t. the action (4) of Γ on $\partial\mathbb{D} \times \mathbb{R}$. Indeed, any h-invariant function on $T^1(\mathbb{D})$ is of the form $F(e^{i\theta}, s)$, and this descends to a function on the surface iff F is invariant under the action (4).

Observe that μ is Γ-invariant. Let $d\mu = \int_Y \mu_y d\pi(y)$ be the ergodic decomposition of μ w.r.t. the Γ-action. For a.e. y, μ_y is a Γ-invariant Radon measure on $\partial\mathbb{D} \times \mathbb{R}$. Consequently, $m_y = \mu_y \times dt$ is a Γ-invariant Radon measure on $(\partial\mathbb{D} \times \mathbb{R}) \times \mathbb{R} \simeq T^1(\mathbb{D})$, and therefore descends to an h-invariant measure on $T^1(\Gamma \setminus \mathbb{D})$. This measure is h-ergodic, because of the ergodicity of μ_y.

It is also non-trivial for a.e. y. Otherwise, there would be a positive measure set of y's for which $m_y(\mathrm{Par}(\Gamma_0) \times \mathbb{R} \times \mathbb{R}) \neq 0$. This can only happen if $m_0(\mathrm{Par}(\Gamma_0) \times \mathbb{R} \times \mathbb{R}) \neq 0$, in which case $\nu[\mathrm{Par}(\Gamma_0)] \neq 0$. But this is impossible, because $\mathrm{Par}(\Gamma_0)$ is countable, and ν is non-atomic.

We may now appeal to Part 1 and see that $m_y = e^{\alpha_y s} d\nu_y ds dt$, where ν_y is a Γ-ergodic and Γ-conformal measure of parameter α_y. It follows that $\mu_y = e^{\alpha_y s} d\nu_y ds$. The identity

$$\mu = e^{-\alpha s_0} \mu \circ g^{s_0} = \int_Y e^{(\alpha_y - \alpha)s_0} \mu_y d\pi(y),$$

in the limit $s_0 \to \pm\infty$ shows that $\alpha_y = \alpha$ for π-a.e. $y \in Y$. Consequently, almost all the ν_y's are Γ-conformal with parameter α. But $\nu = \int_Y \nu_y d\pi(y)$ and ν was assumed to be ergodic, so almost all the ν_y must be equal (uniqueness of the ergodic decomposition [Sch]). It follows that almost all the μ_y are equal, and this can only happen if μ itself is ergodic.

By the discussion at the beginning of the proof, this implies that m_0 descends to an h-ergodic measure. ∎

5.3. Proof of Theorem 3. Now that Theorems 1 and 2 are proved, we can simply follow that argument of [Ba], making the suitable adjustments from the nilpotent case discussed there to the general case.

We start with some general comments on non-trivial normal subgroups Γ of lattices Γ_0 in $\mathrm{M\ddot{o}b}(\mathbb{D})$. Any Γ-conformal measure has parameter larger than or equal to $\delta(\Gamma)$, the critical exponent of the Poincaré series of Γ (Sullivan [Su1], Theorem 2.19). If $\{id\} \neq \Gamma \lhd \Gamma_0$, then $\delta(\Gamma) \geq (1/2)\delta(\Gamma_0)$ (Roblin [Rob], Theorem 2.2.1). The critical exponent of a lattice is equal to 1 ([Su1], Theorem 2.17). Therefore, *any* Γ*-conformal measure has parameter* $\geq 1/2$.

Theorem 2 says that every non-trivial h-e.i.r.m. is of the form $e^{\alpha s} d\nu(e^{i\theta}) ds dt$ with ν non-atomic Γ-conformal and ergodic with parameter α. $F_m(z)$ defined by $(*)$ is a well-defined $\alpha(\alpha-1)$-eigenfunction of $\Delta_{\mathbb{D}}$ (Theorem 4). By Lemma 3 part (2), it is Γ-invariant, and therefore descends to an $\alpha(\alpha-1)$ eigenfunction on $M = \Gamma \setminus \mathbb{D}$.

We claim that this eigenfunction (which we also denote by F_m) is minimal. Suppose F_m dominates another positive $\alpha(\alpha-1)$-eigenfunction F. Then F_m is the average of the two positive eigenfunctions $F_m \pm F$. If $\nu_\pm$ are the Γ-conformal finite measures on $\partial\mathbb{D}$ which represent these functions as in Theorem 4, then $\frac{1}{2}(\nu_+ + \nu_-)$ is another representation of F_m. But the representing measure of F_m is unique (since $\alpha \geq \frac{1}{2}$), so $\nu = \frac{1}{2}(\nu_+ + \nu_-)$. The Γ-ergodicity of ν forces $\nu_\pm$ to be proportional, so $\exists c > 0$ s.t. $F = cF_m$, proving the minimality of F_m.

This shows that $(*)$ is a well-defined map from the collection of h-e.i.r.m. into the collection of minimal eigenfunctions of Δ_M. This map is an injection because of Theorem 4 and the inequality $\alpha \geq 1/2$.

To see that it is a surjection, start with a minimal non-trivial eigenfunction of eigenvalue λ, and let F be its lift to an eigenfunction on $\mathbb{D}$. Write $F(z) = \int_{\partial\mathbb{D}} P(e^{i\theta}, z)^\alpha d\nu(e^{i\theta})$ where $\lambda = \alpha(\alpha - 1)$, $\alpha \geq 1/2$ and ν is some finite measure on $\partial\mathbb{D}$. By Lemma 3 part (2), ν is Γ-conformal with parameter α. Now ν must be Γ-ergodic, otherwise F is not minimal. It follows from Theorem 2 part (1) that $dm = e^{\alpha s} d\nu ds dt$ is a non-trivial h-e.i.r.m. such that $F_m = F$.

We have established the bijection $(*)$ proclaimed in Theorem 3. Property (1) in the statement of this theorem can be checked by direct computation. Property (2) is proved by realizing the deck transformations as elements of Γ_0 (every coset of Γ corresponds to one deck transformation), and this theorem is proved by proceeding as in Lemma 3. ∎

6. Appendix: Proof of some auxiliary results

6.1. Classification of positive eigenfunctions for surfaces of finite area. Let Γ be a lattice in Möb($\mathbb{D}$), and set $M := \Gamma\backslash\mathbb{D}$. We know from §5.3 that the positive eigenfunctions of Δ on $\Gamma\backslash\mathbb{D}$ have eigenvalue $\alpha(\alpha - 1)$ with $\alpha \geq \delta(\Gamma)$, where $\delta(\Gamma)$ is the critical exponent of Γ. The critical exponent of a lattice is equal to one; therefore all the relevant eigenvalues are non-negative. We classify them (see Ji and MacPherson [JM] for a more complicated proof which works in more general situations).

Step 1: Every positive eigenfunction with eigenvalue zero is constant.

Proof: Let $F(z)$ be a positive function such that $\Delta_M F = 0$. Fix some $p \in M$, and denote by B_t the Brownian motion on M started at p. It is a standard fact that $F(B_t)$ is a martingale. Consequently, $F(B_t)$ converges almost surely. On the other hand, it is known that the Brownian motion on a surface of finite area is ergodic and recurrent [Su3]; therefore $F(z)$ must be constant.

Step 2: The number of minimal positive eigenfunctions with a fixed positive eigenvalue is equal to the number of the cusps. These eigenfunctions are trivial.

Proof: Denote the cusps of M by $C_1, \ldots, C_N$. Fix $\lambda > 0$, we construct for every i a trivial λ-eigenfunction E_i which tends to infinity at C_i and to zero at C_j $(j \neq i)$ (compare with the spectral Eisenstein series on the modular surface [Sk]).

Working in the upper half plane, we assume without loss of generality, that C_i is at infinity (otherwise pass to a conjugate of Γ). Let $\Gamma_i \subset \Gamma$ be the stabilizer of ∞. This is an infinite cyclic group of the form $\Gamma_i := \{z \mapsto z + kb : k \in \mathbb{Z}\}$, with b real. Let $s > 1$ be the solution larger than one of $s(s-1) = \lambda$. Noting that Γ_i preserves the imaginary part, we define $\mathrm{Im}[\Gamma_i \gamma \cdot z] := \mathrm{Im}[\gamma(z)]$, and set

$$E_i(z) := \sum_{\Gamma_i \gamma \in \Gamma_i \backslash \Gamma} [\mathrm{Im}(\Gamma_i \gamma \cdot z)]^s.$$

The series converges absolutely (see §1.4 in [Sk]) and:

(1) *E_i is a Γ-invariant positive λ-eigenfunction*, because of Γ-equivariance and $\Delta_{\mathbb{H}}(\mathrm{Im}\, z)^s = [y^2(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2})]y^s = \lambda(\mathrm{Im}\, z)^s$.

(2) *E_i is trivial and minimal.* The map $\varphi(z) = i\frac{1+z}{1-z}$ maps $\mathbb{D}$ onto $\mathbb{H}$, $\varphi(1) = \infty$ and $\mathrm{Im}[\varphi(z)] = P(1, z)$. By (5), if $\Gamma^{\mathbb{D}} := \varphi^{-1}\Gamma\varphi$ and $\Gamma_i^{\mathbb{D}} := \varphi^{-1}\Gamma_i\varphi$, then

$$E_i \circ \varphi = \sum_{\Gamma_i^{\mathbb{D}} g \in \Gamma_i^{\mathbb{D}} \backslash \Gamma^{\mathbb{D}}} |g'(1)|^s P(g(1), z)^s = \int_{\partial\mathbb{D}} P(e^{i\theta}, z)^s d\nu(e^{i\theta}),$$

where ν is supported on $\Gamma^{\mathbb{D}}1$. Since 1 corresponds to a cusp, E_i is trivial. Since ν is $\Gamma^{\mathbb{D}}$-ergodic, E_i is minimal [Ba].

(3) *$E_i(z)$ converges to infinity as $z \to C_i$ and to zero as $z \to C_j$ for $j \neq i$.* See Corollary 3.5 in [Iw].

(4) *For every Γ-invariant positive λ-eigenfunction F, $F(z) = O(E_i(z))$ as $z \to C_i$.* Karpelevich's Theorem and $P(e^{i\theta}, z) \leq P(1, |z|) = \mathrm{Im}\, \varphi(|z|)$ give $(F \circ \varphi)(z) \leq F(\varphi(0))[\mathrm{Im}\, \varphi(|z|)]^s$ $(z \in \mathbb{D})$. Setting $w = \varphi(z)$, $F_0 := F(\varphi(0))$, we see that $F(w) \leq F_0[\mathrm{Im}\, \varphi(|\varphi^{-1}(w)|)]^s$ $(w \in \mathbb{H})$. Noting that $w \in i\mathbb{R}^+ \Rightarrow \varphi^{-1}(w)$ is positive and real, we see that

$$F(w) \leq F_0[\mathrm{Im}\, w]^s \leq F_0 \cdot E_i(w) \quad \text{for all } w \in i\mathbb{R}^+.$$

Now, any $\{\Gamma z_n\}_{n \geq 1}$ which tends to C_i is within hyperbolic distance $o(1)$ from some $\{\Gamma w_n\}_{n \geq 1}$ with w_n imaginary. By Harnack's inequality, $F(\Gamma z_n) \sim F(\Gamma w_n) \leq F_0 E_i(\Gamma w_n) \sim F_0 E_i(\Gamma z_n)$.

We now show that any positive λ-eigenfunction F is a convex combintation of $E_1, \ldots, E_N$. Assume first that F tends to infinity at each of the cusps.

Every cusp C_i is encircled by a one parameter family of closed horocycles. Parametrize these horocycle by $H_i(r)$ in such a way that $H_i(r)$ converge to C_i as $r \to \infty$ (in the coordinate system of the first paragraph, $H_i(r) = \{\Gamma z : z = x + ir, x \in \mathbb{R}\}$). Let Ω_r be the the domain obtained from M by cutting the cusps away at $H_i(r)$, $i = 1, \ldots, N$.

The hyperbolic length of $H_i(r)$ tends to zero as r tends to infinity. By Harnack's inequality, there exists $\varepsilon(r) \xrightarrow[r\to\infty]{} 0$ such that for every positive λ-eigenfunction h $h(z) = e^{\pm\varepsilon(r)}h(w)$ for all $z, w \in H_i(r)$. In particular, $\exists F_j(r), E_{ij}(r)$ such that

$$F = e^{\pm\varepsilon(r)}F_j(r), E_i = e^{\pm\varepsilon(r)}E_{ij}(r) \quad \text{on } H_j(r).$$

Define $\alpha_i(r) := F_i(r)/E_{ii}(r)$ and $\delta(r) := \max\{E_{ij}(r)/F_i(r) : j \neq i\}$. As $r \to \infty$ $\alpha_i(r) = O(1)$ because $F = O(E_i)$, and $\delta(r) = o(1)$, because $F \to \infty$ and $E_j \to 0$ at C_j. Thus,

$$\sum_{j=1}^{N} \alpha_j(r)E_{ji}(r) = \left[1 + \sum_{j\neq i} \alpha_j(r)\frac{E_{ji}(r)}{F_i(r)}\right]F_i(r) = [1 + O(\delta(r))]^{\pm 1}F_i(r).$$

It follows that $F(z) = [1 + o(1)]^{\pm 1}\sum_{j=1}^{N} \alpha_j(r)E_j(z)$ on $\partial\Omega_r$ uniformly in z.

This implies that $F(z) = [1 + o(1)]^{\pm 1}\sum_{j=1}^{N} \alpha_j(r)E_j(z)$ on $\overline{\Omega}_r$ uniformly in z, because of the following general fact:

$$\left.\begin{array}{c} f_1, f_2 \text{ are positive on } \Gamma\backslash\mathbb{H} \\ \Delta_{\mathbb{H}}f_1 = \lambda f_1,\ \Delta_{\mathbb{H}}f_2 = \lambda f_2 \text{ on } \Gamma\backslash\mathbb{H} \\ f_1 \leq f_2 \text{ on } \partial\Omega_r \end{array}\right\} \Rightarrow f_1 \leq f_2 \text{ on } \overline{\Omega}_r. \tag{6}$$

The proof of (6): Karpelevich's Theorem implies that f_1, f_2 are $C^2(\overline{\Omega}_r)$. Therefore $u := f_1 - f_2$ attains its maximum on $\overline{\Omega}_r$ at some point z_0. We claim that $u(z_0) \leq 0$ (proving that $f_1 \leq f_2$ on Ω_r). Otherwise, $u(z_0) > 0$ and z_0 must be in the interior of Ω_r. In the upper half plane model, this implies that $0 < \lambda u(z_0) = \text{Im}(z_0)^2[u_{xx}(z_0) + u_{yy}(z_0)]$ and so at least one of u_{xx}, u_{yy} is positive at z_0. But this is impossible, because z_0 is a point of local maximum.

Since $\alpha_i(r)$ are positive and uniformly bounded, there exists $r_n \to \infty$ such that $\alpha_i(r_n)$ converges as $n \to \infty$, say to α_i. Passing to this limit, we see that

$$F(z) = \sum_{i=1}^{N} \alpha_i E_i(z) \quad \text{on } \bigcup_{n=1}^{\infty} \Omega_{r_n} = M.$$

This proves that any positive eigenfunction which explodes at the cusps is a linear combination with non-negative coefficients of $E_1, \ldots, E_N$.

For a general positive λ-eigenfunction F, we argue as follows: The function $F_0 := F + \sum_{i=1}^N E_i$ explodes at the cusps, and is therefore a linear combination of the E_i's. We use this fact to write $F(z) = \sum_{i=1}^N (\alpha_i - 1)E_i(z)$ for some α_i. But $\alpha_i - 1 \geq 0$ are all positive, since if $\alpha_i - 1$ were negative, then the limit of the right hand side as $z \to C_i$ would have been $-\infty$, whereas the left hand side is positive.

This proves that the cone of positive λ-eigenfunctions is spanned by $E_1, \ldots, E_N$. It follows that there are exactly N minimal positive λ-eigenfunctions, and that these functions are trivial. ∎

6.2. Classification of positive eigenfunctions for cocompact nilpotent periodic surfaces [LP]. Let Γ_0 be a torsion free uniform lattice in $\text{Möb}(\mathbb{D})$ and $\Gamma \lhd \Gamma_0$ a non-trivial subgroup such that $G := \Gamma_0/\Gamma$ is nilpotent. We let G act on $M := \Gamma\backslash\mathbb{D}$ by identifying G with the symmetry group of M.

We show that the set of minimal positive eigenufunctions of Δ_M is equal to $\{cF_\varphi : c > 0, \varphi\colon G \to \mathbb{R} \text{ is a homomorphism}\}$, with $F_\varphi \circ D = e^{\varphi(D)} F_\varphi$ for all $D \in G$. This is a particular case of the much more general theory developed in [LP]. The following proof (a combination of ideas from [Mrg], [CG] and [LS]) is included for completeness.

Step 1: The following holds for all minimal positive eigenfunctions h of the Laplacian of M: $h \circ D \propto h$ for all $D \in Z(G)$, and $h \circ D = h$ for all $D \in Z(G) \cap [G,G]$.

Proof: If $D \in Z(G)$ and d_M denotes hyperbolic distance (on M), then D moves points on M a bounded distance: Choose $K_0 \subset M$ compact s.t. $M = \bigcup_{D\in G} D(K_0)$. Every $z \in M$ can be written as $z = D_0(z_0)$ for some $z_0 \in K_0$, and so

$$\begin{aligned} d_M(z, Dz) &= d_M(D_0 z_0, DD_0 z_0) = d_M(D_0 z_0, D_0 D z_0) \\ &= d_M(z_0, Dz_0) \leq \max\{d_M(w, Dw) : w \in K_0\}, \end{aligned}$$

giving a uniform bound R_0 on $d_M(z, Dz)$. Let K be the closed hyperbolic disc centered at $0 \in \mathbb{D}$ with hyperbolic radius R_0. If h is a positive eigenfunction of $\Delta_{\mathbb{D}}$, so is $h \circ \gamma$ for any hyperbolic isometry γ. Choosing an isometry which moves z to the origin, we see that

$$\frac{h(Dz)}{h(z)} = \frac{(h\circ\gamma^{-1})(\gamma Dz)}{(h\circ\gamma^{-1})(\gamma z)} \leq \sup\left\{\frac{(h\circ\gamma^{-1})(z_1)}{(h\circ\gamma^{-1})(z_2)} : z_1, z_2 \in K, \gamma \in \text{Möb}(\mathbb{D})\right\}.$$

This supremum is finite by Harnack's inequality. It follows that for every $D \in Z(G)$, $h \circ D$ is bounded from above by a multiple of h. By minimality, $h \circ D$ must be proportional to h for every $D \in Z(G)$.

Let c: $Z(G) \to \mathbb{R}$ be the proportionality constant. We show that $c = 1$ on $Z(G) \cap [G, G]$, by extending c to a homomorphism λ: $G \to \mathbb{R}_+$. It will then follow that $c|_{Z(G)\cap[G,G]} = \lambda|_{Z(G)\cap[G,G]} = 1$, because any homomorphism into an abelian group vanishes on the commutator subgroup.

Following Lyons and Sullivan [LS], fix a right invariant mean M on the space of bounded functions on G (a countable amenable group), fix $z_0 \in M$ and set

$$\log \lambda(D) := M\Big[\log \frac{(h\circ\gamma)(Dz_0)}{(h\circ\gamma)(z_0)}\Big].$$

This is well-defined, because $\gamma \mapsto \log \frac{(h\circ\gamma)(Dz_0)}{(h\circ\gamma)(z_0)}$ is bounded, by Harnack's inequality. It is a homomorphism, since

$$\begin{aligned}\log\lambda(D_1D_2) &= M\Big[\log\frac{(h\circ\gamma)(D_1D_2z_0)}{(h\circ\gamma)(D_1z_0)} + \log\frac{(h\circ\gamma)(D_1z_0)}{(h\circ\gamma)(z_0)}\Big]\\ &= \log\lambda(D_1) + M\Big[\log\frac{(h\circ\gamma\circ D_1)(D_2z_0)}{(h\circ\gamma\circ D_1)(z_0)}\Big]\\ &= \log\lambda(D_1) + M\Big[\log\frac{(h\circ\gamma)(D_2z_0)}{(h\circ\gamma)(z_0)}\Big] \quad \text{(right invariance)}\\ &= \log\lambda(D_1) + \log\lambda(D_2) = \log[\lambda(D_1)\lambda(D_2)].\end{aligned}$$

It extends c because for every $D \in Z(G)$,

$$\begin{aligned}\log\lambda(D) &= M\Big[\log\frac{(h\circ\gamma)(Dz_0)}{(h\circ\gamma)(z_0)}\Big] = M\Big[\log\frac{(h\circ D)(\gamma z_0)}{h(\gamma z_0)}\Big]\\ &= M\Big[\log\frac{c(D)h(\gamma z_0)}{h(\gamma z_0)}\Big]\\ &= \log c(D).\end{aligned}$$

Step 2: Suppose that G is nilpotent. Every positive minimal eigenfunction h of Δ_M satisfies $h\circ D = e^{\varphi(D)}h$ ($D \in G$) for some homomorphism φ: $G \to \mathbb{R}$.

Proof: Since G is nilpotent, the sequence $G^{(0)} := G$, $G^{(1)} := [G, G^{(0)}]$, $G^{(2)} := [G, G^{(1)}], \ldots$ terminates at $\{id\}$ after a finite number of steps. Let k be the length of the sequence, i.e., $G^{(k-1)} \neq \{id\}$, $G^{(k)} = \{id\}$. We argue by induction on k.

If $k = 1$, then $[G, G] = \{id\}$ and G is abelian. In this case $G = Z(G)$ and the result follows from Step 1.

Next assume that $k > 1$ and that the statement holds for $k - 1$. Using the invariance properties of the hyperbolic Laplacian it is easy to check that

$$Stab(M) := \{D \in G : h\circ D = h \text{ for all minimal positive eigenfunctions } h\}$$

is a normal subgroup of G. The surface $\widetilde{M} := M/Stab(M)$ is again a cocompact periodic surface with symmetry group $G/Stab(M)$.

We claim that $G/Stab(M)$ is nilpotent of length $k-1$. Observe that $G^{(k-1)} \subset Z(G)$ ($[G, G^{(k-1)}]$ is trivial) so $G^{(k-1)} \subseteq Z(G) \cap [G,G] \subset Stab(M)$ (Step 1). Thus:

$$[G/Stab(M)]^{(k-1)} = G^{(k-1)}/Stab(M) \subseteq Z(G) \cap [G,G]/Stab(M) = \text{trivial},$$

proving that $G/Stab(M)$ is nilpotent of length $\leq k-1$.

Now pick an arbitrary minimal positive eigenfunction h on M. This function is stabilized by $Stab(M)$, and therefore projects down to a minimal positive eigenfunction $\widetilde{h}$: $\widetilde{M} \to \mathbb{R}$. The induction hypothesis implies that $\widetilde{h} \circ D \propto \widetilde{h}$ for all $\widetilde{D} \in G/Stab(M)$. It follows that $h \circ D \propto h$ for all $D \in G$. The proportionality constant depends multiplicatively on D and is therefore of the form $\exp\varphi(D)$ where φ: $G \to \mathbb{R}$ is a homomorphism.

STEP 3: For every homomorphism φ: $G \to \mathbb{R}$ there exists a positive eigenfunction F such that $F \circ D = e^{\varphi(D)}F$ ($D \in G$), and this function is unique up to a constant. This function is also minimal.

Proof: Fix a homomorphism φ: $G \to \mathbb{R}$. The existence and uniqueness of F is equivalent to the existence and uniqueness of a number $\alpha \geq 1$ and a probability measure ν on $\partial\mathbb{D}$ such that

$$\frac{d\nu \circ g}{d\nu} = e^{\varphi(\Gamma g)}|g'|^{\alpha} \quad \text{for all } g \in \Gamma_0. \tag{7}$$

Indeed, (7) implies via (5) that $F(z) = \int_{\partial\mathbb{D}} P(e^{i\theta}, z)^{\alpha} d\nu(e^{i\theta})$ is Γ-invariant and $F \circ D = e^{\varphi(D)}F$ for all $D \in \Gamma_0/\Gamma$. In the other direction, any positive eigenfunction F is represented by a Γ-conformal measure ν on $\partial\mathbb{D}$ with parameter α. This parameter is at least the critical exponent of Γ (Sullivan [Su1]), and for normal subgroups of torsion free lattices with amenable quotients this critical exponent is equal to one (Roblin [Rob]). Since $\alpha \geq 1$, the representing measure of F is unique (Theorem 3). It then follows as the proof of Lemma 3 that $F \circ D = e^{\varphi(D)}F$ ($D \in \Gamma_0/\Gamma$) implies (7).

Consider the Bowen–Series map f_{Γ_0}: $\partial\mathbb{D} \to \partial\mathbb{D}$ associated to the action of Γ_0 on $\partial\mathbb{D}$ (see §3.1). Recall that f_{Γ_0} has a finite Markov partition into intervals $\{I_a\}_{a\in S}$ such that $f_{\Gamma_0}|_{I_a} = g_a$ where $g_a \in \Gamma_0$. Define f'_{Γ_0} to be g'_a on I_a, and set

$$\phi_\alpha(e^{i\theta}) := \alpha \log|f'_{\Gamma_0}(e^{i\theta})| + \varphi(\Gamma f_{\Gamma_0}(e^{i\theta})).$$

It is standard to check, using property (Orb) of f_{Γ_0}, that (7) is equivalent to

$$\frac{d\nu \circ f_{\Gamma_0}}{d\nu} = e^{\phi_\alpha}.$$

The function ϕ_α is Hölder continuous on partition elements. The theory of such equations is well-understood (see, e.g., [Bo]): There exists a unique α for which such a solution exists, and this solution is unique.[12]

Next, we show that the function F we obtained is minimal. Write $F = \int_Y F_y d\pi(y)$ where F_y are positive and minimal eigenfunctions (the 'barycentric representation'). Let φ_y be the homomorphisms associated to F_y (Step 2). We have

$$F = e^{-n\varphi(D^n)} F \circ D^n = \int_Y e^{n[\varphi_y(D) - \varphi(D)]} F_y d\pi(y).$$

Passing to the limit $n \to \pm\infty$ we see that π is supported on the components F_y for which $\varphi_y(D) = \varphi(D)$. Since G is countable, π is supported on the set of components for which $\varphi_y = \varphi$. But we just proved that all these eigenfunctions are proportional to F. It follows that F is minimal. This finishes the proof of Step 3.

Steps 2 and 3 establish the classification of positive minimal eigenfunctions on cocompact nilpotent surfaces mentioned in example 5. An interesting artifact of the proof is that the representing measures of these eigenfunctions are (up to density function) Gibbs measures of the Bowen–Series map. ∎

6.3. Proof of (Fin). Let R be the Ford fundamental domain of Γ_0 (which consists of the closure of the set of points which lie on the external side of all isometric circles of the hyperbolic elements of Γ_0). This is a a Dirichlet domain for Γ_0, and as such is a hyperbolic polygon with finitely many sides $s_1, \overline{s}_1, \ldots, s_n, \overline{s}_n$, and there are side-pairings $g_{s_i}, g_{\overline{s}_i} \in \Gamma_0$ such that $g_{s_i}(s_i) = \overline{s}_i$, $g_{\overline{s}_i}(\overline{s}_i) = s_i$ ([Kat], §3.3 and §3.5). As explained in sections 2 and 3 of [BS], it is possible to assume without loss of generality the **even corners property**: The extension of each of its sides to a complete geodesic lies entirely inside $\mathcal{T} := \bigcup_{g \in \Gamma_0} g(\partial R)$.

We recall the construction of f_{Γ_0} (as described in [Se]). Given a side s of R, let $L(s)$ denote the complete geodesic which contains s, $H(s)$ the hyperbolic

12 Ruelle's Perron–Frobenius theorem provides a unique ν such that $\frac{d\nu\circ f_{\Gamma_0}}{d\nu}$ is proportional to $\exp \phi_\alpha$. The proportionality constant is $\exp P_{top}(-\phi_\alpha)$, where $P_{top}(-\phi_\alpha)$ is the topological pressure of ϕ_α. It is a standard fact that $P_{top}(\phi_\alpha)$ is convex, whence continuous, in α and that $P_{top}(\phi_\alpha) \xrightarrow[\alpha\to\pm\infty]{} \mp\infty$. Consequently, there exists a unique α for which the proportionality constant is equal to one.

half-plane on the side of $L(s)$ which does not contain R and $A(s)$ the boundary of $H(s)$ (an arc in $\partial\mathbb{D}$). It is proved in [BS] that no more than two such arcs intersect. The Bowen–Series map is defined by $f_{\Gamma_0}|_{A(s)} := g_s$. This definition is proper only in the part of $A(s)$ which does not intersect other arcs; on the intersections $A(s) \cap A(s')$, f_{Γ_0} is defined to be one of g_s, $g_{s'}$ (the choice is arbitrary).

Bowen and Series show that $f_{\Gamma_0}(W) \subseteq W$ where W is the set of endpoints of all complete geodesics in $\mathcal{T}$ which pass through a vertex of R. We show below that W partitions $\partial\mathbb{D}$ into a finite or countable collection of arcs $\{I_a\}_{a\in S}$. Since $f(W) \subseteq W$, this partition satisfies (Mar).

STEP 1: The set of accumulation points of W is the set C of the vertices of R which lie in $\partial\mathbb{D}$. In particular:

(1) W partitions $\partial\mathbb{D}$ into a finite or countable collection of intervals;
(2) if Γ_0 is cocompact, then W, whence S, is finite.

Proof: First observe that every vertex in the interior of $\mathbb{D}$ contributes exactly four points to W. Therefore, if Γ_0 is cocompact, then W is finite (in this case the fundamental domain has no vertices in $\partial\mathbb{D}$).

Another trivial consequence is that the set of accumulation points of W is equal to the (finite) union over $v \in C$ of the set $W(v)$ of accumulation points of the endpoints of complete geodesics in $\mathcal{T}$ which pass through v. We prove the step by showing below that $W(v) = \{v\}$.

The vertices in C are divided into vertex cycles; equivalence classes under the Γ-orbit relation. Let $v = v_0, v_1, \ldots, v_k$ be the vertex cycle of v, fix $g_i \in \Gamma_0$ such that $v_i = g_i(v)$, and let L_i, L_i' be the complete geodesics extending the faces of R which terminate at v_i. Denote the stabilizer of v_i in Γ_0 by $Stab_{\Gamma_0}(v_i)$. This is an infinite cyclic group generated by a parabolic $h_i \in \Gamma_0$ ([BM], Proposition 2.17)

Any complete geodesic $L \subset \mathcal{T}$ which terminates at v is the g-image ($g \in \Gamma_0$) of L_i or L_i' for some v_i. If we decompose $g = g_i^{-1}h$ we see that $h \in Stab_{\Gamma_0}(v_i) = \langle h_i \rangle$. It follows that $L \subset \bigcup_{i=0}^{k} \bigcup_{\ell\in\mathbb{Z}} g_i^{-1} h_i^{\ell}(L_i \cup L_i')$. Since h_i is parabolic, $h_i^{\ell}(p) \to v_i$ as $|\ell| \to \infty$ for every $p \in \partial\mathbb{D}$, so $g_i^{-1}h^{\ell}(p) \to v$ for all $p \in \partial\mathbb{D}$, proving that $W(v) = \{v\}$.

STEP 2: Let $N_\varepsilon(C)$ denote the ε-neighbourhood of C. For every $\varepsilon > 0$, $\partial\mathbb{D} \setminus N_\varepsilon(C)$ is covered by finitely many elements of $\{I_a\}_{a\in S}$.

Proof: The endpoints of $\{I_a\}_{a\in S}$ accumulate outside $\partial\mathbb{D}\setminus N_\varepsilon(C)$, so the number of I_a's which intersect $\partial\mathbb{D} \setminus N_\varepsilon(C)$ is finite.

STEP 3: $\exists\varepsilon > 0$ s.t. $\forall x \in \partial\mathbb{D} \setminus \mathrm{Par}(\Gamma_0)$, $\limsup_{n\to\infty} d(f^n_{\Gamma_0}(x), C) > \varepsilon$.

Proof: We begin with the following observations on f_{Γ_0}:

(1) every $v \in C$ has two one-sided neighborhoods J_v, J'_v such that the restriction of f_{Γ_0} to each of these neighbourhoods is an element of Γ_0;

(2) the absolute value of the derivative of this Möbius transformation is strictly larger than one, except at v.

Now choose $\varepsilon > 0$ smaller than $\min\{\mathrm{diam}(J_v), \mathrm{diam}(J'_v) : v \in C\}$ and $\min\{d(v, v') : v, v' \in C, v \neq v'\}$. We claim that if $x \in \partial\mathbb{D}$ and $d(f^n_{\Gamma_0}(x), C) \leq \varepsilon$ for all $n \geq 0$, then necessarily $x \in C$.

For every n there exists $v_n \in C$ such that $f^n_{\Gamma_0}(x) \in J_{v_n} \cup J'_{v_n}$. Let $K_n \in \{J_{v_n}, J'_{v_n}\}$ be the one-sided neighbourhood which contains $f^n_{\Gamma_0}(x)$. If we extend f_{Γ_0} continuously to the endpoints of the K_n from within, and abuse notation by denoting this extension by f_{Γ_0}, we get

$$d(f^n_{\Gamma_0}(x), f^n_{\Gamma_0}(v_0)) \leq \varepsilon \quad \text{for all } n \geq 0.$$

Let k be the length of the vertex cycle of v_0. This cycle is exactly $\{v_i\}_{i=0}^{k-1}$, and $h := f_{\Gamma_0}|_{K_{k-1}} \circ \cdots \circ f_{\Gamma_0}|_{K_0}$ fixes v_0. It follows that h is parabolic. Note that $|h'| > 1$ on $K_0 \setminus \{v_0\}$ (because $|f'_{\Gamma_0}| > 1$ on $K_i \setminus \{v_i\}$). Since h is parabolic, its dynamics is such, that the h-forward orbit of any $y \in K_0 \setminus \{v_0\}$ leaves K_0. But by construction

$$d(h^\ell(x), v_0) < \varepsilon \quad \text{for all } n \geq 0.$$

Therefore $x = v_0 \in C$. This proves that if $d(f^n_{\Gamma_0}(x), C) \leq \varepsilon$ for all $n \geq 0$, then $x \in C$. Step 3 follows, because $C \subset \mathrm{Par}(\Gamma_0)$ ([Kat], Theorem 4.2.5).

We can now finish the proof of (Fin). Pick $\varepsilon > 0$ as in Step 3, and choose a finite $S_0 \subset S$ such that $\partial\mathbb{D} \setminus N_\varepsilon(C) \subset \bigcup_{a\in S_0} I_a =: \Lambda$. Every f_{Γ_0}-forward orbit either hits C and stays there, or leaves $N_\varepsilon(C)$. In the first case, the orbit is contained in $\mathrm{Par}(\Gamma_0)$. In the second case, the orbit enters Λ infinitely many times. ∎

6.4. PROOF OF (FI). Suppose $\Gamma_0\backslash\mathbb{D}$ has finite volume. Any hyperbolic surface with finite volume has a compact subset F which is intersected by any complete geodesic which does not tend to one of the cusps; such a set can be obtained by cutting away each of the cusps along a closed horocycle which encircles it.

Let $F_0 \subset \mathbb{D}$ be a compact subset of the fundamental region of Γ_0 which contains the origin o and which projects to F. Let $\{g_n\}_{n\geq 1}$ be an enumeration of the g in Γ_0 whose axis intersects F_0. Fix some $z_n \in F_0$ on the axis of g_n.

Then $\tau(g_n) = d(z_n, g_n z_n) \geq d(o, g_n o) - 2\,\mathrm{diam}(F_0) \xrightarrow[n\to\infty]{} \infty$, proving that $\{\tau(g_n)\}$ intersects any compact interval finitely many times. Since any $g \in \Gamma_0$ is conjugate to some g_n, $\tau(\Gamma_0)$ intersects any compact interval finitely many times. ∎

References

[ASS] J. Aaronson, O. Sarig and R. Solomyak, *Tail-invariant measures for some suspension semiflows*, Discrete and Continuous Dynamical Systems **8** (2002), 725–735.

[Ad] R. Adler, *F-expansions revisited*, in *Recent advances in topological dynamics* (Proc. Conf., Yale University, New Haven, Conneticut, 1972; in honor of Gustav Arnold Nedlund), Lecture Notes in Mathematics, Vol. 318, Springer, Berlin, 1973, pp. 1–5.

[Ba] M. Babillot, *On the classification of invariant measures for horospherical foliations on nilpotent covers of negatively curved manifolds*, in *Random walks and geometry* Walter de Gruyter GmbH & Co. KG, Berlin, 2004, pp. 319–335.

[BL] M. Babillot and F. Ledrappier, *Geodesic paths and horocycle flows on Abelian covers*, in *Proceedings of the International Colloquium on Lie Groups and Ergodic Theory, Mumbai 1996*, Narosha Publishing House, New Delhi, 1998, pp. 1–32.

[Be] A. Beardon, *The Geometry of Discrete Groups*, (corrected reprint of the 1983 original), Graduate Texts in Mathematics **91**, Springer-Verlag, New–York, 1995.

[BM] M. B. Bekka and M. Mayer, *Ergodic Theory and Topological Dynamics of Group Actions on Homogeneous Spaces*, London Mathematical Society Lecture Note Series, **269**, Cambridge University Press, Cambridge, 2000.

[BE] P. Bougerol and L. Élie, *Existence of positive harmonic functions on groups and on covering manifolds*, Annales de l'Institut Henri Poincaré. Probabilités et Statistiques **31** (1995), 59–80.

[Bo] R. Bowen, *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, Lecture Notes in Mathematics, **470**, Springer-Verlag, Berlin-New York, 1975.

[BS] R. Bowen and C. Series, *Markov maps associated with Fuchsian groups*, Publications Mathématiques de l'Institut des Hautes Études Scientifiques **50** (1979), 153–170.

[Bu] M. Burger, *Horocycle flow on geometrically finite surfaces*, Duke Mathematical Journal **61** (1990), 779–803.

[CG] J.-P. Conze and Y. Guivarc'h, *Propriété de droite fixe et fonctions propres des opérateurs de convolution,* Séminaire de Probabilités, I (Univ. Rennes, Rennes, 1976), Exp. No. 4, Dept. Math. Informat., Univ. Rennes, Rennes, 1976, 22 pp.

[C] Y. Coudene, *Cocycles and stable foliations of Axiom A flows,* Ergodic Theory and Dynamical Systems **21** (2001), 767–775.

[Da1] F. Dal'bo, *Topologie du feuilletage fortement stable,* Annales de l'Institut Fourier (Grenoble) **50** (2000), 981–993.

[Da2] F. Dal'bo, *Remarques sur le spectre des longueurs d'une surface et comptages,* Boletim da Sociedade Brasileira de Matemática (N.S.) **30** (1999), 199–221.

[D] S. G. Dani, *Invariant measures of horospherical flows on noncompact homogeneous spaces,* Inventiones Mathematicae **47** (1978), 101–138.

[DS] S. G. Dani and J. Smillie, *Uniform distributrion of horocycle orbits for Fuchsian groups,* Duke Mathematical Journal **51** (1984), 185–194.

[Eb] P. Eberlein, *Geodesic flows on negatively curved manifolds. II,* Transactions of the American Mathematical Society **178** (1973), 57–82.

[FM] J. Feldman and C. Moore, *Ergodic equivalence relations, cohomology, and von Neumann algebras. I,* Transactions of the American Mathematical Society **234** (1977), 289–324.

[F] H. Furstenberg, *The unique ergodicity of the horocycle flow,* Springer Lecture Notes **318** (1972), 95–115.

[Gr] M. Gromov, *Groups of polynomial growth and expanding maps,* Publications Mathématiques de l'Institut des Hautes Études Scientifiques **53** (1981), 53–73.

[GR] Y. Guivarc'h and A. Raugi, *Products of random matrices: convergence theorems,* in *Random matrices and their applications (Brunswick, Maine, 1984)*, Contemporary Mathematics, **50**, American Mathematical Society, Providence, RI, 1986, pp. 31–54.

[Iw] H. Iwaniec, *Introduction to the spectral theory of automorphic forms,* Biblioteca de la Revista Matemática Iberoamericana, Madrid, 1995, xiv+247 pp.

[JM] L. Ji and R. MacPherson, *Geometry of compactifications of locally symmetric spaces,* Annales de l'Institut Fourier (Grenoble) **52** (2002), 457–559.

[Kai1] V. Kaimanovich, *Ergodic properties of the horocycle flow and classification of Fuchsian groups,* Journal of Dynamic and Control Systems **6** (2000), 21–56.

[Kai2] V. A. Kaimanovich, *Brownian motion and harmonic functions on covering manifolds. An entropic approach,* Doklady Akademii Nauk SSSR **288** (1986), no. 5. Engl. Transl. in Soviet Mathematics Doklady **33** (1986), 812–816.

[Kar] F. I. Karpelevich, *The geometry of geodesics and the eigenfunctions of the Laplacian on symmetric spaces,* Trudy Moskovskogo Matematičeskogo Obščestva **14** (1965), 48–185 (Russian); translated as Trans. Moskov. Math. Soc. **14** (1965), 51–199.

[Kat] S. Katok, *Fuchsian Groups,* Chicago Lectures in Mathematics, University of Chicago Press, 1992.

[LP] V. Lin and Y. Pinchover, *Manifolds with group actions and elliptic operators,* Memoirs of the American Mathematical Society **112** (1994), 78pp.

[LS] T. Lyons and D. Sullivan, *Function theory, random paths and covering spaces,* Journal of Differential Geometry **19** (1984), 299–323.

[Mrc] B. Marcus, *Unique ergodicity of the horocycle flow: variable negative curvature case,* Israel Journal of Mathematics **21** (1975), 133–144.

[Mrg] G. A. Margulis, *Positive harmonic functions on nilpotent groups,* Soviet Mathematics Doklady **166** (1966), 241–244.

[Rat] M. Ratner, *On Raghunathan's measure conjecture,* Annals of Mathematics (2) **134** (1991), 545–607.

[Ro] D. J. S. Robinson, *A Course in the Theory of Groups,* Second edition. Graduate Texts in Mathematics **80**, Springer-Verlag, New York, 1996, xviii+499 pp.

[Rob] T. Roblin, *Un théorème de Fatou pour les densités conformes avec applications aux revêtements galoisiens en courbure négative,* Israel Journal of Mathematics **147** (2005), 333–357.

[Sg] O. Sarig, *Invariant measures for the horocycle flow on Abelian covers,* Inventiones Mathematicae **157** (2004), 519–551.

[Sk] P. Sarnak, *Some Applications of Modular Forms,* Cambridge Tracts in Mathematics **99**, Cambridge University Press, Cambridge, 1990.

[Sch] K. Schmidt, *Cocycles on Ergodic Transformation Groups,* Macmillan Lectures in Mathematics, Vol. 1, Macmillan Company of India, Ltd., Delhi, 1977.

[Se] C. Series, *Geometrical methods of symbolic coding,* in *Ergodic Theory, Symbolic Dynamics, and Hyperbolic Spaces* (T. Bedford, M. Keane, C. Series, eds.), Oxford University Press, 1991, pp. 125–151.

[Su1] D. Sullivan, *Related aspects of positivity in Riemannian geometry,* Journal of Differential Geometry **25** (1987), 327–351.

[Su2] D. Sullivan, *Discrete conformal groups and measurable dynamics,* Bulletin of the American Mathematical Society (N.S.) **6** (1982), 57–73.

[Su3] D. Sullivan, *The density at infinity of a discrete group of hyperbolic motions,* Publications Mathématiques de l'Institut des Hautes Études Scientifiques **50** (1979), 171–202.