C. Lutsko (2021) "Farey Sequences for Thin Groups,"
International Mathematics Research Notices, Vol. 00, No. 0, pp. 1–48
doi:10.1093/imrn/rnab036

Farey Sequences for Thin Groups

Christopher Lutsko*
University of Bristol, Mathematics Department, Bristol, BS8 1QU, UK

**Correspondence to be sent to: e-mail: chris.lutsko@rutgers.edu*

The Farey sequence is the set of rational numbers with bounded denominator. We introduce the concept of a generalized Farey sequence. While these sequences arise naturally in the study of discrete and thin subgroups, they can be used to study interesting number theoretic sequences—for example rationals whose continued fraction partial quotients are subject to congruence conditions. We show that these sequences equidistribute and the gap distribution converges and answer an associated problem in Diophantine approximation. Moreover, for one example, we derive an explicit formula for the gap distribution. For this example, we construct the analogue of the Gauss measure, which is ergodic for the Gauss map. This allows us to prove a theorem about the associated Gauss–Kuzmin statistics.

1 Introduction

Consider the classical *Farey sequence* of height Q:

$$\tilde{\mathcal{F}}_Q := \left\{ \frac{p}{q} \in [0,1) \ : \ (p,q) \in \hat{\mathbb{Z}}^2, 0 < q < Q \right\}, \tag{1.1}$$

where $\hat{\mathbb{Z}}^2$ denotes the set of primitive vectors in $\mathbb{Z}^2$. Naturally, this sequence is a fundamental object in number theory dating back to 1802 with its introduction by Haros and subsequent work by Farey and Cauchy. For example, this sequence has connections

Received October 1, 2020; Revised October 1, 2020; Accepted October 8, 2020

to the Riemann hypothesis (see e.g., [10]) and plays a fundamental role in Diophantine approximation.

In this paper, we generalize the Farey sequence. For concreteness, one example of such a generalized Farey sequence is given by the following: throughout the paper, we use the standard continued fraction notation

$$[a_0; a_1, \dots a_n] = a_0 + \cfrac{1}{a_1 + \frac{1}{a_2 + \frac{\cdots}{a_n}}} \tag{1.2}$$

(see e.g., [9]) then denote

$$\mathcal{Q}_4 := \{[0; a_1, \dots a_k] : k \in \mathbb{N}\ ,\ a_i \in 4\mathbb{Z}_{\neq 0}\ \forall i\}, \tag{1.3}$$

that is, rationals whose continued fraction expansions involve only multiples (*possibly negative*) of 4. The generalized Farey sequence in this context is

$$\widehat{\mathcal{F}}_Q = \{\frac{p}{q} \in \mathcal{Q}_4 : 0 < q < Q,\ \gcd(p,q) = 1\}; \tag{1.4}$$

we return to this example in Section 1.1 where we give a geometric interpretation of these sets. To see some of the points of $\mathcal{Q}_4$, see Figure 1.

There is a geometric interpretation of the classical Farey sequence, which will play an integral role in this paper. Consider the groups $G = \mathrm{PSL}(2,\mathbb{R})$ and $\Lambda := \mathrm{PSL}(2,\mathbb{Z}) < G$. G acts on the hyperbolic half-space, $\mathbb{H}$ via Möbius transformations (see Section 2). As Λ is a lattice, there exists a tessellation of $\mathbb{H}$ into disjoint, finite volume subsets such that Λ acts transitively on them. These *fundamental domains* are not compact as each one contains a point on the boundary $\partial\mathbb{H} = \mathbb{R} \cup \{\infty\}$, at the end of a cusp. The set of such cuspidal points is exactly

$$(\Lambda/\Lambda_\infty)\infty = \mathbb{Q} \tag{1.5}$$

(we use G_x to denote the stabilizer of x in a group G). That is, the set of cuspidal points can be written as the Λ-orbit of the point at $\infty \in \partial\mathbb{H}$, which corresponds to the rationals. Thus, the Farey sequence of height Q can be written

$$\tilde{\mathcal{F}}_Q = \left\{\frac{p}{q} \in (\Lambda/\Lambda_\infty)\infty : (p,q) \in \hat{\mathbb{Z}}^2, 0 < q < Q\right\} \tag{1.6}$$

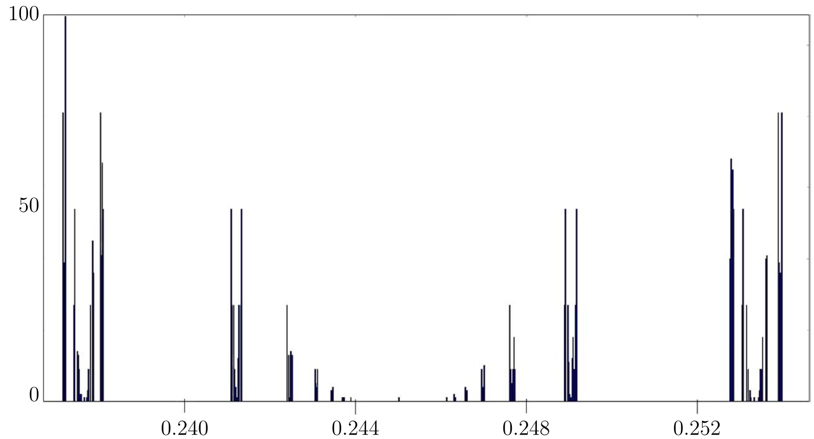

Fig. 1. Above we show some of the points in $\mathcal{Q}_4$. The graph was generated as follows: we generated all words of length 10 (with respect to the two generators applied to ∞). Then separated the interval $[0, 1)$ into bins of size 10^{-5}. The above is a bar chart showing the number of points in each bin. Note that the sequence is supported on a fractal subset of the interval. This does not show $\widehat{\mathcal{F}}_Q$ (as the cut-off is with respect to word length), however will suffice for a qualitative picture.

—the points in the Λ-orbit of the point at $\infty \in \partial\mathbb{H}$ with denominator less than Q. The goal of this paper is to consider a generalization of this setup, where we replace Λ by a general (possibly infinite covolume) discrete subgroup. For our example (1.4), the corresponding subgroup is the Hecke group

$$\widehat{\Gamma} = \left\langle \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\rangle. \tag{1.7}$$

Most of our theorems hold for general subgroups. Hence, let $\Gamma < \mathrm{PSL}(2, \mathbb{R})$ be a *general* non-elementary, finitely generated subgroup in G with critical exponent δ_Γ. In our context, $1/2 < \delta_\Gamma \leq 1$ and δ_Γ is equal to the Hausdorff dimension of the limit set of the subgroup (we introduce these definitions in Section 2). Furthermore, assume Γ has a cusp at ∞ and let $\Gamma^\infty = (\Gamma/\Gamma_\infty)\infty \subset \partial\mathbb{H}$ denote the orbit of ∞. Hence, Γ^∞ is the set of the cusps located at points on the boundary, isomorphic to ∞. Finally, we assume that $\Gamma_\infty = \left\langle \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \right\rangle$. That is, that the fundamental domain is periodic with period 1 along the real line. Note that $\widehat{\Gamma}$ has period 4. However, a scaling could be applied to give it period 1 (in order to preserve the continued fraction description we refrain from doing so).

Let

$$\mathcal{Z} := \{(p,q) \in (0,1)\Gamma\} \subset \mathbb{R}^2 \tag{1.8}$$

denote the analogue of primitive vectors and define

$$\begin{aligned}\mathcal{F}_Q &:= \left\{\frac{p}{q} \in [0,1) : (p,q) \in \mathcal{Z}, 0 < q < Q\right\}\\ &= \left\{\frac{p}{q} \in \Gamma^{\infty} : 0 \leq p < q < Q\right\}.\end{aligned} \tag{1.9}$$

$\mathcal{F}_Q$ is the primary object of study for this article, which we call a *generalized Farey sequence* (occasionally, *gFs*). In Subsection 3.1, we show that asymptotically there exists a constant $0 < c_\Gamma < \infty$ such that

$$|\mathcal{F}_Q| \sim c_\Gamma Q^{2\delta_\Gamma}. \tag{1.10}$$

The goal of the paper is to establish the Theorems in Sections 4–9, which we describe briefly here. As the statements of the theorems require the use of fractal measures, we present them formally only after presenting the necessary notation (readers familiar with Patterson–Sullivan theory may wish to skip ahead and see the theorems now). Sections 2 and 3 present some background and preliminary work. Subsequently, the main results of the paper are as follows:

- Counting primitive points: in Section 4, we present a theorem for the equidistribution of the horocycle flow in infinite volume subgroups (proved by Oh and Shah [17]). Then we show how this equidistribution result can be used to prove a technical theorem about counting primitive points in a sheared set (Theorem 4.3) and another technical theorem about counting primitive points in a rotated set (Theorem 4.5). These theorems generalize the analogous result for lattices in [16].
- Diophantine approximation by parabolics: we prove two theorems in metric Diophantine approximation in Fuchsian groups. These are the analogues of the Erdös–Szüsz–Turán and Kesten problems in the infinite volume setting. In the classical setting, these problems were solved using homogeneous dynamics by Marklof [12, Theorem 4.4] and Athreya and Ghosh [2]. Moreover, Xiong and Zaharescu [24] and Boca [6] solved the problem using number theoretic methods (by applying the BCZ map). Extending classical results

in metric Diophantine approximation to the setting of Fuchsian groups is not new and was done by Patterson [18] who proved Dirichlet and Khintchine type theorems for such parabolic points. More recently, for example Beresnevich *et al.* [4] studied the equivalent problems for Kleinian groups.

In the same section, we show that Theorem 4.5 allows us to prove that there is a limiting distribution for the direction of primitive points, $\mathcal{Z}$, as viewed from the origin. This problem has not been addressed in the Euclidean setting except for lattices [16].

- Equidistribution of gFs: Theorem 6.1 states that the gFs equidistributes over a horospherical section. In a series of papers [13, 14], Marklof showed that the (classical) Farey sequence, when embedded into a horosphere, equidistributes on a particular section. This equidistribution theorem was then used to show that the spatial statistics of the Farey sequence converge. This was followed by work of Athreya and Cheung [1] who (in dimension $d = 2$) were able to construct a Poincaré section for the horocycle flow such that the return time map generates Farey points. We restrict our attention to proving the equidistribution result in this more general setting. Heersink [8] generalized [13] to certain congruence subgroups of Λ (still in the finite covolume setting). Furthermore, the method of [1] has been generalized to more general subgroups such as Hecke triangle groups (e.g., [23]). However, we will not discuss this approach here.
- Convergence of local statistics: Theorem 7.1, as a consequence of Theorem 4.3 and Theorem 6.1, states that two sorts of local statistics converge in the limit. A corollary of one of these is that the limiting gap distribution exists. This distribution in the classical setting was originally calculated by Hall [7] (and is known as the Hall distribution) and has been studied by many people since. The Hall distribution was originally put into the context of ergodic theory in [3].
- An explicit formula for the gap distribution: in Section 8, we restrict to the example $\widehat{\Gamma}$. For this example, we show that the limiting gap distribution can be explicitly written as an integral over a compact region. While the integral involves a fractal measure, this is the 1st time such an explicit formula has been calculated in the infinite volume setting. There is much interest in finding explicit formula for limiting gap distributions for projected lattice point sets and the infinite covolume analogue. The only instance (to our

knowledge) of such explicit examples are those covered in [19]. In that paper, Rudnick and Zhang used the relation between Farey points and Ford circles to produce examples for which they could express the limiting gap distribution explicitly (recovering, in one instance, the Hall distribution). In Section 1.1, we show that the Farey sequence for $\widehat{\Gamma}$ can also be used to generate a (sparse) Ford configuration, which leads to our result.

- Ergodicity of a new Gauss-like measure: continuing to work with the example $\widehat{\Gamma}$, we show that a new fractal measure takes on the role of the Gauss measure (Theorem 9.2). That is, this measure is ergodic for the Gauss map. As an application, we show that this ergodicity implies convergence to an explicit function of the Gauss–Kuzmin statistics in our context. This section takes inspiration from [20] where Series showed how the Gauss measure can be viewed as a projection of the Haar measure on a particular cross-section.

1.1 Ford configurations for $\widehat{\Gamma}$

To give some further intuition for generalized Farey sequences, in this section, we show that the gFs for $\widehat{\Gamma}$ admits a simple geometric interpretation, which we shall return to in Section 8. Returning to our example $\widehat{\mathcal{F}}_Q$—(1.4), note that

$$\widehat{\Gamma}^{\infty} = \mathcal{Q}_4. \tag{1.11}$$

To see this, simply note that the two generators in (1.7) correspond to the maps $f(x) = x + 4$ and $g(x) = \frac{-1}{x}$, which generate these continued fractions.

Consider the action of $\widehat{\Gamma}$ on an initial configuration of circles in the closure $\overline{\mathbb{H}}$:

$$\begin{gathered}\mathcal{K}_0 := (\mathcal{C}_0, \mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3)\\ \mathcal{C}_0 = \mathbb{R} \quad , \quad \mathcal{C}_1 = \mathbb{R} + i \quad , \quad \mathcal{C}_2 = C(i/2, 1/2) \quad , \quad \mathcal{C}_3 = C(i/2 + 4, 1/2)\end{gathered} \tag{1.12}$$

where $C(z, r)$ is a circle located at $z \in \overline{\mathbb{H}}$ of radius r. We are interested in the resulting sparse Ford configuration, $\mathcal{K} := \widehat{\Gamma}\mathcal{K}_0$, shown in Figure 2. Any group element in $\widehat{\Gamma}$ can be decomposed into a composition of circle inversions through vertical lines at 0 and 4 and $C(0, 1)$ and $C(4, 1)$ (these are also shown in Figure 2).

Let $\mathcal{A}_T$ denote the set of tangencies with $\mathcal{C}_0$ in $[0, 1]$ such that the circle tangent to $\mathcal{C}_0$ has diameter larger than T^{-1}. The way we have constructed the packing $\mathcal{K}$, these

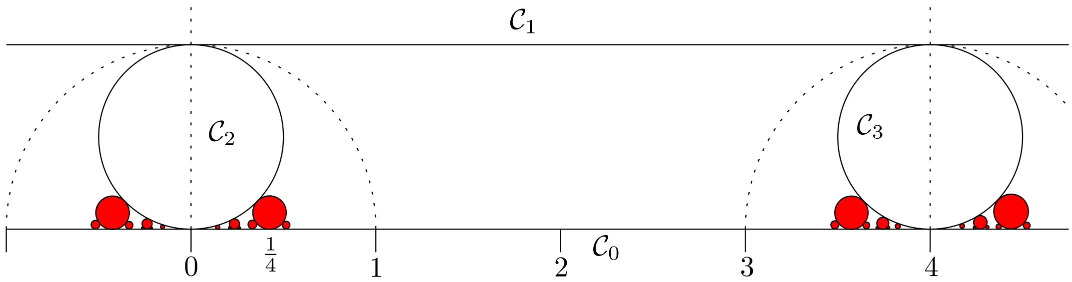

Fig. 2. Diagram of a portion of $\mathcal{K}$. The dotted lines represent the circle inversions corresponding to the subgroup $\widehat{\Gamma}$. The white circles (including the x-axis and horizontal line above) represent the initial configuration $\mathcal{K}_0 = (\mathcal{C}_0, \mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3)$. The filled-in circles represent some of the images.

tangencies are exactly the cuspidal points of the group (i.e., the tangencies are located on the orbit $\widehat{\Gamma}\infty$). Moreover, one can easily show if a circle in this packing is tangent to $\mathcal{C}_0$ at p/q in reduced form then the diameter is given by $1/q^2$. Hence, $\mathcal{A}_{Q^2} = \widehat{\mathcal{F}}_Q$, that is, the set of tangencies of circles with diameter greater than Q^2 is exactly the gFs of height Q.

Given an interval $\mathcal{I} \subset [0,1]$, let $\mathcal{A}_{T,\mathcal{I}} = \mathcal{A}_T \cap \mathcal{I}$. We label the elements of $\mathcal{A}_T = \{x^j_{T,\mathcal{I}}\}_{j=1}^{\#\mathcal{A}_{T,\mathcal{I}}}$ such that $x^j_{T,\mathcal{I}} < x^{j+1}_{T,\mathcal{I}}$ for all j. The gap distribution is then

$$\widehat{F}_{T,\mathcal{I}}(s) := \frac{\#\left\{i \in [1, \#\mathcal{A}_{T,\mathcal{I}}) : T(x^{j+1}_{T,\mathcal{I}} - x^j_{T,\mathcal{I}}) \leq s\right\}}{T^{\delta_{\widehat{\Gamma}}}} \tag{1.13}$$

for $s > 0$.

In Section 8, we show that the limiting gap distribution can be explicitly calculated as a sum of integrals over compact regions involving a fractal measure presented below. This allows us to show that all gaps have size bigger than $s < 2$ (not just in the limiting case) and to say something more about the regularity of F and the growth of the derivative.

Remark. Of course different subgroups generate different sparse Ford configurations and have other interesting relations to continued fractions (and hence Diophantine approximation). We only address this (simplest) example here. That said, our methods generalize without additional effort to any Hecke subgroup of the form $\Gamma_c = \left\langle \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\rangle$ for $c \in \mathbb{R}_{>2}$ (the corresponding continued fraction description will involve c rather than 4 and this loses some elegance for non-integer c).

2 Background—Hyperbolic Geometry

Consider the action of G on $\mathbb{H}$ via Möbius transformations: for $z \in \mathbb{H}$ and $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in G$

$$\begin{aligned} gz &= \frac{az+b}{cz+d} \\ zg &= {}^t gz = \frac{az+c}{bz+d}.gz = \frac{az+c}{bz+d}. \end{aligned} \tag{2.1}$$

Let $X_i \in T^1(\mathbb{H})$ denote the vector pointing upwards based at i. Denote

- $K = \mathrm{Stab}_G(i)$, hence $\mathbb{H} \cong G/K$.
- A—a one parameter subgroup corresponding to the unit speed geodesic flow, $\mathcal{G}_r$, on $T^1(\mathbb{H})$. For X_i, the action of A corresponds to multiplication by $\Phi^t = \left(\begin{smallmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{smallmatrix}\right)$.
- $N_- := \left\{ n_-(x) = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} : x \in \mathbb{R} \right\}$, the contracting horosphere for Φ^t.
- $N_+ := \left\{ n_+(x) = \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix} : x \in \mathbb{R} \right\}$, the expanding horosphere for Φ^t.

We identify points in G with points in $T^1(\mathbb{H})$ via the map $g \mapsto gX_i$ and points in G/K we identify with points in $\mathbb{H}$ via the map $g \mapsto gi$.

2.1 Measure theory on infinite volume hyperbolic manifolds

To construct the appropriate measures, we require the following definitions. For a point $u \in T^1(\mathbb{H})$, denote the forward and backward geodesic projections

$$u^{\pm} = \lim_{r\to\infty} \mathcal{G}_r(u). \tag{2.2}$$

Moreover, for $g \in G$, we denote $g^{\pm} = g(X_i)^{\pm}$. Let $\mathcal{L}(\Gamma) \subset \partial\mathbb{H}$—the *limit set*—denote the set of accumulation points of any orbit under Γ. A classical result in the field states that the Hausdorff dimension of $\mathcal{L}(\Gamma)$ is the critical exponent δ_Γ [21].

Given a boundary point $\xi \in \partial\mathbb{H}$ and two points in the interior $x, y \in \mathbb{H}$, define the *Busemann function* to be

$$\beta_\xi(x, y) := \lim_{t\to\infty} d(x, \xi_t) - d(y, \xi_t), \tag{2.3}$$

where ξ_t is any geodesic such that $\lim_{t\to\infty}\xi_t = \xi$. In words, the Busemann function measures the signed distance between the horospheres containing x and y based at ξ.

Define a Γ-*invariant conformal density of dimension* $\delta_\mu > 0$ to be a family, $\{\mu_x : x \in \mathbb{H}\}$ of finite, Borel measures on the boundary $\partial\mathbb{H}$ such that

$$\gamma_*\mu_x(\cdot) := \mu_x(\gamma^{-1}\cdot) = \mu_{\gamma x}(\cdot), \quad \frac{d\mu_x}{d\mu_y}(\xi) = e^{\delta_\mu \beta_\xi(y,x)}, \tag{2.4}$$

for any $y \in \mathbb{H}$, $\xi \in \partial\mathbb{H}$ and $\gamma \in \Gamma$. Patterson [18] (in dimension 2) and Sullivan [21] (in higher dimensions) constructed a Γ-invariant conformal density of dimension δ_Γ supported on the limit set $\mathcal{L}(\Gamma)$. We denote this conformal density ν_x. Moreover, let m_x denote the G-invariant density of dimension 1 (the *Lesbegue density*).

Given a point $u \in T^1(\mathbb{H})$, let $\pi(u)$ denote the projection to $\mathbb{H}$ and let $s = \beta_{u^-}(i,\pi(u))$. From there, define the following measures:

- The *Burger–Roblin* measure

$$d\mathrm{m}^{BR}(u) = e^{\delta_\Gamma \beta_{u^-}(i,\pi(u))} e^{\beta_{u^+}(i,\pi(u))} d\nu_i(u^-) d\mathrm{m}_i(u^+) ds \tag{2.5}$$

  is supported on $\{u \in T^1(\mathbb{H}) : u^- \in \mathcal{L}(\Gamma)\}$ and is finite on $\Gamma\backslash G$ iff $\Gamma\backslash G$ has finite volume (in which case the Burger–Roblin measure is equal to the Haar measure).
- The *Bowen–Margulis–Sullivan* measure

$$d\mathrm{m}^{BMS}(u) = e^{\delta_\Gamma \beta_{u^-}(i,\pi(u))} e^{\delta_\Gamma \beta_{u^+}(i,\pi(u))} d\nu_i(u^-) d\nu_i(u^+) ds \tag{2.6}$$

  is supported on $\{u \in T^1(\mathbb{H}) : u^\pm \in \mathcal{L}(\Gamma)\}$ and is finite on $\Gamma\backslash G$.

Now define the *Patterson–Sullivan* measure (for N_-) on $\partial\mathbb{H} \simeq \mathbb{R}$ to be

$$d\mu^{PS}(x) := e^{\delta_\Gamma \beta_x(i,i+x)} d\nu_i(x). \tag{2.7}$$

Note that $\mathrm{supp}(\mu^{PS}) = \mathcal{L}(\Gamma)$. We will primarily use this Patterson–Sullivan measure; however, we also use one associated to the expanding horospherical subgroup N_+, defined as

$$d\mu^{PS}_{N_+}(x) := e^{\delta_\Gamma \beta_{\frac{1}{x}}(i,\frac{i}{xi+1})} d\nu_i(\frac{1}{x}). \tag{2.8}$$

3 Preliminary Results

3.1 Proof of (1.10)

Proof of (1.10). A rational $\frac{a}{b}$ belongs to $\mathcal{F}_Q$ if and only if there exists a $\gamma = \begin{pmatrix} a & * \\ b & * \end{pmatrix} \in \Gamma/\Gamma_\infty$ and $0 < a < b < Q$. Using the standard Iwasawa decomposition, one can write

$$\gamma = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} y^{1/2} & 0 \\ 0 & y^{-1/2} \end{pmatrix} \tag{3.1}$$

where $a = \cos\theta y^{1/2}$ and $b = \sin\theta y^{1/2}$. Therefore, the problem is equivalent to counting

$$\#\left\{\gamma \in \Gamma/\Gamma_\infty : (\theta, y) \in \Omega\right\}, \tag{3.2}$$

where $\Omega := \{(\theta, y) : 0 < y^{1/2}\cos\theta < y^{1/2}\sin\theta < Q\}$. Counting the asymptotic number of points in such a sector is the content of [5] (see Theorem 8.5 below).

Below to prove Proposition 8.6, we perform this calculation more carefully (and will calculate the constant in that context; thus, we leave the details till then). ■

3.2 Gauss-type decomposition

Let $M_{\mathbf{y}} := \begin{pmatrix} y_2^{-1} & 0 \\ y_1 & y_2 \end{pmatrix}$, for $\mathbf{y} \in \mathbb{R}^2$. In what follows, we will need the following decomposition of $T^1(\mathbb{H})$.

Proposition 3.1. For any $\phi \in C_c(T^1(\mathbb{H}))$ and any set $\mathscr{A} \subset \mathbb{R}^2$,

$$\int_{N_-\{M_{\mathbf{y}}:\mathbf{y}\in\mathscr{A}\}} \phi(hM_{\mathbf{y}})\, \mathrm{d}m^{BR}(hM_{\mathbf{y}}) = 2\int_{\mathbb{R}\times\mathscr{A}} \phi(n_-(x)M_{\mathbf{y}})y_2^{2\delta_\Gamma-2}\, \mathrm{d}y_2\, \mathrm{d}y_1\, \mathrm{d}\mu^{PS}(x). \tag{3.3}$$

Proof. The goal is to understand the forwards and backwards orbits of $u = hM_{\mathbf{y}}X_i$ (where $h \in N_-$). First, we note that

$$u^- = (hM_{\mathbf{y}}X_i)^- = hX_i^- \tag{3.4}$$

(this follows from the definition of the stable and unstable directions of the geodesic flow). Hence, we can write the following:

$$\begin{aligned} s &:= \beta_{u^-}(i, \pi(u)) \\ &= \beta_{X_i^-}(h^{-1}i, M_{\mathbf{y}}i). \end{aligned} \tag{3.5}$$

Inserting the definition of the Busemann function and using its invariance properties then give

$$\begin{aligned} s &= \lim_{t\to\infty} d(h^{-1}i, \Phi^{-t}i) - d(M_y i, \Phi^{-t}i) \\ &= \lim_{t\to\infty} d(i, \Phi^{-t}i) - d(M_y i, \Phi^{-t}i) + d(h^{-1}i, \Phi^{-t}i) - d(i, \Phi^{-t}i). \end{aligned} \tag{3.6}$$

Now setting $r_0(h) = \beta_{hX_i^-}(i, hi)$ gives

$$\begin{aligned} s &= \lim_{t\to\infty} t - d(\left(\begin{smallmatrix} y_2^{-1} & 0 \\ 0 & y_2 \end{smallmatrix}\right) i, \Phi^{-t}i) + r_0(h) \\ &= \lim_{t\to\infty} t - t + 2\ln y_2 + r_0(h) \\ &= 2\ln y_2 + r_0(h). \end{aligned} \tag{3.7}$$

Thus,

$$ds = \frac{2dy_2}{y_2}. \tag{3.8}$$

Moreover, we note that by definition

$$e^{\delta_\Gamma r_0(n_-(x))} d\nu_i(n_-(x)X_i) = d\mu^{PS}(x). \tag{3.9}$$

Next consider the measure

$$d\lambda_g(z) = e^{\beta_{(hM_y X_i)^+}(i, hM_y i)} dm_i((hM_y X_i)^+), \tag{3.10}$$

with $g = h\left(\begin{smallmatrix} y_2^{-1} & 0 \\ 0 & y_2 \end{smallmatrix}\right)$ and $z = n_+(y_2^{-1}y_1)$. We can write (using the G-invariance of m)

$$\begin{aligned} &= e^{\beta_{(gzX_i)^+}(i, gzi)} dm_i((gzX_i)^+) \\ &= e^{\beta_{(gzX_i)^+}(i, gzi)} dm_{g^{-1}i}((zX_i)^+) \end{aligned} \tag{3.11}$$

and then using the definition of conformal densities:

$$\begin{aligned} &= e^{(\beta_{(gzX_i)^+}(i, gzi) + \beta_{(zX_i)^+}(i, g^{-1}i))} dm_i((zX_i)^+) \\ &= e^{\beta_{(zX_i)^+}(i, zi)} dm_i((zX_i)^+). \end{aligned} \tag{3.12}$$

Hence, $d\lambda_g = d\lambda_e$ and in particular λ_e is N^+-invariant. Hence, it is the Haar measure on N_+. Thus, we have (for y_2 fixed)

$$d\lambda_g(z) = dz = y_2^{-1} dy_1. \tag{3.13}$$

Inserting (3.4), (3.7), (3.8), (3.9), and (3.13) into the definition of the BR-measure we get (3.3). ■

3.3 Global measure formula

The last theorem from the literature we require is the so-called global measure formula stated by Stratmann and Velani [22, Theorem 2], which requires some set up. In actuality, we only use the simpler Corollary 3.3. As stated in [22], there exists a disjoint, Γ-invariant collection of horoballs $\mathscr{H}$ such that $(\mathcal{C}_\Gamma \setminus \mathscr{H})/\Gamma$ is compact, where $\mathcal{C}_\Gamma$ is the convex hull of $\mathcal{L}(\Gamma)$.

We let $\eta \in \mathcal{L}(\Gamma)$ be a *parabolic limit point*. Define η_t to be the unique point along the geodesic connecting i to η whose hyperbolic distance from i is t. And define

$$b(x) = \begin{cases} 0 & \text{if } x \in \mathbb{H} \setminus \mathscr{H} \\ d(x, \partial H_\eta) & \text{if } x \in H_\eta \in \mathscr{H} \end{cases}, \tag{3.14}$$

where H_η is the horoball at η.

Theorem 3.2 ([22, Theorem 2]). There exists a constant $0 < C < \infty$ such that for any $\eta \in \mathcal{L}(\Gamma)$, a parabolic cusp, and for any $t > 0$,

$$C^{-1} e^{-\delta_\Gamma t} e^{b(\eta_t)(1-\delta_\Gamma)} \le \nu_i(\mathcal{B}(\eta, e^{-t})) \le C e^{-\delta_\Gamma t} e^{b(\eta_t)(1-\delta_\Gamma)} \tag{3.15}$$

where $\mathcal{B}(\eta, e^{-t}) \subset \partial\mathbb{H}$ is the ball centered at η of radius e^{-t}

Corollary 3.3. Assume that $\eta \in \mathcal{L}(\Gamma)$ is a parabolic cusp; in a small ball around η, we can approximate the measure:

$$d\nu_i(\eta + h) \le h^{2\delta_\Gamma - 2} dh. \tag{3.16}$$

This corollary follows by differentiating (3.15) with $h = e^{-t}$ and by noting $b(\eta_t) \le t$.

4 Horospherical Equidistribution

Consider an unstable horosphere for the geodesic flow Φ^t, N_+. We parameterize the projection by $n_+ : \mathbb{T} \to \Gamma \cap N_+ \backslash \Gamma N_+$. [11,Theorem 3.3] (which follows from [17, Theorem 3.6]) states

Theorem 4.1. Let λ be a Borel probability measure on $\mathbb{T}$ with continuous density with respect to Lebesgue. Then for every $f : \mathbb{T} \times \Gamma\backslash G \to \mathbb{R}$ compactly supported and continuous

$$\lim_{t\to\infty} e^{(1-\delta_\Gamma)t} \int_{\mathbb{T}} f(x, n_+(x)\Phi^t)\, \mathrm{d}\lambda(x) = \frac{1}{|m^{BMS}|} \int_{\mathbb{T}\times\Gamma\backslash G} f(x,\alpha)\lambda'(x)\, \mathrm{d}\mu^{PS}_{N_+}(x)\, \mathrm{d}m^{BR}(\alpha). \quad (4.1)$$

Furthermore, this theorem can be applied to characteristic functions (this follows in the same way as [11, Corollary 3.5]).

Corollary 4.2. Let λ be a Borel probability measure on $\mathbb{T}$ with continuous density with respect to Lebesgue. Let $\mathcal{E} \subset \mathbb{T} \times \Gamma\backslash G$ be a compact set with boundary of $(\mu^{PS}_{N_+} \times m^{BR})$-measure 0. Then

$$\lim_{t\to\infty} e^{(1-\delta_\Gamma)t} \int_{\mathbb{T}} \chi_{\mathcal{E}}(x, n_+(x)\Phi^t)\, \mathrm{d}\lambda(x) = \frac{1}{|m^{BMS}|} \int_{\mathbb{T}\times\Gamma\backslash G} \chi_{\mathcal{E}}(x,\alpha)\lambda'(x)\, \mathrm{d}\mu^{PS}_{N_+}(x)\, \mathrm{d}m^{BR}(\alpha). \quad (4.2)$$

4.1 Counting primitive points in sheared sets

As a straightforward consequence of Corollary 4.2, we have the following theorem, which (in Sections 5 and 7) we show has a number of important consequences.

Theorem 4.3. Let λ be a Borel probabilty measure on $\mathbb{T}$ with continuous density with respect to Lebesgue. Let $\mathcal{A} \subset \mathbb{R}^2$ be a compact set with boundary of Lebesgue measure 0. Then for every $k \geq 1$:

$$\lim_{t\to\infty} e^{(1-\delta_\Gamma)t} \lambda\left(\{x \in \mathbb{T} : \left|\mathcal{Z} n_+(x)\Phi^t \cap \mathcal{A}\right| = k\}\right) = \frac{C_\lambda}{|\mathrm{m}^{BMS}|} \mathrm{m}^{BR}(\{\alpha \in \Gamma\backslash G : |\mathcal{Z}\alpha \cap \mathcal{A}| = k\}), \quad (4.3)$$

where $C_\lambda = \mu^{PS}_{N_+}(\lambda')$.

Theorem 4.3 is an infinite covolume version of [16, Theorem 6.7]. The proof is a straightforward consequence of Corollary 4.2 and the fact that if $\mathcal{A}$ is compact and has boundary of Lebesgue measure 0, then

$$\{g \in \Gamma\backslash G : \mathcal{Z}g \cap \mathcal{A} = k\} \tag{4.4}$$

is compact and has boundary of volume 0, and the Burger–Roblin measure of a 0 volume set is 0.

Using [15, Theorem 6.10] in the same way, we used [17, Theorem 3.6] to derive Theorem 4.1, we have

Theorem 4.4. Let $\mathcal{A} \subset \mathbb{R}^2$ be a compact set with boundary of Lebesgue measure 0. Then for every $k \geq 1$:

$$\lim_{t\to\infty} \mu^{PS}_{N_+}\left(\{x \in \mathbb{T} : \left|\mathcal{Z}n_+(x)\Phi^t \cap \mathcal{A})\right| = k\}\right) = \frac{|\mu^{PS}_{N_+}|}{|\mathrm{m}^{BMS}|}\mathrm{m}^{BMS}(\{\alpha \in \Gamma\backslash G : |\mathcal{Z}\alpha \cap \mathcal{A}| = k\}). \tag{4.5}$$

In words, each of these two theorems is asking for the limiting probability that a randomly sheared set contains k points. In one instance (Theorem 4.3), we randomly shear the set with measure λ and in the other (Theorem 4.4) we use the measure $\mu^{PS}_{N_+}$.

4.2 Counting primitive points in rotated sets

Similarly to Section 4.1, one can ask about the probability of finding k primitive points in a randomly rotated set (as oppose to a randomly sheared one). In [11, Section 6], we show that similar equidistribution results to Theorem 4.1 and Corollary 4.2 also hold when the horospherical subgroup N_+ is replaced with the rotational subgroup, K. Parameterize the rotation subgroup K by the boundary $\partial\mathbb{H}$ in the natural way $x \mapsto R(x) = \begin{pmatrix} \cos 2\pi x & \sin 2\pi x \\ -\sin 2\pi x & \cos 2\pi x \end{pmatrix}$. Then the rotational Patterson–Sullivan measure is defined to be

$$d\mu^{PS}_K(x) = e^{\beta_x(i,R(x)(ei))}d\nu_i(x). \tag{4.6}$$

Note μ^{PS}_K is supported on $\mathcal{L}(\Gamma)$. Hence, the analogous theorem to Theorem 4.3 follows from [11, Corollary 6.2] (in exactly the same way that Theorem 4.3 follows from Corollary 4.2):

Theorem 4.5. Let λ be a Borel probability measure on $\mathbb{T}$ with continuous density with respect to Lebesgue. Let $\mathcal{A} \subset \mathbb{R}^2$ be a compact subset with boundary of Lebesgue measure 0. Then for every $k \geq 1$

$$\lim_{t\to\infty} e^{(1-\delta_\Gamma)t}\lambda\left(\{x \in \mathbb{T} : \left|\mathcal{Z}R(x)\Phi^t \cap \mathcal{A}\right| = k\}\right) = \frac{D_\lambda}{|\mathrm{m}^{BMS}|}\mathrm{m}^{BR}(\{\alpha \in \Gamma\backslash G : |\mathcal{Z}\alpha \cap \mathcal{A}| = k\}) \quad (4.7)$$

where $D_\lambda = \mu_K^{PS}(\lambda')$.

5 Consequences of Theorems 4.3 and 4.5

5.1 Diophantine approximation in Fuchsian groups

Theorem 4.3 can be used to prove several statements about the set of numbers, which can be approximated by parabolic points in the limit set of the Fuchsian groups studied here. In particular, as discussed in [2], Erdös–Szüsz–Turán (henceforth abbreviated EST) introduced the following problem in Diophantine approximation: what is the probability that a uniformly chosen point, $x \in [0, 1]$, satisfies

$$\left|x - \frac{p}{q}\right| \leq \frac{A}{q^2} \quad (5.1)$$

for $\frac{p}{q} \in \mathbb{Q}$ with $q \in [\theta Q, Q]$ for a fixed triple $(A, \theta, Q) \in \mathbb{R}_{>0} \times (0, 1) \times \mathbb{R}_{>0}$? Hence, if we let $EST(A, \theta, Q)$ be the random variable: the number of solutions to (5.1), the EST problem is to prove the existence of

$$\lim_{Q\to\infty} \mathbb{P}(EST(A, \theta, Q) > 0). \quad (5.2)$$

The limiting distribution for this random variable is given in [2] in great generality. Our goal in this section is to understand the same problem with the rationals replaced by Γ^∞.

Given a triple (A, θ, Q) as above and a number x, define (the analogue of the random variable EST), $E(A, \theta, Q)$ to be the number of solutions, $(p, q) \in \mathcal{Z}$, to

$$|p - qx| \leq \frac{A}{q}, \quad (5.3)$$

with $q < Q$.

Theorem 5.1. Given $(A,\theta) \in \mathbb{R}_{>0} \times (0,1)$. Let λ be a Borel probability measure on $[0,1)$, with continuous density with respect to Lebesgue. Then

$$\lim_{Q\to\infty} Q^{2(1-\delta_\Gamma)}\lambda(\{x \in [0,1) : E(A,\theta,Q) = k\}) = \frac{C_\lambda}{|\mathrm{m}^{BMS}|}\mathrm{m}^{BR}(\{\alpha \in \Gamma\backslash G : |\mathcal{Z}\alpha \cap \mathfrak{C}_{A,\theta}| = k\}), \tag{5.4}$$

where

$$\mathfrak{C}_{A,\theta} := \{(x_1,x_2) \in \mathbb{R}\times\mathbb{R} : |x_1|x_2 \le A : \theta < x_2 < 1\}. \tag{5.5}$$

Moreover,

$$\lim_{Q\to\infty} \mu^{PS}_{N_+}(\{x \in \mathcal{L}(\Gamma) \cap [0,1) : E(A,\theta,Q) = k\}) = \frac{1}{|\mathrm{m}^{BMS}|}\mathrm{m}^{BMS}(\{\alpha \in \Gamma\backslash G : |\mathcal{Z}\alpha \cap \mathfrak{C}_{A,\theta}| = k\}). \tag{5.6}$$

Proof. Write (5.4) as (with $Q = e^{t/2}$)

$$\lim_{t\to\infty} e^{(1-\delta_\Gamma)t}\lambda\left(\left\{x \in [0,1] : \#\left\{(p,q) \in \mathcal{Z} : (p,q)\begin{pmatrix} 1 & 0 \\ -x & 1\end{pmatrix}\begin{pmatrix} Q & 0 \\ 0 & Q^{-1}\end{pmatrix} \in \mathfrak{C}_{A,\theta}\right\} = k\right\}\right)$$
$$= \lim_{t\to\infty} e^{(1-\delta_\Gamma)t}\lambda\left(\left\{x \in [0,1] : \#\left(\mathcal{Z}n_+(-x)\Phi^t \cap \mathfrak{C}_{A,\theta}\right) = k\right\}\right). \tag{5.7}$$

To which we apply Theorem 4.3 to get (5.4).

Equation (5.6) follows in the same way except, in the last step, we apply Theorem 4.4 instead of Theorem 4.3. ■

Moreover, the same proof allows one to prove the *Kesten problem* in our context, stated as follows: for $A > 0$ and Q fixed let $K(A,Q)$ denote the number of solutions to

$$|\alpha q - p| \le \frac{A}{Q} \quad , \quad 1 \le q \le Q. \tag{5.8}$$

In this case, the following theorem holds:

Theorem 5.2. Given $A > 0$, Theorem 5.1 holds with $E(A,\theta,Q)$ replaced by $K(A,Q)$ and $\mathfrak{C}_{A,\theta}$ replaced by

$$R_A = \left\{(x,y) \in \mathbb{R}^2 : |x| \le A, 0 \le y \le 1\right\} \tag{5.9}$$

5.2 Directions of primitive points

Given a point in $\mathbb{R}^2$ (taken here to be the origin, however this is not necessary), one can ask how the directions of primitive points $\mathcal{Z}$ distribute for an observer at that point. The corollary of Theorem 4.5 below answers this question.

Let $\mathcal{D}_t(\sigma, v) \subset S_1^1$ be the interval in the unit sphere with center v and length σe^{-t}, and set

$$\mathcal{N}_t(\sigma, v; \mathcal{Z}) := \#\left\{\mathbf{y} \in \mathcal{Z}_t : \|\mathbf{y}\|^{-1}\mathbf{y} \in \mathcal{D}_t(\sigma, v)\right\}, \tag{5.10}$$

where $\mathcal{Z}_t = \{z \in \mathcal{Z} : \|z\| \leq e^t\}$.

Corollary 5.3. Let λ be a probability measure on $\mathbb{T}$, with continuous density with respect to Lebesgue. For $k \in \mathbb{N}_{>0}$, we have

$$\lim_{t\to\infty} e^{(1-\delta_\Gamma)t}\lambda\left(\{v \in \mathbb{T} : \mathcal{N}_t(\sigma, v; \mathcal{Z}) = k\}\right) = \frac{D_\lambda}{|\mathrm{m}^{BMS}|}\mathrm{m}^{BR}\left(\{\alpha \in \Gamma\backslash G : |\mathcal{Z}\alpha \cap \mathfrak{C}_\sigma| = k\}\right) \tag{5.11}$$

where, in polar coordinates

$$\mathfrak{C}_\sigma = \left\{x = (r, \theta) \in \mathbb{R}^2 : r < 1, |\theta| < \sigma\pi\right\}. \tag{5.12}$$

This corollary follows directly from Theorem 4.5.

6 Equidistribution of gFs

6.1 Statement

In addition to Theorem 4.3, another important consequence of the equidistribution statements in Section 4 is the following theorem, stating that the gFs equidistributes on a horospherical section. This is a generalization of [13, Theorem 6], to the infinite covolume setting.

Theorem 6.1. Let $\sigma \in \mathbb{R}$ and $Q = e^{(t-\sigma)/2}$. Let $f : \mathbb{T} \times \Gamma\backslash G \to \mathbb{R}$ be bounded continuous and supported on a set with finite volume. Then

$$\lim_{t\to\infty} e^{-\delta_\Gamma t}\sum_{r\in\mathcal{F}_Q} f(r, n_-(r)\Phi^{-t}) = \frac{e^{(\delta_\Gamma-1)\sigma}}{|m^{BMS}|}\int_{\mathbb{T}\times\mathbb{T}}\int_\sigma^\infty \tilde{f}(x, n_-(w)\Phi^{-r})e^{\delta_\Gamma r}\,\mathrm{d}r\,\mathrm{d}\mu^{PS}(w)\,\mathrm{d}\mu^{PS}_{N_+}(x) \tag{6.1}$$

where $\tilde{f}(x, \alpha) := f(x, {}^t\alpha^{-1})$.

Remark. Marklof [13] and [14] treat Farey sequences in general dimension. However, in the infinite covolume setting, equidistribution results for $\mathrm{SL}(d,\mathbb{R})$ have not yet been proved (to our knowledge).

6.2 Proof

Proof of Theorem 6.1. The proof will follow the same lines as [13, Proof of Theorem 6] with several exceptions as we are not working with Haar measure.

Note first that by setting $f(x,\alpha) = f_0(x,\alpha\Phi^{-\sigma})$ for f_0 bounded and continuous we may assume that $\sigma = 0$.

Step 1: first, we show that we can reduce the theorem to f compactly supported via a standard approximation argument. Assume the theorem holds for compactly supported functions. Now consider a bounded, continuous function, f supported on a finite-volume set. Fix $\epsilon > 0$ and consider (for some t) the difference

$$\left| e^{-\delta_\Gamma t} \sum_{r\in\mathcal{F}_Q} f(n_-(r)\Phi^{-t})) - \frac{1}{|m^{BMS}|} \int_{\mathbb{T}} \int_0^\infty \tilde{f}(n_-(w)\Phi^{-r}) e^{\delta_\Gamma r} \, \mathrm{d}r \, \mathrm{d}\mu^{PS}(w) \right| . \tag{6.2}$$

Now decompose $f = f_1 + f_2$ such that f_1 is supported on a compact set and f_2 is supported on a set of volume $\varrho > 0$ (as $\mathrm{supp}(f)$ has finite volume ϱ can be chosen arbitrarily small) and both are bounded and continuous. Hence, the difference (6.2) is bounded above by

$$\left| e^{-\delta_\Gamma t} \sum_{r\in\mathcal{F}_Q} f_1(n_-(r)\Phi^{-t})) - \frac{1}{|m^{BMS}|} \int_{\mathbb{T}} \int_0^\infty \tilde{f}_1(n_-(w)\Phi^{-r}) e^{\delta_\Gamma r} \, \mathrm{d}r \, \mathrm{d}\mu^{PS}(w) \right|$$
$$+ \left| e^{-\delta_\Gamma t} \sum_{r\in\mathcal{F}_Q} f_2(n_-(r)\Phi^{-t})) - \frac{1}{|m^{BMS}|} \int_{\mathbb{T}} \int_0^\infty \tilde{f}_2(n_-(w)\Phi^{-r}) e^{\delta_\Gamma r} \, \mathrm{d}r \, \mathrm{d}\mu^{PS}(w) \right| . \tag{6.3}$$

Applying Theorem 6.1 for compact functions implies we can take t large enough that the 1st term is less than $\epsilon/2$.

We may assume that f_2 is supported on the cusp at infinity, that is, $\mathrm{supp}(f_2) = \{z \in \mathbb{H} : \Im(z) > \varrho^{-1}\}$. With that, using the bounded property of f, there exists a $C < \infty$ such that

$$\left| \left|\mathcal{F}_Q\right|^{-1} \sum_{r\in\mathcal{F}_Q} f_2(n_-(r)a_{-t}) \right| \le \frac{C\#\{r \in \mathcal{F}_Q : \Im(\pi_1(n_-(r)a_{-t})) > \varrho^{-1}\}}{\left|\mathcal{F}_Q\right|} \tag{6.4}$$

where π_1 denotes the projection to the fundamental domain above i extending to infinity. This proportion can be upper bounded by $\frac{C|\mathcal{F}_{\varrho Q}|}{|cF_Q|} = C\varrho^{2\delta_\Gamma}$ for some constant $C < \infty$. Thus, by choosing ϱ large enough the summation in the right-hand term in (3) can be bounded by $\epsilon/4$.

Lastly, consider the term

$$\left|\int_{\mathbb{T}}\int_0^\infty \tilde{f}_2(n_-(w)a_{-r})e^{\delta_\Gamma r}\,\mathrm{d}r\,\mathrm{d}\mu^{PS}(w))\right| < \infty. \tag{6.5}$$

As Γ has a cusp, $\delta_\Gamma > 1/2$. Thus, the Patterson–Sullivan measure of $\mathrm{supp}(\tilde{f}_2) \cap \mathcal{L}(\Gamma)$ goes to 0 as $\mathrm{vol}(\mathrm{supp}(\tilde{f}_2))$ goes to 0. Hence, we can choose ϱ such that (6.2) is bounded by ϵ. Thus, Theorem 6.1 for compactly supported f implies the theorem for f with finite volume support.

Henceforth, take f to be compactly supported.

Step 2: note that because f is continuous and has compact support it is uniformly continuous. Hence, for every $\varrho > 0$, there exists a $\epsilon > 0$ such that for all $(x,\alpha),(x'\alpha') \in \mathbb{R}\times G$

$$|x - x'| < \epsilon \qquad d(\alpha,\alpha') < \epsilon \tag{6.6}$$

imply $|f(x,\alpha) - f(x',\alpha')| < \varrho$

Step 3: for $0 \le \theta < 1$ and $\epsilon > 0$, define

$$\mathcal{F}_{Q,\theta} := \left\{\frac{p}{q} \in [0,1) : (p,q) \in \mathcal{Z},\ \theta Q < q < Q\right\} \tag{6.7}$$

$$\mathcal{F}_Q^\epsilon := \bigcup_{r\in\mathcal{F}_{Q,\theta}+\mathbb{Z}} \left\{x \in \mathbb{R} : \|x - r\| < \epsilon e^{-t}\right\}. \tag{6.8}$$

The latter we can write as

$$\mathcal{F}_Q^\epsilon = \bigcup_{\mathbf{a}\in\mathcal{Z}} \left\{x \in \mathbb{R} : (a_1,a_2)n_+(x)\Phi^t \in \mathfrak{C}_\epsilon\right\}, \tag{6.9}$$

where

$$\mathfrak{C}_\epsilon := \{(y_1,y_2) \in \mathbb{R}^2 : |y_1| < \epsilon y_2, \quad \theta < y_2 \le 1\}.$$

Our goal is to write the characteristic function for $\mathcal{F}_Q^\epsilon$ as a sum over simpler characteristic functions, which can write as a disjoint union. Thus, let

$$\mathcal{H}_\epsilon := \bigcup_{\mathbf{a}\in\mathcal{Z}} \mathcal{H}_\epsilon(\mathbf{a}), \qquad \mathcal{H}_\epsilon(\mathbf{a}) := \{\alpha \in G : (a_1, a_2)\alpha \in \mathfrak{C}_\epsilon\}. \tag{6.10}$$

By considering the bijection

$$\Gamma_{N_-}\backslash\Gamma \to \mathcal{Z}, \qquad \Gamma_{N_-}\gamma \mapsto (0,1)\gamma$$

we can write

$$\begin{aligned} \mathcal{H}_\epsilon &= \bigcup_{\gamma\in\Gamma_{N_-}\backslash\Gamma} \mathcal{H}_\epsilon((0,1)\gamma) \\ &= \bigcup_{\gamma\in\Gamma_{N_-}\backslash\Gamma} \gamma\mathcal{H}_\epsilon^1, \end{aligned} \tag{6.11}$$

where

$$\mathcal{H}_\epsilon^1 := \mathcal{H}_\epsilon((0,1)) = H\{M_{\mathbf{y}} : \mathbf{y} \in \mathfrak{C}_\epsilon\}$$

with $M_{\mathbf{y}} := \begin{pmatrix} y_2^{-1} & 0 \\ y_1 & y_2 \end{pmatrix}$.

Step 4: *Claim:* given $\mathcal{C} \subset G$ compact, there exists an $\epsilon_0 > 0$ such that for all $\epsilon < \epsilon_0$

$$\gamma\mathcal{H}_\epsilon^1 \cap \mathcal{H}_\epsilon^1 \cap \Gamma\mathcal{C} = \emptyset, \tag{6.12}$$

for all $\gamma \in \Gamma/\Gamma_{N_-} \neq 1$

Proof of Claim. Equation (6.12) is equivalent to

$$\mathcal{H}_\epsilon((p,q)) \cap \mathcal{H}_\epsilon^1 \cap \Gamma\mathcal{C} = \emptyset, \qquad \forall (p,q) \neq (0,1) \in \mathcal{Z}. \tag{6.13}$$

Consider an $\alpha \in G$ such that $(p,q)\alpha \in \mathfrak{C}_\epsilon$. We can write any such α as

$$\alpha = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}\begin{pmatrix} y_2^{-1} & 0 \\ y_1 & y_2 \end{pmatrix} \tag{6.14}$$

for $b \in \mathbb{R}$ and $y_1 \in \mathbb{R}$.

Therefore, if we assume for the sake of contradiction that $(p,q)\alpha \in \mathfrak{C}_\epsilon$ and $(0,1)\alpha \in \mathfrak{C}_\epsilon$ we have the following four inequalities

$$|y_2^{-1}p + (pb+q)y_1| < \epsilon y_2(pb+q) \tag{6.15}$$

$$\theta < y_2(pb+q) \leq 1 \tag{6.16}$$

$$|y_1| < \epsilon y_2 \tag{6.17}$$

$$\theta < y_2 \leq 1. \tag{6.18}$$

Using (6.16) and (6.17) gives that

$$|(pb+q)y_1| < \epsilon, \tag{6.19}$$

which, plugging that into (6.15), gives

$$|y_2^{-1}p| < 2\epsilon. \tag{6.20}$$

Hence,

$$|p| < 2\epsilon. \tag{6.21}$$

Thus, $p = 0$. Therefore, $(0,q) = (0,1)\gamma$ for some $\gamma \in \Gamma$. However, since $\Gamma_\infty = \left\langle \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \right\rangle$, $q = 1$, which is a contradiction proving the statement. ∎

Step 5: the claim implies that for $\mathcal{C} \subset G$ compact there is an $\epsilon_0 > 0$ such that for all $\epsilon < \epsilon_0$ such that

$$\mathcal{H}_\epsilon \cap \Gamma\mathcal{C} = \bigcup_{\gamma \in \Gamma/\Gamma_{N_-}} (\gamma\mathcal{H}_\epsilon^1 \cap \Gamma\mathcal{C}) \tag{6.22}$$

is a disjoint union. Thus, let χ_ϵ and χ_ϵ^1 denote the characteristic functions of $\mathcal{H}_\epsilon$ and $\mathcal{H}_\epsilon^1$, respectively, then

$$\chi_\epsilon(\alpha) = \sum_{\gamma \in \Gamma_{N_-}\backslash\Gamma} \chi_\epsilon^1(\gamma\alpha) \tag{6.23}$$

for all $\alpha \in \Gamma\mathcal{C}$. Moreover, all of the sets we consider have boundary of BR-measure 0. Set $\tilde{\chi}_\epsilon(\alpha) := \chi_\epsilon({}^t\alpha^{-1})$ and note that $\chi_\epsilon(n_+(x)\Phi^t) = \tilde{\chi}_\epsilon(n_-(-x)\Phi^{-t})$ is the characteristic function for $\mathcal{F}^\epsilon_Q$ Therefore, we write

$$\int_{\mathcal{F}^\epsilon_Q/\mathbb{Z}} f(x, n_-(x)\Phi^{-t})\,\mathrm{d}x = \int_{\mathbb{T}} f(x, n_-(x)\Phi^{-t})\chi_\epsilon(n_+(-x)\Phi^t)\,\mathrm{d}x$$
$$= \int_{\mathbb{T}} \tilde{f}(x, n_+(-x)\Phi^t)\chi_\epsilon(n_+(-x)\Phi^t)\,\mathrm{d}x, \tag{6.24}$$

to which we can apply Theorem 4.1 giving the following:

$$\lim_{t\to\infty} e^{(1-\delta_\Gamma)t} \int_{\mathcal{F}^\epsilon_Q/\mathbb{Z}} f(x, n_-(x)\Phi^{-t})\,\mathrm{d}x = \frac{1}{|m^{BMS}|}\int_{\Gamma\backslash G\times\mathbb{T}} \tilde{f}(x,\alpha)\chi_\epsilon(\alpha)\,\mathrm{d}m^{BR}(\alpha)\,\mathrm{d}\mu^{PS}_{N_+}(x), \tag{6.25}$$

which we write this:

$$= \frac{1}{|m^{BMS}|}\int_{\Gamma_{N_-}\backslash G\times\mathbb{T}} \tilde{f}(x,\alpha)\chi^1_\epsilon(\alpha)\,\mathrm{d}m^{BR}(\alpha)\,\mathrm{d}\mu^{PS}_{N_+}(x),$$
$$= \frac{1}{|m^{BMS}|}\int_{\Gamma_{N_-}\backslash N_-\{M_{\mathbf{y}}:\mathbf{y}\in\mathfrak{C}_\epsilon\}\times\mathbb{T}} \tilde{f}(x,\alpha)\,\mathrm{d}m^{BR}(\alpha)\,\mathrm{d}\mu^{PS}_{N_+}(x). \tag{6.26}$$

Step 6:

Using Proposition 3.1, we write (6.26) as (noting that $(0,1)n_- = (0,1)$)

$$= \frac{2}{|m^{BMS}|}\int_{\mathbb{T}\times\{\mathbf{y}\in\mathfrak{C}_\epsilon\}\times\mathbb{T}} y_2^{2\delta_\Gamma-2}\tilde{f}(x, n_-(w)M_{\mathbf{y}})\,\mathrm{d}y_2\,\mathrm{d}y_1\,\mathrm{d}\mu^{PS}(w)\,\mathrm{d}\mu^{PS}_{N_+}(x), \tag{6.27}$$

which we can write

$$= \frac{2}{|m^{BMS}|}\int_{\mathbb{T}\times\mathbb{T}}\int_\theta^1\int_{\mathcal{B}_{\epsilon y_2}(0)} y_2^{2\delta_\Gamma-2}\tilde{f}(x, n_-(w)M_{\mathbf{y}})\,\mathrm{d}y_2\,\mathrm{d}y_1\,\mathrm{d}\mu^{PS}(w)\,\mathrm{d}\mu^{PS}_{N_+}(x). \tag{6.28}$$

Next we write $D(y_2) := \begin{pmatrix} y_2^{-1} & 0 \\ 0 & y_2 \end{pmatrix}$ and note

$$d(M_{\mathbf{y}}, D(y_2)) = d(n_+(y_2^{-1}y_1), Id) \le \epsilon \tag{6.29}$$

for $\mathbf{y} \in \mathfrak{C}_\epsilon$ (this is the same calculation as [13, (3.42)]). Therefore, using the uniform continuity of Step 2:

$$\begin{aligned}&\left|(6.26) - \frac{2}{|m^{BMS}|}\int_{\mathbb{T}\times\mathbb{T}}\int_\theta^1\int_{\mathcal{B}(\epsilon y_2)}\tilde{f}(x,n_-(w)D(y_2))y_2^{2\delta_\Gamma-2}\,\mathrm{d}y_2\,\mathrm{d}y_1\,\mathrm{d}\mu^{PS}(w)\,\mathrm{d}\mu^{PS}_{N_+}(x)\right|\\ &\quad= \left|(6.26) - \frac{4\epsilon}{|m^{BMS}|}\int_{\mathbb{T}\times\mathbb{T}}\int_\theta^1\tilde{f}(x,n_-(w)D(y_2))y_2^{2\delta_\Gamma-1}\,\mathrm{d}y_2\,\mathrm{d}\mu^{PS}(w)\,\mathrm{d}\mu^{PS}_{N_+}(x)\right| \\ &\quad\le \frac{4\varrho\epsilon|\mu^{PS}|^2}{|m^{BMS}|}\int_\theta^1 y_2^{2\delta_\Gamma-1}\,\mathrm{d}y_2.\end{aligned} \tag{6.30}$$

Evaluating this integral then gives that (6.30) is equal to

$$\frac{2\epsilon\varrho|\mu^{PS}|^2}{|m^{BMS}|\delta_\Gamma}(1-\theta^{2\delta}). \tag{6.31}$$

Now, if we consider the left-hand side of (6.30) and insert (6.25) and finally perform a change of variables writing $y_2 = e^{r/2}$, we conclude that

$$\begin{aligned}&\left|\lim_{t\to\infty} e^{(1-\delta_\Gamma)t}\int_{\mathcal{F}^\epsilon_Q/\mathbb{Z}} f(x,n_-(x)\Phi^{-t})\,\mathrm{d}x\right.\\ &\qquad\left.-\frac{2\epsilon}{|m^{BMS}|}\int_{\mathbb{T}\times\mathbb{T}}\int_0^{2|\ln\theta|}\tilde{f}(x,n_-(w)\Phi^{-t})e^{\delta_\Gamma r}\,\mathrm{d}r\,\mathrm{d}\mu^{PS}(h)\,\mathrm{d}\mu^{PS}_{N_+}(x)\right|\\ &\qquad\qquad< \frac{2\varrho\epsilon|\mu^{PS}|^2}{|m^{BMS}|\delta_\Gamma}(1-\theta^{2\delta_\Gamma}).\end{aligned} \tag{6.32}$$

Step 7: to conclude, consider

$$\lim_{t\to\infty} e^{-\delta_\Gamma t}\sum_{r\in\mathcal{F}_{Q,\theta}} f(r,n_-(r)\Phi^{-t}) \tag{6.33}$$

taking the asymptotic formula (1.10) and using a volume estimate together with uniform continuity (see [13, (3.49)] for details) we can write this as

$$= \lim_{\epsilon\to 0}\lim_{t\to\infty}\frac{e^{(1-\delta_\Gamma)t}}{e^t}\frac{e^t}{2\epsilon}\sum_{r\in\mathcal{F}_{\theta,Q}}\int_{|x-r|<\epsilon e^{-t}} f(x,n_-(x)\Phi^{-t})\,\mathrm{d}x, \tag{6.34}$$

which is equal

$$= \lim_{\epsilon\to 0}\lim_{t\to\infty}\frac{e^{(1-\delta_\Gamma)t}}{2\epsilon}\sum_{r\in\mathcal{F}_{\theta,Q}}\int_{|x-r|<\epsilon e^{-t}} f(x,n_-(x)\Phi^{-t})\,\mathrm{d}x \tag{6.35}$$

Then using the disjoint union in (6.22) we can say

$$= \lim_{\epsilon\to 0}\lim_{t\to\infty}\frac{e^{(1-\delta_\Gamma)t}}{2\epsilon}\int_{\mathcal{F}_Q^\epsilon\backslash\mathbb{Z}} f(x, n_-(x)\Phi^{-t})\,\mathrm{d}x \tag{6.36}$$

and using (32) we thus conclude after taking $\epsilon \to 0$ (and therefore $\varrho \to 0$) this is equal

$$= \frac{1}{|m^{BMS}|}\int_{\mathbb{T}\times\mathbb{T}}\int_0^{2|\ln\theta|}\tilde{f}(x, n_-(w)\Phi^{-r})e^{\delta_\Gamma r}\,\mathrm{d}r\,\mathrm{d}\mu^{PS}(w)\,\mathrm{d}\mu^{PS}_{N_+}(x) \tag{6.37}$$

Taking the limit as $\theta \to 0$ is then possible as

$$\limsup_{t\to\infty}\frac{|\mathcal{F}_Q\backslash\mathcal{F}_{Q\theta}|}{e^{\delta_\Gamma t}} = \theta c_\Gamma \tag{6.38}$$

7 Local Statistics

Theorem 4.3 and Theorem 6.1 can also be used to study the local statistics of $\mathcal{F}_Q$ when viewed as a point process on $[0, 1]$ (note once more we are assuming for notation that Γ^∞ is periodic on $[0, 1]$).

7.1 Statement

For $Q = e^{t/2}$, let $\mathscr{A} \subset \mathbb{R}$ be bounded interval and set $\mathscr{A}_t = \mathscr{A}e^{-t}$. For a bounded $\mathcal{D} \subset \mathbb{T}$, define

$$P_Q(\mathcal{D}, \mathscr{A}, k) = \frac{e^t\,\mathrm{vol}(\{x \in \mathcal{D} : |x + \mathscr{A}_t + \mathbb{Z} \cap \mathcal{F}_Q| = k\})}{\mu^{PS}_{N_+}(\mathcal{D})e^{\delta_\Gamma t}} \tag{7.1}$$

and

$$P_{0,Q}(\mathcal{D}, \mathscr{A}, k) = \frac{|\{r \in \mathcal{F}_Q \cap \mathcal{D} : |r + \mathscr{A}_t + \mathbb{Z} \cap \mathcal{F}_Q| = k\})}{\mu^{PS}_{N_+}(\mathcal{D})e^{\delta_\Gamma t}}. \tag{7.2}$$

Theorem 7.1. Given an interval $\mathscr{A} \subset \mathbb{R}$ and $\mathcal{D} \subset \mathbb{T}$, then for all $k > 0$

$$\lim_{Q\to\infty} P_Q(k, \mathcal{D}, \mathscr{A}) = P(k, \mathscr{A}) \tag{7.3}$$

$$\lim_{Q\to\infty} P_{0,Q}(\mathcal{D}, \mathscr{A}, k) = P_0(k, \mathscr{A}) \tag{7.4}$$

where $P(k, \mathscr{A})$ and $P_0(k, \mathscr{A})$ are given explicitly.

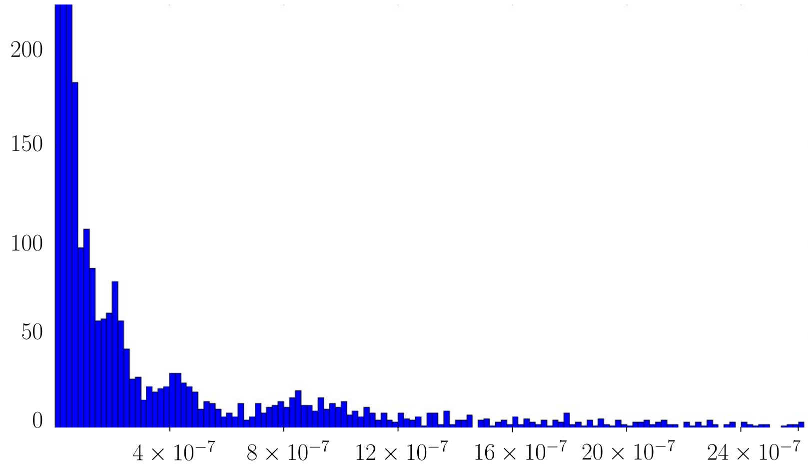

Fig. 3. Above we have shown the gaps in the point set $\widehat{\Gamma}^{\infty}$. The point set is exactly the one shown in Figure 1. We have cut off the image at 240 (thus, the 1st three bars do not have the same height) and the bin size here is 4×10^{-8}. Hence, the bars represent the number of gaps lying in a particular bin.

Remark. In particular, (7.4) implies that the limiting gap distribution exists everywhere.

Remark. Note that the above theorem is restricted to $k > 0$. The reason for this is that the scaling in P_Q and $P_{0,Q}$ is incorrect for the case $k = 0$. For geometrically finite subgroups, the boundary points cluster close together in far apart cluster. This phenomenon was noticed by Zhang [25] and again in [11].

To give another qualitative example, we have graphed the gap distribution for $\widehat{\Gamma}^{\infty}$ in Figure 3.

7.2 Proof

Proof of Theorem 7.1. Theorem 7.1 is a straightforward consequence of Theorem 4.3 and Theorem 6.1. We begin by addressing (7.3), define

$$\mathfrak{C}(\mathscr{A}) := \{(x, y) \in \mathbb{R} \times (0, 1] : x \in \mathscr{A}y\} \subset \mathbb{R}^2 \tag{7.5}$$

and note that

$$\frac{p}{q} \in x + \mathscr{A}_t \quad , \quad 0 < q \leq Q \tag{7.6}$$

is equivalent to

$$\Longleftrightarrow (p,q)n_+(\mathbf{x})\Phi^t \in \mathfrak{C}(\mathscr{A}). \tag{7.7}$$

Therefore, for a given $x \in \mathcal{D}$,

$$P_Q(\mathcal{D},\mathscr{A},k) = \frac{e^{(1-\delta_\Gamma)t}}{\mu^{PS}_{N_+}(\mathcal{D})} \operatorname{vol}(\{x \in \mathcal{D} : \left|\mathcal{Z}n_+(x)\Phi^t \cap \mathfrak{C}(\mathscr{A})\right| = k\}). \tag{7.8}$$

Applying Theorem 4.3 then implies

$$P(k,\mathscr{A}) = \frac{1}{|m^{BMS}|} m^{BR}(\mathcal{S}_k), \tag{7.9}$$

where $\mathcal{S}_k = \{\alpha \in \Gamma\backslash G : |\mathcal{Z}\alpha \cap \mathfrak{C}(\mathscr{A})| = k\}$.

Turning now to (7.4). Write

$$\begin{aligned} P_0(\mathscr{A},k) &= \lim_{t\to\infty} \frac{\left|\{r \in \mathcal{F}_Q \cap \mathcal{D} : |\mathcal{Z}n_+(r)\Phi^t \cap \mathfrak{C}(\mathscr{A})| = k\}\right|}{e^{\delta_\Gamma t}\mu^{PS}_{N_+}(\mathcal{D})} \\ &= \lim_{t\to\infty} \frac{\sum_{r\in\mathcal{F}_Q} \chi_{\mathcal{S}_k}(r, n_+(r)\Phi^t)}{\mu^{PS}_{N_+}(\mathcal{D})e^{\delta_\Gamma t}}. \end{aligned} \tag{7.10}$$

Applying Theorem 6.1 (after extending it to characteristic functions as done in [11]) gives

$$P_0(\mathscr{A},k) = \frac{1}{|m^{BMS}|} \int_{\mathbb{T}\times[0,\infty)} \tilde{\chi}_{\mathcal{S}_k}(n_-(w)\Phi^{-r})e^{\delta_\Gamma r}\,\mathrm{d}r\,\mathrm{d}\mu^{PS}(w). \tag{7.11}$$

Note that the quantity in (7.9) is finite for $k > 0$. This was proven in [11, Proposition 4.3]. This does not hold for $k = 0$ and is the reason for that restriction in the theorem. The integral on the right-hand side of (7.11) is finite whenever the Burger–Roblin measure is finite. Hence, the same [11, Proposition 4.3] implies finiteness of (7.11) as well. ∎

8 Explicit Gap Distribution for $\widehat{\Gamma}$

We now return to the example, $\widehat{\Gamma}$, discussed in Section 1. First, note that Theorem 7.1 implies that, in the limit $T \to \infty$, the gap distribution in (1.13) exists for all $s > 0$. Our goal is to prove the following theorem, which gives a far more explicit formula for the limiting gap distribution:

Theorem 8.1. For $s < s_0 = 7.5$, and $\mathcal{I}$ a closed interval in $[0,1]$, the limiting gap distribution can be written

$$\lim_{T\to\infty} \widehat{F}_{T,\mathcal{I}}(s) =: \widehat{F}_{\mathcal{I}}(s) = F^{1,2}_{\mathcal{I}}(s) + F^{2,3}_{\mathcal{I}}(s), \tag{8.1}$$

where $F^{1,2}_{\mathcal{I}}(s)$ and $F^{2,3}_{\mathcal{I}}(s)$ are explicit integrals over compact regions with respect to a fractal measure (see (8.33)).

The proof follows the methodology of [19]; however, there are significant differences. The plan is to break up the gap distribution into a sum over pairs of circles in the initial configuration $\mathcal{K}_0$. Then, using the following lemma (of Rudnick and Zhang), we can express each term in this sum as an integral over a compact area.

Lemma 8.2 ([19, Lemma 3.5]). Let $M = \begin{matrix} a & b \\ c & d \end{matrix} \in \mathrm{SL}(2,\mathbb{R})$.

(i) If $c \neq 0$, then under the Möbius transform M, a circle $C(x+yi, y)$ is mapped to

$$C\left(\frac{ax+b}{cx+d} + \frac{yi}{(cx+d)^2}, \frac{y}{(cx+d)^2}\right) \tag{8.2}$$

if $cx+d \neq 0$, and to the line $\Im \mathbf{z} = 1/2c^2y$ if $cx+d = 0$. When $c = 0$, the image circle is

$$C\left(\frac{ax+b}{d}, \frac{y}{d^2}\right). \tag{8.3}$$

(ii) If $c \neq 0$, then the line $C = \mathbb{R} + yi$ is mapped to

$$C\left(\frac{a}{c} + \frac{1}{2c^2y}i, \frac{1}{2c^2y}\right), \tag{8.4}$$

and to the line $\mathbb{R} + a^2yi$ if $c = 0$.

8.1 Breaking the gap distribution up

In [19], a fundamental observation is that at a given level T, the two circles corresponding to neighboring tangencies can be mapped by exactly one or two group elements to a pair in the initial configuration. That is not true here; however, the following proposition states that this is the case in the interval $[0, s_0)$.

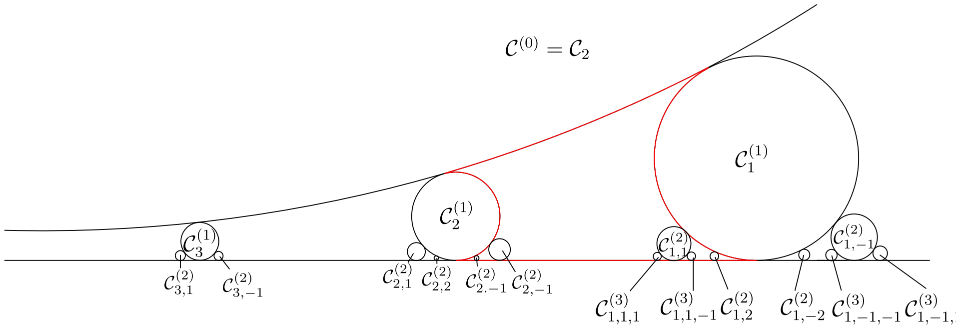

Fig. 4. The labeling used in this section. For clarity, we only show a portion of the interval and a few circles in $\mathcal{K}$. The red section is what we call the rectangle $(\mathcal{C}_1^{(1)}, \mathcal{C}_2^{(1)}, \mathcal{C}^{(0)}, \mathcal{C}_0)$.

Proposition 8.3. For any T and $\mathcal{I}$, suppose $\mathcal{C}$ and $\mathcal{C}'$ are the circles tangent to $\mathcal{C}_0$ at $x^j_{T,\mathcal{I}}$ and $x^{j+1}_{T,\mathcal{I}}$. If $T(x^{j+1}_{T,\mathcal{I}} - x^j_{T,\mathcal{I}}) \leq s$ for $s < s_0$ then there exists a $\gamma \in \Gamma$ such that $\mathcal{C} = \gamma\mathcal{C}_l$ and $\mathcal{C}' = \gamma\mathcal{C}_m$ for $\mathcal{C}_l \neq \mathcal{C}_m \in \mathcal{K}_0$ and neither equal $\mathcal{C}_0$. Moreover, if $\mathcal{C}$ and $\mathcal{C}'$ are not tangent then γ is unique and if they are tangent then there exist exactly two such γ.

Remark. The reason we consider $s < s_0$ in Theorem 8.1 is that Proposition 8.3 fails for larger s. In words, for larger s some of the gaps considered are not the image of a pair in the initial configuration. To get around this, one could consider a larger initial configuration (i.e., consider $\mathcal{K}$ together with the circles tangent at $1/4$ and $4 - 1/4$). This would allow Proposition 8.3 to hold for slightly larger s_0. Therefore, as one considered larger and larger gaps, one would need to consider larger and larger initial configurations and more and more terms in the decomposition below. In this paper, we will stick to the case $s_0 = 7.5$ as it will simplify the following proofs.

For ease of notation, we restrict our attention to circles tangent to $\mathcal{C}_0$ in $[0, 2]$ (i.e., beneath $\mathcal{C}_2$) and adopt the following notation shown in Figure 4: first label $\mathcal{C}_2 = \mathcal{C}^0$ and

- The tangencies are labeled by their continued fraction expansions $\alpha^{(i)}_{k_1,\ldots,k_i} = [0; 4k_1, \ldots 4k_i]$.
- The associated circles are labeled $\mathcal{C}^{(i)}_{k_1,\ldots,k_i}$.
- The diameter of each circle is similarly labeled $h^{(i)}_{k_1,\ldots,k_i}$.

Thus, each circle $\mathcal{C}^{(i)}_{k_1,\ldots,k_i}$ is the *child* of the circle $\mathcal{C}^{(i-1)}_{k_1,\ldots,k_{i-1}}$ (to which it is tangent) and the *parent* of $\mathbb{Z}_{\neq 0}$ children - $\mathcal{C}^{(i+1)}_{k_1,\ldots,k_i,k_{i+1}}$ (to which it is also tangent).

Define a rectangle to be any collection of circles

$$\mathcal{R} = (\mathcal{C}^{(i)}_{k_1,\dots,k_{i-1},k_i}, \mathcal{C}^{(i)}_{k_1,\dots,k_{i-1},k_i\pm1}, \mathcal{C}^{(i-1)}_{k_1,\dots,k_{i-1}}, \mathcal{C}_0) \quad , \quad (k_i \neq 0), \tag{8.5}$$

where $k_i \pm 1 \neq 0$ (see for example the rectangle in Figure 4). A rectangle is thus a pair of neighbors in a generation, the shared parent and the real line. Let $\mathcal{R}_0$ denote the rectangle $(\mathcal{C}_0, \mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3)$ of the initial configuration. The following simple observation is the basis of the proof of Proposition 8.3.

Fact 8.1. For any rectangle $\mathcal{R}$, there exists a unique $\gamma \in \widehat{\Gamma}$

$$\mathcal{R} = \gamma \mathcal{R}_0. \tag{8.6}$$

The configuration $\mathcal{K} = \Gamma\mathcal{K}_0$ where $\mathcal{K}_0$ is the initial configuration. Since circle inversions send circles to circles preserving tangencies, there must be a $\gamma \in \widehat{\Gamma}$ sending $\mathcal{R}_0$ to $\mathcal{R}$. Moreover, the uniqueness follows as we are working in $\mathrm{PSL}(2,\mathbb{Z})$.

Proof of Proposition 8.3. In this proof, given two circles with tangencies α_1 and α_2 and diameters h_1 and h_2, we refer to $|\alpha_1 - \alpha_2|$ as the gap associated to them and to $\min\{h_1, h_2\}^{-1} |\alpha_1 - \alpha_2|$ as the *scaled gap* associated to them. Note that if a scaled gap is larger than s_0, then the gap *will never* contribute to $\widehat{F}_{T,\mathcal{I}}(s)$ for any T. Thus, that gap can be ignored. Fact 8.1 implies that Proposition 8.3 follows if we show that all scaled gaps associated to pairs of circles *not in* rectangles are larger than s_0.

Step 1: the scaled gap associated to a pair of *non-tangent* circles *in a rectangle* has the form

$$\min\{h^{(i)}_{k_1,\dots,k_i}, h^{(i)}_{k_1,\dots,k_i\pm1}\}^{-1} \left|\alpha^{(i)}_{k_1,\dots,k_i} - \alpha^{(i)}_{k_1,\dots,k_i\pm1}\right| \tag{8.7}$$

(again we assume $k_i \pm 1 \neq 0$).

Step 2: we now use some theory of continued fractions to show that (8.7) is bounded above 4. Therefore, the gap arising from non-tangent pairs *in* a rectangle is bounded above 4. Given a tangency $\alpha^{(i)}_{k_1,\dots,k_i} = [0; a_1, \dots a_i]$, let

$$\frac{b_n}{d_n} := [0; a_1, \dots, a_n] \tag{8.8}$$

for $n < i$ where b_n and d_n share no common factors. It is a classical exercise to show (see [9]):

$$b_n = a_n b_{n-1} + b_{n-2}, \qquad b_{-2} = 0, \qquad b_{-1} = 1 \tag{8.9}$$

$$d_n = a_n d_{n-1} + d_{n-2}, \qquad d_{-2} = 1, \qquad d_{-1} = 0 \tag{8.10}$$

and

$$d_n b_{n-1} - d_{n-1} b_n = (-1)^n. \tag{8.11}$$

Hence, we can write the following:

$$\begin{aligned} \min\{h^{(i)}_{k_1,\ldots,k_i}, h^{(i)}_{k_1,\ldots,k_i\pm 1}\}^{-1} \left|\alpha^{(i)}_{k_1,\ldots,k_i} - \alpha^{(i)}_{k_1,\ldots,k_i\pm 1}\right| \\ &= \max\{d_i', d_i\}^2 \left|[1; a_1, \ldots a_i] - [1; a_1, \ldots a_i \pm 4]\right| \\ &= \max\{d_i', d_i\}^2 \left|\frac{a_i b_{i-1} + b_{i-2}}{a_i d_{i-1} + d_{i-2}} - \frac{(a_i \pm 4) b_{i-1} + b_{i-2}}{(a_i \pm 4) d_{i-1} + d_{i-2}}\right| \\ &= \max\{d_i', d_i\}^2 \frac{4}{d_i d_i'} \geq 4, \end{aligned} \tag{8.12}$$

where b_i and d_i are, respectively, the numerator and denominator of $\alpha^{(i)}_{k_1,\ldots,k_i}$ and b_i' and d_i' are the numerator and denominator of $\alpha^{(i)}_{k_1,\ldots,k_i\pm 1}$ (and similarly for all d_j and b_j).

Step 3: suppose $\mathcal{C}^{(i)}_{m_1,\ldots,m_i} = \mathcal{D}_1$ and $\mathcal{C}^{(j)}_{n_1,\ldots,n_j} = \mathcal{D}_2$ are adjacent at time T and do not both belong to a rectangle. For notation, we assume $\alpha^{(i)}_{m_1,\ldots m_i} < \alpha^{(j)}_{n_1,\ldots n_j}$.

- By construction, there is a shared ancestor of $\mathcal{D}_1$ and $\mathcal{D}_2$, $\mathcal{C}^{(k)}_{m_1,\ldots,m_k} = \mathcal{B}_1$ (for $k < \min\{i,j\}$). That is $m_x = n_x$ for all $1 \leq x \leq k$
- At the $k+1$-st generation, $\mathcal{D}_1$ is the descendent of $\mathcal{C}_{m_1,\ldots,m_{k+1}} = \mathcal{B}_3$ and $\mathcal{D}_2$ is the descendent of $\mathcal{C}^{(k+1)}_{n_1,\ldots,n_{k+1}} = \mathcal{B}_2$ (see Figure 5) and $(\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3, \mathcal{C}_0)$ must form a rectangle (otherwise, $\mathcal{D}_1$ and $\mathcal{D}_2$ are clearly not adjacent at any times).
- Lastly, it is evident that $\mathcal{D}_1$ must be the *right-most* descendent of $\mathcal{B}_3$ of its generation. Thus, $|m_l| = 1$ for all $l > k+1$. Moreover, $\mathcal{D}_2$ must be the *left-most* descendent of $\mathcal{B}_2$ in its generation.

Motivated by these three geometric facts, we adopt the following notation (see Figure 5). In each generation l, we label the left-most descendent of $\mathcal{B}_2$ by $\mathcal{B}_{2(l-k)}$. Moreover, we label the right-most descendent of $\mathcal{B}_3$ by $\mathcal{B}_{2(l-k)+1}$. With that notation, all non-tangent adjacent pairs of circles at a given time are of the form $\mathcal{B}_x$, $\mathcal{B}_{x+1}$ for some x.

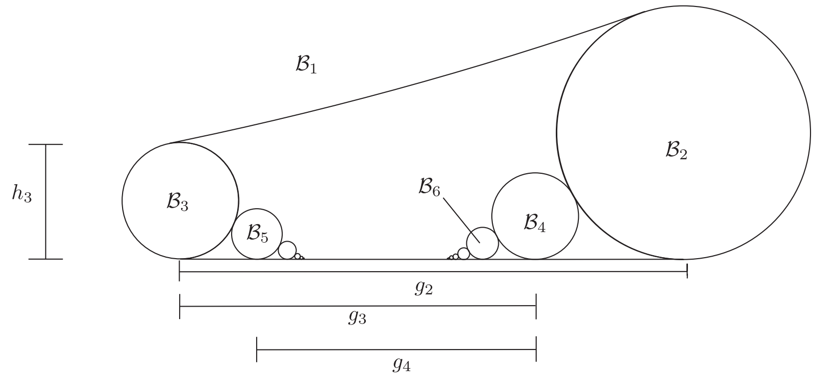

Fig. 5. Above we show the relevant rectangle, circles, and labeling for Step 3. We are only concerned with the ''innermost circles" in the rectangle. The circles are labeled in decreasing order of size.

Label the tangency associated to $\mathcal{B}_i$, α_i. Label the diameter of $\mathcal{B}_i$, h_i. We assume (w.l.o.g) $h_1 > h_2 > h_3$. Label the gap between $\mathcal{B}_i$ and $\mathcal{B}_{i+1}$, $g_i = |\alpha_i - \alpha_{i+1}|$. With this notation, all gaps associated to adjacent (non-tangent) pairs at time T are of the form g_i for $i \geq 2$. We show that $h_{i+1}^{-1} g_i$ (the scaled gap) is larger than 7.5 for all $i > 2$. This will prove the proposition as all gaps associated to non-tangent pairs are of this form.

First, assume $h_3 = 1$ (this is w.l.o.g by a simple scaling argument). Now we collect three facts:

- By (8.10) $h_2 \leq 4h_3 = 4$
- By (8.10) $h_{n+2} \leq \frac{h_n}{3^2}$
- By (8.9)–(8.11) $g_{i+1} \geq g_i - h_{i+2}^{\frac{1}{2}}$

Collecting these facts together lead to the following sequence of inequalities:

$$\begin{aligned}
h_3^{-1} g_2 &\geq 4 \\
h_4^{-1} g_3 &\geq \left(4 - \frac{2}{3}\right)\left(\frac{3}{2}\right)^2 \\
h_5^{-1} g_4 &\geq \left(4 - \frac{2}{3} - \frac{1}{3}\right) 3^2 \\
h_6^{-1} g_5 &\geq \left(4 - \frac{2}{3} - \frac{1}{3} - \frac{2}{9}\right)\left(\frac{9}{2}\right)^2 \\
h_7^{-1} g_6 &\geq \left(4 - \frac{2}{3} - \frac{1}{3} - \frac{2}{9} - \frac{1}{9}\right) 9^2 \\
&\vdots
\end{aligned} \tag{8.13}$$

hence, the gap arising from circles, which do not form the boundary of a rectangle, is at least $7\frac{1}{2}$.

This proves the proposition with $s_0 = 7\frac{1}{2}$ (this may not be sharp). ■

Now that we have established this proposition, the argument to prove Theorem 8.1 follows similar lines to Rudnick and Zhang. Note that Proposition 8.3 implies we can write the gap distribution for $s < s_0$ as

$$\widehat{F}_{T,\mathcal{I}}(s) = F^{1,2}_{T,\mathcal{I}}(s) + F^{2,3}_{T,\mathcal{I}}(s) \tag{8.14}$$

$$F^{i,j}_{T,\mathcal{I}}(s) := \frac{\#\left\{(x^l_{T,\mathcal{I}}, x^{l+1}_{T,\mathcal{I}}) \in \Gamma(\alpha_i, \alpha_j) \,\middle|\, T(x^{l+1}_{T,\mathcal{I}} - x^l_{T,\mathcal{I}}) \leq s\right\}}{T^{\delta_{\widehat{\Gamma}}}}, \tag{8.15}$$

where α_i are the tangencies associated to $\mathcal{C}_i$ in the initial configuration (the contribution from the tangent pair $(1,3)$ has already been counted from the $(1,2)$ pair because of the overcounting in Proposition 8.3 for gaps associated with tangent pairs).

8.2 Geometric description of the gap distribution

The Lemma 8.2 and the Proposition 8.4 play a crucial role in what follows. As these theorems are taken from [19] and are not specific to the subgroup considered, we will not repeat the details here.

We use Lemma 8.2 to provide conditions under which the image of $\mathcal{C}_i$ and $\mathcal{C}_j$ are adjacent at time T. Indeed, it follows from [19, Proposition 4.6] that there exist two regions $\Omega^{1,2}_T$ and $\Omega^{2,3}_T$ such that, for $M = \begin{matrix} a & b \\ c & d \end{matrix}$, the image $M(\alpha_i, \alpha_j)$ is an adjacent pair at time T if and only if $(c,d) \in \Omega^{i,j}_T$ (where $(i,j) = (1,2)$ or $(2,3)$).

We define these two regions as subsets of the cd-plane $\{(c,d) | c \geq 0\}$:

(a) We define $\Omega^{1,2}_T$ to be those $\{(c,d) | c \geq 0\}$ such that

$$c^2 \leq \frac{T}{2} \qquad , \qquad d^2 \leq \frac{T}{2} \tag{8.16}$$

$$(4c + |d|)^2 > \frac{T}{2}. \tag{8.17}$$

(b) We define $\Omega_T^{2,3}$ to be those $\{(c,d)|c \geq 0\}$ such that

$$d^2 \leq \frac{T}{2} \quad , \quad (4c+d)^2 \leq \frac{T}{2}. \tag{8.18}$$

$$\text{If} d(4c+d) < 0 \text{ then} c^2 > \frac{T}{2}. \tag{8.19}$$

Note that $\Omega_T^{i,j}$ is in both cases a union of convex sets and

$$\Omega_T^{i,j} = \sqrt{T}\Omega_1^{i,j}. \tag{8.20}$$

Hence, we have the following restatement of [19, Proposition 4.6] restricted to our context.

Proposition 8.4 ([19, Proposition 4.6]). For $\gamma = \left(\begin{smallmatrix} a_\gamma & b_\gamma \\ c_\gamma & d_\gamma \end{smallmatrix}\right) \in \Gamma$,

(a) the circles $\gamma(\mathcal{C}_1)$ and $\gamma(\mathcal{C}_2)$ are neighbors in $\mathcal{A}_T$ if and only if $(c_\gamma, d_\gamma) \in \sqrt{T}\Omega_1^{1,2}$.
(b) the circles $\gamma(\mathcal{C}_2)$ and $\gamma(\mathcal{C}_3)$ are neighbors in $\mathcal{A}_T$ if and only if $(c_\gamma, d_\gamma) \in \sqrt{T}\Omega_1^{2,3}$.

The relative gap condition in (8.15) can now be written (again following [19, (18)–(20)]):

(a) For $i = 1$ and $j = 2$,

$$c\,|d| \geq \frac{T}{s}. \tag{8.21}$$

(b) For $i = 2$ and $j = 3$,

$$|d(4c+d)| \geq \frac{4T}{s}. \tag{8.22}$$

Thus, we come to the same conclusion as Rudnick and Zhang that

$$F_{T,\mathcal{I}}^{i,j}(s) = \frac{1}{T^{\delta_\Gamma}} \# \left\{ \gamma = \begin{pmatrix} a_\gamma & b_\gamma \\ c_\gamma & d_\gamma \end{pmatrix} \in \Gamma \mid \gamma\alpha_i, \gamma\alpha_j \in \mathcal{I}, (c_\gamma, d_\gamma) \in \Omega_T^{i,j}(s) \right\} \tag{8.23}$$

for $(i,j) = (1,2), (2,3)$, where $\Omega_T^{i,j}(s)$ is defined to be those elements $(c,d) \in \Omega_T^{i,j}$ satisfying (8.21) for $(1,2)$ and (8.22) for $(2,3)$.

Note that $\Omega_T^{i,j}(s)$ are unions of convex, compact sets, and

$$\Omega_T^{i,j}(s) = \sqrt{T}\Omega_1^{i,j}(s). \tag{8.24}$$

8.3 Limiting behaviour

To ease notation and remain consistent with [19], we reparameterize the geodesic flow

$$A := \left\{ \begin{matrix} y^{-\frac{1}{2}} & 0 \\ 0 & y^{\frac{1}{2}} \end{matrix} \;|y > 0 \right\} \tag{8.25}$$

and set

$$A_T := \left\{ \begin{matrix} y^{-\frac{1}{2}} & 0 \\ 0 & y^{\frac{1}{2}} \end{matrix} \;|0 < y < T \right\}. \tag{8.26}$$

Note that this is the *backwards geodesic flow* compared with how we defined it in Section 2. Hence, we have the corresponding Iwasawa decomposition $\mathrm{PSL}(2, \mathbb{R}) = N\_AK$ (note that $N\_$ is *an expanding horosphere* for this flow). In which case, we have the following theorem concerning counting points in the orbits of general discrete subgroups, Γ (as in the rest of the paper), in bisectors due to Bourgain, Kontorovich, and Sarnak.

Theorem 8.5 ([5]). Consider bounded Borel subsets $\Omega_1 \subset N\_$ and $\Omega_2 \subset K$ such that $\mu^{PS}(\partial(\Omega_1(X_i)) = \nu_i(\partial(\Omega^{-1}(X_i^-))) = 0$, then

$$\lim_{T\to\infty} \frac{\#(\Gamma \cap \Omega_1 A_T \Omega_2)}{T^{\delta_\Gamma}} = \frac{1}{\delta_\Gamma \cdot \left|\mathrm{m}^{BMS}\right|} \mu^{PS}(\Omega_1(X_i))\nu_i(\Omega_2^{-1}(X_i^-)). \tag{8.27}$$

This counting theorem then allows us to prove the following:

Proposition 8.6. Let $\mathcal{I}$ be an interval, and let $\Omega \subset \{(c,d) \mid c \geq 0\}$ be a bounded, convex, compact subset with piecewise smooth boundary. Moreover, suppose that in polar coordinates the region Ω is bounded by two piecewise smooth curves $r_1(\theta) \leq r_2(\theta)$

for $\theta \in [\theta_1, \theta_2]$. Then

$$\#\left\{\gamma = \begin{pmatrix} * & * \\ c_\gamma & d_\gamma \end{pmatrix} \in \Gamma_\infty \backslash \Gamma \;\middle|\; x(\gamma) \in \mathcal{I},\ (c_\gamma, d_\gamma) \in \sqrt{T}\Omega\right\}$$
$$\sim \frac{T^{\delta_\Gamma}}{\delta_\Gamma \left|\mathrm{m}^{BMS}\right|} \mu^{PS}(\mathcal{I}(X_i)) \int_{\theta_1}^{\theta_2} \left(r_2^{2\delta_\Gamma}(\theta) - r_1^{2\delta_\Gamma}(\theta)\right) d\nu_i(\theta) \quad (8.28)$$

as $T \to \infty$, where $d\nu_i(\theta) = d\nu_i(k(\theta)X_i)$ and we have written γ in N_AK coordinates as $x(\gamma)a(\gamma)k(\gamma)$.

Proof. The proof follows the same lines as [19, Proposition 5.3]. First, we note that using the Iwasawa decomposition of γ, we have $d_\gamma = y^{1/2}\cos\theta$, $c_\gamma = y^{1/2}\sin\theta$. Therefore, $(y^{1/2}, \theta)$ give a polar coordinate decomposition of the plane. The rest of the argument follows from a Riemann sum approximation, which works equally well when working with ν_i.

Split the interval $I = [\theta_1, \theta_2]$ into separate equally spaced intervals $\{I_i\}_{i=1}^n$. Take $\theta_{1,i}^+$, and $\theta_{1,i}^-$ to be the points in I_i where r_1 is maximized (resp. minimized) and $\theta_{2,i}^+$, and $\theta_{2,i}^-$ to be the points at which r_2 is maximized (resp. minimized). Now define

$$\begin{aligned} \Omega_n^- &= \bigcup_{i=1}^n I_i \times [r_1(\theta_{1,i}^-), r_2(\theta_{2,i}^+)] \\ \Omega_n^+ &= \bigcup_{i=1}^n I_i \times [r_1(\theta_{1,i}^+), r_2(\theta_{2,i}^-)]. \end{aligned} \quad (8.29)$$

Thus, $\Omega_n^- \subseteq \Omega \subseteq \Omega_n^+$ and

$$\begin{aligned} &\lim_{n\to\infty} \sum_{i=1}^n \int_{I_i} \left(r_2^{2\delta_\Gamma}(\theta_{2,i}^+) - r_1^{2\delta_\Gamma}(\theta_{1,i}^-)\right) \mathrm{d}\nu_i(\theta) \\ &\qquad = \lim_{n\to\infty} \sum_{i=1}^n \int_{I_i} \left(r_2^{2\delta_\Gamma}(\theta_{2,i}^-) - r_1^{2\delta_\Gamma}(\theta_{1,i}^+)\right) \mathrm{d}\nu_i(\theta) \\ &\qquad = \int_{\theta_1}^{\theta_2} \left(r_2^{2\delta_\Gamma}(\theta) - r_1^{2\delta_\Gamma}(\theta)\right) \mathrm{d}\nu_i(\theta). \end{aligned} \quad (8.30)$$

For the truncated regions Ω_n^+ and Ω_n^-, the proposition follows readily with the observation that in (8.27), the fact that the conformal density is evaluated at Ω_2^{-1} simply means that the bounds of integration would be $[-\theta_2, -\theta_1]$. However, since our group is symmetric, this is equal the integral over $[\theta_1, \theta_2]$. From, since (28) satisfies finite additivity, the proposition follows. ■

Summarizing: provided $s \leq s_0 = 7\frac{1}{2}$ the gap distribution at time T can be written

$$\widehat{F}_{T,\mathcal{I}}(s) = F^{1,2}_{T,\mathcal{I}}(s) + F^{2,3}_{T,\mathcal{I}}(s). \tag{8.31}$$

Moreover, we can take the limit as $T \to \infty$ and (8.15) becomes

$$\widehat{F}_{\mathcal{I}}(s) = F^{1,2}_{\mathcal{I}}(s) + F^{2,3}_{\mathcal{I}}(s), \tag{8.32}$$

where, for $(i,j) = (1,2), (2,3)$,

$$F^{i,j}_{\mathcal{I}}(s) = \frac{1}{\delta_{\widehat{\Gamma}}|\mathrm{m}^{BMS}|}\mu^{PS}(\mathcal{I}(X_i)) \int_{\theta_1^{i,j}(s)}^{\theta_2^{i,j}(s)} \left(r_2^{i,j}(\theta, s)^{2\delta_{\widehat{\Gamma}}} - r_1^{i,j}(\theta, s)^{2\delta_{\widehat{\Gamma}}}\right) \mathrm{d}\nu_i(\theta), \tag{8.33}$$

where $r_2^{i,j}(\theta, s)\Big|_{\theta\in[\theta_1^{i,j}(s),\theta_2^{i,j}(s)]}$ and $r_1^{i,j}(\theta, s)\Big|_{\theta\in[\theta_1^{i,j}(s),\theta_2^{i,j}(s)]}$ are the curves in polar coordinates forming the boundary of $\Omega^{i,j}(s)$.

For convenience, define the constant

$$\kappa := \frac{1}{\delta_{\widehat{\Gamma}}|\mathrm{m}^{BMS}|}\mu^{PS}(\mathcal{I}(X_i)). \tag{8.34}$$

8.4 Properties of the limiting gap distribution

Looking first at $\Omega^{1,2}_1$ defined by (8.16), (8.17), and (8.21), however since $s < s_0 = 7\frac{1}{2}$, (8.17) can be ignored. Hence, we have the region (in (c, d)-coordinates):

$$\Omega^{1,2}_1(s) = [0, \frac{1}{\sqrt{2}}] \times [-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}] \cap \left\{(c, d) : c \geq \frac{1}{s|d|}\right\}. \tag{8.35}$$

This region is symmetric under reflection across the y-axis and since the conformal density in (8.33) is invariant under this reflection we can consider

$$\tilde{\Omega}^{1,2}_1(s) = [0, \frac{1}{\sqrt{2}}] \times [0, \frac{1}{\sqrt{2}}] \cap \left\{(c, d) : c \geq \frac{1}{s|d|}\right\} \tag{8.36}$$

instead, and the only difference will be a factor of 2.

Regarding $\Omega_1^{2,3}(s)$, from (8.18), we know that $\Omega_1^{2,3}$ is a subset of the triangle

$$-\frac{1}{\sqrt{2}} \le d \le \frac{1}{\sqrt{2}} \quad , \quad 0 \le c < \frac{1}{4\sqrt{2}} - d. \tag{8.37}$$

Moreover, (8.19) implies that when $d < 0$, if $c > -\frac{d}{4}$ then $c > \frac{1}{\sqrt{2}}$, thus $\Omega_1^{2,3} = T_1 \cup T_2$ where

$$T_1 := \left\{(c,d) : c,d \ge 0 \;,\; c < \frac{1}{4\sqrt{2}} - d\right\} \tag{8.38}$$

$$T_2 := \left\{(c,d) : c \ge 0 \;,\; -\frac{1}{\sqrt{2}} \le d \le 0 \;,\; c \le -\frac{d}{4}\right\}. \tag{8.39}$$

Now looking at the condition imposed by (8.22), it is straightforward to see that, for $s < 8$, $\Omega_1^{2,3}(s)$ does not intersect T_2. Hence, for $s < s_0 < 8$,

$$\Omega_1^{2,3}(s) = \left\{(c,d) \in T_1 \;:\; c \le \frac{1}{sd} - \frac{d}{4}\right\}. \tag{8.40}$$

So far we have established that

$$\widehat{F}(s) = \kappa\mu(\Omega_1^{2,3}(s)) + 2\kappa\mu(\tilde{\Omega}_1^{1,2}(s)), \tag{8.41}$$

where, for a general set $A = \left\{(r\cos\theta, r\sin\theta) : r \in [r_1^A(\theta), r_2^A(\theta)], \theta \in [\theta_1^A, \theta_2^A]\right\}$,

$$\mu(A) := \int_{\theta_1^A}^{\theta_2^A} \left(r_2^A(\theta)^{2\delta_\Gamma} - r_1^A(\theta)^{2\delta_\Gamma}\right) \mathrm{d}\nu_i(\theta). \tag{8.42}$$

Thus, $\widehat{F}(s)$ is explicitly calculated in terms of the fractal measure ν_i. Unfortunately, this measure is not itself explicit (in that it is defined as the weak limit of a sequence of measures). However, it does lend itself to simulations (which we will not do here) and one can calculate certain analytic properties of $\widehat{F}$, we present three below:

Proposition 8.7. $\widehat{F}_{\mathcal{I}}(s) = 0$ for all $s < 2$ for any $\mathcal{I}$. Moreover, *all* gaps are larger than 2.

This is a form of level repulsion and follows from the definitions of $\tilde{\Omega}_1^{1,2}(s)$ and $\Omega_1^{2,3}(s)$ and (8.41). Indeed, $\tilde{\Omega}_1^{1,2}(s)$ is empty for $s < 2$ and $\Omega_1^{2,3}(s)$ is empty for $s < 4$.

ν_i is a fractal measure supported on the limit set. Hence, looking at (8.42), if neither θ_1^A nor θ_2^A is in $\mathcal{L}(\Gamma)$ (the support of ν_i). Then the derivative of $\widehat{F}$ will be easy to calculate:

Proposition 8.8. Suppose $\mathcal{S} \subset (2, s_0)$ is a connected subset such that for all $s \in \mathcal{S}$, $\theta_1^{i,j}(s)$ and $\theta_2^{i,j}(s) \notin \mathcal{L}(\Gamma)$ for $(i,j) = (1,2)$ or $(2,3)$, then

$$P(s) = \widehat{F}'(s) = \frac{C_{\mathcal{S}}}{s^{\delta_{\Gamma}+1}}, \tag{8.43}$$

where $0 \le C_{\mathcal{S}} < \infty$ depends on the region $\mathcal{S}$ but not on $s \in \mathcal{S}$ and is explicit.

Proof. Let $s_1 = \inf\{s \in \mathcal{S}\}$, in which case, for $s \in \mathcal{S}$, we separate the integral in (8.41) and write

$$\begin{aligned}\widehat{F}(s) &= \kappa \int_{\theta_1^{2,3}(s)}^{\theta_2^{2,3}(s)} \left(r_2^{2,3}(\theta, s)^{2\delta_{\Gamma}} - r_2^{2,3}(\theta, s)^{2\delta_{\Gamma}}\right) \mathrm{d}\nu_i(\theta) + 2\kappa \int_{\theta_1^{1,2}(s)}^{\theta_2^{1,2}(s)} \left(r_2^{1,2}(\theta, s)^{2\delta_{\Gamma}} - r_2^{1,2}(\theta, s)^{2\delta_{\Gamma}}\right) \mathrm{d}\nu_i(\theta)\\ &= \kappa \int_{\theta_1^{2,3}(s_1)}^{\theta_2^{2,3}(s_1)} \left(r_2^{2,3}(\theta)^{2\delta_{\Gamma}} - r_2^{2,3}(\theta, s)^{2\delta_{\Gamma}}\right) \mathrm{d}\nu_i(\theta) + 2\kappa \int_{\theta_1^{1,2}(s_1)}^{\theta_2^{1,2}(s_1)} \left(r_2^{1,2}(\theta)^{2\delta_{\Gamma}} - r_2^{1,2}(\theta, s)^{2\delta_{\Gamma}}\right) \mathrm{d}\nu_i(\theta)\\ &\quad + R(s, \mathcal{S}),\end{aligned}$$

where we have noted that (by (8.36) and (8.40)) r_2 is independent of s. In fact, since on $\mathcal{S}$, $\theta_1^{i,j}(s)$ and $\theta_2^{i,j}(s)$ are outside $\mathcal{L}(\Gamma)$, $R(s, \mathcal{S})$ is 0 (as the measure is supported away from the range of integration). Hence, taking a derivative:

$$P(s) = -\kappa \int_{\theta_1^{2,3}(s_1)}^{\theta_2^{2,3}(s_1)} \frac{\mathrm{d}r_1^{2,3}(\theta, s)^{2\delta}}{\mathrm{d}s} \mathrm{d}\nu_i(\theta) - 2\kappa \int_{\theta_1^{1,2}(s_1)}^{\theta_2^{1,2}(s_1)} \frac{\mathrm{d}r_1^{1,2}(\theta, s)^{2\delta}}{\mathrm{d}s} \mathrm{d}\nu_i(\theta). \tag{8.44}$$

Moreover, for $s < s_0$, we have that

$$r_1^{1,2}(\theta, s) = \frac{1}{\sqrt{s}}\sqrt{\frac{1}{\cos\theta \sin\theta}} \quad , \quad r_1^{2,3}(\theta, s) = \frac{1}{\sqrt{s}}\sqrt{\frac{1}{\left(\sin\theta\cos\theta + \frac{\cos^2\theta}{4}\right)}}. \tag{8.45}$$

Therefore, for $s \in \mathcal{S}$,

$$P(s) = \frac{\kappa}{s^{\delta_{\Gamma}+1}} \left(\int_{\theta_1^{2,3}(s_1)}^{\theta_2^{2,3}(s_1)} \left(\frac{1}{\left(\sin\theta\cos\theta + \frac{\cos^2\theta}{4}\right)}\right)^{\delta_{\Gamma}} \mathrm{d}\nu_i(\theta) + 2 \int_{\theta_1^{1,2}(s_1)}^{\theta_2^{1,2}(s_1)} \left(\frac{1}{\cos\theta\sin\theta}\right)^{\delta_{\Gamma}} \mathrm{d}\nu_i(\theta) \right). \tag{8.46}$$

■

The final analytic property we calculate for $\widehat{F}$ is the following Lipschitz condition:

Proposition 8.9. $\widehat{F}$ is Lipschitz in a neighborhood of s whenever $s \in [0, 4)$

$$\left|\widehat{F}(s) - \widehat{F}(s+x)\right| \le C_s x \tag{8.47}$$

for some constant $C_s < \infty$.

Proof. $\widehat{F}$ is 0 on $[0, 2)$. Moreover, Proposition 8.8 implies the $\widehat{F}$ is differentiable when both $\theta_1^{1,2}$ and $\theta_2^{1,2}$ are outside $\mathcal{L}(\widehat{\Gamma})$. Hence, we only need to worry about when $\theta_1^{1,2}(s)$ or $\theta_2^{1,2}(s)$ is a parabolic fixed point (since parabolic points are dense in the limit set).

For any $2 \le s < 4$ such that $\theta_1^{1,2}(s)$ or $\theta_2^{1,2}(s)$ is a parabolic fixed point:

$$\begin{aligned}\left|\widehat{F}(s) - \widehat{F}(s+x)\right| \le C \Bigg| &\int_{\theta_2^{1,2}(s)}^{\theta_2^{1,2}(s+x)} \left(r_2^{1,2}(\theta)^{2\delta_{\widehat{\Gamma}}} - r_1^{1,2}(\theta, s)^{2\delta_{\widehat{\Gamma}}}\right) \mathrm{d}\nu_i(\theta) \\ &+ \int_{\theta_1^{1,2}(s+x)}^{\theta_1^{1,2}(s)} \left(r_2^{1,2}(\theta)^{2\delta_{\widehat{\Gamma}}} - r_1^{1,2}(\theta, s)^{2\delta_{\widehat{\Gamma}}}\right) \mathrm{d}\nu_i(\theta)\Bigg|.\end{aligned} \tag{8.48}$$

Plugging in the formula for $r_2^{1,2}$ and $r_1^{1,2}$ and using Corollary 3.3 give that the 1st term on the right-hand side of (48) is less than

$$\le C_s \left| \int_{\theta_2^{1,2}(s)}^{\theta_2^{1,2}(s+x)} \theta^{2\delta_{\widehat{\Gamma}}-2} \left(\left(\frac{1/\sqrt{2}}{\sin\theta}\right)^{2\delta_{\widehat{\Gamma}}} - \left(\frac{1}{(s+x)\cos\theta\sin\theta}\right)^{\delta_{\widehat{\Gamma}}}\right) \mathrm{d}\theta \right| \tag{8.49}$$

in the range with which we are concerned we can bound this integral (by adjusting the constant) by

$$\le C_s \int_{\theta_2^{1,2}(s)}^{\theta_2^{1,2}(s+x)} \theta^{2\delta_{\widehat{\Gamma}}-2}\,\mathrm{d}\theta. \tag{8.50}$$

Evaluating the integral and performing the same analysis on the other term in (48) give

$$\left|\widehat{F}(s) - \widehat{F}(s+x)\right| \le C_s \left(\theta_2^{1,2}(s+x)^{2\delta_{\widehat{\Gamma}}-1} - \theta_2^{1,2}(s)^{2\delta_{\widehat{\Gamma}}-1}\right) + C_s \left(\theta_1^{1,2}(s)^{2\delta_{\widehat{\Gamma}}-1} - \theta_1^{1,2}(s+x)^{2\delta_{\widehat{\Gamma}}-1}\right). \tag{8.51}$$

Inserting the definition of $\theta_2^{1,2}$ and $\theta_1^{1,2}$ then gives

$$\left|\widehat{F}(s) - \widehat{F}(s+x)\right| \le C_s\left(\tan^{-1}(s+x)^{2\delta_{\widehat{\Gamma}}-1} - \tan^{-1}(s)^{2\delta_{\widehat{\Gamma}}-1}\right) + C_s\left(\cot^{-1}(s)^{2\delta_{\widehat{\Gamma}}-1} - \cot^{-1}(s+x)^{2\delta_{\widehat{\Gamma}}-1}\right). \tag{8.52}$$

From here, Taylor expanding gives

$$\left|\widehat{F}(s) - \widehat{F}(s+x)\right| \leq C\left|\left(\frac{\pi}{4} + \frac{x}{4}\right)^{2\delta_{\widehat{\Gamma}}-1} - \left(\frac{\pi}{4}\right)^{2\delta_{\widehat{\Gamma}}-1}\right| + C\left|\left(\frac{\pi}{4}\right)^{2\delta_{\widehat{\Gamma}}-1} - \left(\frac{\pi}{4} - \frac{x}{4}\right)^{2\delta_{\widehat{\Gamma}}-1}\right|. \quad (8.53)$$

Here, expanding again gives us that $\widehat{F}$ is Lipschitz. ■

9 Gauss-Like Measure

As in the previous section, this section is restricted to the example $\widehat{\Gamma}$. The goal for this section is to derive and study the measure

$$m^0(E) = C_0 \int_E \int_{-2}^{2} \frac{\mathrm{d}\mu^{PS}(x)}{|xy-1|^{2\delta_{\widehat{\Gamma}}}} \, \mathrm{d}\mu^{PS}(y), \quad (9.1)$$

where E is a Borel set in $\mathcal{L}(\widehat{\Gamma}) \cap (-2, 2)$ and C_0 is a normalizing constant. In particular, we show that this measure is invariant and ergodic for the Gauss map. Then, as a corollary of this ergodicity, we are able to show that the Gauss–Kuzmin statistics on $\mathcal{Q}_4$ converge to an explicit function. It should be noted that the density in (9.1) is a normalized eigenfunction for the transfer operator associated to the Gauss map. We shall avoid this zeta functions approach here; however, it is a promising avenue for later research.

9.1 Setup

Series [20], for the modular group, shows that one can encode the endpoints of geodesics by a ''cutting sequence", which generates the continued fraction expansions of the endpoints. Moreover, she identifies a cross-section of the unit tangent bundle such that the return map to this cross-section corresponds to the (classical) Gauss map on the end point. As an application of this, she shows that the Gauss measure is simply a projection of the Haar measure onto these end points. Thus, because the Haar measure is ergodic for the geodesic flow, the Gauss measure is ergodic for the Gauss map. The goal for this subsection is to construct the analogous measure in our context (for $\widehat{\Gamma}$). To do this, we will project the BMS measure in the same way and show that the resulting measure is ergodic for the Gauss map (for $\widehat{\Gamma}$). In the end, we will only be working with this measure, however for those interested in the Appendix, we show how to construct the analogous cutting sequences and cross-section in our context (we omit the formal proofs concerning the commuting diagrams as we do not use them and the details are the same as [20]).

Throughout this section, let $(-2,2)^* = (-2,2) \setminus \{0\}$. Consider the restriction of Gauss map to the limit set, $\mathcal{L}(\widehat{\Gamma}) = \overline{\mathcal{Q}_4}$ (where $\overline{\mathcal{Q}_4}$ denotes the closure):

$$\begin{aligned} T :& \mathcal{L}(\widehat{\Gamma}) \to \mathcal{L}(\widehat{\Gamma}) \\ & [0; a_1, a_2, \ldots] \mapsto [0; a_2, \ldots] \end{aligned} \tag{9.2}$$

and its inverse

$$T^{-1}([0; a_1, \ldots, a_{n-1}]) = \bigcup_{k \in 4\mathbb{Z}^*} [0; k, a_1, \ldots, a_{n-1}]. \tag{9.3}$$

The σ-algebra associated to this Gauss map is now the Borel σ-algebra on $\mathbb{R}$ intersected with $\mathcal{L}(\widehat{\Gamma})$. The goal is now to take the Bowen–Margulis–Sullivan measure and project it to obtain a measure on $(-2,2)$. We choose the BMS measure as it is invariant and ergodic under the geodesic flow. Thus, after projecting, we are left with a measure invariant and ergodic under the Gauss map. The following lemma gives the parameterization; this was used in Sullivan's work [21]; however, we include the proof for completeness.

Lemma 9.1. For $u \in T^1(\mathbb{H})$, let z denote the Euclidean midpoint of the geodesic containing u and $t := \beta_{u^-}(z, u)$ (thus, t is the arclength from z to u). Then

$$d\mathrm{m}^{BMS}(u) = \frac{1}{|u^+ - u^-|^{2\delta_\Gamma}} d\mu^{PS}(u^-) d\mu^{PS}(u^+) dt. \tag{9.4}$$

Remark. Note this lemma is not specific to the subgroup $\widehat{\Gamma}$ and holds for any Bowen–Margulis–Sullivan measure associated to a subgroup considered in this paper.

Proof. First (recalling s from the definition of m^{BMS} (2.6)) note

$$\begin{aligned} s &:= \beta_{u^-}(i, u) \\ &= \beta_{u^-}(i, z) + \beta_{u^-}(z, u) \\ &= \beta_{u^-}(i, z) + t \\ &= \beta_{u^-}(i, i + u^-) + \beta_{u^-}(i + u^-, z) + t. \end{aligned} \tag{9.5}$$

Now using the definition of the Busemann function, we note that $\beta_{u^-}(i + u^-, z)$ is the hyperbolic distance (along the vertical geodesic at u^-) between the horoball of height 1 based at u^- and the horoball of height $\left|u^+ - u^-\right|$. Thus,

$$s = t + \beta_{u^-}(i, i + u^-) + \ln\left|u^+ - u^-\right|. \tag{9.6}$$

Similarly,

$$\beta_{u^+}(i, u) = -t + \beta_{u^+}(i, i + u^+) + \ln\left|u^+ - u^-\right|. \tag{9.7}$$

Therefore, writing out the definition of the Burger Roblin measure and inserting (9.6) and (9.7):

$$\begin{aligned} \mathrm{m}^{BMS}(u) &:= e^{\delta_\Gamma s} e^{\delta_\Gamma \beta_{u^+}(i,u)} d\nu_i(u^-) d\nu_i(u^+) ds \\ &= \frac{1}{\left|u^+ - u^-\right|^{2\delta_\Gamma}} (e^{\delta_\Gamma \beta_{u^-}(i,i+u^-)} d\nu_i(u^-))(e^{\delta_\Gamma \beta_{u^+}(i,i+u^+)}) d\nu_i(u^+)) dt \\ &= \frac{1}{\left|u^+ - u^-\right|^{2\delta_\Gamma}} d\mu^{PS}(u^-) d\mu^{PS}(u^+) dt \end{aligned} \tag{9.8}$$

where in the last line we insert the definition of μ^{PS}. ■

To derive the Gauss-type measure (similarly to [20] for the classical Gauss measure), we restrict the BMS measure to the u^- coordinate. Integrating over the u^+ coordinate in $(-2, 2)$ gives

$$\int_{-2}^{2} \frac{\mathrm{d}\mu^{PS}(u^+)}{\left|u^+ - u^-\right|^{2\delta_{\widehat{\Gamma}}}}. \tag{9.9}$$

Thus, for a set $E \subset (-\infty, -2) \cup (\infty, 2)$,

$$\int_E \int_{-2}^{2} \frac{\mathrm{d}\mu^{PS}(x)}{|x - y|^{2\delta_{\widehat{\Gamma}}}} \,\mathrm{d}\mu^{PS}(y) \tag{9.10}$$

is a measure. Changing coordinates and using that $d\mu^{PS}(1/y) = y^{-2\delta_{\widehat{\Gamma}}} d\mu^{PS}(y)$ (this follows from (2.4) and a calculation using the Busemann function) gives, for any set, $E \subset (-2, 2)^*$

$$m^0(E) := C_0 \int_E \int_{-2}^{2} \frac{\mathrm{d}\mu^{PS}(x)}{|xy - 1|^{2\delta_{\widehat{\Gamma}}}} \,\mathrm{d}\mu^{PS}(y), \tag{9.11}$$

where C_0 is a normalizing constant. In the next section, we show that this is indeed T-invariant and ergodic.

9.2 Invariance and ergodicity

Theorem 9.2. On $(-2,2)^*$, m^0 is T-invariant and ergodic.

Proof. To prove invariance, let $E \subset (-2,2)^*$ and consider the measure of its preimage

$$m^0(T^{-1}(E)) = C_0 \int_{T^{-1}(E)} \int_{-2}^{2} \frac{\mathrm{d}\mu^{PS}(x)}{|xy-1|^{2\delta_{\widehat{\Gamma}}}}\, \mathrm{d}\mu^{PS}(y).$$

Plugging in the definition of $T^{-1}(E)$ and changing variables ($d\mu^{PS}(1/y) = y^{-2\delta_{\widehat{\Gamma}}} d\mu^{PS}(y)$) together with the fact that the Patterson–Sullivan measure is invariant under translation by $4n$ gives

$$\begin{aligned} &= C_0 \sum_{n\in\mathbb{Z}^*} \int_{E+4n} \left(\int_{-2}^{2} \frac{\mathrm{d}\mu^{PS}(x)}{|y-x|^{2\delta_{\widehat{\Gamma}}}} \right) \mathrm{d}\mu^{PS}(y) \\ &= C_0 \int_E \sum_{n\in\mathbb{Z}^*} \int_{-2}^{2} \left(\frac{\mathrm{d}\mu^{PS}(x)}{|y-x-4n|^{2\delta_{\widehat{\Gamma}}}} \right) \mathrm{d}\mu^{PS}(y). \end{aligned} \tag{9.12}$$

If we now change the x variable to $x+4n$, this gives

$$= C_0 \int_E \int_{(-\infty,-2)\cup(2,\infty)} \frac{\mathrm{d}\mu^{PS}(x)}{|y-x|^{2\delta_{\widehat{\Gamma}}}}\, \mathrm{d}\mu^{PS}(y).$$

Hence, applying the change of variables $x \mapsto x^{-1}$ gives

$$= C_0 \int_E \int_{-2}^{2} \frac{\mathrm{d}\mu^{PS}(x)}{|xy-1|^{2\delta_{\widehat{\Gamma}}}}\, \mathrm{d}\mu^{PS}(y) = m^0(E).$$

This new measure is ergodic for the Gauss map because the BMS is ergodic for the geodesic flow. However, to see this directly, note first that the density

$$\rho(y) = \int_{-2}^{2} \frac{\mathrm{d}\mu^{PS}(x)}{|xy-1|^{2\delta_{\widehat{\Gamma}}}}$$

is bounded on $\mathcal{L}(\widehat{\Gamma})$. Given $a_1,\dots,a_n$ and writing $\frac{p_i}{q_i} = [0;a_1,\dots,a_i]$, define the cylinder sets

$$\Delta_n := \left\{ \psi_n(t) := \frac{p_n + p_{n-1}t}{q_n + q_{n-1}t} : 0 \le t \le 1 \right\}. \tag{9.13}$$

Note that the sets $\Delta_n \cap \mathcal{L}(\widehat{\Gamma})$ generate the Borel σ-algebra on $\mathcal{L}(\widehat{\Gamma})$.

Now, for any $n > 0$, for $s < t \in [0,1]$, we have that there exists a $\gamma \in \widehat{\Gamma}$ such that

$$\begin{aligned}\mu^{PS}\left(T^{-n}([\frac{s}{4},\frac{t}{4}))\middle|\,\Delta_n\right) &\asymp \nu_i\left(T^{-n}([\frac{s}{4},\frac{t}{4}))\middle|\,\Delta_n\right)\\ &= \frac{\nu_i(\gamma[\frac{s}{4},\frac{t}{4}))}{\nu_i(\gamma[0,\frac{1}{4}))}\\ &= \frac{\nu_i([\frac{s}{4},\frac{t}{4}))}{\nu_i([0,\frac{1}{4}))}.\end{aligned} \tag{9.14}$$

Therefore, as the above-mentioned density is bounded above and below, for any $A \subset \mathcal{L}(\widehat{\Gamma}) \cap (-2,2)^*$ measurable

$$\frac{1}{C}m^0(A) \leq m^0(T^{-n}(A)\big|\,\Delta_n) \leq Cm^0(A). \tag{9.15}$$

To conclude, assume A is T-invariant, then $\frac{1}{C}m^0(A) \leq m^0(A|\,\Delta_n)$. If $m^0(A) > 0$, then $\frac{1}{C}m^0(\Delta_n) \leq m^0(\Delta_n|A)$. Therefore, since the cylinders Δ_n generate the Borel σ-algebra of measurable sets, we have that

$$\frac{1}{C}m^0(B) \leq m^0(B|A)$$

for all B measurable. Setting $B = A^c$ implies that $m^0(A^c) = 0$ and $m^0(A) = 1$. Hence, m^0 is ergodic. ■

9.3 Gauss–Kuzmin statistics

Given a point $x = [0; a_1, a_2, \ldots] \in \mathbb{R}$ ($a_i \in \mathbb{N}$), Gauss considered the following problem (further studied by Kuzmin in 1928): let $\widetilde{P}_{n,k}(x) = \frac{\#(k,n)}{n}$ where $\#(k,n)$ is the number of $a_i = k$ with $i \leq n$. Does there exist a limting distribution for $\widetilde{P}_{n,k}(x)$? Using the ergodicity of the Gauss measure, it is fairly simple to show that for Lebesgue-almost every x

$$\lim_{n\to\infty} \widetilde{P}_{n,k}(x) = \frac{1}{\ln(2)} \ln\left(1 + \frac{1}{k(k+2)}\right). \tag{9.16}$$

This distribution is now known as Gauss–Kuzmin statistics. For a detailed description of the original problem and history, see [9, Section 15]. The problem has an analogue in our setting.

For $[0; a_1, a_2, \ldots] = x \in \overline{\mathcal{Q}_4} \cap (-2,2)$, define $\widehat{P}_{n,k}(x) = \frac{\#(k,n)}{n}$ where $\#(k,n)$ is the number of a_i equal k for $i \leq n$. For simplicity of notation, we assume $k > 0$. In that case, writing

$$\widehat{P}_{n,k}(x) = \frac{1}{n}\sum_{s=0}^{n-1} \chi_{(\frac{1}{k+4},\frac{1}{k}]}(T^s x) \tag{9.17}$$

and applying the Birkhoff ergodic theorem for m^0 imply the following:

Theorem 9.3. For every positive integer k and μ^{PS}-almost every $x = [0; a_1, \ldots] \in \overline{\mathcal{Q}_4} \cap (-2, 2)$

$$\widehat{P}_k(x) = \lim_{n\to\infty} \widehat{P}_{n,k}(x) = m^0\left((\frac{1}{k+4}, \frac{1}{k}] \right). \tag{9.18}$$

Appendix - Cutting Sequences for $\widehat{\Gamma}$

Working with $\widehat{\Gamma}$ the goal of this section is to show that, given a geodesic with right end point in $(-2, 2) \cap \mathcal{L}(\widehat{\Gamma})$ (and left end point in $(-\infty, -2)$), there is a correspondence between the way this geodesic cuts the boundaries of fundamental domains and the continued fraction expansion of the end point. This section is exactly analogous to the Bowen–Series coding for geodesics in $\mathrm{PSL}(2, \mathbb{R})/\mathrm{PSL}(2, \mathbb{Z})$.

Let $\xi \in (-2, 2) \cap \mathcal{L}(\widehat{\Gamma})$ and let γ be any geodesic whose right endpoint is ξ and which intersects the line $x = -2$. As this geodesic moves from left to right, it will cut each fundamental domain. Each fundamental domain has two funnels and a cusp. Thus, the geodesic will separate one of the three from the others. If the geodesic separates a cusp, we write a c. If it separates a funnel, we write an l or an r depending on whether the funnel is to the left or right of the geodesic. See Figure 6.

It is easy to see that the 1st term in the sequence will always be r and the next term will be l/r after that there will be a sequence of cs followed by the same l/r. Thus, we end up with a sequence of the form

$$\xi \mapsto r, q_0, c^{\alpha_0}, q_0, q_1, c^{\alpha_1}, q_1, q_2, c^{\alpha_2}, q_2 \ldots \tag{A.1}$$

(the sequence is finite if the geodesic ends in a cusp) where $q_i = l, r$ and $\alpha_i \geq 0$. With that, it is fairly easy to see that

$$\xi = [0; (-1)^{\eta_0} 4(\alpha_0 + 1), (-1)^{\eta_1} 4(\alpha_1 + 1), \ldots] \tag{A.2}$$

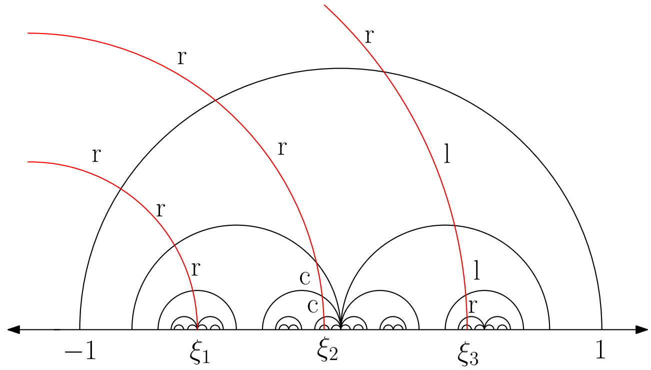

Fig. 6. In this diagram, we show the cutting sequence for three different points ξ_1, ξ_2, ξ_3. For ξ_2, first a funnel is cut off to the *right* of the geodesic, then again a funnel is cut off to the *right*. Then a *cusp* is cut off and then another *cusp*. Thus, the 1st four terms in the cutting sequence are r, r, c, c.

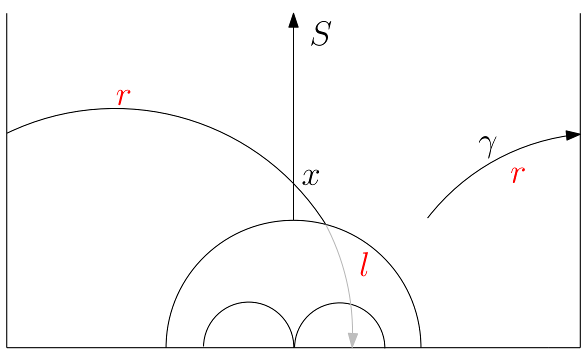

Fig. 7. In this diagram, we show a geodesic in the fundamental domain above i, and a point $x \in S \cap \gamma$ such that the cutting sequence for γ changes type at x. This is because the cutting sequence (pictured in red) with x inserted will read $\dots, r, r, x, l, \dots$.

where

$$\eta_i = \begin{cases} 0 & \text{if} q_i = l \\ 1 & \text{if } q_i = r \end{cases} . \tag{A.3}$$

With that, there is a correspondence between such sequences and geodesics with end points in $(-2, 2)$.

In order to identify the appropriate cross-section of $T^1(\Gamma\backslash\mathbb{H})$, consider the fundamental domain above i and the line connecting i to ∞, call it S. Given a geodesic γ whose left end point is in $(-\infty, -2) \cap \mathcal{L}(\widehat{\Gamma}))$ and whose right endpoint is in $(-2, 2) \cap \mathcal{L}(\widehat{\Gamma}))$

consider a point $x \in \gamma \cap S$. We insert x into the cutting sequence of γ, at its position in the sequence of fundamental domains, resulting in a sequence of the form:

$$r, q_0, c^{\alpha_0}, q_0, q_1, c, c, c, x, c, q_1, \ldots \tag{A.4}$$

We say a cutting sequence *changes type* at x if x lies between a q_i and q_{i+1}.

With that, the cross-section $\mathcal{C} \subset T^1(\Gamma\backslash\mathbb{H})$ are those points, based at $x \in S$ pointed along geodesics whose cutting sequence changes type at x. In that case, the return map to this cross-section corresponds to the Gauss map acting on the end point. For a more formal discussion for the modular group (however, the same details apply here), see [20].

Acknowledgments

The author is very grateful to Jens Marklof for his guidance throughout this project. Moreover, the author thanks Florin Boca, Zeev Rudnick, and Xin Zhang for their insightful comments on early preprints.

Funding

This work was supported by EPSRC Studentship [EP/N509619/1 1793795].

References

[1] Athreya, J. S. and Y. Cheung. "A Poincaré section for the horocycle flow on the space of lattices." *Int. Math. Res. Not. IMRN* 2014, no. 10 (2014): 2643–90.

[2] Athreya, J. S. and A. Ghosh. "The Erdös-Szüsz-Turán distribution for equivariant processes." *Enseign. Math.* 60, no. 2 (2018): 1–21.

[3] Boca, F., C. Cobeli, and A. Zaharescu. "A conjecture of R.R. Hall on Farey points." *J. Reine Angew. Math.* 535 (2001): 207–36.

[4] Beresnevich, V., A. Ghosh, D. Simmons, and S. Velani. "Diophantine approximation in Kleinian groups: singular, extremal, and bad limit points." *J. Lond. Math. Soc. (2)* 98, no. 2 (2018): 306–28.

[5] Bourgain, J., A. Kontorovich, and P. Sarnak. "Sector estimates for hyperbolic isometries." *Geom. Funct. Anal.* 20, no. 5 (2010): 1175–200.

[6] Boca, F. "A problem of Erdös, Szüsz and Turán concerning diophantine approximations." *Int. J. Number Theory* 04, no. 04 (2008): 691–708.

[7] Hall, R. R. "A note on Farey series." *J. Lond. Math. Soc. (2)* s2-2, no. 1 (1970): 139–48.

[8] Heersink, B. "Equidistribution of Farey sequences on horospheres in covers of $\mathrm{SL}(n+1,\mathbb{Z})\backslash\mathrm{SL}(n+1,\mathbb{R})$ and applications." *Ergod Theory Dyn Syst.* 41, no. 2 (2021): 471–93.

[9] Khinchin, A. Y. *Continued Fractions*. Chicago, USA: University of Chicago Press, 1935.

[10] Lagarias, J. and H. Mehta. "Products of Farey fractions." *Experiment. Math.* 26, no. 1 (2017): 1–21.

[11] Lutsko, C. "Directions in orbits of geometrically finite hyperbolic subgroups." *Math. Proc. Cambridge Philos. Soc.* (2020): preprint.

[12] Marklof, J. "The N-point correlations between values of a linear form." *Ergodic Theory Dynam. Systems* 20 (1999): 01.

[13] Marklof, J. "The asymptotic distribution of Frobenius numbers." *Invent. Math.* 181, no. 1 (2010): 179–207.

[14] Marklof, J. "Fine-Scale Statistics for the Multidimensional Farey Sequence." In *Limit Theorems in Probability, Statistics and Number Theory*, 49–57. Berlin, Heidelberg: Springer, 2013.

[15] Mohammadi, A. and H. Oh. "Matrix coefficients, counting and primes for orbits of geometrically finite groups." *J. Eur. Math. Soc.* 17 (2015): 837–97.

[16] Marklof, J. and A. Strömbergsson. "The distribution of free path lengths in the periodic Lorentz gas and related lattice point problems." *Ann. of Math. (2)* 172 (2010): 1949–2033.

[17] Oh, H. and N. A. Shah. "Equidistribution and counting for orbits of geometrically finite hyperbolic groups." *J. Amer. Math. Soc.* 26, no. 2 (2013): 511–62.

[18] Patterson, S. J. "The limit set of a Fuchsian group." *Acta Math.* 136 (1976): 241–73.

[19] Rudnick, Z. and X. Zhang. "Gap distributions in circle packings." *Münster J. Math.* 10 (2017): 131–70.

[20] Series, C. "The modular surface and continued fractions." *J. Lond. Math. Soc. (2)* s2-31, no. 1 (1985): 69–80.

[21] Sullivan, D. "The density at infinity of a discrete group of hyperbolic motions." *Publ. Math. Inst. Hautes Études Sci.* 50 (1979): 171–202.

[22] Stratmann, B. and S. Velani. "The Patterson measure for geometrically finite groups with parabolic elements, new and old." *Proc. London Math. Soc. (3)* s3-71, no. 1 (1995): 197–220.

[23] Taha, D. "The Boca–Cobeli–Zaharescu map analogue for the Hecke triangle groups G_q." (2019): preprint arXiv:1810.10668.

[24] Xiong, M. and A. Zaharescu. "A problem of Erdös–Szüsz–Turán on diophantine approximation." *Acta Arith.* 125, no. 2 (2006): 163–77.

[25] Zhang, X. "The gap distribution of directions in some Schottky groups." *J. Mod. Dyn.* 11 (2017): 477.