Annales Academiæ Scientiarum Fennicæ
Series A. I. Mathematica
Volumen 20, 1995, 359–372

A NOTE ON COUNTING CUSPIDAL EXCURSIONS

Bernd Stratmann

Universität Göttingen, Mathematisches Institut, SFB 170
Bunsenstr. 3–5, D-37073 Göttingen, Germany; stratman@cfgauss.uni-math.gwdg.de

Abstract. For geometrically finite Kleinian groups with parabolic elements we study that part of the Lagrange spectrum which does not lie in the Markov spectrum. Using the ergodicity of the associated geodesic flow with respect to the Liouville–Patterson measure, we obtain an estimate for the asymptotic frequency with which recurrent geodesics enter certain cusp regions. In particular, this allows a quantitative description of the logarithmic affinity of geodesic excursions for the cusps.

1. Introduction

We consider geometrically finite $(N+1)$-manifolds M of constant negative curvature with cusps. These are manifolds which admit a representation by a finite sided, convex, fundamental polyhedron Δ, a subset of the universal covering space. The model for the universal covering space in use is the Poincaré ball model (D^{N+1}, d), where d denotes the hyperbolic metric on the $(N+1)$-dimensional unit ball. M can be thought of as being a compact manifold M_o to which a finite number of 'cusps' ('imploding contact with infinity') and possibly a finite number of 'funnels' ('exploding contact with infinity') are attached. The fundamental group G of M is a discrete subgroup of $\mathrm{Con}(N)$, the group of all orientation preserving diffeomorphisms of D^{N+1}, and is usually referred to as a geometrically finite Kleinian group (we exclude the possibility of G being elementary). For the investigation of Kleinian groups a certain parameter $\delta(G)$ has proved to be of central interest. For any positive s form the Dirichlet series

$$\sum_{g\in G} e^{-sd(0,g0)},$$

then $\delta = \delta(G)$ is defined to be the exponent of convergence of this series and is called the *exponent of convergence of* G.

In this paper we study a certain quantitative aspect of the geodesic dynamic on M. Let $\mathscr{R}(\Delta)$ denote the recurrent part of the geodesic flow on M. $\mathscr{R}(\Delta)$ is a subset of the unit tangent bundle $\mathscr{S}(\Delta)$. Each element of $\mathscr{R}(\Delta)$ gives rise to a recurrent geodesic on M, that is a geodesic which returns infinitely often into the compact part M_o of M, and neither starts nor ends in a funnel. For obvious

---

1991 Mathematics Subject Classification: Primary 30F40.

Research supported by the Sonderforschungsbereich 170, University of Göttingen.

geometric reasons the dynamic inside the funnels, that is, geodesics which either start or end in a funnel, is of minor interest here. A consequence of the existence of funnels is that the relevant 'canonical measure' is no longer comparable to the Lebesque measure. In the geometrically finite case the relevant measure turns out to be the Patterson measure, which we denote in the following by μ (for a construction of μ we refer to [11] and [10]). It was shown by Sullivan ([16], [10, Theorem 4.4.4]), that μ is an ergodic probability measure (with respect to the action of G on $L(G)$). The support of μ is the limit set $L(G)$, the set of accumulation points of the G-orbit of a point x in D^{N+1}. The Patterson measure can be viewed as a measure on the fibre $\mathscr{R}_x(\Delta)$ and gives rise to a geodesic flow invariant, ergodic measure $\tilde{\nu}$ on $\mathscr{R}(\Delta)$. The measure $\tilde{\nu}$ will be called the *Liouville–Patterson measure.*

In [15, Proposition 4.9] we worked out measure theoretical results concerning the 'speed' with which geodesic rays gradually occupy the cusp regions of M. In particular we gave an elementary proof of a generalized version of Sullivan's *logarithmic law for geodesics* (LLG). This law considers the 'distance function'

$$\mathscr{R}_o(\Delta) \times \mathbf{R}^+ \ni (\underline{v}, t) \longmapsto N_t(\underline{v}) \in \mathbf{R}^+,$$

where $N_t(\underline{v})$ measures the hyperbolic distance between a suitable chosen compact region M_o in M and the point one reaches after 'travelling' the hyperbolic distance t along the geodesic ray corresponding to $\underline{v}$. The 'logarithmic law for geodesics' then states that, for μ-almost all $\underline{v}$ in the fibre $\mathscr{R}_o(\Delta)$,

$$\text{(LLG)} \qquad \limsup_{t\to\infty} \frac{N_t(\underline{v})}{\log t} = \frac{1}{2\delta - k_{\max}},$$

where $k_{\max}$ denotes the maximal occurring rank for a cusp in M. This result expresses the affinity of the geodesics for the cusp regions; but it does not give any information concerning the 'rate' with which a geodesic eventually makes its way deeper and deeper into the cusps.

In this paper we add to the qualitative results of [15] a quantitative aspect concerning the 'frequency' with which certain cusp regions are visited by $\tilde{\nu}$-almost all recurrent geodesics. In order to illustrate our results, we assume that the compact part M_o is chosen to be sufficiently large. By C we denote a connected component of $M \setminus M_o$ which represents a cusp of M; and for a positive number λ we let $C(\lambda)$ denote that subregion of C which lies at a hyperbolic distance λ from M_o. We then consider the 'counting function'

$$\mathscr{R}(\Delta) \times \mathbf{R}^+ \times (0,1] \ni (\underline{v}, t, \varepsilon) \longmapsto \alpha_t^p(\underline{v}, \varepsilon) \in \mathbf{N},$$

where $\alpha_t^p(\underline{v}, \varepsilon)$ counts the number of visits to $C\big(\log(1/\varepsilon)\big)$ made after starting at the base point of $\underline{v}$ in D^{N+1} and then 'travelling' a hyperbolic distance t along the geodesic in M which is determined by $\underline{v}$. Generalizing results obtained for cofinite Kleinian groups by Moeckel and Nakada ([8], [9]), we derive a *law for the asymptotic frequency of cusp excursions* (LAF) for the general case of geometrically finite Kleinian groups with parabolic elements. This law states the existence of positive constants c and c', depending on the group G, such that, for all sufficiently small, positive ε and for $\tilde{\nu}$-almost all $\underline{v}$ in $\mathscr{R}(\Delta)$,

$$\text{(LAF)} \qquad c\,\varepsilon^{2\delta-k(p)} \le \lim_{t\to\infty} \frac{\alpha_t^p(\underline{v},\varepsilon)}{t} \le c'\,\varepsilon^{2\delta-k(p)};$$

where $k(p)$ denotes the rank of the cusp p under consideration.

An immediate consequence of this result is an estimate for the 'asymptotic expectation' of the number of cusp visits to a particular parabolic cusp with respect to the number of cusp visits to some other parabolic cusp. To be precise, we deduce the existence of positive constants c_o and c_o' such that, for all sufficiently small, positive ε, for all pairs (p,q) of different cusps and for $\tilde{\nu}$-almost all $\underline{v}$ and $\underline{w}$ in $\mathscr{R}(\Delta)$,

$$c_o\,\varepsilon^{k(q)-k(p)} \le \lim_{t\to\infty} \frac{\alpha_t^p(\underline{v},\varepsilon)}{\alpha_t^q(\underline{w},\varepsilon)} \le c_o'\,\varepsilon^{k(q)-k(p)}.$$

Finally, we convert the statement of the law (LAF) into the language of hyperbolic rays which start at the origin in D^{N+1} and terminate in the limit set of the Kleinian group G. We derive for μ-almost all limit points an asymptotic estimate on the number of 'ε-squeezed standard horoballs' which are intersected after 'travelling' a hyperbolic distance t along those rays.

Also, we would like to point out that the results of this paper have a natural interpretation in terms of elementary number theory. For this we recall some well known facts from the theory of metric Diophantine approximation. For a positive number c, an irrational number θ is called c-approximable if the inequality

$$\left|\theta - \frac{p}{q}\right| < \frac{c}{q^2}$$

can be fulfilled for infinitely many reduced fractions p/q, with q positive. Perron ([12]) first noticed that one can associate to each irrational number θ a positive number

$$\varrho(\theta) := \inf_{c>0}\, \{c : \theta \text{ is } c\text{-approximable}\}.$$

For c greater than or equal to the *Hurwitz number* $1/\sqrt{5}$, each irrational number is c-approximable and thus the *Lagrange spectrum*

$$\mathscr{L} := \{\varrho(\theta) :\ \theta \text{ irrational}\}$$

is a subset of the interval $(0, 1/\sqrt{5}\,]$. Markov showed that the intersection of $\mathscr{L}$ with the interval $(1/3, 1/\sqrt{5}\,]$ comprises a set $\mathscr{M}$ of countably many numbers which accumulate only at the value $1/3$ ([7]). The set $\mathscr{M}$ is called the *Markov* *spectrum.* In [3] it was shown in detail how to derive the continued fraction expansion of an irrational number from a coding for geodesics on the modular surface $PSL_2(\mathbf{R})/PSL_2(\mathbf{Z})$ (see also [13], [14]). In particular it turns out that the set $\mathscr{M}$ corresponds to the set of simple (not self-intersecting) geodesic loops on that surface. The results in this paper thus allow statements concerning the complement of $\mathscr{M}$ in $\mathscr{L}$.

If λ denotes the 1-dimensional Lebesgue measure and if we use the usual notation for the regular continued fraction expansion of an element θ in the unit interval, namely

$$\theta = \cfrac{1}{\theta_1 + \cfrac{1}{\theta_2 + \cdots}} = [\theta_1, \theta_2, \ldots],$$

then the results in this paper give rise to the following number theoretical fact.

For all N in $\mathbf{N}$ and for λ-almost all $\theta = [\theta_1, \theta_2, \ldots]$ in the unit interval we have

$$\lim_{n\to\infty} \frac{\operatorname{card}\{m : 1 \le m \le n \text{ and } \theta_m \ge N\}}{n + \sum_{i=1}^{n} \log \theta_i} = k_o \cdot \log\Big(1 + \frac{1}{N}\Big);$$

where

$$k_o := \left(\log 2 + \sum_{n=1}^{\infty} \log n \cdot \log\Big(1 + \frac{1}{n(n+2)}\Big)\right)^{-1} = (\log 5.2\ldots)^{-1}.$$

We remark that this result may be derived also from an application of the ergodic theorem to the continued fraction map (see also [6]).

The author would like to thank the Sonderforschungsbereich 170 at the University of Göttingen for its support and hospitality.

2. Preliminaries

In this section we recall some concepts of the theory of Kleinian groups and their associated geodesic dynamic. These concepts should provide the reader with the necessary background for an understanding of the following section. As already mentioned in the introduction, $G < \mathrm{Con}(N)$ always denotes a non-elementary, geometrically finite Kleinian group which contains parabolic elements; by Δ we denote a convex, finite sided fundamental polyhedron for the discontinuous action of G on D^{N+1}. If $H\big(L(G)\big)$ denotes the convex hull of the limit set $L(G)$, we may then assume without loss of generality, that the origin 0 in D^{N+1} is an element of $\Delta \cap H\big(L(G)\big)$. Our list of required concepts begins with a brief resumé on the subject of parabolic fixed points.

Parabolic cusps. Let P denote a complete set of inequivalent parabolic fixed points of G. It is well known that P is a finite set of points in S^N. We assume that each element of P lies at the boundary of Δ. If P is chosen in this fashion, then the elements of P are called *basic parabolic fixed points*. To each basic parabolic fixed point p we associate a horoball H_p of radius $r(p)$, that is a Euclidian N-ball in D^{N+1} of radius $r(p)$ which 'touches' S^N at the point p. For p in P we let G_p denote the stabilizer of p in G. Then it is a well known fact that the radii r_p of the 'basic horoballs' H_p can be chosen such that the set $\bigcup_{p\in P}\{g(H_p) : g \in G/G_p\}$ comprises a set of pairwise disjoint horoballs ([2, p. 248]). If in particular the radii $r(p)$ are chosen to be maximal with respect to this disjointness property (which we shall assume from now on), then the G-orbit of these horoballs represents the so called *standard set of horoballs* and this set will be denoted by $\{H_{g(p)}(r_g) : g \in \mathscr{T}_p, p \in P\}$, where $\mathscr{T}_p$, the so called *top representation* of G_p in G, denotes a geometrically chosen set of coset representatives of G/G_p (for more details on this the reader is referred to [15, Chapter 2]). For a positive ε less than or equal to 1 and for p in P, we define $H^p_\varepsilon \subseteq H_p$ to be the horoball at p of radius $\varepsilon\, r(p)$.

A consequence of the Bieberbach theorem is that, for each p in P, the stabilizer G_p contains a subgroup G^*_p of finite index such that G^*_p is isomorphic to $\mathbf{Z}^{k(p)}$ for some $k(p)$ in $\mathbf{N}$. The number $k(p)$ is called the *rank* of p. It then follows that there exists a 'fundamental domain' Q^p_∞ for the action of G^*_p on $S^N \setminus \{p\}$ (Q^p_∞ is the image under the Cayley transformation of the 'fundamental parallelopiped' introduced in [15]). By this we mean an N-dimensional, closed subset of S^N, whose spherical diameter $q_\infty(p)$ is chosen to be minimal with respect to the following properties:

1. $L(G) \setminus \{p\} \subset G^*_p(Q^p_\infty)$;
2. for all distinct g and h in G^*_p,
$$g\big(\operatorname{int}(Q^p_\infty)\big) \cap \big(\operatorname{int}(Q^p_\infty)\big) = \emptyset;$$
3. Q^p_∞ lies 'opposite' p, i.e., every geodesic with initial point in Q^p_∞ and end point p has non-trivial intersection with $\Delta \cap \partial H_p$.

We let $Q^p_\infty \times \{p\}$ denote the set of geodesics which have their initial point in Q^p_∞ and which end in p. Further, for positive ε, we define the subset Q^p_ε of ∂H^p_ε by
$$Q^p_\varepsilon := (Q^p_\infty \times \{p\}) \cap \partial H^p_\varepsilon.$$
The hyperbolic diameter of Q^p_ε will be denoted by $q_\varepsilon(p)$.

Geodesic flows. If $\mathscr{S}(D^{N+1})$ denotes the unit tangent bundle over D^{N+1}, then $\mathscr{S}(D^{N+1})$ is the union of fibres $\mathscr{S}_z(D^{N+1})$, where the union is taken with respect to all base points z in D^{N+1}. The canonical projection $\operatorname{pr} : \mathscr{S}(D^{N+1}) \to D^{N+1}$ maps an element of $\mathscr{S}(D^{N+1})$ to the base point of the fibre of which it is an element.

On $\mathscr{S}(D^{N+1})$ we introduce the metric d^*, which is induced by an additive combination of the hyperbolic metric d and the metric θ, where θ is defined as follows. Consider in $\mathscr{S}(D^{N+1})$ two elements $\underline{v}$ and $\underline{w}$ such that $\operatorname{pr}(\underline{v})$ and $\operatorname{pr}(\underline{w})$

are close. Move $\underline{v}$ from $\mathrm{pr}(\underline{v})$ to $\mathrm{pr}(\underline{w})$ by parallel displacement, which leads to an element $\underline{v}'$. Then, $\theta(\underline{v},\underline{w})$ is defined to be the angle (at most π) between $\underline{v}'$ and $\underline{w}$.

As usual, an element $\underline{v}$ in $\mathscr{S}(D^{N+1})$ is parameterized by a triple (v^-,v^+,s) in $\big((S^N\times S^N)\setminus\{\mathrm{diag}\,.\}\big)\times\mathbf{R}$, where v^-, respectively v^+, denotes the initial respectively end point of the unique geodesic $\gamma(\underline{v})$ determined by $\underline{v}$, and s is the 'signed' hyperbolic distance between the summit of the geodesic $\gamma(\underline{v})$ (i.e., that point on $\gamma(\underline{v})$ which lies at maximal Euclidian distance from the boundary S^N) and the base point $\mathrm{pr}(\underline{v})$; 'signed' means that s is positive if and only if $\mathrm{pr}(\underline{v})$ is an element of that half of $\gamma(\underline{v})$ which lies between v^+ and the summit.

Let Φ_t denote the geodesic flow on $\mathscr{S}(D^{N+1})$. We recall that Φ_t is defined for t in $\mathbf{R}$ and $\underline{v}=(v^-,v^+,s)$ in $\mathscr{S}(D^{N+1})$, by

$$\Phi_t(v^-,v^+,s)=(v^-,v^+,s+t).$$

On $\mathscr{S}(D^{N+1})$ we introduce an equivalence relation '$\sim$' as follows. Two elements $\underline{v}$ and $\underline{w}$ of $\mathscr{S}(D^{N+1})$ are equivalent if and only if there exists an element g in G such that $\gamma(\underline{v})=g\big(\gamma(\underline{w})\big)$ and $\mathrm{pr}(\underline{v})=g\big(\mathrm{pr}(\underline{w})\big)$ are satisfied. The unit tangent bundle over Δ is then defined by

$$\mathscr{S}(\Delta):=\mathscr{S}(D^{N+1})/\sim .$$

Factoring out this equivalence relation, we see that the flow Φ_t on $\mathscr{S}(D^{N+1})$ induces a flow ϕ_t on $\mathscr{S}(\Delta)$, and ϕ_t is called the geodesic flow on $\mathscr{S}(\Delta)$.

In this paper we are mainly interested in a certain subset $\mathscr{R}(\Delta)$ of $\mathscr{S}(\Delta)$. Roughly speaking, the elements of $\mathscr{R}(\Delta)$ give rise to geodesics on the manifold D^{N+1}/G which are recurrent in either their future or their past, but neither start nor end in a funnel. To be precise, $\mathscr{R}(\Delta)$ is defined as follows:

$$\mathscr{R}(\Delta):=\big\{\underline{v}\in\mathscr{S}(\Delta):v^-,v^+\in L(G)\text{ and }\mathrm{pr}(\underline{v})\in\Delta\big\}.$$

If $\mathscr{L}(G)$ denotes the set of loxodromic fixed points of G, then it is known ([1, Theorem 5.3.8]) that $(\mathscr{L}(G)\times\mathscr{L}(G))\setminus\{\mathrm{diag}\,.\}$ is dense in $\big(L(G)\times L(G)\big)\setminus\{\mathrm{diag}\,.\}$. Since each element of $\big(\mathscr{L}(G)\times\mathscr{L}(G)\big)\setminus\{\mathrm{diag}\,.\}$ gives rise to a closed geodesic on the manifold D^{N+1}/G, it follows that the set $\mathscr{R}(\Delta)$ admits an interpretation as the closure of the set of all closed geodesics on D^{N+1}/G.

Local cross sections. In order to define special subsets of $\mathscr{R}(\Delta)$, we recall the notion of a local cross section with respect to some fixed time. For this let ϱ denote some positive number. A closed subset $\mathscr{C}$ of $\mathscr{S}(\Delta)$ which lies transversal to the flow ϕ_t, is called a *local cross section for ϕ_t with respect to the time ϱ* if one has for each element $\underline{v}$ in $\mathscr{C}$,

$$\bigcup_{|t|<\varrho}\phi_t(\underline{v})\cap\mathscr{C}=\underline{v}.$$

In the following section a central part is played by certain subsets of $\mathscr{R}(\Delta)$. The elements of these sets are based on the boundary of 'ε-squeezed' standard horoballs and point inward into these horoballs. To be precise, for positive ε less than or equal to 1 and for p in P, we define the following cross section which is based on $\Delta \cap \partial H^p_\varepsilon$,

$$\mathscr{C}^p(\varepsilon) := \big\{ \underline{v} = (v^-, v^+, s) \in \mathscr{R}(\Delta) : v^- \in Q^p_\infty, \mathrm{pr}(\underline{v}) \in Q^p_\varepsilon \text{ and } s < 0 \big\}.$$

From the construction it is evident that there exists a positive constant $\tau = \tau(G)$, such that, for each p in P and for each positive ε less than or equal to 1, the section $\mathscr{C}^p(\varepsilon)$ is a local cross section for ϕ_t with respect to τ.

Canonical measures. As we have already said in the introduction, we denote by μ the Patterson measure. For geometrically finite Kleinian groups it is well known that μ is an ergodic and 'δ-conformal' probability measure which is supported on the limit set $L(G)$ and which has no atomic part ([10], [16]). By *δ-conformal* we mean that for each Borel subset E of S^N and for each g in G,

$$\mu\big(g(E)\big) = \int_E P\big(g^{-1}(0), \xi\big)^{\delta(G)} \, d\mu(\xi);$$

where the *Poisson kernel* P is defined as usual for z in D^{N+1} and ξ in S^N by $P(z,\xi) := (1 - |z|^2) \cdot |z - \xi|^{-2}$.

For g in $\mathrm{Con}(N)$ and x, y in $D^{N+1} \cup S^N$, an elementary calculation shows that

$$|g(x) - g(y)|^2 = |g'(x)| \cdot |g'(y)| \cdot |x - y|^2.$$

Using this estimate we obtain a $(G \times G)$-invariant measure ν on $\big(L(G) \times L(G)\big) \setminus \{\mathrm{diag}.\}$ by

$$d\nu(\xi, \eta) := \frac{d\mu(\xi)\, d\mu(\eta)}{|\xi - \eta|^{2\delta}}.$$

Using the above mentioned parametrization of $\mathscr{S}(\Delta)$, it is now possible to define on $\mathscr{S}(\Delta)$ a measure $\tilde{\nu}$ by

$$d\tilde{\nu}(\xi, \eta, s) := d\nu(\xi, \eta) \cdot ds.$$

The measure $\tilde{\nu}$ is called the *Liouville–Patterson measure*. For the geometrically finite case it was shown by Sullivan ([16]) that $\tilde{\nu}$ is a finite and ϕ_t-invariant measure, and further that the geodesic flow is ergodic with respect to $\tilde{\nu}$. Because of its finiteness, we can assume in the following without loss of generality that $\tilde{\nu}$ is actually a probability measure on $\mathscr{R}(\Delta)$.

Geometry of horoballs. For ξ in S^N let s_ξ denote the hyperbolic ray in D^{N+1} with initial point the origin and with end point ξ. For positive t, we denote by ξ_t the point on s_ξ which lies at a hyperbolic distance t from the origin.

Further, we let s_{ξ_t} be the geodesic segment on s_ξ which lies between ξ_t and the origin. The 'shadow projection' Π is a map from D^{N+1} onto S^N which is defined for subsets A of D^{N+1} by

$$\Pi(A) := \{\xi \in S^N : s_\xi \cap A \neq \emptyset\}.$$

For the moment let us fix a horoball $H(1)$ in D^{N+1} with diameter 1. Thus in particular 0 is an element of $H(1)$. For positive ε less than or equal to 1 let $H(\varepsilon) \subseteq H(1)$ denote the horoball of radius ε and with the same point of tangency at S^N as $H(1)$. Let $\eta(\varepsilon)$ in S^N be chosen such that $s_{\eta(\varepsilon)}$ is tangent to $H(\varepsilon)$. If further $z(\varepsilon)$ denotes the unique point in D^{N+1} at which $s_{\eta(\varepsilon)}$ intersects $\partial H(1) \setminus \{0\}$, then we derive the following 'horoball excursion formula'.

Lemma 1. *There exist positive constants c_1 and c_2 such that, for each positive ε less than or equal to 1, we have*

$$\text{(HEF1)} \qquad c_1\, \varepsilon^2 \le 1 - |z(\varepsilon)| \le c_2\, \varepsilon^2.$$

Proof. Let $z^*(\varepsilon)$ be the point of tangency of $s_{\eta(\varepsilon)}$ at $H(\varepsilon)$. From the theorem of Pythagoras we deduce that

$$(1-\varepsilon)^2 = |z^*(\varepsilon)|^2 + \varepsilon^2,$$

and thus

$$2\varepsilon = 1 - |z^*(\varepsilon)|^2.$$

On the other hand, it is well known ([1]) that there exist universal, positive constants c and c' such that

$$c\,(1 - |z^*(\varepsilon)|) \le e^{-d(0,z^*(\varepsilon))} \le c' \,\bigl(1 - |z^*(\varepsilon)|\bigr),$$

and thus

$$\tfrac{1}{2}c\,\bigl(1 - |z^*(\varepsilon)|^2\bigr) \le e^{-d(0,z^*(\varepsilon))} \le c'\,\bigl(1 - |z^*(\varepsilon)|^2\bigr).$$

Combining these two estimates, it follows that

$$c\,\varepsilon \le e^{-d(0,z^*(\varepsilon))} \le 2c'\,\varepsilon.$$

Since $d\bigl(0, z(\varepsilon)\bigr) = 2d\bigl(0, z^*(\varepsilon)\bigr)$, we have that

$$c^2\,\varepsilon^2 \le e^{-d(0,z(\varepsilon))} \le 4c'^2\,\varepsilon^2,$$

and hence

$$\frac{c^2}{c'}\,\varepsilon^2 \le 1 - |z(\varepsilon)| \le \frac{4c'^2}{c}\,\varepsilon^2. \ \square$$

For a fixed, positive number q we now partition $\partial H(1)$ into spherical annuli of hyperbolic width q; i.e., we define annuli inductively as follows:

$$A_1 := \{z \in \partial H(1) : d(0,z) \le q\};$$

and for n in $\mathbf{N}$ greater than 1, let

$$A_n := \Bigl\{z \in \partial H(1) \setminus \bigcup_{k=1}^{n-1} A_k : 0 < d(z, A_{n-1}) \le q\Bigr\}.$$

The following 'horoball excursion formula' is easily obtained from Lemma 1 and we shall therefore omit the proof.

Lemma 2. *There exist positive constants c_3 and c_4, depending on q, such that, for each n in $\mathbf{N}$ and z in A_n, we have*

$$\text{(HEF2)} \qquad \frac{c_3}{n^2} \le 1 - |z| \le \frac{c_4}{n^2}.$$

3. Asymptotic frequencies

In this section we shall prove our main result. We begin by giving an estimate for the Liouville–Patterson measure of the local cross sections with respect to the time τ which were introduced in the preceding section. The estimate is required in the proof of our theorem.

Proposition 1. *There exist positive constants k_1, k_2 and ε_o such that, for each p in P and for each positive ε less than ε_o, we have*

$$k_1\, \varepsilon^{2\delta - k(p)} \le \nu\big(\mathscr{C}^p(\varepsilon)\big) \le k_2\, \varepsilon^{2\delta - k(p)}.$$

Proof. Let p in P be given. For convenience, we assume that the origin 0 in D^{N+1} is an element of $H_p \cap H\big(L(G)\big)$ (for the kind of arguments we will give in the following, this simplifying assumption changes only the constants which occur by a factor which depends on the hyperbolic distance between 0 and ∂H_p). It is known that $\big(\partial H_p \cap H\big(L(G)\big)\big)/G_p^*$ is a compact subset of D^{N+1} ([17], [15, Chapter 2]). Using this fact it follows that we can choose a 'horospherical fundamental domain' Q_1^p on ∂H_p, that is a fundamental domain for the action of G_p^* on H_p whose hyperbolic diameter $q_1(p)$ is minimal with respect to this property. We can further assume that Q_1^p is chosen such that it contains the origin. Now, we define horospherical annuli on ∂H_p of constant hyperbolic width $q_1(p)$ as follows.

Let

$$D_1^p := \big\{z \in \partial H_p : d(0, z) \le q_1(p)\big\};$$

and for m in $\mathbf{N}$ greater than 1, let

$$D_m^p := \Big\{z \in \partial H_p \setminus \bigcup_{k=1}^{m-1} D_k^p : 0 < d(z, D_{m-1}^p) \le q_1(p)\Big\}.$$

An elementary Euclidean volume argument then implies the existence of positive constants c_1 and c_2 such that

$$\text{(1)} \qquad c_1\, n^{k(p)-1} \le \operatorname{card}\{g \in G_p^* : g(0) \in D_n^p\} \le c_2\, n^{k(p)-1}.$$

Let g be in G_p^* such that $g(0)$ is contained in D_n^p for some n in $\mathbf{N}$. Using the formula (HEF 2), derived in the preceding section, one obtains the existence of positive constants c_3 and c_4, depending on $q_1(p)$, such that

$$c_3\, n^{-2} \le 1 - |g(0)| \le c_4\, n^{-2}.$$

Combining this estimate with the fact that $|\xi - g^{-1}(0)|$ is bounded from above and below for all ξ in Q^p_∞, it follows that

$$\mu\big(g(Q^p_\infty)\big) = \int_{Q^p_\infty} P(g^{-1}(0),\xi)^\delta \, d\mu(\xi) \asymp \mu(Q^p_\infty)\cdot(1-|g(0)|^2)^\delta \asymp n^{-2\delta}; \tag{2}$$

Here and in the following, the sign '$\asymp$' is used to denote that the quotient of the two related quantities is bounded from below and above.

Combining (1) and (2), we obtain the existence of positive constants c_5 and c_6 such that

$$c_5\, n^{k(p)-1-2\delta} \le \mu\Bigg(\bigcup_{\substack{g\in G_p^*\\ g(0)\in D_n^p}} g(Q^p_\infty)\Bigg) \le c_6\, n^{k(p)-1-2\delta}. \tag{3}$$

Let ε be positive and less than or equal to 1. As mentioned in the preceding section, a cross section of the geodesic flow ϕ_t admits a representation as a subset of $S^N \times S^N$. In particular, there exists a uniquely determined subset R^p_ε of S^N such that $\mathscr{C}^p(\varepsilon)$ is represented in this fashion by $Q^p_\infty \times R^p_\varepsilon$. From the construction it then follows that there exists a positive constant ϱ_o, depending on $q_1(p)$, such that

$$\Pi\big(H_p(\varrho_o\varepsilon)\big) \subset R^p_\varepsilon \subset \Pi\big(H_p(\varrho_o^{-1}\varepsilon)\big). \tag{4}$$

In order to compute the μ measure of R^p_ε, let elements ξ^+ and ξ^- of S^N be chosen such that $s_{\xi^\pm}$ is tangential to $H_p(\varrho_o^{\pm1}\varepsilon)$. An application of the formula (HEF 1) gives that the hyperbolic length of the segment $s_{\xi^\pm}\setminus\{0\}\cap H_p$ is equal to $\log(\varrho_o^{\pm1}\varepsilon)^{-2}$ (apart from an additive constant). This implies that there exist positive numbers $\kappa^+(\varepsilon)$ and $\kappa^-(\varepsilon)$ such that

$$\kappa^\pm(\varepsilon) \asymp \varepsilon^{-1} \tag{5}$$

and

$$s_{\xi^\pm}\setminus\{0\}\cap\partial H_p \in D^p_{\kappa^\pm(\varepsilon)}. \tag{6}$$

Using (3), (5) and (6), we deduce

$$\begin{aligned}\mu\big(\Pi\big(H_p(\varrho_o^{\pm1}\varepsilon)\big)\big) &\asymp \sum_{m=\kappa^\pm(\varepsilon)}^{\infty} \sum_{\substack{g\in G_p^*\\ g(0)\in D_m^p}} g(Q^p_\infty)\\ &\asymp \sum_{m=\kappa^\pm(\varepsilon)}^{\infty} m^{k(p)-1-2\delta} \asymp \big(\kappa^\pm(\varepsilon)\big)^{k(p)-2\delta} \asymp \varepsilon^{2\delta-k(p)}.\end{aligned}$$

Using (4), this implies the existence of positive constants c_7 and c_8 such that

$$c_7\,\varepsilon^{2\delta-k(p)} \le \mu(R_\varepsilon^p) \le c_8\,\varepsilon^{2\delta-k(p)}. \tag{7}$$

Let ε chosen to be sufficiently small; i.e., let ε be positive and less than $\varepsilon_o(p)$, where the positive constant $\varepsilon_o(p)$, depending on $q_1(p)$, is chosen such that

$$R^p_{\varepsilon_o(p)} \cap Q^p_\infty = \emptyset.$$

Now, there exists a positive constant c_9, depending on $\varepsilon_o(p)$ and thus on $q_1(p)$, such that, for all (ξ,η) in $Q^p_\infty \times R^p_\varepsilon$,

$$c_9 < |\xi-\eta| \le \pi. \tag{8}$$

Using (7) and (8), we deduce, for ε less than $\varepsilon_o(p)$,

$$\nu\big(\mathscr{C}^p(\varepsilon)\big) = \int_{Q^p_\infty}\int_{R^p_\varepsilon} \frac{d\mu(\xi)\,d\mu(\eta)}{|\xi-\eta|^{2\delta}} \asymp \mu(Q^p_\infty)\cdot\mu(R^p_\varepsilon) \asymp \varepsilon^{2\delta-k(p)}.$$

If we define $\varepsilon_o := \min\{\varepsilon_o(p) : p \in P\}$, and adjust the constants in order to get rid of the simplifying assumption that 0 is an element of ∂H_p, the proposition follows. □

We may now turn to the main result of this paper. For this, we recall the definition of the 'counting function' under consideration. For $\underline{v}$ in $\mathscr{R}(\Delta)$, p in P, positive ε less than ε_o and for positive t, we define

$$\alpha^p_t(\underline{v},\varepsilon) := \operatorname{card}\big\{s \in [0,t] : \phi_s(\underline{v}) \in \mathscr{C}^p(\varepsilon)\big\}.$$

This function counts, while one is 'travelling' the hyperbolic distance t, the number of visits of the geodesic ray determined by $\underline{v}$ into the particular cusp region which is given by H^p_ε. For this function we have the following result.

Theorem. *For all p in P, for all positive ε less than ε_o and for $\tilde\nu$-almost all $\underline{v}$ in $\mathscr{R}(\Delta)$, we have*

$$k_1\,\varepsilon^{2\delta-k(p)} \le \lim_{t\to\infty} \frac{\alpha^p_t(\underline{v},\varepsilon)}{t} \le k_2\,\varepsilon^{2\delta-k(p)} \tag{LAF}$$

(k_1, k_2 and ε_o are the constants of Proposition 1*).*

Proof. Let p be in P and let ε be positive and less than ε_o. We have seen in the preceding section that $\mathscr{C}^p(\varepsilon)$ is a local cross section for the flow ϕ_t with respect to the time τ. Here, τ is some fixed, positive number which depends on G. We may define the 'τ-flow box' $\tilde{\mathscr{C}}^p(\varepsilon)$ by

$$\tilde{\mathscr{C}}^p(\varepsilon) := \bigcup_{0<u<\tau} g_{-u}\big(\mathscr{C}^p(\varepsilon)\big).$$

Let $\chi_{(\cdot)}$ denote the characteristic function. For $\underline{v}$ in $\mathscr{R}(\Delta)$ and positive t, we have

$$(9) \qquad \tau\cdot\big(\alpha_t^p(\underline{v},\varepsilon)-1\big)\le\int_0^t \chi_{\tilde{\mathscr{C}}^p(\varepsilon)}\big(\phi_s(\underline{v})\big)\,ds\le\tau\cdot\big(\alpha_t^p(\underline{v},\varepsilon)+1\big).$$

Using the ergodicity of the measure $\tilde{\nu}$ and applying Hopf's generalization of the Birkhoff ergodic theorem ([4], [5]), it follows, for $\tilde{\nu}$-almost all $\underline{v}$ in $\mathscr{R}(\Delta)$, that

$$(10) \qquad \lim_{t\to\infty}\frac{1}{t}\int_0^t \chi_{\tilde{\mathscr{C}}^p(\varepsilon)}\big(\phi_s(\underline{v})\big)\,ds=\tilde{\nu}\big(\tilde{\mathscr{C}}^p(\varepsilon)\big).$$

The fact that $\mathscr{C}^p(\varepsilon)$ is a local cross section with respect to τ implies that

$$(11) \qquad \tilde{\nu}\big(\tilde{\mathscr{C}}^p(\varepsilon)\big)=\tau\cdot\nu\big(\mathscr{C}^p(\varepsilon)\big).$$

Combining (9), (10) and (11), we derive, for $\tilde{\nu}$-almost all $\underline{v}$ in $\mathscr{R}(\Delta)$,

$$\lim_{t\to\infty}\frac{\alpha_t^p(\underline{v},\varepsilon)}{t}=\nu\big(\mathscr{C}^p(\varepsilon)\big).$$

Using Proposition 1, it now follows that

$$k_1\,\varepsilon^{2\delta-k(p)}\le\lim_{t\to\infty}\frac{\alpha_t^p(\underline{v},\varepsilon)}{t}\le k_2\,\varepsilon^{2\delta-k(p)}. \ \square$$

An immediate consequence of the preceding theorem is a statement concerning the relative asymptotic frequency of cusps visits. For this we define, for p and q in P, $\underline{v}$ and $\underline{w}$ in $\mathscr{R}(\Delta)$ and positive ε less than ε_o, the *relative asymptotic frequency* $R_\varepsilon^{p,q}(\underline{v},\underline{w})$ of the p-cusp visits of the geodesic corresponding to $\underline{v}$ with respect to the q-cusp visits of the geodesic corresponding to $\underline{w}$, by

$$R_\varepsilon^{p,q}(\underline{v},\underline{w}):=\lim_{t\to\infty}\frac{\alpha_t^p(\underline{v},\varepsilon)}{\alpha_t^q(\underline{w},\varepsilon)}.$$

Corollary. *There exist positive constants k_3 and k_4 such that, for each p and q in P, for each positive ε less than ε_o and for $\tilde{\nu}$-almost all $\underline{v}$ and $\underline{w}$ in $\mathscr{R}(\Delta)$, we have*

$$k_3\,\varepsilon^{k(q)-k(p)}\le R_\varepsilon^{p,q}(\underline{v},\underline{w})\le k_4\,\varepsilon^{k(q)-k(p)}.$$

This corollary reflects the way in which the geodesic dynamic is effected by the interplay of cusps of different ranks. For example, let us consider a geometrically finite 3-manifold with two cusps p and q such that $k(p)=2$ and $k(q)=1$. It is intuitively clear that the proportion of the limit set in the shadow of a standard horoball associated to p is in a certain sense greater than the proportion of the limit set in the shadow of an equally sized standard horoball associated to q. In this particular case, the corollary states that for a recurrent geodesic the 'asymptotic expectation' of visiting the cusp at p is roughly ε^{-1} times higher than the 'asymptotic expectation' of visiting the cusp at q.

Finally, we convert the statement of the Theorem into a statement concerning the asymptotic frequencies with which geodesic rays emanating from the origin in D^{N+1} intersect 'ε-squeezed' standard horoballs. For this we consider the 'counting function' $\beta_t^p(\xi,\varepsilon)$, an analogue of $\alpha_t^p(\underline{v},\varepsilon)$. The function $\beta_t^p(\xi,\varepsilon)$ is defined for ξ in $L(G)$, p in P, positive ε less than ε_o and positive t, by

$$\beta_t^p(\xi,\varepsilon) := \operatorname{card}\bigl\{g \in \mathscr{T}_p : s_{\xi_u} \cap H_{g(p)}(\varepsilon r_g) \neq \emptyset \text{ for some } u \in (0,t]\bigr\}.$$

We obtain the following result.

Proposition 2. *For each p in P, for all positive ε less than ε_o and for μ-almost all ξ in $L(G)$, we have*

$$k_1\,\varepsilon^{2\delta-k(p)} \le \lim_{t\to\infty} \frac{\beta_t^p(\underline{v},\varepsilon)}{t} \le k_2\,\varepsilon^{2\delta-k(p)}$$

(k_1, k_2 and ε_o are the constants of Proposition 1).

Proof. Consider the set of leaves of the strong stable foliation on $\mathscr{S}(\Delta)$ whose images under the projection pr contain the origin in Δ; i.e., consider the set

$$W_o^{ss} := \bigcup_{\underline{v}\in\mathscr{S}_o(\Delta)} W^{ss}(\underline{v}),$$

where $W^{ss}(\underline{v})$ is defined for $\underline{v}$ in $\mathscr{S}(\Delta)$ by

$$W^{ss}(\underline{v}) := \bigl\{\underline{w} \in \mathscr{S}(\Delta) : \lim_{t\to\infty} d^*\bigl(\phi_t(\underline{v}),\phi_t(\underline{w})\bigr) = 0\bigr\}.$$

It is clear that if (LAF) is satisfied for some $\underline{v}$ in $\mathscr{S}_o(\Delta)$, then (LAF) is also satisfied for each $\underline{w}$ in $W^{ss}(\underline{v})$. On the other hand, if some $\underline{w}$ in $\mathscr{S}(\Delta)$ satisfies (LAF), then (LAF) also holds for $\phi_t(\underline{w})$ for all t in $\mathbf{R}$. Thus we have

– if (LAF) holds for $\underline{v}$ in $\mathscr{S}_o(\Delta)$, (LAF) holds for each $\underline{w}$ in $\bigcup_{t\in\mathbf{R}}\phi_t\bigl(W^{ss}(\underline{v})\bigr)$.

The proposition now follows from our theorem when we recall that 0 is assumed to be an element of $H\bigl(L(G)\bigr)$ and that the following inclusion is satisfied:

$$\mathscr{R}(\Delta) \subseteq \bigcup_{t\in\mathbf{R}} \phi_t(W_o^{ss}). \ \square$$

References

[1] Beardon, A.F.: The Geometry of Discrete Groups. - Springer-Verlag, New York, 1983.

[2] Greenberg, L.: Finiteness theorems for Fuchsian and Kleinian groups. - In: Discrete Groups and Automorphic Functions, edited by W.J. Harvey, Academic Press, 1977, 199–255.

[3] Haas, A.: Diophantine approximation on hyperbolic Riemann surfaces. - Acta Math. 156, 1986, 33–82.

[4] Hopf, E.: Statistik der geodätischen Linien in Mannigfaltigkeiten negativer Krümmung. - Sitzungsber. Sächs. Akad. Wiss. Leipzig Math.-Natur. Kl. 91, 1939, 261–304.

[5] Hopf, E.: Ergodic theory and the geodesic flow on surfaces of constant negative curvature. - Bull. Amer. Math. Soc. 77:6, 1971, 863–877.

[6] Khintchine, A.: Kettenbrüche. - B.G. Teubner Verlagsgesellschaft, Leipzig, 1956.

[7] Markov, A.A.: Sur les formes quadratique binaires indefinies. - Math. Ann. 17, 1880, 379–400.

[8] Moeckel, R.: Geodesics on modular surfaces and continued fractions. - Ergodic Theory Dynamical Systems 2, 1982, 131–154.

[9] Nakada, H.: On metrical theory of diophantine approximation over imaginary quadratic field. - Acta Arith. 51, 1988, 393–403.

[10] Nicholls, P.J.: The ergodic theory of discrete groups. - London Math. Soc. Lecture Note Ser. 143, 1989.

[11] Patterson, S.J.: The limit set of a Fuchsian group. - Acta Math. 136, 1976, 241–273.

[12] Perron, O.: Über die Approximation irrationaler Zahlen durch rationale. - Sitzungsber. Heidelb. Akad. Wiss. Math.-Natur. Kl. Abh. 4, 8, 1921.

[13] Series, C.: The modular surface and continued fractions. - J. London Math. Soc. (2) 31, 1985, 69–80.

[14] Series, C.: The geometry of Markoff numbers. - Math. Intelligencer 7, 1985, 20–29.

[15] Stratmann, B., and S. Velani: The Patterson measure for geometrically finite groups with parabolic elements, new and old. - Proc. London Math. Soc. 1994 (to appear).

[16] Sullivan, D.: The density at infinity of a discrete group. - Inst. Hautes Études Sci. Publ. Math. 50, 1979, 171–202.

[17] Tukia, P.: On isomorphisms of geometrically finite Möbius groups. - Inst. Hautes Études Sci. Publ. Math. 61, 1985, 171–214.

Received 15 March 1994