The distribution of geodesic excursions into the neighborhood of a cone singularity on a hyperbolic 2-orbifold

Andrew Haas

Abstract. A generic geodesic on a finite area, hyperbolic 2-orbifold exhibits an infinite sequence of penetrations into a neighborhood of a cone singularity of order $k \geq 3$, so that the sequence of depths of maximal penetration has a limiting distribution. The distribution function is the same for all such surfaces and is described by a fairly simple formula.

1. Introduction

Many of the concepts of number theory can be realized in a geometric setting. This is especially true in the field of classical diophantine approximation, where much of the theory has been reformulated as statements about the depth of penetration of a geodesic into the cusp, or non-compact, end of the modular surface.

The theorem of Bosma, Jager and Weidijk [6], called the Lenstra Conjecture, describes the distribution of values $\theta_n = q_n|q_n x - p_n|$, for almost all real numbers x, where p_n/q_n is the nth continued fraction convergent to x. Bosma proved a closely related result for approximation by mediants [5]. In [8] we proved analogues of these results which describe the distribution of the depth of maximal penetration of a generic geodesic into the cusp end of a finite area, non-compact hyperbolic 2-orbifold. These purely geometric results were then applied to prove Fuchsian group versions of [6] and [5]. The distribution of the θ_n, defined with respect to a finite area Fuchsian group, is given independent of the group and agrees with the classical result.

In this paper we look at how a generic geodesic on an orbifold surface "approximates" a cone singularity of the surface. This type of investigation was begun in [9]. In that paper, the geometric Markoff theory was generalized to this setting. Here it is shown that the sequence of depths of maximal penetration of a generic geodesic into a neighborhood of a cone singularity has a limiting density which is explicitly

computable and the same for all finite area surfaces and most orbifold surfaces. As in [8], the resulting densities bear a striking similarity to the ones found in [6] and [5], although they are somewhat more complex and arcane. The similarity becomes particularly clear when the distributions are expressed in terms of area.

1.1. The main results. We shall begin with a heuristic treatment of some of the definitions and then state the main results of the paper. A hyperbolic 2-orbifold S is represented as the quotient of the hyperbolic plane $\mathbb{H}$ by a Fuchsian group G. We shall assume that S has finite area and, if S is compact, G is not a triangle group. A point p on S is a cone point of order $k \geq 3$ if, under the projection from $\mathbb{H}$ to S, p is covered by a point $\tilde{p}$ which has non-trivial stabilizer of order k in G.

An r-excursion e, of a geodesic ray γ into a neighborhood of the cone point, is an arc of γ associated to a point of maximal penetration of the ray γ into the radius r neighborhood, $B_r(p)$, of the cone point. The depth of the excursion e, denoted $d(e)$, is the "distance" from e to p. When r is sufficiently small, the excursions are simply the arcs $\gamma \cap B_r(p)$ and $d(e)$ is the actual distance. The precise definition becomes more complicated with larger values of r. An r-excursion e, which is associated with a closed arc of γ that loops about p, is called an approximating r-excursion. e is called an approximating excursion if it is an approximating r-excursion for some value r. These definitions parallel the geometric analogues of mediant and continued fraction approximations, as they appear on the classical modular surface.

The r-excursions of a ray γ naturally form a sequence $\{e_j\}$, which goes out the end of γ if the sequence is infinite. In a similar fashion the approximating excursions of γ form a sequence $\{e_j^*\}$. For a generic ray γ, both sequences are infinite for all $r > 0$.

Let #A denote the cardinality of the set A. For $r > 0$, $0 \leq z \leq r$ and a geodesic ray γ with infinitely many r-excursions e_j into the radius r neighborhood of a cone point p of order k, the distribution of the depths of r-excursions is defined as

$$\mathrm{dist}_k(r, z)(\gamma) = \lim_{n \to \infty} \frac{1}{n} \#\{j \mid 1 \leq j \leq n, d(e_j) \leq z\}, \tag{1}$$

if the limit exists. Similarly, if e_j^* is an infinite sequence of approximating excursions of γ into the neighborhood of a cone point p of order k, then the distribution, $\mathrm{dist}_k^*(z)(\gamma)$, of the depths of approximating excursions is defined just as the function (1) above, with e_j^* replacing e_j in the definition, and $0 \leq z \leq r_k$, where $r_k = \sinh^{-1}(\cot \frac{\pi}{k})$. Note that the function does not depend on a parameter r. In Corollary 2, of Section 4.2.2, we show that every approximating excursion is an r-excursion for some $r < r_k$.

We are now in a position to state the main results of the paper.

Theorem 1. *For almost all geodesic rays γ, for $r > 0$ and $0 \leq z \leq r$ the distribution $\mathrm{dist}_k(r, z)(\gamma)$ converges to a function $\mathrm{dist}_k(r, z)$. When $r \leq r_k$,*

$$\mathrm{dist}_k(r, z) = \frac{\sinh z}{\sinh r}.$$

When $r > r_k$,

$$\mathrm{dist}_k(r, z) = \begin{cases} \dfrac{\frac{\pi}{k} \sinh z}{\sinh r \tan^{-1}\left(\frac{1}{\sinh r}\right) + \log(\sin \frac{\pi}{k} \cosh r)} & \text{if } z \leq r_k, \\[4ex] \dfrac{\sinh z \tan^{-1}\left(\frac{1}{\sinh z}\right) + \log(\sin \frac{\pi}{k} \cosh z)}{\sinh r \tan^{-1}\left(\frac{1}{\sinh r}\right) + \log(\sin \frac{\pi}{k} \cosh r)} & \text{if } z > r_k. \end{cases}$$

The result for approximating excursions is given in

Theorem 2. *For almost all geodesic rays γ and for $0 \leq z \leq r_k$ the distribution $\mathrm{dist}_k^*(z)(\gamma)$ converges to a function $\mathrm{dist}_k^*(z)$, which is given by the formula*

$$\mathrm{dist}_k^*(z) = \begin{cases} \dfrac{\frac{\pi}{k} \sinh z}{\log\left(2 \cos \frac{\pi}{k}\right)} & \text{if } z \leq \delta_k, \\[4ex] \dfrac{\sinh z \tan^{-1}(\varphi_k(\sinh z)) + \log\left(2 \cosh z \sin \frac{\pi}{k} \cos \frac{\pi}{k}\right)}{\log\left(2 \cos \frac{\pi}{k}\right)} & \text{if } z > \delta_k, \end{cases}$$

where $\delta_k = \sinh^{-1}\left(\cot \frac{2\pi}{k}\right)$ and $\varphi_k(x) = \frac{1 - x \tan \frac{\pi}{k}}{x + \tan \frac{\pi}{k}}$. When $k = 3$ or 4, $\delta_k \leq 0$ and the first case does not occur.

1.2. Outline of the paper. There are two main pieces to the proofs of Theorems 1 and 2. First in Theorem 3, we prove the existence of the distributions and describe their values by certain integrals. In Theorem 4, we complete the proof by computing the values of these integrals.

Section 2 is concerned with background material and a precise treatment of excursions. The statement of Theorem 3 is in Section 3.2 and the proof occupies the remainder of Section 3. Theorem 4 is the focus of Section 4. In the last section, Section 5, we see how the results look if area, rather than distance, is used to define the depth of an excursion. Written in terms of area, it is easier to see how the work with cone points relates to the more classical case.

2. Excursions into a cone neighborhood

2.1. Basic notions and definitions. The quotient of the Poincaré upper-half plane $\mathbb{H}$ by a Fuchsian group G is called a hyperbolic 2-orbifold. Let $\pi : \mathbb{H} \to \mathbb{H}/G = S$ denote the projection from the upper-half plane to the orbifold S. A point $p \in S$ is called a *cone point of order k* if there is a preimage $\tilde{p}$ of p in $\mathbb{H}$ so that the stabilizer of $\tilde{p}$ in G is generated by a transformation T_k of order k.

Throughout this paper we shall assume that S is a finite area hyperbolic 2-orbifold with a cone point p of order $k \geq 3$ and that, if S is compact then G is not a triangle group. One important consequence of this last hypothesis is that, if S is compact, then there is a simple closed geodesic on S that does not pass through p. This follows from the fact that such a surface will contain a homotopically non-trivial, simple loop in the complement of the cone points on S, which does not go around a cone point. We should mention that in the special case where G is a group with signature $(0; 2, 2, 2, k)$, the simple closed geodesic degenerates to a geodesic ray going back and forth between order two cone points.

For $r > 0$, let $\tilde{B}_r(\tilde{q})$ denote the open hyperbolic disc in $\mathbb{H}$ of radius r centered at $\tilde{q}$. When $r = 0$ set $\tilde{B}_0(\tilde{q}) = \{\tilde{q}\}$. For $\tilde{p}$ covering the cone point p, $\tilde{B}_r(\tilde{p})$ projects to what we shall call the *cone neighborhood of p of radius r*, written $B_r(p)$. When G is not a triangle group, if $r < r_k$ then for $g \in G$, $g(\tilde{B}_r(\tilde{p})) \cap \tilde{B}_r(\tilde{p}) \neq \emptyset$ if and only if g is in the stabilizer of $\tilde{p}$, [4]. Consequently, the projection of $\tilde{B}_r(\tilde{p})$ is exactly k-to-1 in the complement of $\tilde{p}$. For larger values of r, the projection of $\tilde{B}_r(\tilde{p})$ can be considerably more complicated.

Let $\gamma : (-\infty, \infty) \to S$ be a geodesic parameterized by arc length. A lift $\tilde{\gamma}$ of γ to $\mathbb{H}$ has endpoints $\tilde{\gamma}_+$ and $\tilde{\gamma}_-$ in the extended real line $\widehat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$, representing respectively the limits at infinity and minus infinity. If the domain of γ is restricted to $[0, \infty)$ then we shall refer to it as a geodesic ray. A lift of the ray $\tilde{\gamma}$ has the single endpoint $\tilde{\gamma}_+$ at infinity. Henceforth, all geodesics shall be parameterized by arc length.

2.1.1. The distinguished ray λ. It is necessary to choose a distinguished, simple geodesic ray λ on S with initial point $\lambda(0) = p$, which is used to catalogue the excursions of a geodesic on S into a cone neighborhood of p. The two cases, S non-compact and S compact, are considered separately.

If S is non-compact then it has a cusp end, also called a puncture, and there is at least one simple geodesic rays with initial point p that eventually go out the cusp end. Choose such a ray and call it λ.

Now suppose S is compact. We have assumed that G is not a triangle group and therefore, on S there is a simple closed geodesic β that does not contain p. We shall choose λ to be a simple geodesic ray that twirls asymptotically into β. It can be constructed as follows. Let δ be a minimal length, and therefore simple, geodesic

arc with its initial point at p and its terminal point on β. Let $\tilde{\beta} \colon \mathbb{R} \to \mathbb{H}$ be a full connected preimage of β. Then there are lifts $\tilde{p}$ of p and $\tilde{\delta}$ of δ, so that $\tilde{\delta}$ has its initial point at $\tilde{p}$ and its terminal point on $\tilde{\beta}$. Let $\tilde{\lambda}$ be the geodesic ray from $\tilde{p}$ to $\tilde{\beta}_+$, one of the endpoints of $\tilde{\beta}$. Then one easily shows that for $g \in G$, either $g(\tilde{\lambda}) \cap \tilde{\lambda} = \emptyset$ or else the intersection is the point set $\{\tilde{p}\}$ and g is in the stabilizer of $\tilde{p}$. The ray $\tilde{\lambda}$ thus projects to a simple geodesic ray on S with initial point p.

Observe that, given a sequence of values $t_j \to \infty$, in the non-compact case $\lambda(t_j)$ has no limit points in S, while in the compact case the sequence $\lambda(t_j)$ will limit exactly at points on the geodesic β.

2.1.2. Normalization of the covering group.

For the sake of clarity and computational simplicity we shall choose a nicely normalized Fuchsian group G defining the orbifold S. First, let the complex number $i \in \mathbb{H}$ be our point $\tilde{p}$, covering the cone point p. Then the transformation

$$T_k(\zeta) = \frac{\cos \frac{\pi}{k} \zeta + \sin \frac{\pi}{k}}{-\sin \frac{\pi}{k} \zeta + \cos \frac{\pi}{k}}$$

generates the stabilizer of i in G.

We may further normalize G so that the geodesic λ has a lift $\tilde{\lambda}_0$ with initial point i and terminal point $0 \in \mathbb{R}$. Then $\tilde{\lambda}_1 = T_k(\tilde{\lambda}_0)$ and $\tilde{\lambda}_{-1} = T_k^{-1}(\tilde{\lambda}_0)$ are two other lifts of λ beginning at i. The point $\tan \frac{\pi}{k} = a_k$ is the endpoint at infinity of $\tilde{\lambda}_1$ and $-a_k$ is the endpoint of $\tilde{\lambda}_{-1}$.

We fix some notation. Define the sets

$$I^+ = \left\{ (x, y) \in \mathbb{R}^2 \mid x \in (0, a_k), \ y \in \left(-\frac{1}{x}, 0 \right) \right\},$$

$$I^- = \left\{ (x, y) \in \mathbb{R}^2 \mid x \in (-a_k, 0), \ y \in \left(0, -\frac{1}{x} \right) \right\},$$

$$J^+ = \{ (x, y) \in I^+ \mid y < -a_k \} \quad \text{and} \quad J^- = \{ (x, y) \in I^- \mid y > a_k \}.$$

Set $I = I^+ \cup I^-$ and $J = J^+ \cup J^-$.

We shall use the shorthand $\tilde{B}_r$ for the disc $\tilde{B}_r(i)$ and B_r for the disc $B_r(p)$.

2.2. Excursions

2.2.1. The definition.

Given a cone neighborhood B_r of p and a geodesic ray γ on S, we are concerned with the "excursions" of γ into B_r. We saw in the introduction that these can be regarded as certain arcs of γ, although their definition is complicated by the self-overlap of B_r for larger values of r. Hence, we would like to distinguish the fine structure of the excursions as γ intersects different overlapping pieces of B_r.

The above considerations dictate the precise, but non-intrinsic, definition of an excursion which follows. Suppose $r > 0$. Define an *r-excursion e* of a geodesic γ into the cone neighborhood B_r to be a lift $\tilde{\gamma}$ of γ to $\mathbb{H}$ with $(\tilde{\gamma}_+, \tilde{\gamma}_-) \in I$ so that $\tilde{\gamma} \cap \tilde{B}_r \neq \emptyset$. In particular, $\tilde{\gamma} \cap \tilde{\lambda}_0 \neq \emptyset$.

If β is a geodesic in $\mathbb{H}$ that intersects $g^{-1}(\tilde{\lambda}_0)$ for some $g \in \Gamma$, then for some unique $0 \leq n \leq k - 1$, $(T_k^n \circ g(\beta_+), T_k^n \circ g(\beta_-)) \in I$. Consequently, associated to each intersection $\gamma \cap B_r \neq \emptyset$ with $(\gamma \cap B_r) \cap \lambda \neq \emptyset$, there is a lift $\tilde{\gamma}$, with endpoints in I; that is, there is an r-excursion associated with the intersection. Conversely, each r-excursion is associated with such an intersection. Rather than trying to make this relationship more precise, we shall work with the definition given in the preceding paragraph.

The r-excursion $e = \tilde{\gamma}$ is called an *approximating r-excursion* if $(\tilde{\gamma}_+, \tilde{\gamma}_-) \in J$. More generally, call e an *approximating excursion* if e is an approximating r-excursion for some $r \geq 0$. In Section 4.2.2 we shall prove that every approximating excursion is an r-excursion for some $r < r_k = \sinh^{-1}\left(\cot \frac{\pi}{k}\right)$.

Let $E(\gamma, r)$ be the set of r-excursions of γ and let $A(\gamma)$ be the set of approximating excursions of γ.

2.2.2. Ordering the excursions along γ. To each r-excursion $e = \tilde{\gamma}$ there is a single intersection of the geodesics $\tilde{\gamma}$ with the ray $\tilde{\lambda}_0$. We may therefore associate to e the unique real parameter t_e, for which $\tilde{\gamma}(t_e) \in \tilde{\lambda}_0$. Given a positive real value r and a geodesic γ the identification $e \rightarrow t_e$ defines a function $\psi : E(\gamma, r) \rightarrow (-\infty, \infty)$. Let $E^+(\gamma, r)$ be the subset of $e \in E(\gamma, r)$ so that $\psi(e) \geq 0$. ψ also induces a map on $A(\gamma)$ and we similarly define $A^+(\gamma)$. The range of ψ shall be called the excursion parameters, written $\{t_e\}$. We then have

Proposition 1. *If e and e' are two distinct r-excursions along γ, then $\psi(e) \neq \psi(e')$.*

Proof. Write $e = \tilde{\gamma}$, $e' = \tilde{\gamma}'$ and suppose $t_e = t_{e'}$. Since $\tilde{\gamma}$ is distinct from $\tilde{\gamma}'$, there is a non-trivial $g \in G$ so that $g(\tilde{\gamma}) = \tilde{\gamma}'$ and $g(\tilde{\gamma}(t_e)) = \tilde{\gamma}'(t_{e'})$. It follows that $g(\tilde{\lambda}_0) \cap \tilde{\lambda}_0 \neq \emptyset$. This is only possible if $g = T_k^n$ for some $n = 1, \ldots, k - 1$. But then because we have $(\tilde{\gamma}_+, \tilde{\gamma}_-) \in I$, applying the transformation gives $(g(\tilde{\gamma}_+), g(\tilde{\gamma}_-)) = (\tilde{\gamma}'_+, \tilde{\gamma}'_-) \notin I$, which is impossible. $\square$

Define an order on $E(\gamma, r)$ by stipulating that $e < e'$ if $\psi(e) < \psi(e')$. This ordering does not depend on the parametrization of γ. The next proposition shows that this ordering of E is very well behaved.

Proposition 2. *Given a geodesic γ, which is distinct from the simple closed geodesic β used to define λ, suppose that the set of r-excursions $E^+ = E^+(\gamma, r)$ is infinite. Then there is a unique map from $\mathbb{N}$ onto E^+, making E^+ into a sequence $\{e_j\}_{j=1}^\infty$ so that $e_j < e_k$ if and only if $j < k$. Furthermore, $\lim_{j \to \infty} \psi(e_j) = \infty$.*

Proof. As G is discrete, E^+ must be countable. Therefore, it will suffice to show that for any sequence $\{e_j\}_{j=1}^{\infty}$ in E^+, which is ordered as in the proposition, $\lim_{j\to\infty} \psi(e_j) = \infty$. We argue by contradiction. Suppose there is such a sequence of r-excursions for which the excursion parameters, which we shall write $\psi(e_j) = t_j$, do not diverge to ∞. By passing to a subsequence, of the same name, we may suppose $t_j \to t^* < \infty$.

Write $e_j = \tilde{\gamma}_j$. Choose a lift $\tilde{\gamma}$ of γ. For each $j \in \mathbb{N}$ there is a transformation $g_j \in G$ with $g_j(\tilde{\gamma}_j) = \tilde{\gamma}$. These transformations must be distinct, since they respect the parameterization; in particular, $g_j(\tilde{\gamma}_j(0)) = \tilde{\gamma}(0)$. Then $g_j(\tilde{B}_r) \cap \tilde{\gamma} \neq \emptyset$ and $\tilde{\gamma}(t_j) = g_j(\tilde{\lambda}_0) \cap \tilde{\gamma}$. We take the liberty to define, hopefully without confusion, $g_j(\tilde{\lambda}_0) = \tilde{\lambda}_j$. The rays $\tilde{\lambda}_j$ intersecting $\tilde{\gamma}$ in a sequence of points $\tilde{\gamma}(t_j)$ converging to $\tilde{\gamma}(t^*)$.

$\tilde{\lambda}_j$ has endpoints $g_j(i) \in \mathbb{H}$ and $b_j \in \hat{\mathbb{R}}$. Since the interiors of the rays are all disjoint and the sequence $\{t_j\}$ is increasing, the sequence $\{b_j\}$ can limit at, at most two points in $\in \hat{\mathbb{R}}$, on opposite sides of $\tilde{\gamma}$. By restricting to one side, we may further stipulate that the sequence $\{e_j\}$ was chosen so that $b_j \to b^* \in \hat{\mathbb{R}}$.

The initial point $g_j(i)$ of the ray $\tilde{\lambda}_j$ is the center of the translate $g_j(\tilde{B}_r)$ of the disc $\tilde{B}_r$. Since G is discrete, the points $g_j(i)$ will limit on $\hat{\mathbb{R}}$ and not in $\mathbb{H}$. Consequently the discs $g_j(\tilde{B}_r)$ have Euclidean radii going to zero and they also will limit on $\hat{\mathbb{R}}$. Recall that for each j, $g_j(\tilde{B}_r) \cap \tilde{\gamma} \neq \emptyset$. Therefore the points $g_j(i)$ all lie in the radius r neighborhood of $\tilde{\gamma}$ and must limit at an endpoint of $\tilde{\gamma}$, denoted by $\tilde{\gamma}_*$.

If the rays λ_j accumulate in $\mathbb{H}$, it must be at a lift of β. The limit point b_* cannot be an endpoint of $\tilde{\gamma}$ distinct from $\tilde{\gamma}_*$, for then the λ_j would accumulate on $\tilde{\gamma}$, contradicting the hypothesis. Thus if b^* is an endpoint of $\tilde{\gamma}$ then $b_* = \tilde{\gamma}_*$.

Putting this all together: we have the sequence of geodesic rays $\tilde{\lambda}_j$, with the endpoints on $\hat{\mathbb{R}}$ converging to b^* and the endpoints in $\mathbb{H}$ converging to $\tilde{\gamma}_*$. In order for this to happen the sequence of intersections $\tilde{\lambda}_j \cap \tilde{\gamma} = \tilde{\gamma}(t_j)$ must converge to $\tilde{\gamma}_*$, contrary to the assumption that the excursion parameters $\psi(e_j) = t_j$, do not diverge to ∞. $\square$

If the ray γ contains infinitely many approximating excursions then they are ordered as a subsequence of the r_k-excursions. We write $\{e_j^*\}_{j=1}^{\infty}$ for the sequence of approximating excursions in $A(\gamma)$.

2.2.3. The depth of an excursion. Given a geodesic $\alpha \in \mathbb{H}$, define $\tilde{d}(\alpha)$ to be the distance from α to the point i. Define the *depth* of the r-excursion $e = \tilde{\gamma}$ of γ, to be the value $d(e) = \tilde{d}(\tilde{\gamma})$. For a pair of points $(x, y) \in I$, let α be the geodesic with endpoints $(\alpha_+, \alpha_-) = (x, y)$ and define $D(x, y) = \tilde{d}(\alpha)$.

3. The application of ergodic theory

3.1. The geodesic flow. The unit tangent bundle $T_1\mathbb{H}$ over $\mathbb{H}$ may be identified with $\mathbb{H} \times S^1$. Define the measure $\tilde{\mu} = A \times \theta$, where A and θ are respectively, area measure on $\mathbb{H}$ and lebesgue measure on S^1. The measure $\tilde{\mu}$ is invariant under the geodesic flow $\tilde{G}^t$ on $T_1\mathbb{H}$.

A full measure subset of $T_1\mathbb{H}$ can be modeled by triples $(\psi, \zeta, t) \in \mathbb{R}^3$ where $\psi \neq \zeta$. Let α be a geodesic in $\mathbb{H}$ parameterized so that $\alpha(0)$ is the Euclidean midpoint of the semi-circle $\alpha(\mathbb{R})$. If $\psi = \alpha_+$ and $\zeta = \alpha_-$, then (ψ, ζ, t) corresponds to the unit tangent vector $\dot{\alpha}(t) \in T_1\mathbb{H}$. It is clear that the correspondence defines an injection onto the subset of non-vertical vectors in $T_1\mathbb{H}$. In these coordinates the invariant measure $\tilde{\mu}$ has the form $\tilde{\mu} = \frac{1}{(\psi - \zeta)^2}\, d\psi d\zeta dt$, up to a constant multiple.

The measure $\tilde{\mu}$ projects to a measure μ on $T_1 S$, which is invariant under the geodesic flow G^t on $T_1 S$. It is well known that when S is a finite area surface, G^t acts ergodically with respect to μ. Every orbifold has a finite branched cover that is a hyperbolic surface. Thus ergodicity of the flow on the surface implies the ergodicity of the flow on the orbifold. See [1], [10] and [11] for details.

3.2. Convergence of the distribution. Every $v_q \in T_1 S$ uniquely determines a geodesic (or a geodesic ray) γ where $\gamma(0) = q$ and $\dot{\gamma}(0) = v_q$, the unit vector tangent to γ at q. Conversely, every geodesic determines a unit tangent vector at the parameter $t = 0$. We say that a property holds for almost all geodesics on S if there is a set $A \subset T_1 S$, of full μ-measure, so that the property holds for every geodesic determined by a vector in A, as above.

There are two classes of subsets of I and J that play particularly important roles in what follows. For $0 \leq z < \infty$ set $\Omega(z) = \{(x, y) \in I \mid D(x, y) \leq z\}$ and for $0 \leq z \leq r_k$ set $\Omega_*(z) = \{(x, y) \in J \mid D(x, y) \leq z\}$. As a consequence of Corollary 2 in Section 4.2.2, $\Omega_*(z) = J$ for $z \geq r_k$. At times we shall use the letters ψ and ζ in place of x and y. Then for appropriate values of z define the integrals

$$\Lambda(z) = \int_{\Omega(z)} \frac{1}{(\psi - \zeta)^2}\, d\psi d\zeta \quad \text{and} \quad \Lambda^*(z) = \int_{\Omega_*(z)} \frac{1}{(\psi - \zeta)^2}\, d\psi d\zeta.$$

After the next theorem, the proofs of Theorems 1 and 2 become computations in hyperbolic geometry and plane integration.

Theorem 3. *For almost all geodesic rays γ on S, for $r > 0$ and $0 \leq z \leq r$, the distribution* $\text{dist}_k(r, z)(\gamma)$ *converges to* $\text{dist}_k(r, z)$, *where*

$$\text{dist}_k(r, z) = \frac{\Lambda(z)}{\Lambda(r)}.$$

For almost all geodesic rays γ on S and for $0 \leq z \leq r_k$, $\mathrm{dist}_k^*(z)(\gamma)$ converges to $\mathrm{dist}_k^*(z)$, where

$$\mathrm{dist}_k^*(z) = \frac{\Lambda^*(z)}{\Lambda^*(r_k)}.$$

The remainder of Section 3 shall be devoted to the proof of this theorem. The idea is to produce a measure transverse to the geodesic flow, on which there is an ergodic action related to the sequence of excursions along a generic geodesic on S. The distributions can be expressed as limiting sums with respect to this action which then, using the Birkhoff Ergodic Theorem, become the above integrals.

3.3. A cross-section of G^t over $T_1\lambda$. Let $T_1\lambda$ denote the subset of the unit tangent bundle over the geodesic λ, consisting of the vectors v_q with $q \in \lambda - \{p\}$. We shall define subsets of $T_1\lambda$ associated with r-excursions and separately, with approximating excursions. Most of the work takes place in $\mathbb{H}$, on the preimage $T_1\tilde{\lambda}_0$ of $T_1\lambda$.

Given $r > 0$ define the set of vectors $\tilde{l}^r \subset T_1\tilde{\lambda}_0$, where $v_q \in \tilde{l}^r$ if there is a geodesic α in $\mathbb{H}$ with

(1) $(\alpha_+, \alpha_-) \in I$,

(2) $\alpha(0) = q \in \tilde{\lambda}_0 - \{i\}$ and $\dot{\alpha}(0) = v_q$,

(3) $\alpha \cap \tilde{B}_r \neq \emptyset$, and

(4) the geodesic ray $\pi \circ \alpha$ on S contains infinitely many r-excursions.

Similarly, define $\tilde{l}^* \subset T_1\tilde{\lambda}_0$, where $v_q \in \tilde{l}^*$ if there is a geodesic α in $\mathbb{H}$ with

(1) $(\alpha_+, \alpha_-) \in J$,

(2) $\alpha(0) = q \in \tilde{\lambda}_0 - \{i\}$ and $\dot{\alpha}(0) = v_q$, and

(3) the geodesic ray $\pi \circ \alpha$ on S contains infinitely many approximating excursions.

In order to treat the various cases simultaneously, we adopt the convention that τ denotes either the parameter $r > 0$ or the symbol $*$. The set of vectors $\tilde{l}^\tau$ projects to a set of vectors $l^\tau \subset T_1\lambda$. Essentially, l^τ consists of the tangent vectors of the form $\dot{\gamma}(t_e)$, associated to r-excursions (or approximating excursions) e, with the caveat that the ray γ contains infinitely many such excursions. Since λ is simple, the projection $\pi_*: T_1\tilde{\lambda}_0 \to T_1\lambda$ is bijective and consequently we have the following.

Proposition 3. *The projection $\pi_*: \tilde{l}^\tau \to l^\tau$ is a bijection.*

We let $*$-excursion be another term for approximating excursion. The next proposition asserts that l^τ is a cross-section for the geodesic flow on S.

Proposition 4. *For each value τ, there is a full measure, G^t-invariant subset $M_\tau \subset T_1 S$ so that l^τ is a cross-section for the flow G^t acting on M_τ. In particular, for each $v \in M_\tau$, there is a geodesic ray γ, possessing an infinite sequence of τ-excursion parameters $\{t_{e_j}\}$, so that $\dot{\gamma}(0) = v$ and $G^t(v) \in l^\tau$ for $t > 0$ if and only if $t = t_{e_j}$ for some $j \in \mathbb{N}$.*

Proof. We argue when $\tau = r$. Let M_r be the set of vectors $v \in T_1 S$ so that a geodesic ray γ with $\dot{\gamma}(0) = v$ contains infinitely many r-excursions, none of which intersect the cone point p. It follows from Proposition 2 that M_r is G^t-invariant. In particular, $l^r \subset M_r$.

In order to see that M_r has full μ-measure, choose a value u so that $0 < u < r_k$ and $u \leq r$. As a consequence of the Poincaré Recurrence Theorem [7], there is a set of full measure $N_u \subset T_1 S$, with the property that for each $v \in N_u$ the geodesic γ with $\dot{\gamma}(0) = v$ returns infinitely often to B_u. It is proved in Corollary 3 of Section 4.2.2, that every intersection of a geodesic γ with B_u produces a u-excursion. Therefore, such a geodesic γ must contain an infinite sequence of u-excursions. The subset N_u^* consisting of $v \in N_u$ so that the G^t-orbit of v does not pass through p, has full measure in N_u. As $u \leq r$, $N_u^* \subset M_r$, showing that M_r has full measure in $T_1 S$.

Given $v \in M_r$, let γ be the geodesic with $\dot{\gamma}(0) = v$. Then γ contains an infinite sequence of r-excursions $\{e_j\}$ with excursion parameters $\psi(e_j) = t_{e_j}$. As before write $e_j = \tilde{\gamma}_j$ with $((\tilde{\gamma}_j)_+, (\tilde{\gamma}_j)_-) \in I$. Thus for each $j \in \mathbb{N}$, $G^{t_{e_j}}(v) = \pi_*(\dot{\tilde{\gamma}}_j)(t_{e_j}) \in l^r$.

If for some $t > 0$ and $v \in M_r$, $G^t(v) \in l^r$, then there is a lift $\tilde{v}$ of v and $\tilde{\gamma}$ of γ so that $\dot{\tilde{\gamma}}(t) = \tilde{v} \in T_1 \tilde{\lambda}_0$, $(\tilde{\gamma}_+, \tilde{\gamma}_-) \in I$ and $\tilde{\gamma} \cap \tilde{B}_r \neq \emptyset$. Therefore $e = \tilde{\gamma}$ is an r-excursion. The argument for approximating excursions is similar. $\square$

3.4. The first return map. Let $F_\tau : l^\tau \to l^\tau$ be the first return map under the geodesic flow $G^t : M_\tau \to M_\tau$. In other words, given $v \in l^\tau$ let γ be a geodesic in S with $\dot{\gamma}(0) = v$. Then $F_\tau(v) = \dot{\gamma}(t)$, where $t > 0$ is the first value with $\dot{\gamma}(t) \in l^\tau$. The invariant measure μ for the geodesic flow on S induces an F_τ-invariant measure ν_τ on l^τ for which F_τ is ergodic, [1, 2]. Under the identification of l^τ with $\tilde{l}^\tau$, F_τ lifts to a map $\tilde{F}_\tau : \tilde{l}^\tau \to \tilde{l}^\tau$ which is invariant and ergodic with respect to the lifted measure $\tilde{\nu}_\tau$. Furthermore, this construction produces an isomorphism between the dynamical systems $(l^\tau, F_\tau, \nu_\tau)$ and $(\tilde{l}^\tau, \tilde{F}_\tau, \tilde{\nu}_\tau)$.

It will be easier to work with $\tilde{F}_\tau$ in a form associated to the (ψ, ζ, t) coordinates. Let Ω_τ denote $\Omega(r)$ if $\tau = r$ and J is $\tau = *$. There is a map $B_\tau : \tilde{l}^\tau \to \Omega_\tau$, taking v_q to the endpoints (α_+, α_-) of the geodesic α with $\alpha(0) = q$ and $\dot{\alpha}(0) = v_q$. B_τ is injective. If we represent $v_q = (\psi_q, \zeta_q, t_q)$ in coordinates, then B_τ becomes the projection $B_\tau(\psi_q, \zeta_q, t_q) = (\psi_q, \zeta_q)$, onto the first two coordinates. Let $\omega_\tau = B_\tau(\tilde{l}_\tau)$ and define the transformations $R_\tau : \omega_\tau \to \omega_\tau$ by $R_\tau \circ B_\tau = B_\tau \circ \tilde{F}_\tau$. Note that ω_τ is precisely the subset of Ω_τ so that if α is a geodesic with $(\alpha_+, \alpha_-) = (\psi, \zeta) \in \Omega_\tau$

the geodesic ray α projects to a geodesic ray on S that contains infinitely many r-excursions or approximating excursions, not through p, as appropriate.

Arguing as in [11], we can use a theorem of Ambrose [3] to show that the geodesic flow can be built up, in a nice way, from the first-return map on the cross section. This theorem allows us to conclude that the G^t-invariant measure μ on S has the form $\mu = \nu_\tau \times dt$ in local coordinates, where dt represents lebesgue measure on the flow lines. Then $\mu = \nu_\tau \times dt$ pulls back to the $\tilde{G}^t$-invariant measure $\tilde{\mu} = \tilde{\nu}_\tau \times dt$ on $T_1\mathbb{H}$. It follows that in (ψ, ζ) coordinates, $\tilde{\nu}_\tau = \frac{1}{(\psi-\zeta)^2}d\psi\,d\zeta$ is the induced invariant measure for R_τ. We shall write $\tilde{\nu}_\tau = \tilde{\nu}$ since, in these coordinates, the measure is independent of τ.

Since M_τ is full measure in $T_1 S$ and G^t-invariant, the cross-section l^τ is of full measure in $T_1\lambda$. It follows, in particular, that ω_τ is full measure in Ω_τ. In light of this, the transformation $R_\tau : \Omega_\tau \to \Omega_\tau$ is well defined up to sets of measure zero, is B_τ-conjugate to $\tilde{F}_\tau$, and has invariant measure $\tilde{\nu}$.

It is also important to note that, as a consequence of Proposition 4, for almost all $v \in T_1 S$, there are infinitely many values $t_j \to \infty$ so that the geodesic γ with $\dot{\gamma}(0) = v$ has $\dot{\gamma}(t_j) \in l^\tau$. Therefore, for almost all $v \in T_1 S$ the geodesic γ with $\dot{\gamma}(0) = v$ has a lift with its endpoints in ω_τ. In other words, almost every geodesic on S has a lift with its endpoints in ω_τ.

3.5. The proof of Theorem 3. As usual, we shall prove the theorem for r-excursions. The details for approximating excursions are similar. Let χ_Y denote the characteristic function of the set Y. First observe that for $(\psi_0, \zeta_0) \in \Omega(r)$ we have the equality

$$\lim_{n\to\infty} \frac{1}{n} \#\{j \mid 1 \le j \le n, D \circ R_r^j(\psi_0, \zeta_0) \le z\} \tag{2}$$

$$= \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^n (\chi_{[0,z]} \circ D)(R_r^j(\psi_0, \zeta_0)). \tag{3}$$

The measure $\tilde{\nu}^* = \frac{1}{\Lambda(r)}\tilde{\nu}$ is an R_r-invariant probability measure on $\Omega(r)$, with respect to which R_r is ergodic. By Birkhoff's Ergodic Theorem and the fact that ω_r has full $\tilde{\nu}^*$-measure in $\Omega(r)$, for almost all $(\psi_0, \zeta_0) \in \Omega(r)$ the limit (3) is equal to

$$\int_{\Omega(r)} (\chi_{[0,z]} \circ D)\tilde{\nu}^* = \frac{1}{\Lambda(r)} \int_{\Omega(z)} \tilde{\nu} = \frac{\Lambda(z)}{\Lambda(r)}.$$

Almost all geodesics γ on S have lifts with endpoints in ω_τ. Let γ be one of these geodesics. Then there is an infinite sequence $e_j = \tilde{\gamma}_j$, $j = 0, 1, \ldots,$ of r-excursions along the ray γ, with associated excursion parameters $t_j = \psi(e_j)$. Let $v_j = \dot{\gamma}(t_j) \in l^\tau$. By Proposition 4, $F_\tau(v_j) = v_{j+1}$ and therefore $F_\tau^j(v_0) = v_j$. The lift of v_j to $\tilde{\lambda}_0$ is $\tilde{v}_j = \dot{\tilde{\gamma}}_j(t_j)$ and, as above, $\tilde{F}_\tau^j(\tilde{v}_0) = \tilde{v}_j$.

Write $\tilde{\gamma} = \tilde{\gamma}_0$. To each $j \in \mathbb{N}$ there is a unique $g_j \in G$ so that $\tilde{\gamma}_j = g_j(\tilde{\gamma})$. Then in (ψ, ζ, t) coordinates, $\tilde{v}_j = ((\tilde{\gamma}_j)_+, (\tilde{\gamma}_j)_-, s_j) = (g_j(\tilde{\gamma}_+), g_j(\tilde{\gamma}_-), s_j)$, for some values s_j that are not easily related to the t_j. Then if we set $(\psi_0, \zeta_0) = (\tilde{\gamma}_+, \tilde{\gamma}_-) \in \omega_\tau$ we get

$$
\begin{aligned}
R_\tau^j(\psi_0, \zeta_0) &= B_\tau \circ \tilde{F}_\tau^j \circ B_\tau^{-1}(\tilde{\gamma}_+, \tilde{\gamma}_-) = B_\tau \circ \tilde{F}_\tau^j(\tilde{v}_0) \\
&= B_\tau(\tilde{v}_j) = B_\tau(g_j(\psi_0), g_j(\zeta_0), s_j) = (g_j(\psi_0), g_j(\zeta_0)).
\end{aligned}
\tag{4}
$$

Turning to the distribution we have

$$
\begin{aligned}
\operatorname{dist}_k(r, z)(\gamma) &= \lim_{n \to \infty} \frac{1}{n} \#\{j \mid 1 \leq j \leq n, d(e_j) \leq z\} \\
&= \lim_{n \to \infty} \frac{1}{n} \#\{j \mid 1 \leq j \leq n, \tilde{d}(\tilde{\gamma}_j) \leq z\} \\
&= \lim_{n \to \infty} \frac{1}{n} \#\{j \mid 1 \leq j \leq n, D(g_j(\tilde{\gamma}_+), g_j(\tilde{\gamma}_-)) \leq z\} \\
&= \lim_{n \to \infty} \frac{1}{n} \#\{j \mid 1 \leq j \leq n, D(g_j(\psi_0), g_j(\zeta_0)) \leq z\}.
\end{aligned}
\tag{5}
$$

By the sequence of equalities (4), the limit (5) is equal to the limit (2). Therefore, we may conclude that for almost all geodesics γ on S, $\operatorname{dist}(r, z)(\gamma)$ converges to $\operatorname{dist}(r, z) = \frac{\Lambda(z)}{\Lambda(r)}$, which proves Theorem 3 for r-excursions. $\square$

Remark. In the argument for approximating excursions the roles of I and $\Omega(z)$ are taken by J and $\Omega_*(z)$ and the proof proceeds as above.

4. The computation of Λ and Λ^*

4.1. The values of Λ and Λ^*. In light of Theorem 3, the proofs of Theorems 1 and 2 will be completed by computing the values of Λ and Λ^*. Recall that $r_k = \sinh^{-1}\left(\cot \frac{\pi}{k}\right)$, $\delta_k = \sinh^{-1}\left(\cot \frac{2\pi}{k}\right)$ and $\varphi_k(x) = \frac{1 - x \tan \frac{\pi}{k}}{x + \tan \frac{\pi}{k}}$.

Theorem 4. *For $z \geq 0$ we have*

$$
\Lambda(z) = \begin{cases}
\frac{2\pi}{k} \sinh z & \text{if } z \leq r_k, \\[2ex]
\sinh z \tan^{-1}\left(\frac{1}{\sinh z}\right) + \log(\sin \frac{\pi}{k} \cosh z) & \text{if } z > r_k,
\end{cases}
$$

$$\Lambda^*(z) = \begin{cases} \frac{2\pi}{k} \sinh z & \text{if } z \leq \delta_k, \\[2mm] \sinh z \tan^{-1}(\varphi_k(\sinh z)) + \log\left(2 \cosh z \sin \frac{\pi}{k} \cos \frac{\pi}{k}\right) & \text{if } \delta_k < z < r_k, \\[2mm] \log\left(2 \cos \frac{\pi}{k}\right) & \text{if } z \geq r_k. \end{cases}$$

When $k = 3$ or 4, $\delta_k \leq 0$ and only the second two cases occur for the value of $\Lambda^*(z)$.

The rest of the section is devoted to the proof of this theorem.

4.2. Hyperbolic geometric considerations

4.2.1. The geodesic tangent to a disc. In order to turn Λ and Λ^* into easily computable double integrals, we shall derive several formulae that will be of use in determining the limits of integration. Given distinct points $a, b \in \hat{\mathbb{R}}$, let $\overline{ab}$ denote the geodesic α in $\mathbb{H}$ with $\alpha_+ = a$ and $\alpha_- = b$.

Theorem 5. *Given $x \geq 0$, the point $w \in [-\frac{1}{x}, x)$, for which the geodesic $\overline{xw}$ is tangent to the hyperbolic disc $\tilde{B}_\rho$, is determined by the formula*

$$w = W_\rho(x) = \frac{x \sinh \rho - 1}{x + \sinh \rho}.$$

Similarly, given $x \leq 0$ the point $w \in (x, -\frac{1}{x}]$ for which the geodesic $\overline{xw}$, is tangent to the hyperbolic disc $\tilde{B}_\rho$ is given by the formula $w = -W_\rho(-x)$.

There are two geodesics in $\mathbb{H}$ tangent to $\tilde{B}_\rho$, with an endpoint at x. The condition $w \in [-\frac{1}{x}, x)$ is equivalent to stipulating that i lies in the unbounded region of the upper-half plane in the complement of $\overline{xw}$

Let Δ denote the unit disc model for the hyperbolic plane, which we envision as sitting in the complex plane. Write $\partial\Delta$ for the unit circle, which is its boundary at infinity. The proof of Theorem 5 involves computations in Δ and the transformation of those results to $\mathbb{H}$. The following lemma can be proved by an easy computation using the hyperbolic metric in Δ.

Lemma 1. *In Δ, the circle of hyperbolic radius ρ and center 0 has Euclidean radius $r = \tanh \frac{\rho}{2}$*

If $z, \xi \in \partial\Delta$ then, as in the previous case, we let $\overline{z\xi}$ be the hyperbolic geodesic in Δ with endpoints z, ξ. Similarly, for $z \in \partial\Delta$ let $\overline{z0}$ be the geodesic with endpoint z passing through the origin. This geodesic is a Euclidean straight line. We use the same notation $\tilde{B}_\rho(0)$ to denote the hyperbolic disc of center 0 and radius ρ in Δ. To avoid confusion we shall continue to write $\tilde{B}_\rho$ for $\tilde{B}_\rho(i)$.

Lemma 2. *Suppose that $z, \xi \in \partial\Delta$ are two points so that the geodesic $\overline{z\xi}$ is tangent to the hyperbolic disc $\tilde{B}_\rho(0)$. Then $|z - \xi| = 2\operatorname{sech}\rho$.*

Proof. There is no loss of generality in supposing that $z = x + iy$ with $x, y \geq 0$ and $\xi = \overline{z}$. Then we have $|z - \xi| = 2y$ and it remains to show that $y = \operatorname{sech}\rho$. Let C be the circle perpendicular to $\partial\Delta$ whose intersection with Δ is the geodesic $\overline{z\xi}$. Since C is invariant under the reflection $z \to \frac{1}{\overline{z}}$ which fixes $\partial\Delta$, it crosses the real axis in points $0 < r < 1$ and $\frac{1}{r}$. By Lemma 1, since C is tangent to the circle of radius ρ, $r = \tanh\frac{\rho}{2}$.

As a consequence of z lying on C, it satisfies the equation $|z - \frac{1}{2}(\frac{1}{r} + r)|^2 = [\frac{1}{2}(\frac{1}{r} - r)]^2$. Also, since $z \in \partial\Delta$, $|z| = 1$. Solving these equations simultaneously gives $y = \frac{1-r^2}{1+r^2}$. Setting $r = \tanh\frac{\rho}{2}$ and simplifying, gives the result. $\square$

Lemma 3. *Suppose z is a point in the upper half of the unit circle $\partial\Delta$. Let $\xi \in \partial\Delta$ be the point which lies to the right of the line $\overline{z0}$ so that $\overline{z\xi}$ is tangent to the disc $\tilde{B}_\rho(0)$. Then $\xi = z(c - i\sqrt{1 - c^2})$ where $c = 1 - 2\operatorname{sech}^2\rho$. The above also holds for $z = 1$, where ξ is then a point in the lower half of $\partial\Delta$.*

Proof. From the previous lemma we have $|z - \xi| = 2\operatorname{sech}\rho$, which simplifies to $\operatorname{Re} z\overline{\xi} = 1 - 2\operatorname{sech}^2\rho$. For simplicity write $c = 1 - 2\operatorname{sech}^2\rho$. Note that as ρ increases on the interval $[0, \infty)$, c increases on $[-1, 1)$.

The point ξ may be written in the form $\xi = \eta z$, for a point η of modulus one. Let $\eta = u + iv$. One see by a simple computation that $u = \operatorname{Re} z\overline{\xi} = c$ and then $w = \pm\sqrt{1 - c^2}$. With the choice of the minus sign the points $\eta = c - i\sqrt{1 - c^2}$ fill out the lower half of the unit circle. As a function of ρ, η moves counter-clockwise through its arc as ρ increases from zero to infinity. Thus, treating the point $\xi = \eta z$ as a function of ρ, we see that it moves counter-clockwise around the circle from $-z$ to z as ρ increases from zero to infinity. It follows that for all values of ρ, the point $\xi = \eta z$ lies to the right of the line $\overline{z0}$, or in the lower half-plane if $z = 1$. With the choice of a plus sign, η fills out the upper half of the circle and the points ηz lie to the left of $\overline{z0}$, or in the upper half-plane if $z = 1$. That completes the proof. $\square$

Proof of Theorem 5. The Möbius transformation $g(z) = \frac{-iz+i}{z+1}$ maps Δ isometrically to $\mathbb{H}$, taking the disc $\tilde{B}_\rho(0)$ to the disc $\tilde{B}_\rho$. g also defines a one-to-one correspondence between $\partial\Delta$ and $\hat{\mathbb{R}}$, where 1 goes to 0, -1 goes to ∞ and the upper and lower semi-circles respectively, are mapped to the positive and negative real axes.

As in the previous argument let $\eta = c - i\sqrt{1 - c^2}$, where $c = 1 - 2\operatorname{sech}^2\rho$, and define the transformation $W_\rho(x) = g(\eta g^{-1}(x))$. Simplifying we get

$$W_\rho(x) = g(\eta g^{-1}(x)) = \frac{ix\left(\frac{1+\eta}{1-\eta}\right) - 1}{x + i\left(\frac{1+\eta}{1-\eta}\right)}.$$

A further computation gives

$$i\left(\frac{1+\eta}{1-\eta}\right) = -\frac{2\operatorname{Im}\eta}{|1-\eta|^2} = \frac{\sqrt{1-c^2}}{1-c} = \sinh\rho,$$

which shows that the formula for W_ρ is the one asserted in the theorem. Now we need to see that W_ρ does what it is claimed to do.

If $x \geq 0$ then $g^{-1}(x) = z$ lies in the upper-half of $\partial\Delta$ or is equal to 1. Then $\xi = \eta z$ is the point of $\partial\Delta$ for which the geodesic $\overline{z\xi}$ satisfies the hypotheses of the previous lemma. The image of this geodesic, $g(\overline{z\xi})$, is the hyperbolic geodesic in $\mathbb{H}$ which is tangent to $g(\tilde{B}_\rho(0)) = \tilde{B}_\rho$ and has endpoints x and $W_\rho(x)$. Recall from the proof of Lemma 3 that $\xi = \eta z$ lies in the counter-clockwise arc of $\partial\Delta$ between $-z$ and z. Since the image under g of this arc is the interval $[-\frac{1}{x}, x)$, the proof is complete for $x \geq 0$.

Now suppose $x \leq 0$. Consider the isometry $h(z) = -\bar{z}$ of $\mathbb{H}$. h maps the geodesic tangent to $\tilde{B}_\rho$, with endpoints $-x$ and $W_\rho(-x) \in [\frac{1}{x}, -x)$ to the geodesic tangent to $\tilde{B}_\rho$ with endpoints x and $-W_\rho(-x) \in (x, -\frac{1}{x}]$. Thus $-W_\rho(-x)$ is as asserted in the proposition. $\qquad\square$

4.2.2. Consequences of the tangency computations

Corollary 1. *Given $x, z \geq 0$, $D(x, y) \leq z$ for $y \in [-\frac{1}{x}, x)$ if and only if $y \in [-\frac{1}{x}, W_z(x)]$.*

Proof. For fixed $x \geq 0$ and for $y \in [-\frac{1}{x}, x)$, define $H(y) = D(x, y)$. Then $H(y)$ is the value ρ for which the geodesic $\overline{xy}$ is tangent to the disc $\tilde{B}_\rho$. We have shown in the proofs of Lemma 3 and Theorem 5 that H is an increasing function of y, mapping $[-\frac{1}{x}, x)$ onto $[0, \infty)$. Also recall that, for x, y as above, $\overline{xy}$ is tangent to $\tilde{B}_z$ precisely when $y = W_z(x)$. It follows that $D(x, y) = H(y) \leq z$ if and only if $y \in [-\frac{1}{x}, W_z(x)]$. $\qquad\square$

Corollary 2. *When $k \geq 5$, every z-excursion with $z \leq \delta_k$ is an approximating excursion. Moreover, every approximating excursion is a z-excursion for some $z < r_a$.*

Proof. First observe that the equation $W_z(a_k) = -a_k$ has the solution

$$z = \sinh^{-1}(\varphi_k^{-1}(a_k)) = \sinh^{-1}\left(\frac{1-\tan^2\frac{\pi}{k}}{2\tan\frac{\pi}{k}}\right) = \sinh^{-1}\left(\cot\frac{2\pi}{k}\right) = \delta_k.$$

This is positive if and only if $k \geq 5$, which we henceforth suppose to be the case. For $x, z \geq 0$, $W_z(x)$ is an increasing function in both variables. Therefore for $z \leq \delta_k$ and $x \in (0, a_k)$, $W_z(x) < W_z(a_k) \leq W_{\delta_k}(a_k) = -a_k$.

Let $e = \tilde{\gamma}$ be a z-excursion for some $z \le \delta_k$. Then $x = \tilde{\gamma}_+ \in (0, a_k)$ and, as a consequence of the above, $\tilde{\gamma}_- = W_z(x) < -a_k$. This says that e is an approximating excursion, proving the first assertion of the corollary.

Now observe that $z = D(0, -a_k)$ is the value for which $W_z(0) = -a_k$. Solving for z we get $z = \sinh\left(\cot\left(\frac{\pi}{k}\right)\right) = r_k$. Thus the geodesic α with endpoints 0 and a_k is tangent to the disc $\tilde{B}_{r_k}$.

Let $e = \tilde{\gamma}$ be an approximating excursion. Then $\tilde{\gamma}_+ \in (0, a_k)$ and $\tilde{\gamma}_- \in [-\frac{1}{x}, -a_k)$. It follows that $\tilde{\gamma}$ lies above α in $\mathbb{H}$ and therefore, $\tilde{\gamma} \cap \tilde{B}_{r_k} \ne \emptyset$. The second part of the corollary is proved. $\square$

Corollary 3. *Suppose γ is a geodesic on S and $\gamma \cap B_{r_k} = \epsilon \ne \emptyset$. Then there is a lift $\tilde{\gamma}$ of γ and a lift $\tilde{\epsilon}$ of ϵ contained in $\tilde{\gamma}$, so that $\tilde{\epsilon} \cap \tilde{B}_{r_k} \ne \emptyset$ and $e = \tilde{\gamma}$ is a z-excursion for some $z < r_k$.*

Proof. Choose a lift $\tilde{\gamma}$ so that $\tilde{\epsilon} \cap \tilde{B}_{r_k} \ne \emptyset$ and $\tilde{\gamma}_+ \in (0, a_k)$. Without loss of generality we suppose that $\tilde{\gamma}_- < \tilde{\gamma}_+$. Since $W_{r_k}(a_k) = 0$, the geodesic α with endpoints 0 and a_k is tangent to the disc $\tilde{B}_{r_k}$. If $\tilde{\gamma}_- > 0$ then $\tilde{\gamma}$ lies entirely below α and $\tilde{\gamma} \cap \tilde{B}_{r_k} \ne \emptyset$, which is impossible. Therefore $\tilde{\gamma}_- < 0$ and $e = \tilde{\gamma}$ is a z-excursion for some $z < r_k$.
                                                                    $\square$

4.3. The proof of Theorem 4. We shall write the integrals as sums of iterated double integrals. Once in this form, their actual evaluation is a elementary.

4.3.1. The computation of $\Lambda(z)$. Consider fixed values $z \ge 0$ and $x \in (0, a_k)$. If $W_z(x) < 0$ then, by Corollary 1, $D(x, y) \le z$ for $y \in [-\frac{1}{x}, 0)$ precisely when $y \in [-\frac{1}{x}, W_z(x)]$, whereas if $W_z(x) \ge 0$ then $D(x, y) \le z$ for all $y \in [-\frac{1}{x}, 0)$. As a result of this observation, for fixed $z > 0$, there are two cases to be considered.

The first case is when $z \le r_k$. This condition is equivalent to $a_k \le \frac{1}{\sinh z}$ or $W_z(a_k) \le 0$. Since $W_z(x)$ is an increasing function of x, this occurs if and only if $W_z(x) < 0$ for all $x \in (0, a_k)$. Then, as observed above, for $x \in (0, a_k)$, $D(x, y) \le z$ for $y \in [-\frac{1}{x}, 0)$ if and only if $y \in [-\frac{1}{x}, W_z(x)]$. By symmetry, we have that for $x \in (-a_k, 0)$, $D(x, y) \le z$ for $y \in (0, -\frac{1}{x}]$ if and only if $y \in [-W_z(-x), -\frac{1}{x}]$. Consequently, when $z \le r_k$ we get

$$\Lambda(z) = \int_0^{a_k} \int_{-\frac{1}{x}}^{W_z(x)} \frac{1}{(x-y)^2}\, dy dx + \int_0^{-a_k} \int_{-W_z(-x)}^{-\frac{1}{x}} \frac{1}{(x-y)^2}\, dy dx$$

In the second case $z > r_k$ and we will have $W_z(x) = 0$ for $x = \frac{1}{\sinh z} \in (0, a_k)$. Then the set of y values for which $D(x, y) \le z$ will take two different forms. For $x \in (0, \frac{1}{\sinh z})$, we have $D(x, y) \le z$ for $y \in [-\frac{1}{x}, 0)$ if and only if $y \in [-\frac{1}{x}, W_z(x)]$;

whereas, for $x \in \left[\frac{1}{\sinh z}, 0\right)$ we get $D(x, y) \leq z$ for all $y \in \left[-\frac{1}{x}, 0\right)$. Similar results again hold, by symmetry, for $x \in (-a, 0)$. Then for $z > r_k$ we get

$$\Lambda(z) = \int_0^{\frac{1}{\sinh z}} \int_{-\frac{1}{x}}^{W_z(x)} \frac{1}{(x-y)^2} \, dy \, dx + \int_{\frac{1}{\sinh z}}^{a_k} \int_{-\frac{1}{x}}^0 \frac{1}{(x-y)^2} \, dy \, dx$$

$$+ \int_{-\frac{1}{\sinh z}}^0 \int_{-W_z(-x)}^{-\frac{1}{x}} \frac{1}{(x-y)^2} \, dy \, dx + \int_{-a_k}^{-\frac{1}{\sinh z}} \int_0^{-\frac{1}{x}} \frac{1}{(x-y)^2} \, dy \, dx.$$

4.3.2. The computation of $\Lambda^*(z)$. The same considerations and cases come into play in the computation of $\Lambda^*(z)$. In addition there is a third special case, which is a consequence of the fact that every approximating excursion is an r-excursion for some $r < r_k$. The three cases depend on the position of the points $W_z(x)$ with respect to $-a_k$, rather than with respect to 0, as in the earlier computations. The cases are also described by the position of the particular point x_z, for which $W_z(x_z) = -a_k$.

In the first case $z \leq \delta_k$. As observed in the proof of Corollary 2, this is equivalent to $W_z(x) < -a_k$ for all $x \in (0, a_k)$. Since $\delta_k \leq 0$ for $k = 3, 4$, this case does not arise for those values of k. Then, by Corollary 1, for all $x \in (0, a_k)$, $D(x, y) \leq z$ for $y \in [-\frac{1}{x}, -a_k)$ if and only if $y \in [-\frac{1}{x}, W_z(x)]$, and by symmetry, we have that for all $x \in (-a_k, 0)$, $D(x, y) \leq z$ for $y \in (a_k, -\frac{1}{x}]$ if and only if $y \in [-W_z(-x), -\frac{1}{x}]$. This is the same situation we had computing the integral in the earlier first case. Therefore for $z \leq \delta_k$,

$$\Lambda(z) = \Lambda^*(z) = \frac{2\pi}{k} \sinh z.$$

In the second case $\delta_k < z < r_k$, where we let $\delta_k = 0$ for $k = 3, 4$. Consider the equation $W_z(x) = -a_k$, whose solution is a point $x_z = \varphi_k(\sinh z)$. x_z is a decreasing function of z. We have seen that $x_{\delta_k} = a_k$ and $x_{r_k} = 0$. Therefore, for $\delta_k < z < r_k$ we have $W_z(x) = -a_k$ for the point $x = x_z \in (0, a_k)$. As in case two in the computation of Λ, the set of y values for which $D(x, y) \leq z$ will take two different forms. For $x \in (0, x_z)$, we have $D(x, y) \leq z$ for $y \in [-\frac{1}{x}, -a_k)$ if and only if $y \in [-\frac{1}{x}, W_z(x)]$ and for $x \in [x_z, 0)$, we get $D(x, y) \leq z$ for all $y \in [-\frac{1}{x}, -a_k)$. Similar results again hold, by symmetry, for $x \in (-a_k, 0)$. It follows that for $\delta_k < z < r_k$

$$\Lambda^*(z) = \int_0^{x_z} \int_{-\frac{1}{x}}^{W_z(x)} \frac{1}{(x-y)^2} \, dy \, dx + \int_{x_z}^{a_k} \int_{-\frac{1}{x}}^0 \frac{1}{(x-y)^2} \, dy \, dx$$

$$+ \int_{-x_z}^0 \int_{-W_z(-x)}^{-\frac{1}{x}} \frac{1}{(x-y)^2} \, dy \, dx + \int_{-a_k}^{-x_z} \int_0^{-\frac{1}{x}} \frac{1}{(x-y)^2} \, dy \, dx.$$

The last case to consider is when $z \geq r_k$. By Corollary 2, $\Omega_*(z) = J$ for all $z \geq r_k$. Therefore, for all $x \in (0, a_k)$, $D(x, y) \leq z$ for all $y \in [-\frac{1}{x}, -a_k)$. Consequently, in

this last case we have, for $z \geq r_k$

$$\Lambda^*(z) = \int_0^{a_k} \int_{-\frac{1}{x}}^{-a_k} \frac{1}{(x-y)^2} \, dy dx + \int_0^{-a_k} \int_{a_k}^{-\frac{1}{x}} \frac{1}{(x-y)^2} \, dy dx$$

5. Depth as a function of area

We have been dealing with excursion depth as expressed in terms of the radius of the largest cone neighborhood disjoint from the excursion. In [8], where we looked at excursion into a cusp, it was just not possible to express the depth of an excursion as the distance to the cusp end, since the cusp is off at infinity. Instead we considered the area of the largest neighborhood of the cusp disjoint from the excursion. Suppose we were to take the same approach here.

In the hyperbolic plane $\mathbb{H}$ the area of a hyperbolic disc of radius r is $2\pi(\cosh r - 1)$ [4]. Let G be a Fuchsian group representing S and let $\pi : \mathbb{H} \to \mathbb{H}/G = S$ denote the natural projection map. Suppose p is the cone point on S and $\tilde{p}$ is a preimage in $\mathbb{H}$. The actual area of a disc $B_r(p)$ is at most $1/k$ times the area of its preimage $\tilde{B}_r(\tilde{p})$, with equality if the projection is precisely k-to-1. In the case of equality, if A is the area of $B_r(p)$, then $\tilde{B}_r(\tilde{p})$ has area $kA = 2\pi(\cosh r - 1)$. Even if there is overlap and the projection is not k-to-1, in order to be consistent with [8], we define the A-depth of an excursion e, written $A(e)$, to be $\frac{2\pi}{k}(\cosh(d(e)) - 1)$. Define an R-excursion e to be a usual r-excursion where $r = \cosh^{-1}\left(\frac{k}{2\pi}R + 1\right)$, the radius of the disc of area R about p. The definition of an approximating excursion is unchanged.

Suppose $\{e_j\}$ is a sequence of R-excursions along a geodesic γ on S. Then define

$$A \operatorname{dist}_k(R, Z)(\gamma) = \lim_{n \to \infty} \frac{1}{n} \#\{j \mid 1 \leq j \leq n, \ A(e_j) \leq Z\}.$$

As a consequence of the definition of $A(e)$, this is equal to

$$\lim_{n \to \infty} \frac{1}{n} \# \left\{ j \mid 1 \leq j \leq n, d(e_j) \leq \cosh^{-1}\left(\frac{k}{2\pi}Z + 1\right) \right\}.$$

For almost all geodesics γ on S this last limit was shown to converge to

$$\operatorname{dist}_k\left(\cosh^{-1}\left(\frac{k}{2\pi}R + 1\right), \cosh^{-1}\left(\frac{k}{2\pi}Z + 1\right)\right) = \frac{\Lambda\left(\cosh^{-1}\left(\frac{k}{2\pi}Z + 1\right)\right)}{\Lambda\left(\cosh^{-1}\left(\frac{k}{2\pi}R + 1\right)\right)}.$$

For $R \leq \frac{2\pi}{k}(\cosh(r_k) - 1)$ and $Z \leq R$, the above takes the form

$$\sqrt{\frac{\left(Z + \frac{2\pi}{k}\right)^2 - \left(\frac{2\pi}{k}\right)^2}{\left(R + \frac{2\pi}{k}\right)^2 - \left(\frac{2\pi}{k}\right)^2}},$$

which is almost linear in Z and which approaches Z/R as $k \rightarrow \infty$. This last expression in the variable Z is the distribution function for an R-excursions along a geodesic with respect to a cusped end of a surface, when $R \leq 2$, [8]. Note also that as $k \rightarrow \infty$, R_k approaches 2.

For approximating excursion, the corresponding distribution in terms of area has the form

$$\frac{\sqrt{\left(Z + \frac{2\pi}{k}\right)^2 - \left(\frac{2\pi}{k}\right)^2}}{2 \log \left(2 \cos \frac{\pi}{k}\right)},$$

for $Z \leq \frac{2\pi}{k}(\cosh(\delta_k) - 1)$. Again as $k \rightarrow \infty$ the distributions limit at $Z/2 \log 2$, the value of the distribution of approximating excursions alone geodesics out a cusped end when $Z \leq 1$.

Clearly, the other cases look rather nasty. Nevertheless, one can verify that they limit at the corresponding distributions on surfaces with cusped ends.

References

[1] R. L. Adler, Geodesic flows, interval maps and symbolic dynamics. In *Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces* (ed. by T. Bedford, H. Keane, C. Series), Oxford Sci. Publ., Oxford University Press, New York 1991, 93–123. Zbl 0752.58007 MR 1130174

[2] R. L. Adler and L. Flatto, Cross-section maps for geodesic flows. I. The modular surface. In *Ergodic Theory and Dynamical Systems II* (ed. by A. Katok), Progr. Math. 21, Birkhäuser, Boston, MA, 1982, 103–161. Zbl 0496.58009 MR 0670077

[3] W. Ambrose, Representation of ergodic flows. *Ann. Math.* (2) **42** (1941), 723–739. Zbl 0025.26901 MR 0004730

[4] A. F. Beardon, *The Geometry of Discrete Groups*. Grad. Texts in Math. 91, Springer-Verlag, New York 1983. Zbl 0528.30001 MR 0698777

[5] W. Bosma, Approximation by mediants. *Math. Comp.* **54** (1990), 421–432. Zbl 0697.10042 MR 0995207

[6] W. Bosma, H. Jager and F. Wiedijk, Some metrical observations on the approximation of continued factions. *Indag. Math.* **45** (1983), 281–299. Zbl 0519.10043 MR 0718069

[7] I. P. Cornfeld, S. V. Fomin and Ya. G. Sinai, *Ergodic Theory*. Grundlehren Math. Wiss. 245, Springer-Verlag, New York 1982. Zbl 0493.28007 MR 0832433

[8] A. Haas, The distribution of geodesic excursions out the end of a hyperbolic orbifold and approximation with respect to a Fuchsian group. *Geom. Dedicata* **116** (2005), 129–155. Zbl 1090.53069 MR 2195445

[9] A. Haas, Diophantine approximation on hyperbolic orbifolds. *Duke Math. J.* **56** (3) (1988), 531–547. Zbl 0653.10022 MR 0948532

[10] P. Nicholls, *The Ergodic Theory of Discrete Groups*. London Math. Soc. Lecture Note Ser. 143, Cambridge University Press, Cambridge 1989. Zbl 0674.58001 MR 1041575

[11] C. Series, The modular surface and continued fractions. *J. London Math. Soc.* (2) **31** (1985), 69–80. Zbl 0545.30001 MR 0810563

Andrew Haas, Department of Mathematics, The University of Connecticut, Storrs, CT. 06269-3009, U.S.A.

E-mail: haas@math.uconn.edu